Supporting File Sharing Through Improved Awareness

by

Tara Whalen

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

at

Dalhousie University
Halifax, Nova Scotia
August 2008

© Copyright by Tara Whalen, 2008

DROPBOX 1017

Library and Archives Canada

Published Heritage Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et Archives Canada

Direction du Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-43953-1
Our file Notre référence
ISBN: 978-0-494-43953-1

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

DALHOUSIE UNIVERSITY

To comply with the Canadian Privacy Act the National Library of Canada has requested that the following pages be removed from this copy of the thesis:

Preliminary Pages
- Examiners Signature Page (pii)
- Dalhousie Library Copyright Agreement (piii)

Appendices
- Copyright Releases (if applicable)

Dedication

To DGC

Table of Contents

List of Tables ..... **xiv**

List of Figures ..... **xix**

Abstract ..... **xxi**

List of Abbreviations and Symbols Used ..... **xxii**

Acknowledgements ..... **xxiii**

Chapter 1. Introduction **1**

1.1 What is File Sharing? 2

1.2 Research Approach 3

1.3 Document Overview 4

Chapter 2. Research Design **6**

2.1 Overview: Research Phases 6

2.2 Phase One: Characterizing File Sharing Practice 6

2.3 Phase Two: Development of a Theoretical Framework for File Sharing Awareness 7

2.4 Phase Three: Applying the Framework: Designing Tools for Managing Shared Files 8

2.5 Phase Four: Refining the Design Elements 8

2.6 Phase Five: Evaluating the Shared File Management Tools 9

2.7 Summary: Research Design 10

Chapter 3. File Sharing Surveys **12**

3.1 Study One: Fundamentals of File Sharing 12

3.1.1 Summary and Objective 12

3.1.2 Methods 12

3.1.3 Instrument 13

3.1.4 Participants 13

3.1.5 Data Analysis 14

3.1.6 Results 14

3.1.6.1 Types of Files Shared 14

3.1.6.2 Methods of Sharing 14

3.1.6.3 Most Frequently-Used Methods 15

3.1.6.4 Sensitive Files 16

3.1.6.5 Managing File Permissions 17

3.1.6.6 Inconveniences Arising from Access Control 18
3.1.6.7 Discontinuing Sharing: Removal and Reconfiguration 18
3.1.7 Final Comments 19
3.1.8 Summary of Key Findings 20
3.1.8.1 Complexity of File Sharing 20
3.1.8.2 User Needs: Security and Information 20
3.2 Study Two: Further Investigations into File Sharing 21
3.2.1 Overview of Second Research Study 22
3.2.2 Research Objectives 22
3.2.3 Methods 23
3.2.4 Instrument 23
3.2.5 Respondents 24
3.2.6 Data Analysis 24
3.2.7 Results 25
3.2.7.1 File Sharing 25
3.2.7.2 Sharing Groups 26
3.2.7.3 Sensitive Documents 27
3.2.7.4 Managing Limits on Shared Files 28
3.2.7.5 File Sharing Knowledge 28
3.2.8 Discussion 29
3.2.8.1 Key Findings 29
3.2.9 Comparison to Prior Studies 30
3.2.9.1 Comparison to Previous Survey (Study #1) 30
3.2.9.2 Comparison to Findings of Voida et al. 31
3.3 Summary 33
Chapter 4. Related Work 35
4.1 File Sharing Research 35
4.1.1 Characterizing File Sharing 35
4.1.2 Usability Studies on File Sharing 37
4.1.2.1 Music-Sharing Application Usability Study (KaZaa) 37
4.1.2.2 Ad Hoc Collaboration Usability Study (Impromptu) 37
4.2 Computer-Supported Collaboration 39
4.2.1 Collaboration Theory 39

4.2.2 Collaboration Systems 41
4.3 Usable Security and Access Control 42
4.3.1 Usable Security 43
4.3.2 Access Control 44
4.3.3 Summary: Usable Security and Access Control 45
4.4 Personal and Group Information Management 45
4.5 Summary: Current Findings and Future Directions 48
Chapter 5. File Sharing Awareness Framework 50
5.1 Definitions of Awareness 50
5.2 Neisser's Perception-Action Cycle 51
5.3 Situation Awareness 52
5.4 Workspace Awareness Framework 53
5.4.1 Types of Workspace Awareness Information 54
5.4.2 Mechanisms Used to Gather Workspace Awareness Information 55
5.4.3 Uses of Workspace Awareness in Collaboration 55
5.4.4 Summary of WA Framework 55
5.5 Asynchronous Workspace Awareness 57
5.6 Summary of Awareness Definitions and Frameworks 61
5.7 File Sharing Awareness Framework 62
5.7.1 Information about Shared Files 63
5.7.2 Elements of File-Sharing Awareness 65
5.7.2.1 Elements Relating to the Present 66
5.7.2.2 Elements Relating to the Past 67
5.7.2.3 Elements Relating to the Future 68
5.7.2.4 Summary: Awareness Information 69
5.7.3 Gathering Perceptual Information for File Sharing 70
5.8 Summary 72
Chapter 6. Designing for File Sharing Awareness 73
6.1 Selecting a Tool for Presenting File Sharing Information 73
6.2 Characteristics of a File Manager 73
6.3 Information Displays in File Managers 74
6.4 File Sharing Information and Features to Include in File Manager 77
6.5 Adapting a File Manager Design: Windows Explorer for XP 79

6.5.1 File Listings: Icons for File Sharing Information 80
6.5.2 File Manager Popups 81
6.5.3 Supplementary Window: Sharing Console 83
6.5.3.1 Query Feature 83
6.5.3.2 Detailed File Information: History and Permissions 83
6.5.3.3 Alert Feature 84
6.6 Synopsis of Preliminary Design and Unresolved Design Issues 86
6.7 File Sharing Icons and Terminology 87
6.7.1 File Sharing Icons 87
6.7.2 File Sharing Terminology 88
6.8 Summary: Refining the Design for Future Evaluation 89
Chapter 7. Validating and Refining the Preliminary Design 91
7.1 Study One: Group Design Activity 91
7.1.1 Research Objectives 91
7.1.2 Study Overview 92
7.1.3 Instruments 93
7.1.3.1 Instrument 1: Group Card Sort 93
7.1.3.1.1 Card Sorting: An Overview 94
7.1.3.1.2 Card Sorting: File Sharing Icons 94
7.1.3.2 Instrument 2: Labelled Icon Cards 95
7.1.3.3 Instrument 3: Computer-Based Questionnaire 98
7.1.3.3.1 Questionnaire Preamble: Demographic Information 98
7.1.3.3.2 Questionnaire Section One: Icon Critique 98
7.1.3.3.3 Questionnaire Section Two: Push-Oriented and Pull-Oriented Terminology 99
7.1.3.3.4 Questionnaire Section Three: File Manager Design Questions 100
7.1.3.3.5 Questionnaire Section Four: Sharing Console Design Questions 102
7.1.3.4 Instrument 4: Group Discussion 102
7.1.4 Study Environment 102
7.1.5 Participants 103
7.1.6 Procedure 105
7.1.7 Data Analysis 109
7.1.7.1 Group Card Sort 109

7.1.7.2 Icon-Labelling 109
7.1.7.3 Questionnaire 109
7.1.7.4 Group Discussion 110
7.1.8 Results 110
7.1.8.1 Results: Group Card Sort 110
7.1.8.2 Results: Icon Labels 112
7.1.8.3 Results: Questionnaire 114
7.1.8.3.1 Icon Critique 114
7.1.8.3.2 Push-Oriented and Pull-Oriented Terminology 115
7.1.8.3.3 File Manager Questions (Main Listing Window) 117
7.1.8.3.4 File Manager Icon Preferences 117
7.1.8.3.5 Sharing Console Design Questions 118
7.1.8.4 Results: Group Discussion 118
7.1.8.5 Discussion 119
7.1.8.5.1 Group Card Sort 119
7.1.8.5.2 Icon Labelling 119
7.1.8.5.3 Questionnaire 121
7.1.8.5.4 Group Discussion 124
7.2 Study Two: Consensus for Group Icon 124
7.2.1 Research Objective 124
7.2.2 Instrument: Questionnaire 124
7.2.3 Participants 125
7.2.4 Results 126
7.2.5 Discussion 126
7.3 Study Three: File Sharing Terminology Survey 126
7.3.1 Research Objective 126
7.3.2 Instrument: Online Survey 127
7.3.3 Participants 128
7.3.4 Data Analysis 128
7.3.5 Results 128
7.3.6 Discussion 130
7.4 Summary of Results 130
7.5 Resulting Designs for File Sharing Tools: Prototypes 131

7.5.1 File Manager Prototype 131
7.5.2 Sharing Console Prototype 133
7.5.3 Using the File Manager and Sharing Console 134
7.6 Summary 135
Chapter 8. Evaluation Part One: Design and Methodology 136
8.1 Study Overview 136
8.2 Research Design and Hypotheses 137
8.2.1 Hypotheses and Research Questions 137
8.2.2 Experimental Design 138
8.2.2.1 Use of Multiple Experiment Sessions 139
8.2.2.2 Task Ordering 140
8.3 Simulating the Workplace: Scenario and Experimental Environment 142
8.3.1 Scenario 142
8.3.2 Experimental Environment 144
8.4 Digital Workspace: Email and File Collections 147
8.4.1 Real-World Email: The Enron Dataset 147
8.4.2 Creating a Realistic File Environment 147
8.4.2.1 Workspace File Study 148
8.4.2.2 Windows Desktop files 150
8.4.2.3 Workplace Document Folders 151
8.4.3 Implementation of the File Collection: Email and Documents 153
8.4.3.1 Creating the Mailbox 153
8.4.3.2 Workplace Documents 154
8.5 Tasks 155
8.5.1 Mapping of Tasks to Awareness Framework 155
8.5.2 Task Types 158
8.5.2.1 Tasks Completed with Both File Managers 158
8.5.2.2 Tasks Completed Only with the Experimental File Manager 160
8.5.2.3 Addendum: Tasks Completed with Microsoft Outlook 161
8.6 Variables 162
8.6.1 Independent Variables 162
8.6.2 Dependent Variables (Measures) 162
8.6.2.1 Task Completion Time 162

8.6.2.2 Correctness of Question Responses 163
8.6.2.3 Usability Scores 163
8.6.2.4 User Preference 165
8.6.3 Additional Measures: Experimental File Manager Only 166
8.6.3.1 Evaluation of New Features and Functionality 166
8.6.3.2 Intention to Use 169
8.6.3.2.1 Technology Acceptance Model Questionnaire 169
8.6.3.2.2 Follow-Up Email 170
8.7 Participants 170
8.7.1 Eligibility 170
8.7.2 Recruitment 171
8.7.3 Sample Size 171
8.7.4 Participant Characteristics: Demographics and Experience 172
8.8 Procedures 172
8.8.1 Session One 172
8.8.2 Session Two 175
8.9 Summary 179
Chapter 9. Evaluation Part Two: Results 180
9.1 Overview of the Presentation of Results 180
9.2 Comparison between File Managers 181
9.2.1 Task Completion Times 182
9.2.1.1 Task Completion Times: Data Analysis 182
9.2.1.2 Task Completion Times: Results 182
9.2.2 Correctness of Question Responses 182
9.2.2.1 Question Responses: Data Analysis 183
9.2.2.2 Question Responses: Results 183
9.2.3 Usability Scores 185
9.2.3.1 Usability Scores: Data Analysis 186
9.2.3.2 Usability Scores: Results 186
9.2.4 Preference Data 188
9.2.4.1 Preference Data: Data Analysis 189
9.2.4.2 Preference Data: Results 189
9.3 Evaluation of New Features and Functionality 192

9.3.1 Task Completion Times: Experimental File Manager-Only Tasks 192
9.3.1.1 Results 192
9.3.2 Correctness of Question Responses: Experimental File Manager-Only Tasks 193
9.3.2.1 Results 193
9.3.3 Event Notification: Alert Message 194
9.3.3.1 Results 194
9.3.4 Subjective Ratings: Information Display and Ease of Learning 195
9.3.4.1 Results 195
9.3.5 Suggestions for Improvement 196
9.3.5.1 Results 196
9.3.5.1.1 File Manager Columns 196
9.3.5.1.2 Search Feature (Console) 197
9.3.5.1.3 File Details (Console) 197
9.3.5.1.4 Alert 197
9.3.5.1.5 Additional features 197
9.4 Intention to Use 197
9.4.1 Technology Acceptance Model 198
9.4.1.1 Technology Acceptance Model Results 198
9.4.2 Follow-up Email 199
9.4.2.1 Follow-up Email Results 199
9.5 Discussion 200
9.5.1 Comparison Between Experimental and Control File Managers 200
9.5.1.1 Task Completion Times 200
9.5.1.2 Question Response Scores 202
9.5.1.3 Usability Scores 203
9.5.1.4 User Preference 204
9.5.1.5 Summary of Comparative Results 206
9.5.2 Evaluation of New Features and Functionality of Experimental File Manager 207
9.5.2.1 Task Completion Times: Experimental File Manager-Only Tasks 207
9.5.2.2 Question Response Scores: Experimental File Manager-Only Tasks 207
9.5.2.3 Evaluation of the Notification Feature 208

9.5.2.4 Subjective Ratings: Information Display and Ease of Learning 209
9.5.2.5 Suggestions for Improvement 210
9.5.3 Intention to Use Experimental File Manager 210
9.6 Summary 211
Chapter 10. Conclusion 213
10.1 Research Contributions 213
10.1.1 Theoretical Contributions 213
10.1.1.1 Characterizing File Sharing Practice 213
10.1.1.2 Creation of a File Sharing Awareness Framework 213
10.1.1.3 Creation of Iconography and Terminology for File Sharing Concepts 214
10.1.2 Applied Contributions 214
10.1.2.1 Design of File Manager to Support File Sharing Awareness 214
10.1.2.2 Demonstration of the Value of Awareness for File Sharing 214
10.2 Research Limitations 215
10.2.1 Limitations of File Sharing Surveys 215
10.2.2 Limitations of Laboratory Evaluation 215
10.2.3 Limitations of Prototype Functionality 216
10.2.4 Limitations of Integration 216
10.3 Future Work 216
10.3.1 Addressing Limitations: Extending the File Manager 217
10.3.2 Integrating the File Manager with Search Tools 218
10.3.3 Additional Sharing Information for the File Manager 218
10.3.4 Tailoring the File Manager for Specialized Domains 218
10.4 Conclusion 219
References 221
Appendix A: Questions, Survey One 231
Appendix B: Questions, Survey Two 247
Appendix C: Group Design Activity: Script, Handouts, and Questionnaires 263
Appendix D: Terminology Survey 290
Appendix E: Specification for Sharing Console 298
Appendix F: Materials from Laboratory Study 309

List of Tables

Table 1: The five major research phases (and the preliminary phase), including research objectives and descriptions of the results 11

Table 2: Types of materials shared 14

Table 3: File sharing methods for work and personal files 15

Table 4: Most frequently-used sharing methods. The top three choices for each category are marked (e.g., #1 indicates the top choice). 16

Table 5: Types of sensitive digital information 16

Table 6: Methods of protecting sensitive data 17

Table 7: Reasons for setting file permissions 17

Table 8: Problems resulting from access control configuration and management 18

Table 9: Reasons for resetting file permissions to terminate sharing 19

Table 10: Most frequently-selected methods for file sharing 25

Table 11: The most commonly-used file sharing methods 25

Table 12: Advantages and disadvantages of various file sharing methods 26

Table 13: Types of sensitive files that people use and share 27

Table 14: Reasons for resetting file permissions 28

Table 15: Awareness of shared files and activity on shared files 29

Table 16: Similarities between the two file sharing surveys 31

Table 17: Differences between the two file sharing surveys 31

Table 18: Similarities between findings of Voida et al. (2006) and the second file sharing survey 32

Table 19: Differences between findings of Voida et al. (2006) and the second file sharing survey 32

Table 20: Elements of WA knowledge relating to the present, from Gutwin and Greenberg (2002, 2004) 58

Table 21: Elements of WA knowledge relating to the past, from Gutwin and Greenberg (2002, 2004) 58

Table 22: Elements of asynchronous workspace awareness, from Tam and Greenberg (2006) 59

Table 23: File-sharing awareness elements relating to the present. Adapted from Tam and Greenberg (2006). 64

Table 24: File-sharing awareness elements relating to the past. Adapted from Tam and Greenberg (2006). 65

Table 25: File-sharing awareness elements relating to the future 69

Table 26: Relationship between exploration and information gathering, from file s haring awareness framework. These elements were presented earlier in Figure 5 70

Table 27: Terms that can be used to describe pull-oriented and push-oriented file activities 89

Table 28: File sharing icons used in group card sorting activity 95

Table 29: Set of 30 terms used to label icons 97

Table 30: Terms that participants could choose from to describe pull-oriented and push-oriented file activities 99

Table 31: Demographics for design activity participants (n=18) 105

Table 32: Categorization of icons, including labels, created by each participant group. Note that the elements in each category are listed as numbers (rather than graphically as icons); consult the key in the top row for the mapping between numbers and icons. 111

Table 33: Overall clusters of icons, based on participant groupings. The sets in Column One consist of icons that were grouped together (at least once) with at least one other icon in that set. Items in different rows were never categorized together by any participant groups. 112

Table 34: Counts of how often a label was applied to each icon. The first set of rows (first 19 white rows) are intended to represent file-activity concepts, while the last set (11 grey rows) are intended to represent user-oriented concepts. 113

Table 35: New labels created by participants 114

Table 36: How participants described icons, according to the checklist. Numbers in each row represent the total number of times a descriptor was chosen for that particular icon. Columns that are shaded grey indicate negative descriptors. The bold cells indicate the descriptor with the highest count for each icon; there are some ties. 115

Table 37: Top choices for push-oriented and pull-oriented terms. Additional suggestions also listed; any that appeared twice are marked as (x2). The *negative count* refers to the number of times that a term was marked as "I would never use this word", each occurrence of which led to a point deduction. 116

Table 38: Top choices for icons for the file manager, with counts and percentages for each (n=18). 118

Table 39: Icons that participants interpreted in at least three ways: as *users*, *group*, or *groups*. A checkmark appears under any icon which was labelled with that term (e.g. *users*) at least once; the actual count is listed beside each checkmark. 121

Table 40: Results of followup group icon study (for paper version, n=12; for online version, n=15). 126

Table 41: Sets of terms used in followup survey 127

Table 42: Highest and lowest scoring terms for pull-oriented and push-oriented activities. The top three items for each category are listed, along with the three lowest-ranked term. The *negative count* refers to the number of times that a term was marked as "I would never use this word," each occurrence of which led to a one-point deduction. The highest possible total score was 325. 129

Table 43: Summary of icons and terms resulting from user study, included in the final design. Note that the colour of the "accessed" icon was changed to better match the rest of the icon set. 131

Table 44: Task ordering for Session One and Session Two. (The tasks are described in Section 8.5 below). Odd-numbered participants (grey rows) started with the control file manager, while even-numbered participants (white rows) started with the experimental manager. 141

Table 45: Summary of workplace file organization. Counts are shown for the number of files and folders in the various categories, grouped into ranges. 150

Table 46: Structure and number of workplace documents. Each subfolder (e.g., "drafts") contained only files (no subfolders). This structure was used in Session One; for Session Two, another 20 files were added (15 to "chris's docs" and five to "Sharing Folder"). 152

Table 47: Information about files and folders used in the workspace environment for testing; numbers of files varied for each of the two sessions. 152

Table 48: Mailbox folder sizes, including numbers of email attachments, for both testing sessions. 154

Table 49: The questions from the Session One tasks, mapped to questions from the awareness framework 156

Table 50: The questions from the Session Two tasks, mapped to questions from the awareness framework 157

Table 51: Paired tasks performed in laboratory study. Each row contains a pair of equivalent tasks (such as finding a list of email recipients), such as A and F, o r J and P 159

Table 52: Tasks performed only in the Helium file manager condition for both sessions 161

Table 53:Usability questionnaire used to generate usability scores: CSUQ and ASQ... 164

Table 54: User preference questions, between control "Carbon" and experimental "Helium" file managers, for Session One. 166

Table 55: User preference questions, between control "Carbon" and experimental "Helium" file managers, for Session Two. 167

Table 56: Evaluation of information display and ease of learning, for experimental file manager. Note that the odd-numbered questions are reverse-worded. 168

Table 57: Information about study participants 173

Table 58: Mapping of the testing variables to usability measures, showing how each was assessed in evaluation. 180

Table 59: Results of task completion times; mean times in seconds are presented with standard deviation. Mann-Whitney U test results are shown; single asterisk indicates p<.05; double asterisk indicates p<.01 183

Table 60: Results for question response scores for both control and experimental file managers. The mean score is shown with the maximum possible score for that question, and the standard deviation. Independent samples t-test results are shown; single asterisk indicates p<.05; double asterisk indicates p<.01. (n=28) When both results have an SD of zero, the t-test cannot be run; these items are marked as SD=0. A description of each question (with its ID) is provided for reference. 184

Table 61: CSUQ and ASQ scores for the control and experimental file managers, for each session. The mean score and standard deviation are listed for each factor. Independent samples t-test results are shown; single asterisk indicates p<.05; double asterisk indicates p<.01. (n=28) 187

Table 62: CSUQ and ASQ scores, compared for each file manager on the basis of order effect Independent samples t-test results are shown; single asterisk indicates p<.05 and double asterisk indicates p<.01. (n=28) 187

Table 63: Descriptive statistics and chi-square result for preference data on tasks (Session One). "Exp." refers to the experimental file manager. A single asterisk indicates p<.05; double asterisk indicates p<.01 189

Table 64: Preference data for forced-choice questions (Session One), expressed as descriptive statistics and chi-square results. Descriptive results shown only for experimental files manager (the control represents the remainder of the total). A single asterisk indicates p<.05; double asterisk indicates p<.01 190

Table 65: Descriptive statistics and chi-square result for preference data on tasks (Session Two). A single asterisk indicates p<.05; double asterisk indicates p<.01. 190

Table 66: Preference data for forced-choice questions (Session Two), for either control or experimental file manager, presented as descriptive statistics and chi-square results. A single asterisk indicates p<.05; double asterisk indicates p<.01 191

Table 67: Task completion times (in seconds) for experimental file manager-only tasks, presented as means. Standard deviation is included for each measure. 193

Table 68: Results for question correctness (experimental file manager-only tasks). Response results presented in categories (correct, partially correct and incorrect), with counts for each category; percentages are based on n=28. 194

Table 69: Responses to questions about the file sharing information and the ease of learning to use the experimental file manager. The responses are means, based on a 7-point Likert scale (where "1" means "strongly agree"). For the first three statements, a lower score is better; for the last four statements, a higher score is better (because of the negative wording). 196

Table 70: Responses to TAM questionnaire items. U1-U6 and E1-E6 are from Davis (1989); X1 is from Laitenberger and Dreyer (1998). Each item used a 7-point Likert scale. Mean (M) and standard deviation (SD) are listed in the third column. 198

Table 71: Cronbach's alpha measure of consistency, for the twelve items in the TAM questionnaire. Values of α above 0.7 are considered to have high inter-item correlation. 199

Table 72: Categories of positive attributes of the experimental file manager, derived from post-study emails about real world use. 200

Table 73: Summary of hypothesis results, for comparisons between experimental and control file managers 206

List of Figures

Figure 1: Screenshot of Impromptu, a system for ad hoc collaboration (Rode et al., 2006, p. 146) 38

Figure 2: Using Recent Shortcuts to compose an email with an attachment (from Tang et. al, 2007) 47

Figure 3: Neisser's perception-action cycle, as described by Neisser (1976) and Gutwin and Greenberg (2002) 51

Figure 4: Workspace awareness (WA) model, adapted from Gutwin and Greenberg (2002). 54

Figure 5: Awareness framework for file sharing 63

Figure 6: Details view in Windows Explorer for XP, showing six columns of file information 75

Figure 7: Example of pop-up (infotip) in Windows Explorer for XP 76

Figure 8: Supplementary window in Windows Explorer for XP, displaying the file properties information 77

Figure 9: Early design sketches for file manager, showing a variety of possible icons for file events and permissions. 81

Figure 10: Paper prototype of the file manager, illustrating the basic design. Note the pop-up that provides details about the most recent "sent" event. 82

Figure 11: Preliminary design for Sharing Console 85

Figure 12: Notification feature: editing window (left) and notification window (right) .. 86

Figure 13: File sharing icons used in group card sorting activity 88

Figure 14: Example of a card used in the card sorting task 95

Figure 15: Example of a labelled icon card. 96

Figure 16: Sample question for analysing icons 99

Figure 17: The design activity environment, in the Groupwork Lab. 104

Figure 18: Video stills from group design activity, showing interaction between group members 107

Figure 19: Questionnaire for group icon followup study 125

Figure 20: The file manager prototype developed for testing, based on *Jeff's File Manager.* 132

Figure 21: Sample XML entry for file history and settings information, for the file "chapter1.pdf" 133

Figure 22: Sharing console prototype developed for testing 133

Figure 23: Excerpt from sample XML file, showing queries and alerts 134

Figure 24: Right-click (context) menu for launching sharing console from file manager 135
Figure 25: The experimental environment, designed to replicate an office setting 145
Figure 26: Control room, used for observation and for data collection. 146
Figure 27: Questions about file organization 149

Abstract

File sharing is an integral part of computer-based collaboration that requires careful management of access, to permit teamwork while protecting confidentiality. Existing file managers provide limited functionality to support user awareness regarding shared files, leading to unchecked errors and difficulties in keeping track of shared items. The question is: how can we provide better support for awareness, to assist with file sharing?

A review of the research literature revealed that little information existed about sharing practices. Two surveys were conducted to characterize file sharing. The surveys found that people need information about file sharing activity and to keep track of dynamic file settings, and confirmed that privacy and security are important. From past work, it is evident that users are unaware of many activities that occur on their shared files, which can hinder both security and collaboration.

Providing file sharing information is a two-fold challenge: 1) what information needs to be presented about shared files? 2) how can that information be presented in a clear and understandable way? These questions were addressed through a file sharing awareness framework, adapted from an existing awareness model, which provided the theoretical foundation for design. It identified elements (*feedthrough*, *individualized views*, and *notifications)* essential for improving awareness of shared files. These elements were instantiated in a design for a file manager that displays file sharing information. Using a series of design studies, a set of icons and terms was established to represent file-sharing concepts, which were integrated in the final design.

The file-sharing manager was compared with a control file manager (Windows Explorer-like) in a within-subjects design laboratory experiment. Participants (n=28) completed tasks, which involved answering questions about shared files, using both the control and the file-sharing managers. The file-sharing manager led to greater efficiency and effectiveness for a number of tasks, and provided new functionality that allowed users to discover information about shared files that was previously unavailable to them. Participants were more satisfied with the file-sharing manager, and demonstrated likely intention-to-use. These research findings suggest that presenting clear, readily-accessible information about shared files is a beneficial approach for supporting shared file management.

List of Abbreviations and Symbols Used

ACL	Access Control List
ASQ	After-Scenario Questionnaire
CSCW	computer-supported collaborative work
CSUQ	Computer System Usability Questionnaire
GIM	group information management
HCI	human-computer interaction
INFOQUAL	information quality
INTERQUAL	interface quality
M	mean
P2P	peer-to-peer
PIM	personal information management
SA	situation awareness
SD	standard deviation
SNARF	Social Network and Relationship Finder
SYSUSE	system usefulness
TAM	Technology Acceptance Model
USB	Universal Serial Bus
WA	workspace awareness
WebDAV	Web-based Distributed Authoring and Versioning
XML	eXtensible Markup Language

Acknowledgements

First and foremost, I must thank my co-supervisors: Dr. Elaine Toms and Dr. Jamie Blustein. I am profoundly grateful for your guidance, particularly as you came to the research in mid-stream. Your expertise and support were vital in bringing this thesis to completion.

I also thankfully acknowledge the assistance of Dr. Carolyn Watters, who served as a keen and thoughtful reader. Your feedback on this research provided innumerable improvements.

Thank you as well to Professor Ann Blandford (University College London) who agreed to serve as my external reviewer.

I would like to thank Dr. Kori Inkpen (FCS) for guidance and support in the first part of my doctoral research.

Thank you to all the past and present members of the EDGElab and the Centre for Management Informatics who assisted with this research in countless large and small ways, from brainstorming research ideas to helping with study logistics:

EDGElab members and affiliates:

Ritchie Argue, Tim Burrell, Jeff Cressman, Dave Dearman, Vicki Ha, Dr. Kirstie Hawkey, Dr. Regan Mandryk, Bonnie MacKay, J. Karen Parker, Derek Reilly, Malcolm Rodgers, Dr. Stacey Scott, Greg Smolyn, Michael Nunes, Jim Wallace, and Ryder Ziola

CMI members:

Margie Clow-Bohan, Emilie Dawe, Chris Jordan, Tayze MacKenzie, Alexandra MacNutt, Lori McCay-Peet, Jodi McLaughlin, Heather O'Brien, and Sandra Toze

A special mention must be made to CMI research assistants Alexandra MacNutt, Tayze MacKenzie, and Emilie Dawe, who provided key support during two of my user studies: many thanks for ensuring these studies ran smoothly.

Thanks to Dr. Diana Smetters (PARC) and Dr. Elizabeth Churchill (Yahoo) for their role in the first file sharing survey I conducted at the beginning of this thesis work. I benefitted tremendously from your expertise and assistance.

Thanks to Vicki Ha (University of British Columbia) and Christopher Dean (NSCAD University) for their assistance in developing file sharing icon designs.

Thanks to Dr. Holly Etchegary (IWK Health Centre Interdisciplinary Research) and Vaneeta Grover (Department of Mathematics and Statistics) for statistical advice.

Thank you to Dana Echtner (FCS Tech Support) for her assistance with the Enron email dataset.

Thanks to Steve Voida (University of Calgary and Georgia Tech) for providing detailed information about his file sharing research study.

Thanks to my pseudonymous contract programmer for implementing my prototype requirements.

Thank you to those who provided administrative support for my graduate work: Angie Boliver and Menen Teferra (FCS); Janet Lord (Faculty of Management); and FCS Graduate Coordinators Srinivas Sampalli and Malcolm Heywood.

Thanks to Dr. Dirk Balfanz (Google) for coordinating my PARC summer internship, and Alia Wilcox (PARC summer intern) for her assistance with my research project.

Thanks to Dr. Richard J. Nowakowski and Paul Ottaway (Department of Mathematics and Statistics) for combinatorial assistance.

Thanks to John Bullock (UCIS) for answering questions about NTFS file permissions.

Thanks also to other researchers who provided advice and suggestions at various stages of this work: Dr. Maryanne Fisher (Saint Mary's University), Dr. Carrie Gates (Dalhousie and CAlabs), Dr. John McHugh (FCS), and Dr. W. Keith Edwards (Georgia Tech).

Thanks to Doug Cuff for expert editorial assistance.

To all my research study participants: thanks so much for volunteering your time to assist me with this work.

I also gratefully acknowledge the financial support of the Natural Sciences and Engineering Council (NSERC), the Network for Effective Collaboration Technologies through Advanced Research (NECTAR, an NSERC Research Network), the Killam Trusts, the Faculties of Graduate Studies and Computer Science, the Dalhousie Student Union, and the Computer Science Society.

Chapter 1. Introduction

File sharing is an integral part of computer-based collaboration, and is a common task within organizations. It is an important way for groups to distribute information: files are exchanged among employees for information, for modification, and for further distribution to others both within and outside the organization. The files may have differing levels of confidentiality and security requirements, with strict limits concerning with whom the content may be shared and to what extent it may be modified. At the same time, when working collaboratively, people need access to the objects and resources that are required to complete their joint tasks. If a shared file is made inaccessible to a colleague, this becomes an obstacle to effective teamwork. Thus sharing requires careful management of access: users want to ensure that people have access to specific shared items, but they may need to limit access by other people and to other files.

File sharing can lead to security problems, as illustrated by recent incidents in which people have unwittingly revealed private data. For example, one study of a peer-to-peer file sharing system showed that some people had inadvertently shared the contents of their entire hard drive, when they intended only to share their music files (Good & Krekelberg, 2003). The U.S. Senate Judiciary Committee experienced a scandal when it was found that Republican members had taken advantage of errors on a shared server to gain access to private Democratic files for more than a year (Savage, 2004). Microsoft leaked confidential corporate documents when they were accidentally included on the same server as the publicly-downloadable patches and drivers (McWilliams, 2002).

In addition to causing security problems, the mismanagement of shared files can also hinder collaboration. When performing collaborative tasks, people need access to the objects and resources that are required to complete their tasks. There is evidence that people forget which items they have shared, and with whom, which interferes with effective collaboration (Voida, Edwards, Newman, Grinter, & Ducheneaut, 2006). A shared file that is inaccessible is an obstacle to effective teamwork. Bellotti theorized that collaboration necessitates control over access to shareable data, and that when users engage in file sharing, they often forget what their settings are and unwittingly make items too public or too private (Bellotti, 1996).

The complexities of managing shared files have been identified as a research challenge within the domain of group information management (GIM). This field deals "with how personal information is shared amongst a group, with an emphasis on the norms that underlie that sharing, and the ways in which participants negotiate those norms in response to a variety of tensions" (Erickson, 2006, pp. 74-75). Researchers in this field have outlined many difficult challenges, given the complexities of managing shared information, and identified a need for better understanding of how groups share digital information (Jones & Bruce, 2005). Hence file sharing, which is one aspect of group information management, is a research domain that needs to be better understood.

These examples illustrate how advances in file sharing support can lead to improved security and collaboration, as well as more effective information management. The question is: *how* can we provide better support for file sharing? This question forms the basis for this thesis research, which approached this problem from a user-centered design perspective.

1.1 What is File Sharing?

Before detailing the specifics of the research studies, it is helpful to define the term *file sharing*, a concept that is at the core of this thesis work. A *file* is a named collection of data, organized as a single unit, and located on a data-storage medium (such as a disk). The focus of this work is on a file as an object, with specific attributes (e.g., size, type, permissions); the information contained within the file is not relevant to this research. File sharing means to make specified file(s) available to specified individuals or groups, with the option of granting specific rights (e.g., read, write) over those files; this task requires some interaction with a set of files, or the file system. An example would be to upload a file into a shared folder, and giving all group members the right to read and write to that file. This task is dynamic as rights and recipients change over time.

This definition also includes the mode in which a file is shared, such as emailing files through attachments, or placing them on public web servers. In all these cases, an action is taken to grant access to a file, such as providing a copy to a person who did not have that file, or modifying permissions so that an additional person could read a file in a shared folder. The definition also allows for flexible scope on sharing: the "specified group" could be the public, or it could be two colleagues.

This definition excludes such actions as simply telling someone where to find a file, without placing the file there oneself; for example, emailing someone the URL for a public file on the Web. Although this information is necessary for a person to locate a specific file, it does not include any interaction with the file itself, or the file system, and therefore is excluded from this definition of file sharing. There is a distinction between the sharing of a file, and the sharing of the information within that file (although the two concepts are closely related). File sharing, which is one specific type of information sharing, forms the focus of this project. The concept of information sharing is much broader, and may include exchanging paper copies of documents or conducting an oral discussion of a report. The research findings of this project may have application beyond the scope of file sharing, into the broader realm of information sharing, but the scope here is limited to digital files.

1.2 Research Approach

Olsen, in an article on interaction and information management, singled out file sharing as an example of a task requiring better support for collaboration, stating that, "It is important that collaborative mechanisms such as...file sharing become as important, visible, and convenient as files and folders are in the current environment" (Olsen, 1999). Providing this type of support requires a more complete understanding of file sharing management, which can inform designs that support users' needs in this domain. At the outset of this research, little information was available about file sharing, and the problem space was not well-defined. Therefore, the initial phase of this thesis research consisted of two exploratory studies on file sharing, to be used as a foundation for further investigation. These studies gathered fundamental data on user sharing behaviour and file-sharing concerns.

Once the file sharing data was collected and analysed, this provided a more clearly-defined research space, which directed a literature review of related domains. The survey results, along with the review of the literature, indicated that *awareness* support was integral to better shared file management: that is, developing designs that help users to be aware of vital file sharing information.

The application of awareness to file sharing was informed in part by my earlier research in usable security, which explored how security information could be clearly presented to users (Whalen & Inkpen, 2004, 2005).

Adopting research from computer-supported collaborative work (CSCW) led to a file sharing awareness framework, which was used to create a prototype file manager. This prototype design was refined through user studies, with a particular emphasis on how to best communicate file sharing concepts through icons and text. The resulting prototype was then formally evaluated through a laboratory study, to determine how well it helped users to manage shared files. This study demonstrated that the new file manager helped users to be more aware of the status of their shared files, and the presentation of information allowed them to complete file sharing tasks more efficiently, more effectively, and with greater satisfaction.

1.3 Document Overview

Chapter 2 specifies the research design, which introduces all the phases of the research and shows how each contributed to the overall body of work. Chapter 3 details the surveys used in the exploratory research phase for characterizing file sharing practices.

Chapter 4 provides an overview of the literature relevant to the research presented in this thesis, which includes the domains of CSCW, usable security, and information management. The placement of this chapter is unusual, but is due to the requirement for exploratory studies (the surveys in Chapter 3) to be completed prior to the literature review, because the research works relevant to this domain could not be identified without a clearer understanding of the file sharing problem space.

Chapter 5 presents the file sharing awareness framework, which was derived from the preliminary explorations of file sharing as well as the research literature. Chapter 6 introduces the file manager design, which was created by applying the awareness framework. Chapter 7 describes a series of studies that were used to refine the initial file manager design; in particular, these studies focused on identifying appropriate terminology and iconography for file sharing concepts. Chapter 8 details the evaluation of the refined file manager prototype in a laboratory setting, with a focus on the experimental design. Chapter 9 presents the results, including both subjective and

objective measures of the effectiveness of the file manager design. Chapter 10 concludes the thesis, with a summary of the contributions and some directions for future work.

Chapter 2. Research Design

This chapter outlines the overall research design, which took place in several phases, each of which contributed to the thesis results. In order to demonstrate how these parts worked together, the entire research design is presented, starting from the exploratory stage (characterizing user practice) and concluding with the formal evaluation of the shared file management tools.

2.1 Overview: Research Phases

The thesis research consisted of five major phases; each is listed below and is described in detail in the following sections:

- Phase One: Characterizing File Sharing Practices
- Phase Two: Development of a Theoretical Framework for File Sharing Awareness
- Phase Three: Applying the Framework: Designing Tools for Managing Shared Files
- Phase Four: Refining the Design Elements
- Phase Five: Evaluating the Shared File Management Tools

2.2 Phase One: Characterizing File Sharing Practice

At the outset of this research, a preliminary review of the research literature revealed that there was little information about the specifics of file sharing. Therefore, the first phase of this research was the collection of basic information around sharing practices, which could be used to refine further exploration, and to determine what design choices would be most appropriate for users, based on the practices and concerns that they communicated. This initial research phase was in two parts, each consisting of online surveys. The first was a preliminary exploration that was very broad in scope. It was conducted within a research institute, and explored such factors as the types of files shared, the methods used to share them, and the ways in which users managed limits around their shared files.

The initial results of this survey revealed elements that needed to be further explored, and a second survey was developed to collect additional data (as well as to validate the initial findings by replicating questions from the first survey). Some of the new aspects explored in the second study were determining why certain file sharing

methods were more prevalent than others, and identifying the frequency of such tasks as file sharing and file permission management.

The results of these studies allowed for a broad characterization of file sharing behaviour, which could be used to inform design as well as to direct more refined explorations of the research domain. This led to the next stage of the research, which was to review the literature in related domains (including security and collaboration), to see how these fields might be applied to file sharing concerns.

2.3 Phase Two: Development of a Theoretical Framework for File Sharing Awareness

During this phase of the research, the preliminary survey findings were considered in conjunction with the research literature, to determine how to provide better support for shared file management. First, a review of related research domains was conducted, which incorporated usable security, collaboration, and information management. A common thread among these areas was the provisioning and presenting of information, so that users know what is happening in a system. Within the domain of computer-supported cooperative work, this is considered to be part of *awareness* research.

A focused review of the awareness research (within CSCW) unearthed theoretical frameworks that could be applied to file sharing awareness. The basic framework that was adopted in this thesis research was the Workspace Awareness framework of Gutwin and Greenberg (2002). More specifically, a later version of this framework, developed by Tam and Greenberg (2006) for asynchronous awareness, formed the core of the file sharing awareness framework. There are two key elements of this framework, which are applied to the presentation of file sharing information. The first is the gathering of perceptual information; this information is presented through feedthrough on artifacts, individualized views, and notification of events. The second element is the exploration of the environment (information gathering), which users perform through background awareness, explicit queries for information, and selective attention to events. This framework was used to create a preliminary design for a file manager to support awareness, one that displayed the perceptual information and supported the information-gathering methods outlined above.

2.4 Phase Three: Applying the Framework: Designing Tools for Managing Shared Files

Because file management tasks often involve the use of a file manager, the framework elements were used to identify what features a file manager should have for supporting shared files. The design was based on Windows XP (a well-known tool), which was augmented with some new columns for displaying file sharing information. A supplementary window, called the "sharing console," was designed for additional support, both for displaying greater amounts of detail and to allow for querying and alerts.

The outcome of this design was a high-fidelity prototype, but this prototype had some elements that needed to be refined to ensure that it clearly communicated the relevant file sharing information to users. Because this was a new design, there was a lack of standardized icons and terminology for referring to file sharing concepts. In order to identify the most appropriate icons and terminology, and to gather feedback on the overall design, a user-centered design activity was carried out in the next phase of research.

2.5 Phase Four: Refining the Design Elements

This research phase consisted of three sub-components. The main component was a user-centered group design activity, which was followed by two smaller follow-up studies (a survey and a questionnaire). The group design activity had two main goals: to identify which icons and terms most clearly expressed the file sharing concepts that were displayed by the file manager, and to collect feedback about design improvements for incorporation in the final prototype (used for evaluation). Ensuring clarity of the experimental prototype was crucial for effective evaluation: the design elements (including the terms and icons) needed to clearly communicate information about file sharing concepts, in order for the file manager to be useful. (If, for example, users could not find or understand information they needed when carrying out tasks involving shared files, then the design would be ineffective.) Thus, the design was refined to ensure its clarity prior to the evaluation phase.

The icons and text were central components of the file manager: the new columns for this manager required headings (text) for sharing activities, as well as icons to

indicate activities (e.g., "sent") and file permission information (e.g., "shared with multiple users"). Participants completed a series of design activities that indicated their preferred terms and icons, as well as highlighted problems with interpretation. These activities consisted of a group card sorting activity, as well as an icon-labelling task and a multi-part questionnaire. The questionnaire also solicited feedback about other aspects of the design, such as the query feature in the console, and participants also provided comments in a group discussion.

Two short follow-up studies were performed after this group design activity. The first was a questionnaire used to determine icon preference for the concept "group" (within the context of "shared with group"). This was carried out because no strong preference was expressed during the earlier design activity, and thus additional data was used to make an informed choice.

The second follow-up study was a terminology survey, which replicated some questions from the design activity questionnaire: in particular, to find the best terms for "push" activity versus "pull" activity (see Chapter 7 for definitions). This survey was designed to validate the findings from the group design, by collecting data from a larger sample. The results from the survey were used to determine the most appropriate terminology choices, which were used in the column headings for the final prototype.

2.6 Phase Five: Evaluating the Shared File Management Tools

After the icon and terminology elements had been finalized, and the overall file manager design was refined, the prototype was ready for formal evaluation. In particular, the evaluation was designed to determine whether the new file manager helped people to manage shared files, through greater efficiency, effectiveness, or satisfaction? The focus of this study was on finding and interpreting file sharing information, by answering questions about specific files in a desktop environment.

This evaluation took place in a laboratory setting. McGrath (1995), writing on methodology, discusses how the researcher must trade off between three factors: generalizability of evidence; precision of measurement; and realism of the setting. For this evaluation, precision of measurement was optimized; this approach allowed for analysis of specific variables (such as task completion time), that would identify what ways (if any) that the new design provided improvements over existing technology.

To mitigate the artificiality of the laboratory setting, the scenario approach was used (Borlund, 2003); participants carried out tasks in a simulated office environment, playing the role of an employee filling in for an absent worker. This approach allowed the participant to act "as if" they were truly completing file-sharing tasks, despite the controlled nature of the experiment. Participants used both a control file manager (similar to Windows Explorer) as well as the new file manager.

Participants completed tasks in two sessions, and several types of data were collected, including task completion time, question responses (about the shared files), usability scores, and subjective preference data. These data were used to compare the control and new file managers, as well as to assess the innovative features of the new file manager. In addition, participants indicated the likelihood of using the new file manager in the real world, through the Technology Acceptance Model questionnaire, as well as a post-study email that solicited any examples for which they felt the new file manager features would have been helpful.

2.7 Summary: Research Design

A summary of these five phases is presented in Table 1 below. This includes a description of data collection and analysis in each phase, and the research objectives. (The preliminary phase of this research—the literature review mentioned in Section 2.2—is listed as "Phase 0".) The first phase of this research—the surveys used to characterize file sharing practices—is presented in the following chapter.

Phase	research objectives	method	outcome: results
0	preliminary exploration of file sharing issues	review of research literature	determination that little information existed about file sharing practices
1	characterizing file sharing practices	two surveys (self-administered online questionnaires)	descriptive statistics and coded comments describing prevalent file sharing tasks and concerns
2	development of theoretical awareness framework for file sharing	review of research literature	file sharing awareness framework, adapted from CSCW awareness models
3	applying framework to design (tools for managing shared files)	synthesis of framework and data from Phase 1 surveys	initial prototype of file manager
4	refining the design elements	three user studies: i. group design activity ii. icon questionnaire iii. online survey	establishment of iconography and terminology for file sharing concepts refined version of file manager, to be used in formal laboratory evaluation
5	evaluating shared file management tools	two user studies: i. characterizing realistic workplace file organization ii. scenario-based laboratory evaluation	 i. data about representative workplace file systems (numbers of files and folders, folder names used) ii. evidence that the new file manager with file sharing features led to greater efficiency, effectiveness, and satisfaction for managing shared files (compared to a current file managers), with high intention-to-use demonstrated

Table 1: The five major research phases (and the preliminary phase), including research objectives and descriptions of the results

Chapter 3. File Sharing Surveys

In order to find ways to provide better support for file sharing, we need to understand existing user practices. This is a fundamental requirement for developing appropriate designs: one needs to know what types of information are most useful within a sharing context, and how best to present and manage them. For file sharing specifically, it is necessary to understand, amongst other things, what the common sharing tools and methods are, and the types of files being shared, and the groups with which they were being shared. It is also necessary to know how people manage sharing, their security concerns, and any specific problems that users face when they managed shared files.

To achieve these goals, two surveys were carried out; the research objective was to characterize file sharing. These surveys composed Phase One of the thesis research (as outlined in the research design in the previous chapter). These two surveys are detailed in the following sections. The research described in this chapter was approved by the Social Sciences and Humanities Human Research Ethics Board at Dalhousie University.

3.1 *Study One: Fundamentals of File Sharing*

3.1.1 Summary and Objective

The first study was a preliminary exploration of file sharing, and was very broad in scope, because there was little fundamental data about this task. The objective of this initial study was to gather a variety of information on user attitudes, behaviours, and requirements about file sharing. These results would provide a starting point for later, more refined studies. This initial study consisted of an online survey, in the form of a self-administered questionnaire.

3.1.2 Methods

The survey was conducted at a U.S.-based research institution in the summer of 2005. At the time, the institution had about 200 employees, mainly technical researchers in areas broadly related to information technologies, along with technical and administrative support staff and management. The institution was heavily involved in establishing and maintaining contracts with external collaborators and customers.

3.1.3 Instrument

The survey was maintained on an internal web server for two weeks. It had 84 questions, and took about one hour to complete. The survey was primarily quantitative in nature, although it included some open questions (with free-form responses) as well as categorical questions. The full set of questions is provided in Appendix A.

The questions were intended to explore sharing and access control, focusing mainly on the following issues:

File sharing methods and materials

- How many different methods of file sharing (e.g., email attachments, shared folders) are people using?
- Which methods are most commonly used?
- What kinds of files are shared?

Managing limits on shared files

- Do people need to manage limits around their shared files and computers? If so, what kinds of limits (access control) do they set?
- Do they need to reconfigure these limits?
- What problems do people encounter as a result of access controls?

Handling sensitive files

- Do people have sensitive items on their computers?
- If so, what are they (e.g., email), and how do they keep these items private?

3.1.4 Participants

Participants were recruited through email and through a notice posted on the organization's intranet. All employees of the organization, 18 years of age and older, were eligible to participate in the survey. There were 53 respondents, 34 male (64%) and 19 female (36%), representing approximately 25% of the employee population (of 200).

The largest single group of respondents was researchers (45%); 22% were students or interns; 15% were managers. The respondents had a great deal of computer experience and expertise: 47% of them had over 20 years of computer experience; 41% had 11-20 years. (No respondents had less than 5 years of computer experience.) The respondents were technically adept: as a weekly task, 35 programmed computers (66%), 50 installed software (94%) and 41 performed system maintenance (77%).

3.1.5 Data Analysis

For categorical questions, results are reported using descriptive statistics; primarily, the percentages who selected each response. For open questions, I coded the free-form responses, and identified dominant themes within this data.

3.1.6 Results

3.1.6.1 Types of Files Shared

Respondents were given a list of types of digital materials that they might have been sharing in the course of their work and leisure; they chose as many as applied to them. The participants could also add more types if they chose; fifty-three participants chose to add to this list. The results are listed in Table 2.

type of material	#occurrences
written documents (e.g., Word files, PDF)	51 (96%)
photos or graphics	47 (89%)
slides (PowerPoint)	42 (79%)
spreadsheets/numeric data	30 (57%)
source code	28 (53%)
audio (e.g., music)	23 (43%)
video	19 (36%)
none	0 (0%)
other (11): executables (5); Web pages/URLs (6); calendars (1); varied data files (1)	

Table 2: Types of materials shared

On average, each respondent shared 4.79 types of materials ($\bar{x}$ =4.79, σ = 1.53).

3.1.6.2 Methods of Sharing

Respondents identified the methods they used to share files, in a variety of circumstances (e.g., work files with local and remote work colleagues, and personal files with friends and family). They chose from a set of 11 methods, including email attachments, peer to peer networks, and portable drives; respondents selected all options that applied, or specified "other." The results are presented in Table 3. For brevity, only the most frequently mentioned methods for each category are reported; percentages reflect the number of respondents who selected that specific method. Table shows the most prevalent methods for file sharing; note that email attachments dominate for all categories (e.g., personal and work, and with all groups). (The proprietary document

		Sharing methods in most *widespread* use		
	shared with:	1st	2nd	3rd
work files	local colleagues (n=53)	Email attachments: 53 (100%)	Network file share: 42 (79%)	Document mgmt. system: 24 (45%)
	remote colleagues (n=48)	Email attachments: 48 (100%)	Network file share: 20 (42%)	Web: 16 (33%)
	outside associates (n=50)	Email attachments: 50 (100%)	Document mgmt. system: 14 (28%)	Disks, Web (tied): 13 (26%)
personal files	**shared with:**			
	spouse/partner (n=45)	Email attachments: 43 (96%)	Portable drive: 27 (60%)	Web: 16 (36%)
	family (n=45)	Email attachments: 44 (98%)	Web: 20 (44%)	Disks: 8 (18%)
	close friends (n=49)	Email attachments: 48 (98%)	Web: 20 (41%)	Portable drive: 15 (31%)
	anybody online (n=32)	Email attachments: 20 (63%)	Web: 18 (56%)	Peer-to-peer network: 4 (13%)

Table 3: File sharing methods for work and personal files

management system, used only for work files, is Xerox Docushare, which was deployed within this organization.) Also, with the exception of "anybody online," four methods occupied the top three spots for all personal groups: email, portable drives, the Web, and writeable disks. Overall, when counted across all categories of work and personal sharing, people reporting using an average of 5.43 methods to share files ($\bar{x}$ =5.43, σ = 1.63). No one used fewer than three methods, and 16 people (30.2%) reported using seven or more methods to share files; the largest number reported was nine methods.

3.1.6.3 Most Frequently-Used Methods

Respondents identified their three *most-frequently* used methods for sharing in two categories: sharing work files with colleagues, and personal files (with anyone). The same list of methods as before was provided, with the option of adding others. The results are presented in Table 4. For work files, the top three most frequently-mentioned choices were email (52 people [98%]), network file sharing (33 people [62%]), and, at a distant third, portable hard drives (14 people [26%]). Note there was a large gap among choices; for instance, the third place choice was chosen much less frequently than the second

For personal sharing, email was still the most prevalent, being chosen by 48 respondents (92%). There was a large gap between this choice and all other choices; portable drives were second, with 19 (37%), closely followed by the Web at 18 (35%) and writable disks at 17 (33%).

Most frequently-used sharing methods	Work files (n=53)	Personal files (n=52)
Email attachments	52 (98%) #1	48 (92%) #1
Network file sharing	33 (62%) #2	4 (8%)
Portable hard drives/tokens	14 (26%) #3	19 (37%) #2
Proprietary mgmt. system (Docushare)	12 (23%)	1 (2%)
Web transfer	11 (21%)	18 (35%) #3
Writable disks	7 (13%)	17 (33%)
Other network transfer (ftp, scp)	6 (11%)	4 (8%)
Other (less than 10% for both categories): direct connection between machines; infrared synchronization software, peer-to-peer networks, "print digital picture and snail mail"		

Table 4: Most frequently-used sharing methods. The top three choices for each category are marked (e.g., #1 indicates the top choice).

3.1.6.4 Sensitive Files

Respondents identified which files, if any, they did not want others to see; the question specified "the people who have legitimate access" to a machine (as opposed to a random hacker, for example). Thirty-five out of 44 respondents (80%) stated they did have such items; 24 provided free-form responses describing these items. These items are categorized in Table 5; note that many respondents provided more than one type of material, making 44 responses in total.

sensitive material	#occurrences (n=24)
Personal email	9 (20%)
Passwords/encryption keys; financial data	7 (16%) x 2
Confidential work data	5 (11%)
Personnel files; photos; "personal stuff" (unspecified)	3 (7%) x 3
Contacts; web history	2 (4.5%) x 2
Other: calendar; jokes; creative writing.	1 each (2.3% x 3) : 7%
	Total: 44

Table 5: Types of sensitive digital information

Respondents next described what they did to protect these items, if anything. Thirty-nine people responded; this included the 24 who provided information in the

previous question, plus 15 people who skipped that question. There were a range of responses, including both technical and social approaches. These responses are categorized in Table 6.

Method of protection	**#occurrences (n=39)**
Passwords	10 (19%)
Politeness/trust/social conventions	8 (15%)
File permissions/access control lists	7 (13%)
Encryption	6 (11%)
Obscurity: hidden directories, innocuous filenames	4 (7.5%)
Limit physical access to computer	4 (7.5%)
Logging out/screensaver lock	3 (5.7%)
Keep separate accounts	3 (5.7%)
Don't know/nothing	3 (5.7%)
Keep material off shared machine/shared folders	3 (5.7%)
Other: run stateless computer; delete temporary files	1 each: (2% x 2): 4% Total: 53

Table 6: Methods of protecting sensitive data

3.1.6.5 Managing File Permissions

File permissions are sometimes used as a means for controlling access to shared materials. When asked if they had ever set file permissions, 36 out of 52 (68%) respondents said yes. Thirty-one of the 36 people (86%) who set permissions provided details about the circumstances in which this occurs; the five most commonly-cited reasons are shown in Table 7. The most frequently-cited reason was limiting access to a specific group or individual at work.

context for setting permissions	**#occurrences (n=31)**
Limit access to only a specific group or individual at work	9 (27.3%)
Grant access for others in work group	5 (15.2%)
Prevent all changes to a file/document	5 (15.2%)
Disallow access to personal files	5 (15.2%)
Prevent inadvertent modification/deletion	4 (12.1%)

Table 7: Reasons for setting file permissions

When asked whether they had ever set temporary permissions, 25 of 52 respondents (48%) said yes; when asked if they had gone back and reset the permissions afterwards, 23 of 25 people (92%) said "yes".

3.1.6.6 Inconveniences Arising from Access Control

Respondents were also asked to describe instances in which they could not get work done due to problems with access control settings. Thirty-eight people (72%) reported having problems, and provided detailed examples of their frustrations; these responses are categorized in Table 8 below. Four respondents provided multiple instances of problems, making a total of 44 responses. The most commonly-cited problem (with 8 instances, or 18% of responses) was waiting on another person to grant access.

Problem encountered	**#occurrences (n=38)**
Waiting on another person (file owner, system administrator) to grant access	8 (18%)
Need root permission to complete task	5 (11%)
New person on project, permissions not updated	5 (11%)
[insufficient detail to categorize: "could not access files"]	4 (9%)
Cross-platform problems (e.g., Mac/Windows)	3 (7%)
Lack of technical expertise: can't figure out how to modify permissions	2 (4.5%)
Restricted to one specific machine in order to gain access to files	2 (4.5%)
Mobile work problems (remote access)	2 (4.5%)
Permission defaults set incorrectly	2 (4.5%)
Denied access to shared network folders	2 (4.5%)
Other (9) examples: user interface makes permission management difficult; lack of notification to access control changes; forgot password [1 each]	9 x 2.3% = 20.7%
	Total: 44

Table 8: Problems resulting from access control configuration and management

3.1.6.7 Discontinuing Sharing: Removal and Reconfiguration

Sharing contexts are subject to change: some materials are not, or should not, be shared any longer (e.g., when the file to be shared is no longer needed). Twenty-one of 53 respondents (40%) had changed permissions to restrict sharing (rather than deleting the file). The various reasons they gave for this change are categorized in Table 9. The most

commonly-cited reason (with 5 instances, or 28% of responses) was because the access required was only temporary.

reason for revoking permissions	#occurrences (n=21)
Only temporary access was required	5 (28%)
Change in team members	3 (17%)
Change in organization	2 (11%)
Security/confidentiality concerns	2 (11%)
Protect personal files	2 (11%)
Correct permission errors	2 (11%)
Other: testing; change in content	1 each: 2x 5.5%=11%
	Total: 18

Table 9: Reasons for resetting file permissions to terminate sharing

3.1.7 Final Comments

At the end of the survey, respondents were provided with the opportunity to supply additional comments about file sharing. While only four respondents provided comments, their statements underscore file-sharing issues that were significant enough to merit special mention; these indicate problems of particular concern. These comments are presented in full below:

> P8 It is very irritating that there isn't a single place to go to see all this stuff. When one asks oneself the question, "what possible security threats do I suffer from in my current state?" I have to go to many places and ask many people to help me figure it out. I'm not technically savvy, or, worse, just not technically intrigued enough to go poking around. I have other things on my mind and out of sight is out of mind as far as security status goes. I think that's a pretty standard attitude, sad as that may seem.

> P10 In many cases, I believe that people don't know what they need to restrict access to. For example, few people are aware that anyone inside the [parent organization] network has access to files stored on [local organization's web] and in ~/projects/ *unless* specific actions are taken to restrict access. I regularly "hop on" wireless access points using my work laptop, knowing that there's some risk. The need for access outweighs what I know about the risks; but what I know about the risks may not be complete.

> P11 Using/accessing a new system usually ends up with multiple layers. They are not necessarily co-ordinated by/known by the same person/department, so it usually is a major hassle figuring out all the bits that need to be in place to access the resources you need. True at [my university], true [at this institute].

P27 Biggest day-to-day issue is probably management of passwords for access to array of different accounts mostly on web; Second is probably figuring out a sharing arrangement in any particular situation

3.1.8 Summary of Key Findings

Results indicated that respondents shared a wide variety of files (approximately five types of data, on average). Note that every respondent shared at least one type of file: in fact, no one shared fewer than three different kinds of files. A wide variety of methods were used to share files: on average, five different methods. Email attachments were the most common, and the most widespread, method for file sharing, in both work and personal contexts. Another common sharing method in work contexts was shared network folders; 62% of respondents used this method. Other sharing methods were not as prevalent, but a number were used by over 25% of respondents: for work files, portable drives (26%); and for personal files, portable drives (37%), the Web (35%), and writeable disks (33%).

3.1.8.1 Complexity of File Sharing

What this data reveals is a complex picture of file sharing: multiple types of files, and multiple methods of sharing. If people only used a single file-sharing method, or a few files, this would likely make shared file management fairly simple. However, that does not reflect the user behaviour revealed in the survey: different methods are appropriate for different situations and needs, and therefore users rely on a diverse set of sharing mechanisms. A file might be used with many different applications when it is shared with different recipients. This evidence suggests that keeping track of shared files is complicated. For example, a single user might email some files to multiple recipients, and place others in shared folders. In order to know which files are available to one's colleagues, one would have to search through an email archive as well as look through different folders.

3.1.8.2 User Needs: Security and Information

The data confirms that the majority of respondents (80%) did maintain sensitive documents on their computers, including both work and personal items. People engage in a variety of behaviours to protect these items from those who have legitimate access to their machines. Therefore, respondents indicated that they have a need to protect data, and that they take steps to ensure that the data is protected. Along with technical

methods, such as access control lists and encryption, there was also the use of "security through obscurity" (hiding files), and a reliance on normative behaviour to keep people from prying. Note that users cannot easily tell whether other people have accessed their private files; this information is not readily available, and therefore people must simply hope that prying has not taken place.

People need to manage limits on file sharing, and in many cases, they use file permissions to do so: 68% had experience with setting permissions, which included both granting and removing access to items. Almost half of respondents (48%) needed to do this temporarily, after which they needed to reset the permissions. Similarly, 41% of respondents encountered situations in which they needed to reset permissions to limit sharing. Therefore, the management of limits is a dynamic process: people do not generally set permissions once, but instead must revisit and reconfigure them as their circumstances dictate.

Because permissions need to be updated, it is possible that people do not always make the required changes. Sometimes people might forget to make the necessary updates. This possibility was suggested by the survey data, in which people described how access control interfered with work: one common complaint was waiting on another person to make a necessary change to the system.

Finally, the concluding comments—although from only four people—indicated an overall lack of information about sharing. These participants stated that "out of sight is out of mind" for security status, and that there is a lack of knowledge about what files are accessible. One participant also stated that it can be difficult to work out appropriate sharing arrangements; this may be due to a lack of knowledge about the file sharing environment. The results of this study were published in (Whalen, Smetters, & Churchill, 2006).

3.2 *Study Two: Further Investigations into File Sharing*

It is evident from the initial study that file sharing is a complex, and real, workplace problem. The results of the study revealed elements that needed to be further explored, and a number of unanswered questions remained about sharing practices. Collecting additional data on basic sharing, including methods, files, and groups, would also allow for validation of the initial data set with another user population.

There were also some new avenues to explore. Some of the results from the initial survey that required further study were the technological methods used in sharing: why were certain approaches, such as email attachments, so prevalent? In addition, what were the disadvantages of these methods, such that users would seek alternative mechanisms for sharing?

Another missing element was the frequency of file sharing and of file permission management, which would indicate the magnitude of the problem. The initial data showed that people perform these tasks, but the survey did not ask about their frequency or the number of groups that people shared with. As well, data about file permission errors was lacking: what types of errors did users encounter, and how easily were such errors detected? Finally, to design a solution, it was necessary to discover how people managed file limits (individually or in groups).

3.2.1 Overview of Second Research Study

A second study was developed to explore these issues. While this study was being designed, a publication on file sharing affordances was published by Voida, Edwards, Newman, Grinter, & Ducheneaut (2006). Their study, on a similar population as in my initial survey, also presented some data on file sharing behaviours and needs. Voida et al.'s survey included a set of questions on knowledge about file sharing, and the perceived importance of setting limits around shared files.

As this topic was relevant to my own work, a subset of these questions was replicated in my second survey. Specifically, these questions focused on issues of knowledge of, and information about, sharing activities: perceived knowledge about shared files, and what types of sharing information were most important to present to users. Findings from Voida et al.'s study are the subject of Section 3.2.9.2. A more detailed review of that work appears in Chapter 4 (Related Work).

3.2.2 Research Objectives

There were three main goals for this study:

- to assess the validity of results from prior studies on file sharing
- to expand what was known about file sharing, building a more complete understanding of user behaviours and needs, and

- to investigate what types of information, about shared files and activity on shared files, might be helpful to users.

3.2.3 Methods

A web-based survey was conducted at a Canadian university; this university had a population of approximately 15,000 students and 1,100 faculty. The self-administered survey was maintained on a secure commercial survey Web site for two weeks. Recruitment was carried out through email; all students, faculty, and staff of the university were eligible to participate in the survey.

3.2.4 Instrument

The survey had 72 questions, although the number varied depending on responses to conditional questions. It took about forty minutes to complete. A subset of questions was selected for analysis, which is described in this section; the full set of questions is provided in Appendix B.

Some questions in this survey were based on the initial survey described in the previous section. Most of these questions dealt with personal file sharing practices, the methods used for sharing, managing limits and restrictions on use, and handling of sensitive information.

Eight questions on the following topics were repeated from my previous study, with minor modifications:

- How many different methods of file sharing (e.g., email attachments, shared folders) are people using? Which methods are most commonly used?
- Do people need to manage limits around their shared files and computers? If so, what kinds of limits (access control) do they set?
- What percentage of people have sensitive items on their computers? What are those items (e.g., email)?

Twelve new questions on the following topics were also included in this survey:

- How often are people sharing files? (Is sharing a frequent task?)
- How many different groups do people share with?

- What are the advantages and disadvantages of different sharing methods? (What factors influence people's choice of a sharing mechanism in a given situation?)
- How often do people need to manage file permissions?
- What types of permission errors do people encounter? How long does it take them to detect these errors?
- Do people put limits on individual shared items, or on folders?

Finally, six questions (including two with minor modifications) were replicated from the study of Voida et al. (2006) on the following topics:

- Are people aware of their shared files and sharing activities?
- What types of file sharing activity might people want to know about?

3.2.5 Respondents

Seventy-nine respondents (25 female (32%), 54 male (68%)) completed the survey, some of whom did not respond to all questions. The majority of respondents were students: 35 undergraduates (44%) and 23 graduates (29%). Four faculty members (5%) responded, as well as 14 staff members (18%) and six people (8%) who categorized themselves as "other" (namely, alumni and affiliated staff). Overall, the respondents were highly competent in the use of computers. Many student respondents were from computer science: 51% of undergrads and 65% of graduates. The staff members were also technically-oriented, with 64% of them in a computer-related occupation (such as system administrator). In addition, 62% of respondents had over 10 years of computer experience, and 34% had between five and ten years of experience.

3.2.6 Data Analysis

The survey was primarily quantitative in nature, although it included some open questions (with free-form responses) as well as categorical questions. For categorical questions, results are reported using descriptive statistics: primarily, the percentages who selected each response. For open questions, I coded the free-form responses and identified dominant themes within this data.

3.2.7 Results

3.2.7.1 File Sharing

Respondents provided details about their file sharing activities: how often they shared; the file sharing methods they used; the types of files they shared; and the people with whom they shared files. File sharing was a common professional or educational activity for most respondents; 49% shared files daily, and 25% shared files weekly. (Only one person reported never sharing files in the workplace.) Similarly, personal file sharing was also common, with 30% of people sharing daily and 47% sharing weekly.

Several file sharing methods were used: respondents were provided with a list of methods, choosing all that they had used, and additionally had the option of listing other methods. Table 10 shows the methods most frequently selected from the list; only those options selected by over 20% of respondents are included.

Sharing method	#occurrences (n=69)
email attachments	68 (99%)
physical devices (e.g., USB token, portable drive, disc, CD)	67 (97%)
network file share (e.g., Windows file sharing, NFS)	56 (81%)
instant messenger (e.g., MSN, Yahoo, AIM, ICQ, etc)	53 (77%)
Web server (e.g., Web page, blog, wiki)	49 (71%)
peer-to-peer (P2P) program (e.g., torrent client, KaZaa, Limewire)	48 (70%)
file copy protocol (e.g., scp, ftp)	46 (67%)

Table 10: Most frequently-selected methods for file sharing

On average, each person people used approximately six sharing methods ($\bar{x}$ =6.01, σ = 1.59). No one used fewer than three methods; the greatest number reported was 10. Respondents then identified their most *commonly* used sharing method. Email attachments were by far the most prevalent (43%), distantly followed by network file sharing (15%).

Sharing method	#occurrences (n=68)
Email attachments	29 (42.7%)
Network file share	10 (14.7%)
Peer-to-peer program	7 (10.3%)
File copy protocol	7 (10.3%)
Instant messenger	6 (8.8%)
Web server	4 (5.9%)
Physical devices (e.g. USB token)	4 (5.9%)
iTunes music sharing	1 (1.5%)

Table 11: The most commonly-used file sharing methods

Respondents also described the positive and negative features of the sharing method that they selected, using free-form responses which were coded and categorized. Table 12 below lists both the positive and negative features for the five most commonly used choices.

Most Common Sharing Method	Positive features n=64	Negative features n=60
Email attachments n=29	• Convenient/easy to use • Widespread availability • Understandable • Fast • Ensure item goes to intended recipient(s) • Well-suited to documents/comments • Provides a backup • Suited to organizational structure/needs • Supports additional security • Reliability	• File size limitations • Person needs specific application to use/open file • Can't control forwarding • Requires lot of space on mail server • Virus strips off attachment • Inconvenient for multiple files; lack of built-in encryption; firewall issues; multiple versions of document; slow; poorly suited to collaboration; interoperability issues; don't know when file read
Network file share n=10	• Convenience /ease of use • Provides control/access control • Fast • Provides security • Suited to organizational structure/needs • Suited to home networks	• Problem crossing organizational boundaries • Limited file space • Slow • Operating system incompatibilities; poorly suited to sharing across internet; file permission problems; remote file system may be insecure
File copy protocol n=7	• Easy to configure, use • Fast • Reliability/stability • Accessibility • Provides access control • Good for multiple users; flexible	• Need special application • Backing up is complex; slow
peer-to-peer (P2P) program n=7	• Variety of files available • Ease of use • Fast • Avoid digital rights management; obtain free media; reliability; ease of search	• Insecure/untrusted sources • Slow • Unfair use (others don't share) • Consumes bandwidth; spyware installed; limited file selection; need client program; legal issues
Instant messenger n=6	• Ease of use • Fast • "User friendly"; know who gets files	• Slow • Unreliable • Doesn't work for all users; synchronous delivery required

Table 12: Advantages and disadvantages of various file sharing methods

3.2.7.2 Sharing Groups

Respondents indicated how many distinct groups they shared with, such as friends and colleagues. Over 69% of respondents shared with two to four groups, and another 25% shared with five to 20. (Three respondents (4%) shared with over 20 distinct groups.)

Respondents optionally named these groups; common examples included colleagues, family, and friends.

Respondents also indicated how often they needed to share the same item with multiple groups, using a five-point scale ranging from "never" (1) to "very frequently" (5). Sixty-four percent of respondents selected "sometimes" or "often" ($\bar{x} = 3.86$). Fifty-nine respondents also described a recent example of their file sharing, in terms of what they shared, with whom, and how. In three instances (5%), people reported sending the same material to two distinct groups, such as a boss and a family member. Also, in two cases (2%), people used different methods (e.g., email as well as a shared folder) to share the same file with different sets of recipients

3.2.7.3 Sensitive Documents

Respondents described their handling of sensitive materials; first, whether they had any sensitive material that they wished to keep private. Out of 60 respondents, 48 (80%) did have some sensitive material. A list of possible sensitive materials was presented; respondents selected all applicable items, and indicated whether they *shared* this type of material with others. Results are in Table 13 below. The type of sensitive item stored by the most people was personal information (such as email), with 85% of respondents reporting that they had this item on their computer. The type of material that was most commonly *shared* was confidential organizational documents, with 44% of respondents reporting that they shared this sensitive item.

Type of sensitive material	**I *have* this material (n=48)**	**I *share* this material (n=48)**
personal items containing socially-awkward information that you wish to keep private (e.g., personal email, photos)	41 (85%)	11 (27%)
personal financial or medical information (e.g., credit card numbers)	38 (79%)	4 (11%)
professional data or documents that are sensitive to your work organization (e.g., proprietary documents, trade secrets, grades)	36 (75%)	16 (44%)
professional data or documents that are governed by law/regulation (e.g., medical records, personnel files)	25 (52%)	7 (28%)

Table 13: Types of sensitive files that people use and share

Second, respondents indicated to what degree security and privacy were a concern when sharing files, using a five-point scale ranging from "very frequently"(5) to "never"

(1). Seventy-nine respondents answered, with 64% stating that this was a concern at least "sometimes" and 22% being concerned "very frequently" ($\bar{x} = 3.13$).

3.2.7.4 Managing Limits on Shared Files

Respondents also provided information about managing shared files, and in particular, on setting boundaries around sharing. Eighty percent of respondents had configured file permissions; 63% configured permissions at least once per month, and 42% did so every week.

Respondents also described their reasons for reconfiguring permission settings; they were given a list of reasons and they could choose as many as were applicable. The most common reasons are listed in Table 14 below:

Reason for change	**#responses (n=42)**
the previous settings were for temporary sharing and I needed to restrict access again	27 (18%)
an item that was not being shared needed to be shared	25 (16%)
an error was made and I had to correct it	22 (14%)
an item that was being shared shouldn't be shared anymore	22 (14%)

Table 14: Reasons for resetting file permissions

Forty-six respondents indicated the type of sharing error—too restricted or too open—that they most recently corrected. Twenty-nine people (63%) had permissions set too restrictively (preventing sharing), while 16 (35%) people had permissions that were too open. They were next asked how long the delay was in discovering and fixing this error; they were provided with a list of times, ranging from "immediately" to "years", as well as "unknown." Twenty-three percent stated that they did not know when the error was made, and a further 30% reported a delay of at least hours.

3.2.7.5 File Sharing Knowledge

Respondents indicated the types of sharing information they would like to know about, and the degree to which they felt they knew what was being shared (and with whom). These questions included five-point Likert scales; "strongly agree" corresponded to a score of five and "strongly disagree" corresponded to one. The questions, with the mean for each, are presented in Table 15.

Questions on awareness	mean (n=79)
I am generally aware of all of the files that I am sharing and with whom I am sharing them.	4.0
When I share my files, it is important to me to know exactly who has access to the files.	3.8
I would like to be able to know who was using each file.	3.3
When I share my files, it is important to me to know when others are using the shared files.	3.0

Table 15: Awareness of shared files and activity on shared files

3.2.8 Discussion

3.2.8.1 Key Findings

Not surprisingly, file sharing is a common activity, with over 70% of respondents sharing work and personal files at least once per week. Many different types of sharing tools were used: on average, about six types. As in the previous study, the most popular of these is email; networked folders were a distant second, followed by file copy protocols and peer-to-peer networks (tied for third place).

Results show that there are a number of factors influencing the choice of sharing method used. A major factor is convenience and ease of use, which was a feature cited across several different types of sharing mechanisms. When speaking of email specifically, users also liked the fact that it was readily available (both to self and recipient), and fast. Speed was also a determining factor, and for those concerned about security, access control was an advantage provided by networked folders and file copy protocols.

The respondents shared with multiple groups: the vast majority reported sharing with between two and four groups. In addition, most respondents needed to share the same item with multiple groups: 71% of them reported doing this at least "sometimes."

Sensitive materials were commonly stored, with 80% of respondents having such files on their computers. These materials were also shared, to some degree: 44% of respondents with professional files shared them with others, although only 11% of people shared their financial or medical information.

File permission management was a task that most respondents engaged in, 80% had experience with setting permissions, with the majority of them (63%) performing this task at least once per month. Permission settings were also dynamic; the most common reason for resetting permissions was because of a temporary sharing situation, after

which sharing was restricted. Respondents reported encountering errors with permission settings; approximately half of respondents needed to correct such an error. Permission errors were also not always quickly detected: 23% of respondents who corrected an error stated that they did not know when the error was made, and a further 30% reported a delay of at least hours.

The responses to the questions on file sharing knowledge indicated that users wanted to know about shared file activity. In particular, knowing *who* had access was rated as important ($\bar{x}$ =3.8 out of 5). Although respondents self-reported a high degree of knowledge ($\bar{x}$ =4.0 out of 5) about what was shared, and with whom, this may not be an accurate result. For example, many respondents (53%) discovered errors in their settings after some delay, which suggests that their knowledge was incomplete.

3.2.9 Comparison to Prior Studies

How do these results compare with previous research, specifically my earlier survey and the results of Voida et al. (2006)? (The latter study will be described in more detail in Chapter 4). There are significant similarities between these studies, but also some differences that should be noted.

3.2.9.1 Comparison to Previous Survey (Study #1)

The majority of the findings from the initial study were confirmed in the second study; these are summarized in Table 16. In both cases, the top three most prevalent sharing methods were the same: email, network file shares, and physical devices (USB tokens, CDs). There are also similar results for the overall number of methods used on average to share files: between five and six in both cases. No one used fewer than three methods in either case, and the highest number of different methods used was nine in the first study and 10 in this follow-up study. In both studies, the majority of respondents reported setting file permissions: 68% and 80% respectively. Although the initial study respondents expressed a stronger overall need for security when dealing with sharing (likely owing to the types of proprietary materials they shared), both sets of respondents worked with sensitive documents that they needed to protect: 80% in both cases.

result	Study #1 (n=53)	Study #2 (n=79)
most prevalent sharing methods (top three)	1. email 2. network folders 3. physical devices (portable drives, CDs)	1. email 2. physical devices (portable drives, CDs) 3. network folders
average number of methods used to share files	$\bar{x} = 5.43$	$\bar{x} = 6.01$
lowest number of methods used to share files	3	3
highest number of methods used to share files	9	10
percentage of respondents who managed file permissions	68%	80%
percentage of respondents who handled sensitive files	80%	80%

Table 16: Similarities between the two file sharing surveys

There are also some differences in the results of these two studies; these are summarized in Table 17. The research institute respondents reported greater use of a proprietary management system (see Section 3.1.6.3); this is likely due to the fact that one system has been deployed throughout the institution, and such a system is not in place at the university (where the second study was conducted). Another difference is the use of peer-to-peer networks, which is far more common at the university. The most likely explanation of this is the difference in populations: most of the university respondents were young students, who constitute the heaviest users of peer-to-peer media file sharers (Pew Internet & American Life Project, 2006).

result	Study #1 (n=53)	Study #2 (n=79)
percentage of respondents who used a proprietary management system for file sharing	45%	4%
percentage of respondents who used a peer-to-peer network for file sharing	4%	70%

Table 17: Differences between the two file sharing surveys

3.2.9.2 Comparison to Findings of Voida et al.

Voida et al. (2006) carried out a file-sharing study at a research institution, which consisted of a survey and interviews with ten participants. Voida et al.'s study is described in detail in Chapter 4; here, a brief summary of findings is presented for the sake of comparison.

The results of my second survey confirmed, for the most part, those of Voida et al. A comparative summary between these surveys is found below: similarities are

presented in Table 18 and differences in Table 19. Note that these comparisons include the second survey only (and not the initial survey), because this second survey incorporated elements from Voida et al.'s study.

result	Voida et al. (n=10)	Study #2 (n=79)
most prevalent sharing methods (top three)	email network folders Web	email physical devices (portable drives, CDs) network folders
evidence of multiple methods used to share files?	yes (13%)	yes (2%)*
level of self-assessment of file sharing awareness	$\bar{x} = 3.7$ out of 5	$\bar{x} = 4.0$ out of 5
evidence of errors caused by lack of awareness	yes: participants reported forgetting what files they had shared and with whom, and forgot to reset permissions	yes: 53% found permission error after lengthy or unknown delay
Notes: * the percentages in this row (multiple sharing methods) are not *directly* comparable, as the Voida et al. percentage is derived from multiple file sharing events over time, and the percentage in Study #2 is derived from only the *most recent* event (and is thus likely to be lower).		

Table 18: Similarities between findings of Voida et al. (2006) and the second file sharing survey

Both studies (in fact, all *three* studies) showed that email attachments are the most popular and heavily-used sharing method by far, and both studies included networked folders in the top three most-prevalent methods. As with Voida et al., I found that people sometimes used multiple methods to share the same material with different recipients. Also, as in Voida et al.'s study, respondents self-reported being aware of their file sharing, but in fact commented upon many instances in which they encountered problems that may have been caused by lack of awareness (e.g., forgetting to reset permissions).

result	Voida et al. (n=10)	Study #2 (n=79)
number of methods used for file sharing	"few" (no number reported)	$\bar{x} = 6.01$; largest number: 10
reasons for using multiple methods	primarily to compensate for lack of notification functionality	many reasons cited, including convenience, suitability to task, or required to reach an entire set of recipients

Table 19: Differences between findings of Voida et al. (2006) and the second file sharing survey

There are some differences between these studies, however. Voida et al. reported that respondents used only a few different methods for sharing; my studies found that they used between five and six methods on average. In addition, Voida et al. proposed that people used multiple methods mainly to compensate for lack of notification (e.g., people needed to use email to tell someone a new file had been put in the shared folder). My survey uncovered a variety of reasons why people preferred one method over another; a user may need to utilize multiple methods simply to reach all recipients. This finding indicates that people will continue to use a variety of methods, due to the complexity of their sharing situations and the limitations of available technologies.

3.3 *Summary*

The first survey uncovered some of the major issues with file sharing; in particular, it showed that multiple methods were necessary to cope with the various circumstances in which people need to share files. It also confirmed that privacy and security were important within an organization, and that sharing poses a risk to sensitive data when it is incorrectly managed. The second survey delved into the technical means which people used to manage limits. It reconfirmed the use of multiple sharing methods, within a different institutional context, and showed that sharing errors were being made. It also demonstrated that people want to have information about file sharing activity, and that people need to keep track of dynamic settings as their collaborations change.

One of the major threads running through these studies is how people gather and understand information about shared files and file sharing activities. There are two components to this issue. The first is that very little information about file sharing is readily available to the user. Some of this information is available within the system, but is inconvenient to access, or is so obscure that users do not even know it is there. File permissions are one example of this; in systems such as Windows XP, this information can be located, but requires that the user take considerable effort to find it (Maxion & Reeder, 2005). It is also important to note that some information is not available at all: activities on shared files, for example, are not generally recorded. This means that users are unaware of many activities that occur on their shared files, which can be problematic for both security and for collaboration.

The second component of information about sharing is that when such information is presented, it must be clear and understandable, so that a person can make use of it as they complete their tasks. Obviously, as described above, it is vital to make the information available. However, the presentation of this information is also a key factor in its usefulness. For example, the information needs to be accessible, but not so distracting that it interferes with productivity. In addition, the user must be able to comprehend what is being communicated. When people are presented with security information (in the case of file permissions), they often find it confusing. The challenges of "usable security" have been well-documented in the literature (e.g., Whitten & Tygar, 1999; Wu, Miller, & Garfinkel, 2006). These works demonstrate that considerable design effort must be made to assist users in coping with the complexities of computer security.

Chapter 4. Related Work

There are a number of different research areas that are directly relevant to, or closely influence, file sharing issues. The related literature is detailed in this chapter. Section 4.1 presents usability studies that involve file sharing. Section 4.2 focuses on the collaborative aspects that affect file sharing, as presented through relevant computer-supported cooperative work (CSCW) literature; this includes both theory and practice. Because security is an important aspect of file sharing, there is much pertinent work in the field of usable security: this is detailed in Section 4.3. Finally, issues of file management (including shared file management) have been discussed in the personal information management field; some germane topics are presented in Section 4.4, followed by a summary discussion.

4.1 *File Sharing Research*

There is a large body of research on file sharing, but little of it is focused on everyday practice or the problems of collaboration and security. For example, there has been much interest in peer-to-peer file sharing issues, but these works are often concerned with topics unrelated to this thesis research, such as the legality of media sharing (e.g., Gal, Singer, & Popkin, 2003), fairness in resource sharing (e.g., a, Lee, Lui, & Yau, 2004) or network and protocol issues (e.g., Legout, Liogkas, Kohler, & Zhang, 2007). In contrast, my focus is in two areas: broadly characterizing the task of file sharing overall, and looking at the usability issues of file sharing. Research in these two areas is presented in the following sections.

4.1.1 Characterizing File Sharing

There have been few user studies that incorporated broad user practice around file sharing. The previous chapter presented three such studies: two are my own work, and the third was by (Voida, Edwards, Newman, Grinter, & Ducheneaut, 2006). The results of Voida et al.'s work were presented in summary in Chapter 3; as it occurred between my two studies, it is discussed in more detail in this section.

There were two major components to Voida et al.'s work. The first was to explore current practices: how users share; what the breakdowns in sharing are; and how these breakdowns reflect a gap in desired sharing practices and the practices supported by

current sharing tools. The second was to expose and describe the "dimensions of file sharing" that are implicit in current file sharing tools. The authors surveyed ten employees at a research institution, and conducted follow-up interviews. They asked respondents to provide detailed examples of work-related information sharing as well as file sharing. Respondents were also asked to describe situations in which file sharing failed or did not work out as planned. Voida et al. also asked a series of questions focusing on awareness of file sharing and activities on shared files.

Their study provided some useful information about file sharing (although the population was small and therefore the results have limited validity). The ten respondents shared many different types of information (34 kinds on average), using multiple sharing tools. Email was most common (45%), followed by shared network folders (16%) and posting to the web (11%). Additionally, breakdowns in file sharing consisted of three main types: forgetting what files had been shared and with whom; difficulties in finding a suitable sharing mechanism that all recipients had access to; and problems knowing when new content was made available. Voida et al. also found that although respondents felt that they were aware of what files they were sharing (and with whom), the majority of the sharing breakdowns that the respondents experienced were based on a lack of awareness. Examples included forgetting what had been shared, or forgetting to reset permissions. The researchers used their results to develop a specialized sharing tool: the Sharing Palette. This tool included features to facilitate file sharing, such as support for simple group management; it also displayed information about activities on shared files. Although implementing this tool was a useful step, it does not fully address the complexity of the file sharing behaviours observed in my studies. For example, my results, as well as Voida et al.'s, showed that people needed to ensure everyone had access to the files for collaboration. Few people used a specialized file sharing application to reach all recipients, in part because such a tool might not be available to all parties (e.g., installed on their computer). It should also be noted that the Sharing Palette was not formally evaluated, and therefore it is unclear which of its features are useful or whether its basic design is effective in supporting shared file management.

4.1.2 Usability Studies on File Sharing

In addition to the broad study described above, there have been a few studies that focused specifically on the usability of file sharing applications and designs. These studies have mainly been concerned with security issues arising from poor usability design.

4.1.2.1 Music-Sharing Application Usability Study (KaZaa)

One of the first studies that concretely examined the security issues of file sharing was carried out in 2003 by Good and Krekelberg (2003). The authors performed a cursory search of the Kazaa peer-to-peer filesharing network, and discovered a number of files that users likely did not intend to share: email files, Web browser caches and cookies, and financial data. (Although it is possible that these files were deliberately shared, the private nature of their content makes this an unlikely supposition.) The authors then performed a cognitive walkthrough of the Kazaa interface, in order to determine how these misconfigurations might have taken place. They concluded the interface made it easy to inadvertently make too much material publicly available, and that it was difficult for people to determine the extent to which they were sharing files. In a follow-up laboratory study, they found that only two out of twelve users could correctly identify which files were being shared with the public.

Based on their findings, the authors developed a set of criteria that defined whether or not a peer-to-peer file sharing application was safe and usable. These include guidelines that emphasize the clear presentation of information: if users "are clearly made aware of what files are being offered for others to download", and "are comfortable with what is being shared with others and confident that the system is handling this correctly." Good and Krekelberg stress the need to assist users in comprehending the scope of their sharing, and to know what is happening in the system. This research provided a valuable starting point, but the authors did not create any specific designs; in addition, their results were focused only on peer-to-peer systems, rather than other sharing tools.

4.1.2.2 Ad Hoc Collaboration Usability Study (Impromptu)

One research project that implemented and evaluated usability aspects of file sharing was the Impromptu system, which is designed for ad hoc collaboration (de Paula et al., 2005; Rode et al., 2006). The authors designed this application to support

visualization of system activity, as a means for aiding secure interaction among participants. This application was designed to allow people to easily configure permissions; an example of the Impromptu interface is depicted in Figure 1.

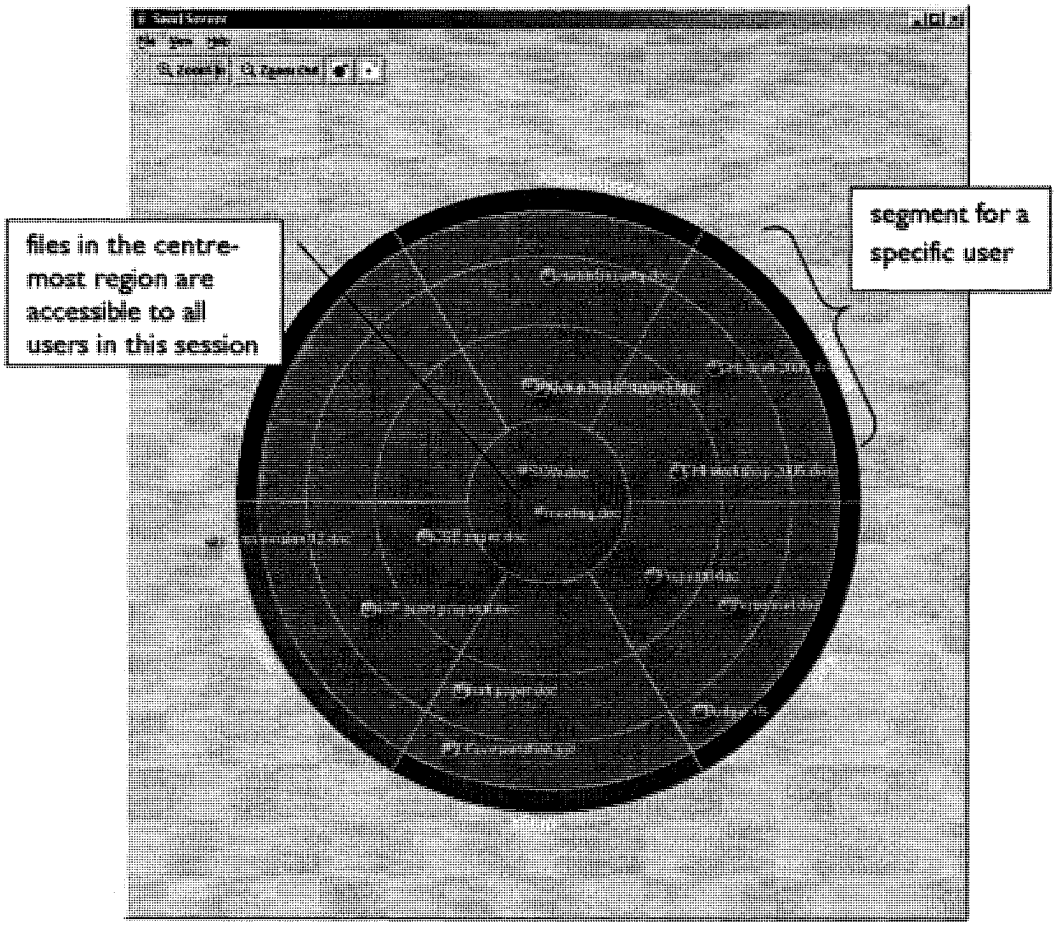


Figure 1: Screenshot of Impromptu, a system for ad hoc collaboration (Rode et al., 2006, p. 146)

Users are presented with a "dartboard"-like interface, divided into regions; contained within the regions are files, which can be shared with others in the ad hoc session. Each user controls a "slice" within the circle; the closer to the centre an object is, the more "shared" it is. To configure file permissions (e.g., to allow others to read a file), users move files into different regions within the sharing space; they can make a file more or less private, depending on their requirements for collaboration or confidentiality. In addition, the interface provides detailed information about activity; "ripples" show actions on a file, and users can view one individual's actions on a specific file.

The advantages of the Impromptu application are that it allows a person to easily see what items are shared, and what permissions those items have. In addition, changes to permissions are made simply, by moving the file to a new region in the interface. This interface seems well-suited to an ad hoc environment, in which people will want to complete their interaction quickly and simply.

An evaluation of Impromptu, using 24 graduate informatics students (in eight groups of three), indicated that it effectively revealed activity on shared objects. For example, users could see who had joined a group, and would make objects more or less public, depending on what was appropriate. This system provides a very useful example for further work: it appears to be the only application in which detailed sharing information (including permissions) is specifically supported within the design, and for which any evaluation has been performed.

4.2 *Computer-Supported Collaboration*

File sharing involves interaction between people, through a shared object; this interaction might involve a specific group of colleagues, for example, or a person and the public at large. As such, file sharing encompasses aspects of collaboration and computer-supported cooperative work (CSCW); there has been some broad theoretical work in this area involving shared materials. In addition, some collaborative environments have been developed that include support for shared files. These theories and tools are described in the following sections.

4.2.1 Collaboration Theory

Computer-supported collaborative work incorporates ways in which technology can be effectively used to facilitate interpersonal interaction. Erickson and Kellogg (2000) developed a CSCW approach for supporting social interaction in digital domains by making the actions of users visible to others, through what is termed "social translucence." Social translucence consists of three major components: visibility, awareness, and accountability. *Visibility* refers to the ability to make socially relevant information (such as actions) perceptible. *Awareness* is a response to visible actions: people behave in specific ways in response to how they see others behave. Finally, *accountability* describes the interaction between actors: a person's behaviour is constrained by social norms, because they know their actions are observed by other

people. The authors illustrate this with the example of a glass window in a door: the window lets one see people coming (visibility), which allows one to avoid opening it in someone's face (awareness); the knowledge that someone can see who opens the door prevents one from rudely opening it in someone's face (accountability).

Erickson and Kellogg describe how we might use social translucence in a shared space; they applied it to a conversation space (the Babble software environment), although it has much wider application to other collaborative domains. Activities on shared files, for example, could fit this model very well. For example, if we show who has been working on a document, this can better support awareness and aid in coordinating collaboration. In addition, we may be able to enforce social norms: for example, an employee might not snoop through sensitive files, because these actions will be visible to others.

Bellotti outlines a number of fundamental issues surrounding sharing practices, specifically within the context of CSCW systems (1996). Bellotti postulates that with collaboration comes a need for better control over access to shareable data (*informational access control*). As a means for supporting access control, Bellotti recommends that users be provided with feedback that reveals the accessibility of their information: who has access, to what items, and what actions have taken place on these items. Bellotti posits that when users engage in file sharing, they often forget what their settings are, and unwittingly make items too public or too private; therefore, feedback about access control settings is very important. (Note that this work (Bellotti, 1996) is strictly theoretical; it does not include any user studies, evaluations or implementations.)

Bellotti's concepts are very similar to those of social translucence. In particular, feedback allows for accountability and "social access controls" within a shared space, which could allow for more flexible approaches for access. For example, a person might need to allow access to information in an emergency; they may be willing to maintain a relaxed set of controls, so long as they were notified whenever people made access to those items. The expectation is that people can provide a reasonable account of *why* they accessed certain files; in this way, they are accountable for their actions. In addition, Bellotti suggests that users should be informed when someone fails to gain access to an item. This could be due to attempted misuse, but it could also be due to a colleague being

locked out due to permission errors. In either case, making users aware of the information gives them a chance to take appropriate action. Fundamentally, such an approach requires visible feedback: there needs to be a trace of activity upon those shared objects in order to know what actions are being performed on them by other users.

4.2.2 Collaboration Systems

There are a number of special-purpose collaboration applications and systems that include support for file sharing. One early prototype model, GroupDesk, was designed to provide "shared awareness" by displaying events on shared artifacts (including files); this process involves presenting information on past and ongoing actions of collaborating users (Fuchs, Pankoke-Babatz, & Prinz, 1995). GroupDesk was neither evaluated nor developed into a functional collaborative system (remaining at the prototype stage), but did provide a model for the design of other collaborative tools.

A number of commercial networked collaboration systems have been developed that use specialized software to create a dedicated workspace. These tools usually support file sharing, and provide users with information about shared files, primarily notification of updates to documents and simple version control. These systems include Groove Virtual Office (Groove Networks, 2005) and NextPage (NextPage Incorporated, 2006). These systems are somewhat limited, in terms of the information they provide and, in some cases, the types of files that one can use with them. For example, Groove provides only update notification (there is no archive of activities over time), and NextPage only supports a few specific Microsoft file types (namely, Word, Excel, and Powerpoint).

An alternative to these systems is web-based collaboration, which provides a dedicated online workspace for access through a web browser. As with the tools described above, these systems usually include notifications of updated files and version control. Examples of these web-based systems with simple file sharing features include Basecamp (37signals LLC, 2005) and GoPlan (Webreakstuff, 2007). Certain web-based workspaces provide more detailed sharing information, including a history of activity on shared files; these include Copper (Element Software Pty Ltd, 2003) and BSCW (Bentley et al., 1997). These systems log and display all events on files, including all viewings and modifications by users.

These solutions provide a contribution to further work in file sharing support, but they are in themselves incomplete. One problem with most of these tools, in terms of their value to research, is that they are proprietary or commercial (or both); there have been no publications describing their design rationale, or publicly available user studies evaluating their effectiveness. Although they can provide some ideas for design, they cannot in themselves be used as a reliable research foundation. (BSCW (Appelt, 1999) is an exception, as it is part of a research project; however, its developers have not performed a formal evaluation of the tool or its design elements.). A greater limitation is that these are *dedicated* workspaces for collaboration: these tools are specifically deployed and maintained for particular groups and projects. This has two implications.

First, as my survey data indicates, these systems are not widely used, particularly within a personal context. Therefore, the majority of file sharing activities take place outside of these dedicated systems (such as sending email attachments). Any solution that takes into account actual sharing practice needs to encompass more than special-purpose workspaces.

Second, dedicated collaborative systems are much simpler than most desktop environments, in terms of the people with whom objects are shared and the overall complexity of file management. When a dedicated space is created, people are specifically invited into that space, which dramatically limits the scope of collaboration. One implication of this is that the security management is simplified: the general public, for example, will not have access to these workspaces. In addition, a limited number of shared files are placed into the workspace: those that are relevant to the particular project. In everyday practice in a desktop environment, shared files are co-located with private files, and the same file may go to many different recipients through different applications. This makes the file management problem much more complex: in such an environment, users need specialized support (e.g., through specially-designed interfaces) to determine what they have shared with whom.

4.3 *Usable Security and Access Control*

Another fundamental aspect of file sharing is the management of limitations: although some sharing is very open and public, there are also many examples of placing boundaries on the items shared, and the parties with whom they are shared. In this way,

file sharing is influenced by the field of security, which provides the tools and techniques that people use when managing technical limitations on their files. The emergent field of "usable security" focuses on the human factors of security, in an attempt to help users to better manage the complexities and challenges of computer security (Cranor & Garfinkel, 2005). Research from this domain, which is pertinent to file sharing, is discussed in the following section.

4.3.1 Usable Security

It has been recognized that computer security is difficult to use for many people; user studies on such diverse tasks as using encrypted email (Garfinkel & Miller, 2005; Gaw, Felten, & Fernandez-Kelly, 2006; Whitten & Tygar, 1999), online financial transactions (Schechter, Dhamija, Ozment, & Fischer, 2007; Wu, Miller, & Little, 2006), and password management (Adams & Sasse, 1999; Renaud, 2005; Yan, Blackwell, Anderson, & Grant, 2005) reveal that computer security is often error-prone, cumbersome, and complicated. In an attempt to ameliorate these problems, researchers have developed guidelines to inform the design of security tools, systems, and techniques. Some of these guidelines are domain-specific (e.g., for web browser design), but others can be applied more broadly.

Smetters and Grinter (2003; 2002) suggested that making the security state visible to the user (by supporting implicit security in applications) could help to avoid mistakes and facilitate better security decisions. Similarly, Whitten and Tygar (1999) outlined a number of properties that make designing for usable security difficult. One of the challenges that they identified is lack of feedback: security configurations are often complex and difficult to encapsulate in an interface, but clear and comprehensible feedback is fundamental in preventing users from making dangerous errors.

Good and Krekelberg's (2003) study of peer-to-peer file sharing (described in Section 4.1.2.1) inspired work by DiGioia and Dourish, who described how one might apply *social navigation* to usable security (2005). The concept of social navigation was first presented by Dourish and Chalmers (1994). It refers to the idea that people might use the cues of others' actions in order to navigate through information spaces. For example, one might choose to view objects that many other people had read, on the basis that those items might be interesting. This would require that traces of activity be recorded on

objects; in the security paper, the authors state that "the central concept—that an artifact can display the accumulated pattern of activities that have been performed over it—is a key element of the social navigation approach" (DiGioia & Dourish, 2005). They describe how this information might be helpful in managing security. For instance, a person could more easily detect deviations from normal behaviour, such as certain files being unusually active; this could alert them to the fact that they have not configured their sharing settings correctly. The authors present some design ideas (such as making a "messy pile" of frequently accessed files), which could be applied to such applications as KaZaa, to provide feedback that assists users in sharing the right set of files.

These ideas of usable security, which focus primarily on the presentation of information about system and activities, can be applied to many areas of usable security, including file sharing. There is also more focused research that is more specific to the security aspects of file sharing: in particular, on the topic of access control. Research results in access control are presented in the following section.

4.3.2 Access Control

Access control is a means of managing "...who can have access to a resource, under what conditions access can occur, and what those accessing the resource are allowed to do)" (Stallings, 2006, p. 17). In the context of sharing files, this often refers to the setting of file permissions on folders and individual files. This is only one aspect of file sharing, but studies on the usability of file permissions demonstrate some of the complexities and problems that people encounter when they perform this task.

Cao and Iverson (2006) investigated the usability problems of access control lists (ACLs) in the WebDAV (Web-based Distributed Authoring and Versioning) system, which controls access to a distributed file system. (An *access control list* is a set of permissions; in this case, file permissions.) An analysis of the WebDAV access control mechanism indicated problems with the conceptual usability of the system: users expended significant effort to determine how to configure the desired access rules. Cao and Iverson developed an "intentional access model" prototype, which included additional feedback for users. A preliminary lab study comparing this prototype to the WebDAV system (using ten participants) indicated that the prototype decreased errors, improved speed and accuracy, and improved user confidence and satisfaction.

Another preliminary research project on file permission interfaces led to the RubaViz permission visualization prototype (Montemayor, Freeman, Gersh, Llanso, & Patrone, 2006). RubaViz was designed to display multiple representations of access control rules, particularly to help system administrators to interpret the outcome of setting specific permissions. The prototype has not been fully implemented or evaluated, but it does reinforce the opinion that visual information in a file permission interface can improve security.

4.3.3 Summary: Usable Security and Access Control

Research in the areas of usable security and access control has provided tools and techniques that can be applied to file sharing issues. In particular, this research demonstrates the value of providing clear information to users, to assist them in understanding the complexities of security. This approach has been grounded in theory, and applied in practice, which provides a solid foundation for its application in the file sharing domain.

Beyond the security issues of file sharing, there are also issues of digital information management: that is, how people manage their files, particularly within a group context. Research within this field, and its relevance to file sharing, is described in the following section.

4.4 *Personal and Group Information Management*

Personal Information Management (PIM) is a growing area of research that involves "the practice and the study of the activities people perform in order to acquire, organize, maintain, and retrieve information for everyday use" (Jones & Bruce, 2005, p. 2). Management of files, including shared files, is a major part of this domain. One set of PIM activities involves the mapping of information to need; PIM researcher William Jones calls these *m-level* activities, such as "maintaining" and "measuring" (2005, p. 14). One collaborative *m-level* activity Jones describes is "managing privacy, security, and the distribution of items in a PIC [Personal Information Collection]" (2005, p. 15). Jones indicates that people need to, and want to, distribute some personal information, but they also need control over this distribution. Thus, Jones suggests that one beneficial element for security and privacy for PIM would be to provide user interfaces to help users maintain appropriate control, by guiding their choices.

The collaborative aspects of information management are significant enough to have been recognized as an emergent area termed Group Information Management (GIM). GIM involves "how personal information is shared with a group, emphasizing the norms that underlie sharing, as well as the ways participants negotiate these norms in response to the tensions that sharing inevitably produces" (Erickson, 2006). A number of challenges with managing shared information in a group have been identified, a number of which relate to shared file management. These include coping with dynamic group membership and the multiple contexts in which people share the same information (Erickson, 2006; Jones & Bruce, 2005). As with the privacy and security aspects of PIM described above, there is speculation as to how users can be provided with better support for their GIM tasks; in particular, is it feasible (and beneficial) to provide "awareness of usage, correction of errors, [and] retraction upon completion of task" (Erickson, 2006). This research area is still largely unexplored, which means that there are many unanswered questions and few solutions. Further work on shared file management, particularly on interface issues, may help illuminate these theoretical ideas.

There have been some PIM projects that incorporate elements of shared file management. The "Recent Shortcuts" tool (Tang, Lin, Pierce, Whittaker, & Drews, 2007) is designed to give users access to recently accessed objects (including files) across multiple applications. Users can display a list of people with whom they have interacted (through email or instant messaging), files recently received in email, and files that they have recently accessed. Recent Shortcuts is illustrated in Figure 2, from Tang et al. It shows a sample listing: the user is selecting a file to attach to an outgoing email (2007). (The tool was created after an observational study of knowledge workers; a prototype received preliminary evaluation.) Recent Shortcuts is intended to simplify coordination. Users often need multiple applications to complete a task, and Recent Shortcuts allows users to carry a recently-used object, such as a file or a username, from one context to another. Recent Shortcuts focuses primarily on temporal aspects of user tasks, rather than the specifics of file sharing, but the project demonstrates the value of supporting objects across multiple contexts. My file sharing surveys (see Chapter 3) demonstrated that this is definitely needed in a file sharing context, given the use of multiple applications for sharing.

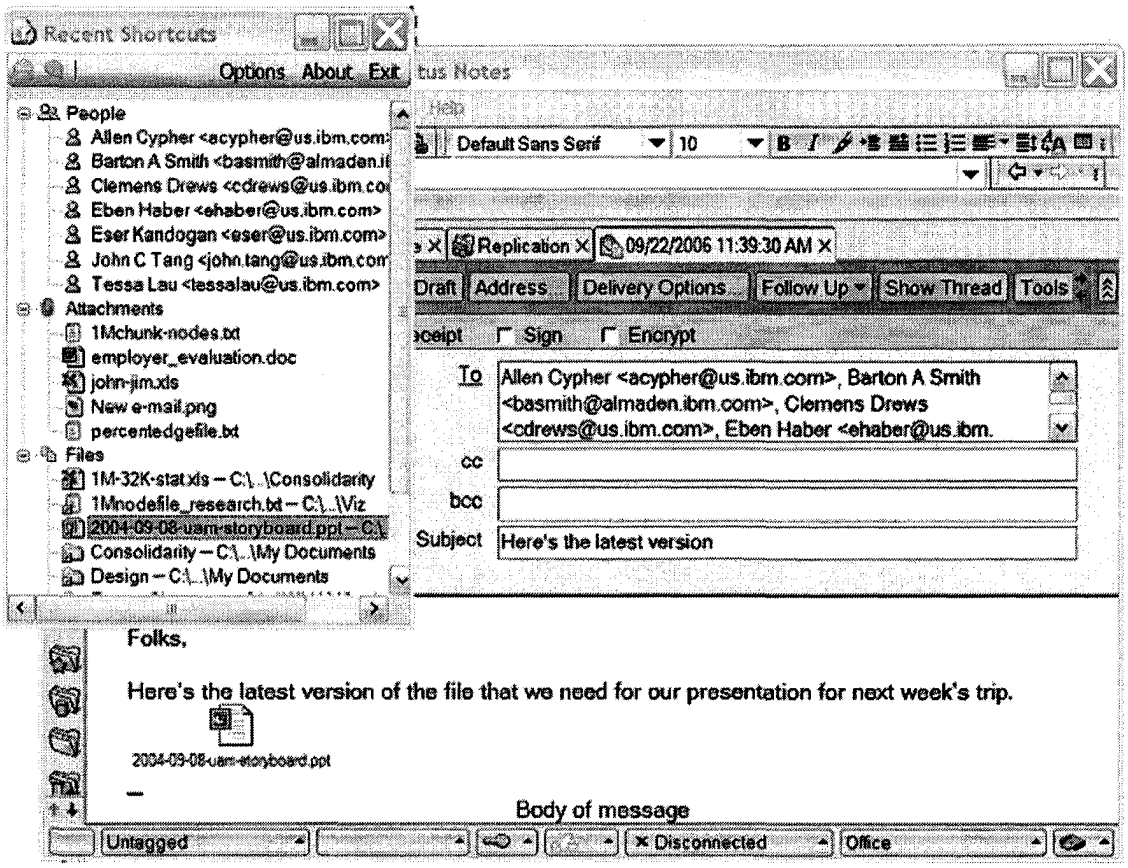


Figure 2: Using Recent Shortcuts to compose an email with an attachment (from Tang et. al, 2007)

Another PIM tool with relevance to file sharing is the Social Network and Relationship Finder, or "SNARF" (Neustaedter, Brush, Smith, & Fisher, 2005). SNARF is designed to help users prioritize their email handling, by providing a visualization of their email-based relationships. The SNARF prototype aggregates social metadata about correspondents, such as the number of emails sent to a recipient or the number of unread emails from a particular sender. The researchers hypothesized that providing social information would help people to manage large amounts of email. Although the prototype itself has not been formally evaluated, its design was based on a survey of email users, which suggested that presenting social relationship information would be beneficial (Neustaedter, Brush, & Smith, 2005). In a similar vein, presenting information about "file sharing relationships" could help people to manage their shared objects; it could also reveal information about their collaborative patterns (such as the popularity of a specific shared file). In addition, shared file information could be used as a metric for

such tools as SNARF, as one more component of its "social sorting" data (e.g., how often do I provide files to this person?).

4.5 *Summary: Current Findings and Future Directions*

My survey data, as described in Chapter 3, suggested that there is insufficient information about file sharing being provided to the user. This idea is echoed in the work of Voida et al. (2006), who indicated that people have trouble remembering which files have been shared with whom. Other file sharing usability studies (de Paula et al., 2005; Rode et al., 2006) also indicated that providing better information could improve the usability of file sharing applications. In particular, they highlighted how users require appropriate feedback in order to understand what is happening when they share files.

The idea of providing clear feedback also resonates in the CSCW and usable security fields, as shown in both theory (e.g., Bellotti, 1996) and practice (e.g., Cao & Iverson, 2006). Similarly, work in personal information management showed how additional activity cues can assist users to carry out file-related tasks (e.g., Tang et al., 2007), particularly in group contexts (e.g., Erickson, 2006).

There are a number of ways in which file-sharing information could be presented to users, including data about activity and system settings. The survey respondents indicated a desire to know about activity on shared files; particularly, to know who used them. In addition, people wanted to be sure that they knew who could access their files. People also interacted with different groups; this suggests that it could be helpful to support viewing sets of files associated with different groups, in such a way that is was made clear which files each group had access to.

Maintaining a history of sharing on an object has benefits both for security and for collaboration. For example, if a file is erroneously made too public, it would be helpful to know whether anyone had accessed that file while it was exposed; the degree of "damage control" depends on the degree of exposure. In collaboration, it can be useful to see what sorts of group sharing took place over time: if a new member joins a group, it may be necessary to determine which files were shared over the last year, to ensure that the newcomer is given access to all the relevant documents. (That set of documents might include, for example, some files sent over email, as well as files in shared folders.)

Other information that could be useful is notification, to allow users to keep track of changes in the system. For example, survey respondents demonstrated a use of temporary sharing. In these cases, it might be beneficial to send an alert after a certain action was detected (such as a file being downloaded from a shared directory); the user could then have a reminder about a necessary reset. A person might also want to monitor specific sensitive files, so that they could be aware of unexpected activity. Another useful notification would be when access was denied: this could be due to a mistake, or to an intrusion.

Presenting sharing information in a clear fashion is also important. There may also be many different items shared, in multiple contexts; it could be helpful to allow a user to create a customized view that shows only the most relevant information: for example, "all files shared with Bob", or "all accesses on report.doc in the past 2 months."

Finding a way to present relevant information, so that users know what is happening in a system (e.g., activities and settings), has been addressed in research under the term *awareness*. That is, if users are made more aware of what is happening to shared files, and what the access control settings are on shared files, this could help them to ensure that shared files are managed correctly. Thus, this thesis research focuses on how we can support awareness, in the context of file sharing.

The next step in developing a solution for file sharing that involves awareness is to specify exactly what is meant by the term *awareness* in this context, and what the components of file sharing awareness should be. To this end, a review of the literature in the domain of awareness and CSCW was performed, which led to a framework to be used in design. This framework, along with its foundations, is presented in detail in the following chapter.

Chapter 5. File Sharing Awareness Framework

The role of awareness in file sharing is explored in this chapter. First, the term *awareness* will be defined, as it applies to this thesis research. Next, the model of awareness described by the *perception-action cycle* will be presented, along with an outline of *situation awareness*. These concepts have been incorporated into *workspace awareness*, which originates in the CSCW domain. *Frameworks of workspace awareness* have been developed for both synchronous and asynchronous collaboration; two such frameworks will also be described in detail. These workspace awareness frameworks were used to develop a *framework of awareness for file sharing*, which is described in the final section of this chapter.

5.1 *Definitions of Awareness*

As a first step, it is useful to define exactly what is meant by the term *awareness*. There have been several definitions proposed, because this term is quite fluid and has been adopted by many sub-domains within CSCW and beyond (Robertson, 2002; Schmidt, 2002). One simple definition of awareness was developed by Endsley (1995): "knowing what is going on". This is an inclusive definition, but overly broad. Gutwin and Greenberg (2002), in their work on workspace awareness, have synthesized Endsley's work along with concepts of Norman (1993) and Adams et al. (1995) to derive four characteristics of awareness:

1. Awareness is knowledge about the state of an environment bounded in time and space.
2. Environments change over time, so awareness is knowledge that must be maintained and kept up to date.
3. People interact with and explore the environment, and the maintenance of awareness is accomplished through this interaction.
4. Awareness is a secondary goal in the task—that is, the overall goal is not simply to maintain awareness but to complete some task in the environment.

These characteristics help to define awareness more specifically, although these elements are general enough to be applied to multiple domains (including file sharing). It is helpful to consider these definitions in conjunction with theoretical model of awareness, in order to illustrate its core concepts. One such theoretical model, originating

from cognitive psychology, is the *perception-action cycle*, which describes how people maintain awareness of their environment.

5.2 *Neisser's Perception-Action Cycle*

Neisser's model describes the interaction between a person and their environment, in terms of information gathering. It is depicted in Figure 3 below; this is a combination of a representational figure from Gutwin and Greenberg (2002) as well as the original by Neisser (1976).

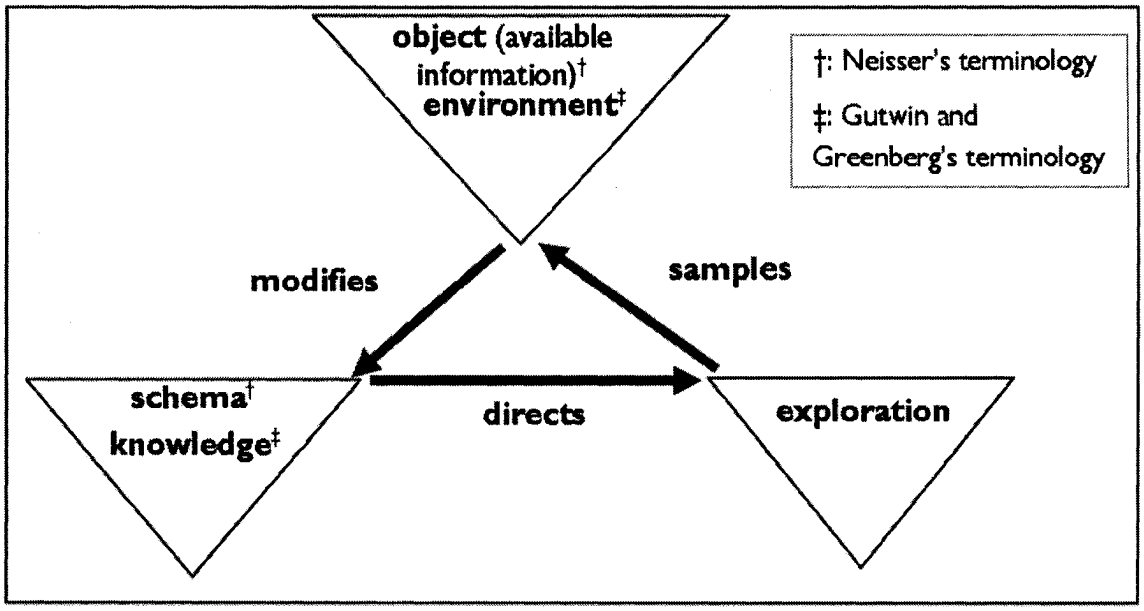


Figure 3: Neisser's perception-action cycle, as described by Neisser (1976) and Gutwin and Greenberg (2002)

The perception-action cycle is a "cognitive framework for the interdependence of memory, perception, and action" (Adams et al., 1995). Neisser's model illustrates the interrelationship between a person's existing knowledge and their gathering of new information (*exploration*), as well as the interaction between a person and their environment. Note that the model is cyclical: a person explores the new environment and updates their knowledge, and their new knowledge affects the ways in which their exploration proceeds.

Neisser's model is useful for describing how people maintain awareness: they begin to interact with their environment, with some understanding of the context and an initial idea of what to look for. They explore the environment, and re-interpret their existing knowledge based on what they discover, incorporating changes that they

encounter (new knowledge and changes to the environment itself). This is how they "know what is going on," to use Endsley's (1995) words, and how they can make predictions on what will happen in the near future. The perception-action cycle has proven to be a constructive model for much human-computer interaction (HCI) research, having been applied to topics ranging from ambient displays (Wisneski et al., 1998) to multimodal interaction (Wolff, De Angeli, & Romary, 1998).

In addition to Neisser's model, CSCW awareness research also draws from the field of *situation awareness*. As with the perception-action model, situation awareness research identifies, categorizes, and defines the ways in which people perceive and interact with their environment. These concepts are explored in the following section.

5.3 *Situation Awareness*

Endsley has developed much of the foundational work in the field of situation awareness (SA), and provides a basic definition with three major components: "Situation awareness is the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the perception of their status in the near future." (1995, p. 36). These three elements can be briefly described as:

1. **perception of the elements in the environment within a volume of time and space**: people must be able to gather information from their environment, and attend to the items that are most relevant to their task
2. **comprehension of their meaning:** people must be able to formulate meaningful conclusions from the environment, integrating their existing knowledge with new data
3. **perception of their status in the near future:** people must be able to develop predictions in the near term, anticipating changes to the environment and to data

A similar definition of situation awareness (in the domain of aviation) comes from Hourizi and Johnson (2004), who break down the process of awareness into four sub-components:

> "...raw data from the environment has been
> 1. available...
> 2. perceived
> 3. attended to in some manner, and
> 4. subject to further, higher-level cognitive processing." (p. 160)

Situation awareness has been most commonly applied to highly-dynamic environments with high information load, such as air traffic control and tactical systems

(Adams et al., 1995; Endsley, 1995). However, its central concepts can be applied to other domains: Gutwin and Greenberg (2002, 2004) adapted them in defining workspace awareness: "up-to-the moment understanding of another person's interaction with the shared workspace." Gutwin and Greenberg (2002) outline two major differences between situation awareness and workspace awareness. First, workspace awareness is a specific type of SA, which is mainly concerned with *cooperation between people*; situation awareness generally refers to a single user interacting with a workspace to complete a domain-specific task. Second, most collaborative tasks within shared workspaces do not involve extremely dynamic environments or information overload. However, despite these differences, both these domains require that users be given additional information to help them make sense of the environment in which they operate. In the case of computer-supported collaboration, people need additional information to tell them about the activities of others; this includes awareness information that would be readily available if the people were working together in a face-to-face environment.

Because file sharing represents a means of collaboration, the workspace awareness research of Gutwin and Greenberg is well-suited for developing a framework for awareness of file sharing. File sharing involves basic aspects of situation awareness (e.g., setting within the computing environment) as well as those specific to collaboration (e.g., group activities on shared objects). Gutwin and Greenberg (2002) created a framework for workspace awareness, which was developed primarily from the two concepts described above: situation awareness, and the perception-action cycle. Their framework is described in detail in the following section.

5.4 *Workspace Awareness Framework*

Gutwin and Greenberg's (2002) workspace awareness framework is composed of three main components: the types of information that make up workspace awareness (WA), the mechanisms people use to gather WA, and the ways in which people use WA in collaboration. These three elements are described below, and are depicted in Figure 4.

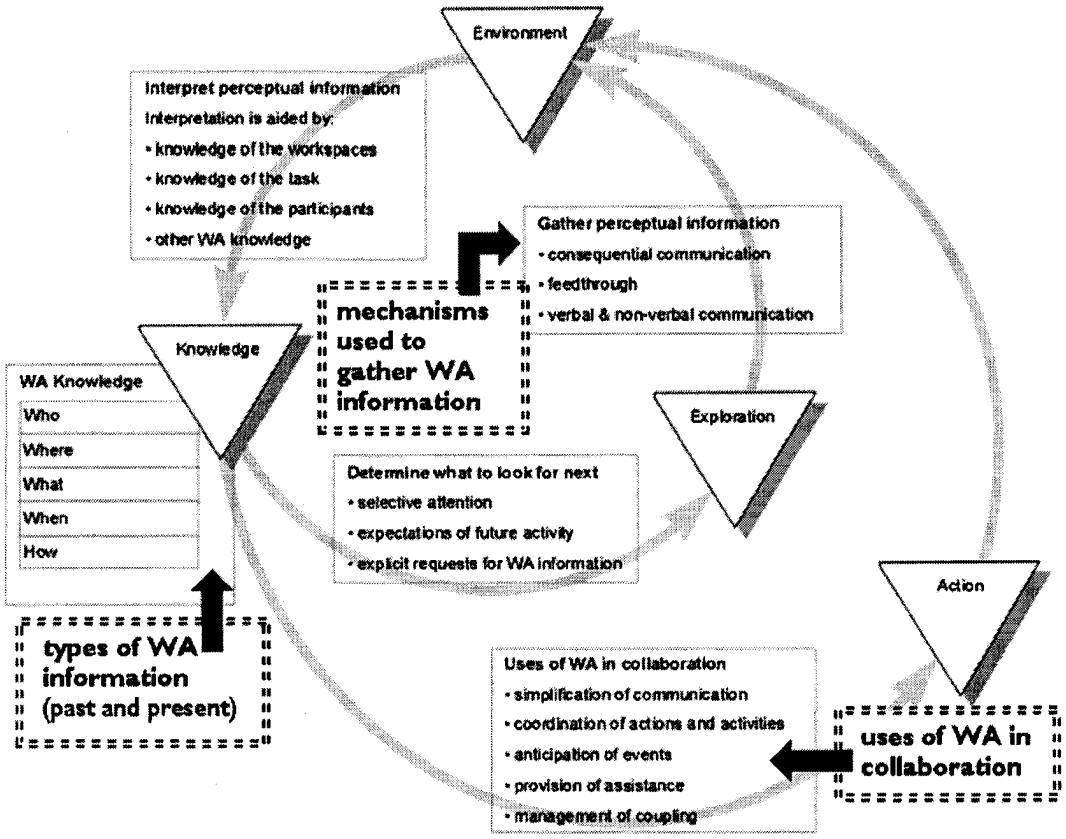


Figure 4: Workspace awareness (WA) model, adapted from Gutwin and Greenberg (2002).

5.4.1 Types of Workspace Awareness Information

This is the "knowledge" component of the framework: what the user knows, and can interpret from the environment. Gutwin and Greenberg (2002) describe this as the "who, where, what, when, and how" questions. Note that these elements exist in both the past and the present: collaboration is carried out over time, and there must be consideration of the temporal aspects of the environment.

For example, one question about "who" in the present would be "Is anyone in the workspace?" A question about the past would be "Who was here, and when?" Gutwin and Greenberg (2002, 2004) provide detailed examples of these questions in their framework discussion; these are presented in Section 5.5, in Table 20 and Table 21. The inclusion of temporal information is also central to file sharing, as this type of interaction takes place over varying periods of time (in both long-term and short-term collaborations).

5.4.2 Mechanisms Used to Gather Workspace Awareness Information

A collaborative environment provides different types of awareness information, which the user must gather and interpret. Gutwin and Greenberg (2002, 2004) identify three major elements of information gathering in collaboration, from both face-to-face (physical) and computer-supported methods of interaction. The first is *consequential communication*: information transferred through observing a person's interactions within an environment (Baker, Greenberg, & Gutwin, 2001; Segal, 1994). The second is through artifacts—the items with which people interact, which can carry cues about the environment. When artifacts change (e.g., through interaction), the observed change provides information about actions that are occurring; this process is referred to as *feedthrough* (Dix, Finlay, Abowd, & Beale, 2003). The third source is through verbal and non-verbal communication, which is referred to as *intentional communication*. These are the methods by which people gather information from the environment, which is interpreted in conjunction with knowledge of such factors as the workspace, task, and participants (see Figure 4). This updated knowledge feeds into the cyclical exploration process, as people determine what they need to look for next as they complete their task.

5.4.3 Uses of Workspace Awareness in Collaboration

The final element of this workspace awareness framework is the set of ways in which people use workspace awareness to carry out specific collaborative goals and tasks. Gutwin and Greenberg (2002) have expanded upon Niesser's model by adding a new "action" link, as shown in Figure 4; this indicates that "people take action based on their knowledge as well as exploring the environment." Examples of this include assistance (helping others carry out a task) and coordination of actions (e.g., facilitating division of labour, preventing conflicts).

5.4.4 Summary of WA Framework

Gutwin and Greenberg (2004) have provided a comprehensive framework that describes the components of workspace awareness and their relationship with one another. This framework is suitable for many types of collaboration, although it most closely describes synchronous, tightly-coupled interaction, with a small group within a shared workspace. Examples of this type of collaboration would be a team editing a document, or a group creating a page layout for typesetting. This type of interaction is

fairly dynamic, requiring that awareness be maintained throughout the task in order to promote coordination among team members. When considering the task of file sharing, it is evident that it shares some aspects with this basic scenario: there is an environment (containing files, with different properties), which changes as people interact with it (e.g., sending, accessing, and updating files). There are aspects of collaboration, where people interact mainly through the shared artifact (the file), and in some cases, using other communication along with the file transfer (e.g., an email message with an attachment). There are also temporal aspects: people may interact over a short or long period, and a particular file may undergo many changes over time. However, because file sharing is a very limited type of interaction, and because it is mainly asynchronous, the framework of Gutwin and Greenberg needs to be modified in order to more accurately describe file sharing in particular.

There are two main differences to be considered. The first is that the artifacts are extremely important in file sharing. In fact, they are the mechanism through which people are interacting: the *sharing* is the focus of interaction. Therefore the *knowledge* component of the framework (e.g., who, what, and how) is extremely important, because these items describe the shared artifacts and the environment in which they are shared. This element is already present in the basic framework, but for file sharing, it becomes central; there is less emphasis on the "uses in collaboration", for example, as this aspect is much less important in a file-sharing situation.

The second major difference is the asynchronous nature of file sharing. In some cases, files are sent with the expectation that some type of further collaboration might immediately take place, but this is not always the case. In many instances, the sharing is one time only, and is a one-way interaction. For example, a person might send a document to a colleague to read, rather than to update; once the transfer is successfully made, the collaboration is complete. In other cases, files may be updated, but this update is not made immediately; rather, the file is sent and an updated version is returned, after some time has passed. In many file sharing scenarios, there is no need to maintain synchronous coordination between the senders and the recipients; what is of paramount concern is that the item is successfully shared with the desired recipients.

The asynchronous aspects of collaboration were explored by Tam and Greenberg (2006), who adapted the WA framework to incorporate asynchronous tasks. They started from Gutwin and Greenberg's original framework, modifying it to make it more suitable for asynchronous interaction; for example, people editing a document over multiple sessions with several iterations. This asynchronous model of awareness has many aspects in common with file sharing, and therefore provides a useful foundation; it is described in detail in the following section.

5.5 *Asynchronous Workspace Awareness*

Tam and Greenberg's (2006) work focuses on asynchronous change of artifacts (*change awareness*), defined as "the ability of individuals to track the asynchronous changes made to a collaborative document or surface by other participants" (p. 584). Their focus is mainly on how people become aware of changes to a shared artifact that are visible through inspection of that artifact, rather than through other types of communication (e.g., an email message stating what changes have been made). As their focus is on artifacts, this suggests their framework may be applicable to shared files.

One requirement for change awareness that Tam and Greenberg identified was to determine what types of information were necessary for people to comprehend change in the shared workspace. These information elements correspond to the "workspace awareness knowledge" in Gutwin and Greenberg's model (see Figure 4), although modified to be more suitable for asynchronous collaboration. Workspace awareness (WA) knowledge, as described by Gutwin and Greenberg, consists of a set of elements that people need to track in the workspace. These elements were detailed in their framework papers (Gutwin & Greenberg, 2002, 2004); the set of elements incorporates both past and present occurrences, and answer questions relating to the "who, what, where, when, why, and how" of interaction. These elements are summarized in Table 20 and Table 21 below:

Category	Element	Specific Questions
who	presence	Is anyone in the workspace?
	identity	Who is participating? Who is that?
	authorship	Who is doing that?
what	action	What are they doing?
	intention	What goal is that action a part of?
	artifact	What object are they working on?
where	location	Where are they working?
	gaze	Where are they looking?
	view	Where can they see?
	reach	Where can they reach?

Table 20: Elements of WA knowledge relating to the present, from Gutwin and Greenberg (2002, 2004)

Category	Element	Specific questions
who	presence history	Who was here, and when?
what	action history	What has a person been doing?
where	location history	Where has a person been?
how	action history	How did that operation happen?
	artifact history	How did this artifact come to be in this state?
when	event history	When did that event happen?

Table 21: Elements of WA knowledge relating to the past, from Gutwin and Greenberg (2002, 2004)

Tam and Greenberg (2006) note that although the basic workspace awareness framework is a useful starting point, there is little information about the past, given that the original framework was more focused on the present (for synchronous collaboration). For asynchronous collaboration, a person needs to be kept up-to-date with changes that have happened while they have been away from the workspace, and therefore support for *history* becomes much more critical. Tam and Greenberg (2006) provide a set of high-level questions that need to be answered, such as "Who has made the changes?" and "When did the changes take place?" which are similar to those in the original WA framework. These questions are summarized in Table 22 below.

Category	Information Elements	Specific Questions		
		artifact-based view	person-based view	workspace view
where	location history	Where was this artifact (when I left)?	Where in the workspace has a person visited?	Where have people been in the workspace? Where were artifacts in the workspace?
	gaze history	Where is the artifact now?	Where in the workspace has a person looked at?	Which parts of the workspace have people looked at?
	edit history	Where has this artifact been during the time that I have been away?	Where in the workspace has a person made changes?	Which parts of the workspace have people made changes in?
who	presence history	Who has looked at this artifact?	Who has this person interacted with?	Who has been in the workspace?
	identity	Who has changed this artifact?		
	readership history		Who made changes with this person?	Who has looked at the workspace?
	authorship history			Who has made changes to the workspace?
what	action history	What changes have been made to the artifact?	What artifacts has a person looked at? What artifacts has a person changed? What activities has a person engaged in?	What changes have occurred in the workspace? What artifacts were viewed? What artifacts were changed?
how	process history outcome history	How has this artifact changed?	How has a person changed things?	How has the workspace changed?
when	event history	When was this artifact changed? When was a particular change to this artifact made? In what order were changes made to this artifact?	When did a person make changes? When did a person make a particular change? In what order did this person make changes?	When were changes made to the workspace? When did a particular change in the workspace occur? In what order did changes to the workspace occur?
why	cognitive history motivational history	Why was this artifact changed?	Why did a person make that change?	Why was that change made in the workspace?

Table 22: Elements of asynchronous workspace awareness, from Tam and Greenberg (2006)

However, Tam and Greenberg (2006) also add a novel dimension to the framework, by including support for different *views* of the workspace. They describe how a person may need to gather information on changes by querying the workspace from different perspectives, in particular, from:

- the artifacts that exist in it (**artifact-based** view)
- the people who work within it (**person-based** view), or
- the workspace as one locale or a collection of related locales (**workspace-based** view)

The idea of locales and the workspace-based view require further explanation: a *locale*, which originates in the work of Fitzpatrick (2003), is defined as a "place that arises from the use of space and its resources by and for a particular group of people." Fitzpatrick developed a locales framework to be used in CSCW. A central component of this framework is support for individual views: because people work in many different social worlds, with different tasks and groups, there is a need to support a "viewset" that is organized according to a person's current task and needs (Greenberg, Fitzpatrick, Gutwin, & Kaplan, 1999). In addition to locales, Tam and Greenberg refer to Berlage and Sohlenkamp's (1999) artifact metaphor. This metaphor incorporates a "common artifact" which gathers and displays activity information in order to support group awareness. The types of information in the artifact are quite similar to the "WA knowledge" elements, such as "who is doing something."

Starting from the original WA framework, and related work on locales and the artifact metaphor, Tam and Greenberg (2006) created a set of elements for asynchronous WA. Because these elements are highly relevant for developing a file sharing framework, they are presented in full in Table 22.

These elements can be applied in designing a variety of CSCW systems. Tam and Greenberg (Tam, 2002; Tam & Greenberg, 2006) used them to provide awareness support in a collaborative drawing system, which displays such information as where a figure has been moved from (before its current location), and who performed this movement. The drawing system also provides filters to give users an individualized view that includes the information of most interest, such as all changes made by a certain person (person-based view) or all changes to one drawing element (artifact-based view).

5.6 *Summary of Awareness Definitions and Frameworks*

There are several related threads running through the research summarized above, all of which relate to awareness and many of which involve collaboration. The major ideas from that work will now be recapitulated, in order to emphasize how each piece of work contributes to the development of awareness frameworks. This outline will be followed by a discussion of how these elements can be applied specifically to support awareness in the file sharing domain.

Through the perception-action cycle of Neisser (1976), and the field of situation awareness (particularly the work of Endsley (1995)), the core elements of awareness have been identified. In particular, people perceive elements of the environment, which are gathered through exploration. A person's existing knowledge is modified through interaction with the environment, and this cycle continues as new knowledge directs further exploration. In this way, people maintain awareness of the environment with which they interact.

Gutwin and Greenberg (2002) expanded this basic awareness model to include collaboration, which requires awareness of *other people*. In their framework (called workspace awareness, or WA), several facets of awareness were identified. Specifically, they described types of WA knowledge, mechanisms used to gather WA knowledge, and ways in which WA is used in collaboration. Of these elements, the one with the greatest relevance to file sharing is WA knowledge: it incorporates artifacts (such as files) and events on artifacts. Another vital aspect is that WA knowledge includes both the past and the present, which applies to the temporal nature of file sharing.

Tam and Greenberg (2006) adapted the workspace awareness model so that it was more suitable for *asynchronous* collaboration. The aspect that they focused on was that of WA knowledge, in terms of change awareness: noting and interpreting changes that have occurred when a person returns to a shared workspace after an absence. Two central elements in this model were the use of history, and support for individualized views. Being able to view history helps users to maintain awareness in asynchronous collaboration. In addition, individualized views (in particular, artifact-, person-, and workspace-based views) allow a person to group related items in a way that helps them to make sense of events, and to view only the most relevant details.

Because file sharing is a collaborative activity that is often asynchronous, this model can be readily adapted for file sharing awareness. In particular, a shared file gathers a "history" over time, when shared with multiple people in multiple contexts. In addition, individualized views allow a person to group and filter files in different ways, such as showing details for one file (artifact-based), or viewing the set of files shared with a specific colleague (person-based).

There are several valuable components and ideas within this research that can be applied to a framework for file sharing. Of greatest relevance are the ideas of improving knowledge by presenting data from the environment (such as the workspace ideas of "who" and "when"); of supporting both the past and the present (temporal aspects); of providing extensive history information (such as event and artifact history); and of supporting individualized views. These ideas are incorporated into the file sharing awareness framework described in the next section.

5.7 *File Sharing Awareness Framework*

Adapting the previous research on awareness to the domain of file sharing leads to the framework depicted in Figure 5 below. This framework, as well as the accompanying figure, borrows heavily from Gutwin and Greenberg's workspace awareness model (from Figure 4). Three of the components—environment, knowledge, and exploration—are from Neisser's model, with the addition of "action" from Gutwin and Greenberg's workspace awareness model. The user maintains awareness through a cycle: as shown in Figure 5. This model incorporates **knowledge**, which is **gathered** and **interpreted**, and **exploration** of an **environment**, through which the user determines **what to look for next**. These elements constitute awareness, which is **used** in managing file sharing.

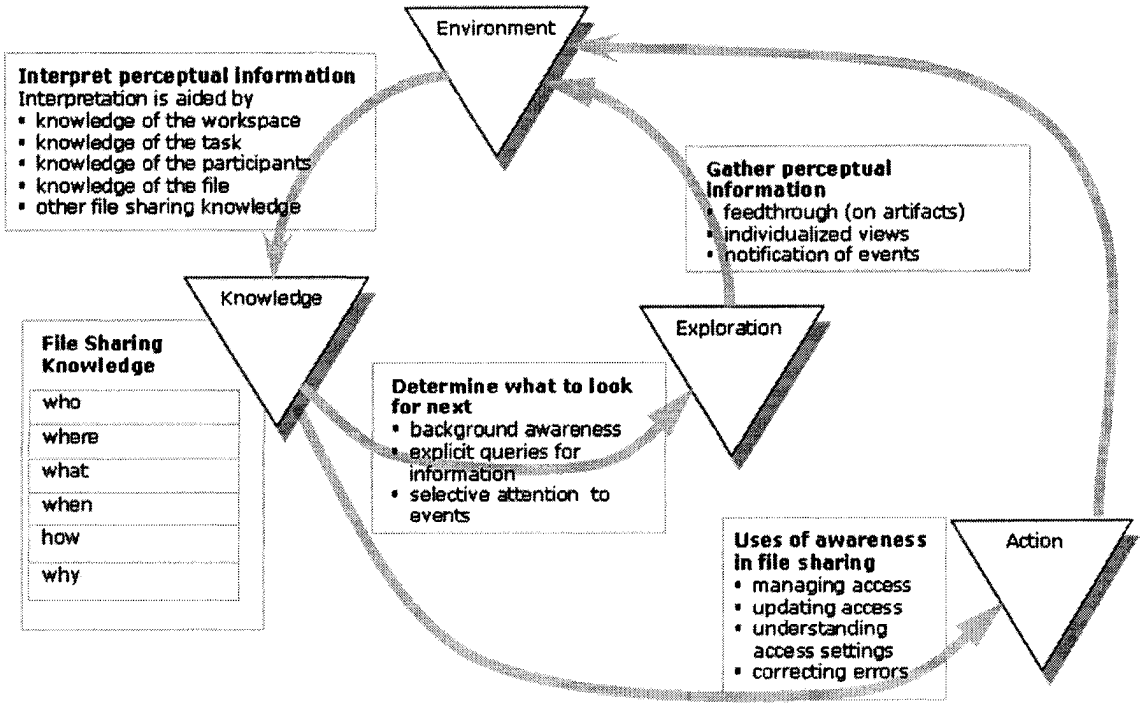


Figure 5: Awareness framework for file sharing

There are two major components to the file sharing awareness framework that require more detailed discussion. These are the first and second elements of Gutwin and Greenberg's framework (see Figure 4): 1. WA information, and 2. ways of gathering that WA information. These are the components (from Gutwin and Greenberg's framework) that have a significant influence on a framework for file sharing. The "uses" component is task-driven and domain-specific, and therefore outside the scope of this discussion. The focus here is on the other two components, which are strongly influenced by the shared artifacts of file sharing: the *information* about shared files, and the *ways of gathering* that information. These two aspects are described in detail in the following sections.

5.7.1 Information about Shared Files

The elements of workspace awareness, as described above, can be applied to the types of information that people need to know about their shared files. This information includes elements of the computing environment (e.g., properties of shared files) and activities of other people on common shared files. There is also a temporal nature to file sharing awareness, including aspects of both the past (history of activities on a shared file) and the present (current settings and activities). Because a file may persist over time,

a person might forget what interactions they had with it, such as who they sent it to, and therefore a file's *history* becomes important. As in the work of Tam and Greenberg (2006), individualized views are supported, including artifact-based and person-based. *Artifact* in this case refers to a file, or set of files, within the context of sharing. A set of files may be related to one another in a variety of ways: a person may wish to group and view all files that are part of one project or originating from one company, or all those that include the keyword "confidential." Because all the artifacts in this framework are files, the term *file* will be used instead of artifact, simply for the sake of clarity (i.e., there are no other kinds of artifacts to consider). *Person*, within the file sharing context, can also be a group of people (e.g., a file is emailed to four people). Note that the singular terms *file* and *person*, rather than files and group, will be used in the framework, with the understanding that the underlying examples may be extended to the plural case.

The resulting awareness framework has two types of customizable views (file and person) and two temporal aspects (past and present)[1]. These components are summarized in Table 23 and Table 24 below, followed by a more detailed description of each element.

		Specific Questions	
Category	Information Elements	file-based view	person-based view
where	location	Where is this file?	
who	accessibility	Who can access this file?	
	presence	Who is accessing this file?	
what	accessibility		What files can this person access?
	action	What actions can be taken on this file?	What actions can this person take on this file?

Table 23: File-sharing awareness elements relating to the present. Adapted from Tam and Greenberg (2006).

[1] There is also a "future" aspect, which is a minor part of this framework and deferred for later discussion.

		Specific Questions	
Category	Information Elements	file-based view	person-based view
where	location history	Where has this file been located?	
who	access history	Who has accessed this file? Who has a copy of this file?	What files has a person accessed?
what	action history	What actions have taken place on this file?	What actions has this person performed on this file?
how	distribution history	How has this file been distributed?	How has this person received files? What files have been distributed to this person?
	property history	How has this file's properties changed?	How has a person changed a file's properties?
when	event history	When was this file distributed? When has this file been accessed? When was this file created? When have this file's properties been changed? When was a particular change to this file made?	When was a file sent to this person? When did this person access a file? When did this person create files? When did a person change a file's properties? When did a person make a particular change?
why	cognitive history	Why was this file shared? Why were this file's properties changed?	Why was this file accessed? Why did this person change a file's properties?

Table 24: File-sharing awareness elements relating to the past. Adapted from Tam and Greenberg (2006).

These tables provide a high-level overview of the types of information that can be collected and displayed about file sharing: properties of the file themselves, the people with whom they are shared, and activities on shared files. This set of elements was derived by taking the framework of Tam and Greenberg (2006) and adapting it to the domain of file sharing. These elements are described in detail in the following section.

5.7.2 Elements of File-Sharing Awareness

The elements in the above tables are divided temporally into the past and present. One way in which this difference is vital is when considering the approach used to share files: "push-oriented" versus "pull-oriented" approaches (Voida, Edwards, Newman, Grinter, & Ducheneaut, 2006). A *push* approach is when the sender actively delivers the

file to the recipient, such as an email attachment or on a USB key. A *pull* approach is when the recipient retrieves the file from an accessible location; an example would be a file placed in a shared folder. Files that are distributed in a *push* manner generate a history of discrete sharing events, for each instance of the file being *pushed*; this means that the *past* is the place to look for this sharing information. In contrast, the *pull* approach requires that files be left in an accessible location to be downloaded in future, which makes the *present* settings of great relevance. Of course, there is some overlap between the categories, where the same files may be both pushed and pulled, but these basic distinctions help to illustrate the types of information that are most relevant in each case.

5.7.2.1 Elements Relating to the Present

The first set of awareness elements is concerned with the *present* state of file sharing: what files are currently accessible to others, and what activities are currently occurring. These elements are as follows:

location: knowing where a file is can help a person determine whether or not it is accessible to the right set of recipients. For example, a file on an internal file server that is blocked from external access will only be available to internal employees.

access: the question of "who can access this file" is crucial to file sharing; answering it is the means by which we determine whether our files are sufficiently public or private. The term *access* is used here in an operating systems sense: for a user to perform an action on a file, the file permissions for that file must be configured to allow that user to *access* it (Sandhu & Samarati, 1994). The set of people with access to a given file can range from "no one else" to "everyone", with many different subgroups in between. This question can be looked at in two ways: given a file, who can access that file; and given a person, what files can that person access?

presence: if another person is currently using a file, this indicates that access was successful (whether or not it was intended by the sender). In addition, if another person is using a file, then that file cannot generally be modified by other people; if changes were necessary (such as file permission updates), these changes would have to wait until the file was inactive again.

action: in this case, actions on files are those that can be performed by the file system, and are very coarse-grained. Examples include read, write (modify), and delete. (This does not include specific actions on the content of the file, such as editing a specific portion of the file.)

5.7.2.2 Elements Relating to the Past

The second set of elements is related to the *past* state of file sharing: how a file has been shared over time, and with whom, and what sharing-related actions have taken place on that file. There are several more questions regarding the past than the present, owing to the fact that events over time involve more detailed information than a snapshot of the present does. As in Tam and Greenberg's (2006) asynchronous model, the elements are all described in terms of a file's "history."

location history: knowing where a file has been stored (e.g., which directories or file servers) can indicate whether it had ever been accessed by the appropriate groups (or by unwanted parties). When this information is used in conjunction with access and event history (below), it provides feedback as to whether the file had been appropriately shared over time. For instance, if the file was on a public server for five minutes, but was not accessed during that period (before it was removed), then accidental file leakage is unlikely because the mistake was caught in time.

access history: the access history reveals who is using shared files. This information helps a person to determine whether sharing has been correctly configured: for example, if more people are accessing the file than expected (too public), or fewer people than expected (too private), then the settings could be repeatedly reconfigured. Access history is also key to supporting "social access control" (Bellotti, 1996), in which a person may be willing to configure lax controls so long as they know who is accessing their files.

action history: the action history involves simple events on shared files, such as read and write. This is closely tied to access history, but can provide more detail about the specific action that was taken. For example, when a file is accessed, it may be read, but not modified. Actions can also include failed actions, such as an attempt to write; this could indicate that permissions are set too tightly (if a colleague was unable to write), or that they are correct (if an unknown person attempted to write).

distribution history: given that files are sent in many ways, it may be difficult to remember all the instances of file sharing and its associated context. Sometimes, when people send a file, they send some associated information with it; an email attachment, for example, may contain a conversation that relates to the shared file (Duchenaut & Bellotti, 2001). Knowing that a file was emailed, or sent over IM, could cue to the user to refer to the associated conversation (e.g., email message) for further information.

property history: this is a lesser source of information, and refers to changes that have been made to the basic file properties over time. This includes items such as file permissions; the history would show, for example, when a file was made more or less accessible to others. It would also show when *other* people changed the properties; if this change was unexpected, or an error was made, then this information would indicate who should be contacted for clarification or corrections.

event history: this is a broad set of information relating to the "when" aspect of file sharing, revealing the timestamps on actions. Many of these events have been described above. In many cases, the exact time is not relevant: it is more important to know whether or not a specific event happened at all. However, time can sometime be useful for clarity; it allows for ordering of events (useful for cognitive history, below) and can show how actions relate to one another.

cognitive history: the final element from adapted from Tam and Greenberg's (2006) model is cognitive history. This is a high-level concept, by which a person attempts to answer the question "why" in relation to file sharing. This is not reflected directly in the environment, but is partially derived from information gathered from that environment. For example, to answer to the question "why were this file's properties changed?," one might see that a file read access failed (access history), followed by a change of file permissions (property and event history). This series of events shows that the original settings were too restrictive and therefore the person needed to open them for access.

5.7.2.3 Elements Relating to the Future

Note that both the past and the present have been described in this framework, but that the future has not. The reason for this is that the future is quite different from the past and present: if events have not yet taken place, and there is no explicit information about them. Users may anticipate that certain actions will occur, but these expectations are not

actual events in the environment. For the most part, the anticipation of future events is based on the current state of the environment; as time passes, the future eventually becomes the present. Therefore, the file sharing awareness elements relating to the present accommodate much of the future as well.

However, one additional aspect of the future can be incorporated into this framework. In the "push" model, a file is placed in a location with the expectation that it will be accessed in the future. This would be asynchronous interaction: there is a gap between the sender making the file accessible, and the recipient actually accessing the file. If users wish to be notified of this future event, they would set an alert. This alert is created in the present, for a future time. Unlike the aspects of the past and present, this is quite a simple item, and therefore it does not require individualized views. Instead, the user can simply be shown a list of future notifications: that is, events that they wish to be informed about, which have not yet occurred. This element is listed in Table 25.

Information Element	Specific Question
future events	What sharing-related events do I wish to be informed of in future?

Table 25: File-sharing awareness elements relating to the future

5.7.2.4 Summary: Awareness Information

The above model encompasses many elements that work together to support awareness of file sharing. It is based predominantly on the CSCW work of Gutwin and Greenberg (2002, 2004), and Tam and Greenberg (2006) in their research on workspace awareness. The above elements form one part of the basic framework: the *information* about shared files, and interaction with shared files, which can be made visible in the interface. Making this information readily available allows users to answer questions such as "What files has a person accessed?" and "What actions can be taken on this file?". This information has a temporal aspect, including the past, present, and future. For past and present information, details are presented via individualized views that reflect the perspectives of artifacts (files) and people. The future is also incorporated (although it is a minor aspect), by allowing users to view a set of planned notifications of future events that are relevant to file sharing.

Returning to the cyclical awareness model (as depicted earlier in Figure 3), we see that these elements work together to provide an up-to-date view of file sharing settings and events, through their presentation of the environment and the user's explorations of that environment. There is a second component to the framework: *how the user gathers* this perceptual information, as shown in the original WA framework in Figure 4. This component will be described in the following section.

5.7.3 Gathering Perceptual Information for File Sharing

When people interact with information from the environment, there are two aspects of the framework that are involved: the means by which people explore the environment, and the ways in which they gather perceptual information. The exploration is directed by "what to look for next", which influences the types of perceptual information that are sought from the environment. There are three major ways to explore file sharing information, with a corresponding set of three methods to gather perceptual information. These methods were outlined in Figure 5 and are presented again in Table 26 for clarity. The items listed in column 1 were presented in Figure 5 as "determine what to look for next"; those in column 2 were presented as "gather perceptual information."

When exploration is directed by:	then information is gathered through:
background awareness	feedthrough on artifacts
explicit queries	individualized views
selective attention to events	notifications

Table 26: Relationship between exploration and information gathering, from file sharing awareness framework. These elements were presented earlier in Figure 5.

To avoid confusion, a clarification of the term *exploration* is provided here, because this word refers to different concepts in different research domains. For this framework, the term is used in the cognitive science sense, because it is rooted in a cognitive model: the perception-action cycle (Neisser, 1976). Within that model, *exploration* means to gather new information from an environment (Neisser, 1976). This is similar (although not identical) to the information retrieval sense of "exploratory search," in which a person is collecting information for learning and investigation

(Marchionini, 2006). However, in the awareness framework, explicit queries (to retrieve information) represent only one of the possible ways of gathering new information; thus, *exploration* is more broadly applied.

Exploration is directed by background awareness (i.e., general knowledge of what is going on in the environment), by asking explicitly for certain information about the environment, or by having attention drawn to specific events. There is a corresponding set of ways to perceive the environment for each of these factors.

To maintain background awareness, the user can rely on feedthrough (change in artifacts), which indicates relevant changes in the environment on an ongoing basis. For a person who is trying to maintain a general awareness of file sharing (e.g., keeping on top of events and settings), feedthrough provides a means of tracking changes and presenting simple information about the state of the environment. For example, suppose a colleague made a file accessible to the rest of the team. If this change is reflected in the environment, through a change of file icon, for example, then the user could note this change of icon when exploring the file system. In this case, no explicit request for information was required; awareness was supported simply by noting changes in the environment, which could happen when engaged in other tasks.

For explicit queries, individualized views provide a means for clearly presenting the requested information, such as file-based or person-based views. If a user wanted a specific piece of information, they could use individualized views (in conjunction with a query system) to view the relevant data. One example would be to check which files were being shared with Karen (i.e., to see whether she has access to the file that I wanted her to download). This query would produce a list of such files, which would answer the user's question about file sharing information. The individualized views allow for flexibility in creating sets of files matching certain criteria, which is necessary given that files are shared in many contexts. As shown in the set of file sharing awareness elements (e.g., Table 23–Table 25), possible queries might include person-based questions ("what files can this person access?") or file-based questions ("when was this file distributed?").

Finally, for selective attention to events, notifications can be provided (e.g., alerts) when an event of note occurs. For example, if a user needed to know when Karen downloads the file, they could define an alert for notification of the event. (They may, for

example, wish to know this event happened so that they can delete the file from a shared area.) Because file sharing is dynamic, and often temporary (as shown by the survey results in Chapter 3), having a means for highlighting events of particular importance could assist users in making required changes.

5.8 *Summary*

This chapter has described theories underlying awareness research, particularly in the CSCW domain, which enable an understanding of how awareness can be supported within a file sharing context. Existing research has been adapted to develop a framework for file sharing awareness, which incorporates information about shared files as well as the means for gathering that information.

Within this framework, three elements were identified as the means through which file sharing information is gathered: *feedthrough*, *individualized views*, and *notifications*. These elements are not currently well-supported in the everyday contexts in which most people share files. One goal of this thesis research is to provide awareness mechanisms and determine how well they assist people in managing shared files. In order to do this, these elements (e.g., feedthrough) need to be instantiated into a specific design, which can be formally evaluated. Further research identified how these items could be operationalized through particular application designs, including a file manager and a specialized file sharing console. These designs are described in the following chapter.

Chapter 6. Designing for File Sharing Awareness

The file sharing awareness framework outlined some elements that could help users to keep track of their shared files, such as feedthrough. The challenge is how to incorporate those elements into tools, using a fully-realized design. The framework was used to create a preliminary design for file sharing tools: an augmented file manager and a specialized console. These two components (the file manager and console) are discussed in this chapter. The first section discusses the reasons for adapting a file manager as the means for supporting shared file management. Next, a description of core file manager characteristics is presented, followed by a description of how file sharing features can be incorporated into file managers. The preliminary design for the file manager (including a separate console window) is illustrated, which includes these file sharing features. Finally, a set of icons and terms is introduced, a subset of which were selected for inclusion in the final version of the file manager design.

6.1 *Selecting a Tool for Presenting File Sharing Information*

The framework elements for file sharing—feedthrough, individualized views, and notifications—could be implemented in a number of different ways. The first step in applying these elements to a concrete design was to decide how to implement these features. One approach would be to create a new application with these features, which could be designed from scratch to include file sharing functionality. However, software tools have already been created that provide a foundation to build on: specifically, file managers. Leveraging existing file management tools has advantages. In particular, users already manage their files using these tools, and therefore the additional features can be added to a familiar tool, used frequently to carry out file-related tasks. Adding file sharing features to a file manager would make this new functionality convenient, and would not require users to learn to use a completely new type of technology.

6.2 *Characteristics of a File Manager*

A file manager is a core tool for interacting with the file system, being a program "which provide[s] file management services, allowing users to delete, copy, move, name, back up, search for, and view files."(file management, 2006). Examples of file managers

include Windows Explorer (for Microsoft operating systems), Finder (for Apple operating systems) and Dolphin (for Unix/Linux).

Each of these programs displays filenames in a list or graphical representation (e.g., icons), along with information about that file, such as its size and modification date. File managers are generally customizable, allowing the user to select which details are displayed. For example, Windows Explorer for XP provides a "Details" view that displays multiple columns of file information; thirty-three column types are built-in, including "status", "attributes", and "owner", and this basic set can be extended by installing file system utilities for additional information (such as folder size). A screenshot of this view is shown in Figure 6.

Because file managers are the core tools that people use for interacting with files, they are the central source for file information. Therefore, additional information, such as file sharing information, could be added to the list of items that a file manager displays. Additionally, users interact with a file manager when browsing through their files; for example, to locate specific information or a specific file. These browsing events give opportunities for supporting background awareness: when users are navigating the file system, file sharing information could be automatically displayed, allowing them to see this information "at a glance," without even having to explicitly search for it. Augmenting a file manager with file sharing features is therefore in keeping with user practice, and allows for easy access to information in a relevant context.

6.3 *Information Displays in File Managers*

There are a few ways in which information is displayed in file managers. Primarily, information is presented as text (including numbers), and through graphics (usually icons or thumbnails). In some file managers, this information is listed in columns. For instance, in the Windows Explorer example in Figure 6, there are six columns of information, primarily text, with file type icons to the left of each filename. More information can be displayed by adding columns, although the number of columns that are visible at once is limited by the width of the screen.

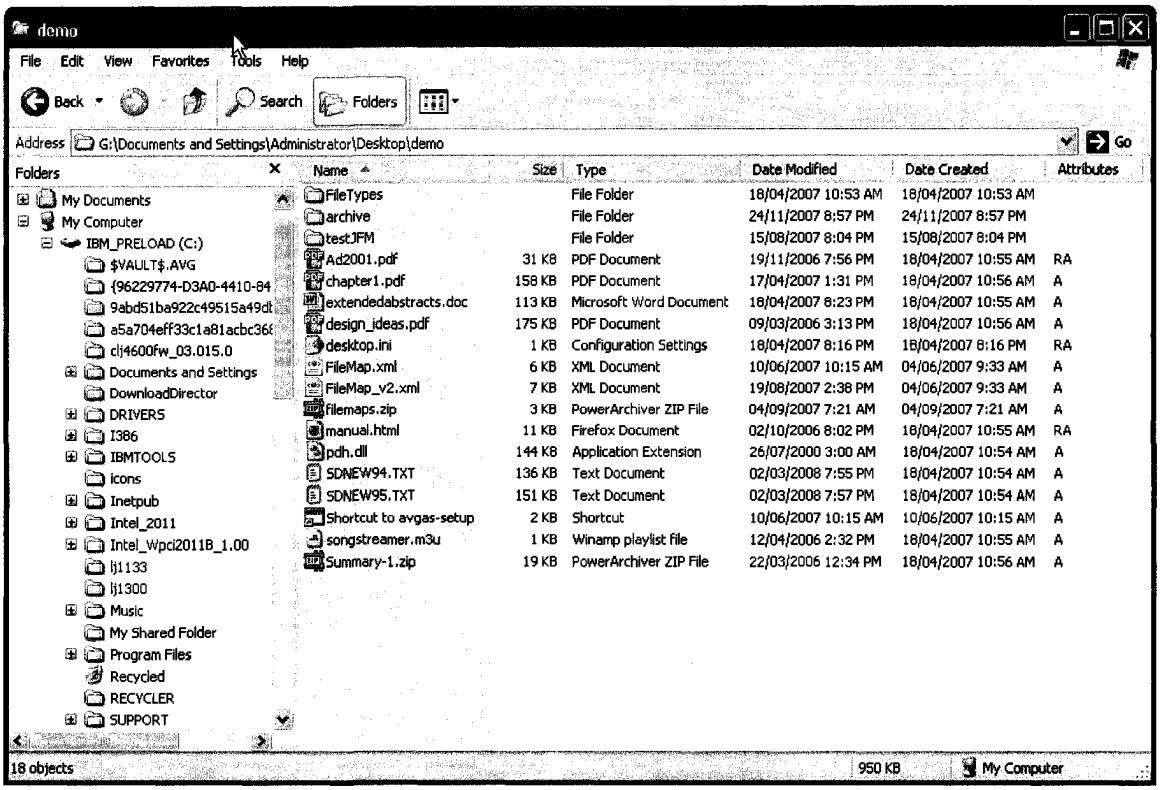


Figure 6: Details view in Windows Explorer for XP, showing six columns of file information

In addition to columnar information, pop-up windows can be used to temporarily display small amounts of file information. These pop-ups (also called *infotips)* are small windows with text that appear only when the user hovers the mouse pointer over a specific object, such as an icon (Microsoft Corporation, 2008); this window is depicted in Figure 7. The pop-up window can occlude other items in the file manager window, making it suitable only for temporary (and small) displays of information. However, this feature does provide an alternative means of presenting additional file information.

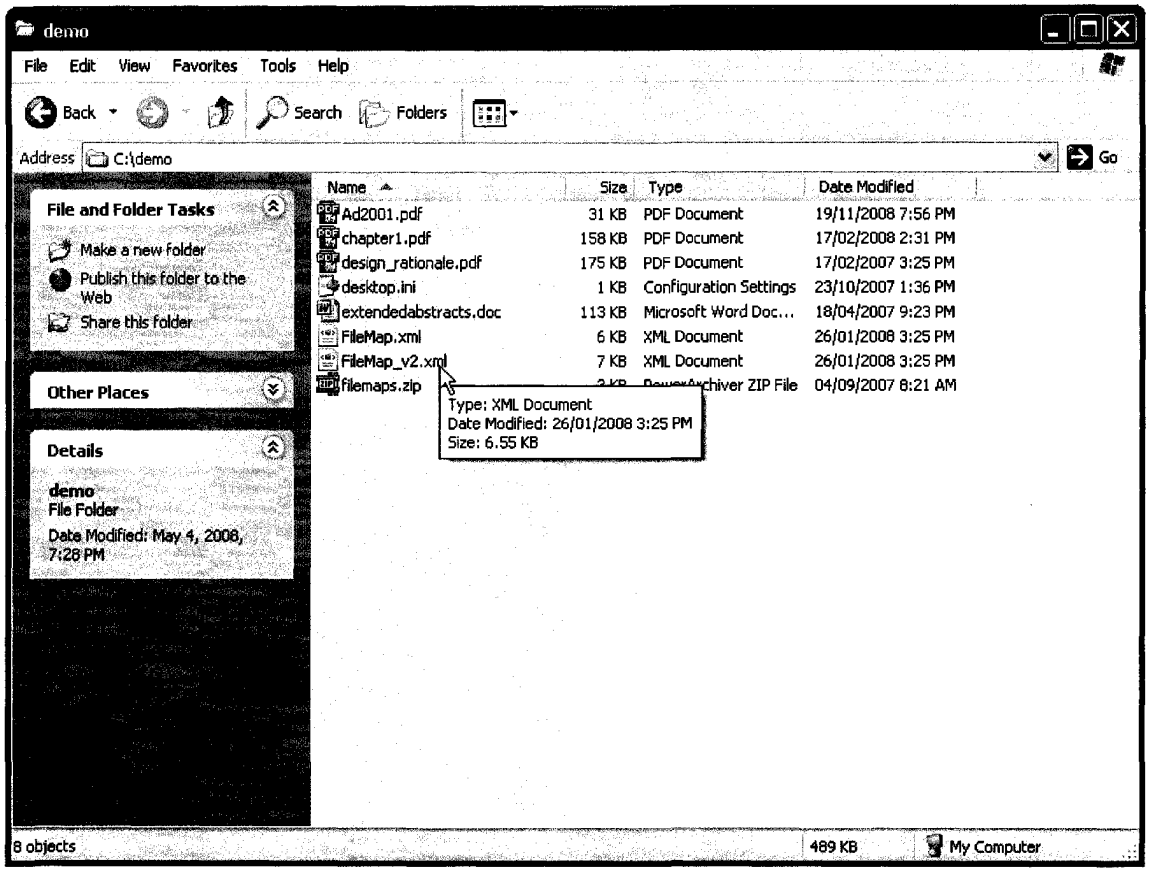


Figure 7: Example of pop-up (infotip) in Windows Explorer for XP. In this example, the infotip lists the file type, the date modified, and the size

File managers can also be used with supplementary windows, either for displaying more information than can fit in the main file listings window, or to support additional file management features. One example of this is the "Properties" feature in Windows Explorer: by right-clicking on a file and selecting "Properties" from a menu, a new window is launched that lists detailed file property information, as depicted in Figure 8. This window also provides features for performing file tasks, such as changing the default application used to open this type of file.

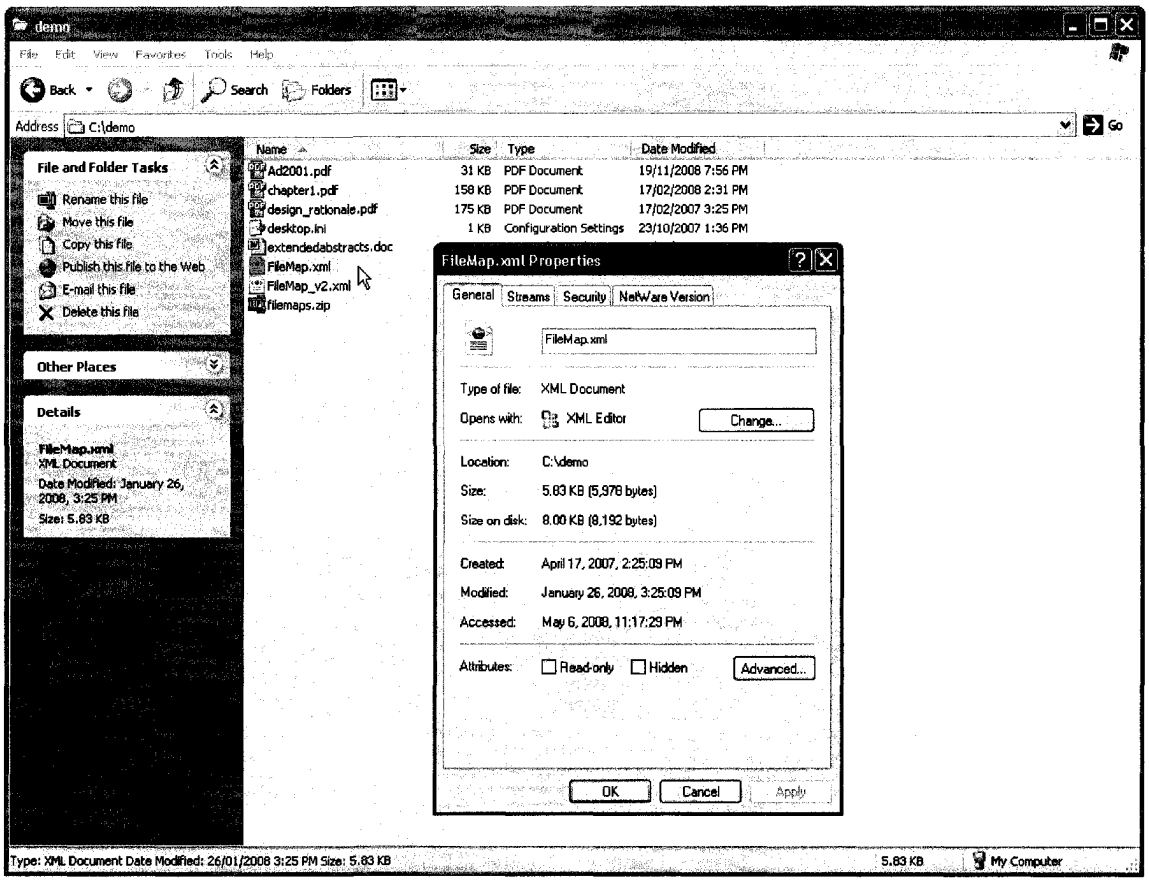


Figure 8: Supplementary window in Windows Explorer for XP, displaying the file properties information

These three ways of presenting information—listings of text and icons, pop-ups, and supplementary windows—are the means by which file sharing information can be displayed, when it is incorporated into existing file managers.

6.4 *File Sharing Information and Features to Include in File Manager*

The previous section described *how* information is displayed in a file manager; it is also necessary to choose *what* file sharing information should be displayed, and what features to add. In keeping with the file manager model described above, the main listing (of columns) contains information, while any new functions (such as search) would be included in a supplementary window.

The survey results, and the related work on file sharing, indicate what types of information are most useful to present in the file manager. Survey respondents (as described in Chapter 3) indicated a need to know who has access to a shared file (in the present), and who *has* had access to it (in the past). As discussed in Chapter 5, file

sharing includes push-oriented and pull-oriented approaches, such as copying a file to others over email, versus placing a file in a folder and setting permissions to allow sharing. However, these means of sharing are quite different, and therefore the events and settings for *push* and *pull* must be presented distinctly. For example, a *push* event (e.g., emailing an attachment) creates a *copy* of a file, which may be distributed outside the local file system. In contrast, for a *pull* event, all actions take place on the same instance of the file (in the local file system). Separate columns for *push* and *pull* oriented events serve to make that distinction clear. Additionally, to know who currently has access, and what actions they can take, users need to be able to view the file permissions.

Therefore, to provide a complete account of who has access (past and present), it is necessary to display the *push* events as well as the *pull* events, and the file permissions, all of which indicate who has access to shared files (and the file actions that can be performed), along with information about past accesses to shared files.

Note that some of this information may not be available in a single file manager listing, because file managers only list information about files in a specific folder. If, for example, the user wants to view all files shared with a certain group, those files may be stored in multiple folders. As described in the awareness framework, providing a feature for individualized views would address this problem. A query feature would allow users to search for all files matching certain criteria (such as "shared with group X"), which would display that exact set of files, no matter where they were stored.

This functionality is best addressed through a supplementary window, for two reasons. First, the main file manager listing (e.g., the columns) is designed to display files stored in the same folder: displaying a mix of files from different folders would likely be confusing, unless this display were presented in a separate window. Second, creating the set of files requires a query feature; as described above, additional file management features are generally supported in supplementary windows, where the appropriate interface features can be provided.

One additional feature, which could also appear in a supplementary window, is notifications. As the survey data indicated, file sharing is dynamic, and users sometimes need to have events brought to their attention. In keeping with the awareness framework, this function is realized by notifications. The file manager can allow users to create and

edit alerts for events of interest; a pop-up message could appear when that event happened, to notify the user.

The previous sections outlined the means for presenting information in file managers, as well as the types of file sharing information and features that would be most beneficial to provide in a file manager. In order to apply these elements in practice, they must be realized as part of a fully-formed file manager design. This process was completed to create a file manager design based on Windows Explorer for XP; the design is detailed in the following sections.

6.5 *Adapting a File Manager Design: Windows Explorer for XP*

For this research, the design of a file manager for file sharing was based on Windows Explorer for XP, mainly for purposes of formal user evaluation. Creating an entirely new file manager would require that study participants would need to learn to use an entirely new application, making it extremely difficult to isolate and evaluate *only* the file sharing features of that file manager. Thus, it was decided to adapt a commonly-used, established file manager. At the time the preliminary design was created (June 2007), Windows XP was the most prevalent operating system: it had an estimated 82.25% of the market share, according to a leading market researcher (Net Applications, 2007). As well as being the most commonly-used operating system, Windows XP was also well-established, having been released in October 2001. Therefore, the built-in file manager for this operating system—Windows Explorer for XP—was the best application on which to base the new design. Note, however, that the design choices made here are not limited simply to Windows Explorer (or even the Windows XP operating system): they can be readily adapted for other file managers and systems.

The "Details" view of Windows Explorer (as depicted in Figure 6) provided a good model to build on, as it allows users to view multiple details about files, including file sharing information. Only a limited amount of information can be seen at a glance in the Details view. (The exact size depends on screen size and resolution; the height of each row is 17 pixels.) With these constraints, priorities must be set as to which information is most important to include in the display. This information may change depending on the exact context of use; it is beneficial to allow users to customize this display, choosing the information of greatest importance to them. However, for the

purposes of evaluation, a standard design prototype was developed, one that presents identical information displays to all study participants.

As described in Section 6.4, information about *who* has access should be displayed, which includes *push* and *pull* events (file history), and file permissions (who currently has access). Three new columns were created for Windows Explorer in order to display this information: one each for *pushed* (called "sent") and *pulled* (called "accessed") events, and one for file permissions.

6.5.1 File Listings: Icons for File Sharing Information

A complete file events history cannot not be shown in the limited space provided by Windows Explorer: instead, it can display visual indications of the *amount* of interaction that has occurred over the lifetime of the file. This approach was inspired by the social navigation ideas of DiGioia and Dourish (2005), who suggested that "digital wear" on files be displayed, to indicate the degree to which those files have been used by others. For this file manager design, there is a simple pair of indicators: a number, representing the number of events (e.g., sends), along with an icon for that activity. (If no icon or text appears for a specific file, then no sharing action has occurred.)

The reason for including a graphical icon with the number is to promote background awareness. Findings from information awareness displays indicate that graphical displays better than text-only displays for helping users to notice and recall information (Plaue, Miller, & Stasko, 2004). Therefore, given than the file manager needs to convey information "at a glance", graphical indicators of activity should help attain this goal. There are a variety of visual means for conveying this file activity information; an example of an early paper prototype design is shown in Figure 9 , which includes some sketches of preliminary icon ideas.

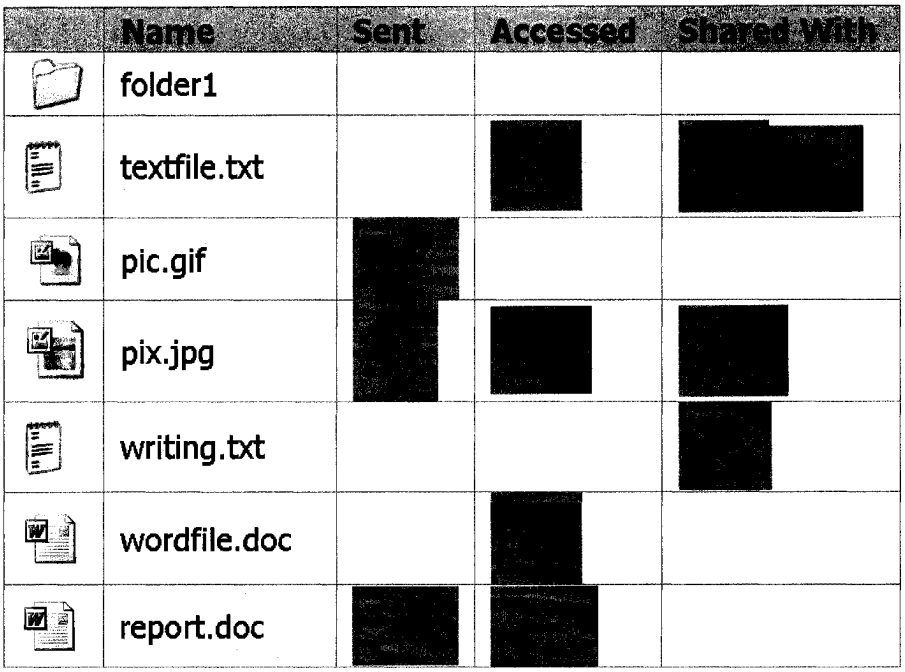


Figure 9: Early design sketches for file manager, showing a variety of possible icons for file events and permissions.

In addition to indicating file activity, icons can also be used to provide a summary of file permissions. In this design, they are used to indicate the scope, or degree, of sharing: that is, to communicate the general size of the pool of potential recipients: a single user or group, multiple users or groups, or everyone (also referred to as "all" or "others"). These five general groupings—user, users, group, groups, everyone/all—correspond to the fundamental permission settings for accounts in most operating systems (Allison, Hawley, Borr, Muhlestein, & Hitz, 1998). For a shared file, between one and three icons could be shown, depending on how permissions can be configured (e.g., one user, and two groups). Displaying these users/group icons summarizes the degree of sharing, allowing users to maintain a general sense of how public or private a particular file is.

6.5.2 File Manager Popups

A small amount of information can also be provided through pop-ups. Pop-ups are used in the expanded file manager to display two types of information: recent file events, and file permission details. If users hover over the icon for "sent" or "accessed", they can view details about the most recent sharing event on that file. For "sent", this

consists of the date of the event, the method used to send the file (e.g., instant messenger), recipient information (e.g., email address), and, if applicable, some contextual information from the sending application (e.g., subject line of the email). For "accessed", information is displayed about the date and time of the most recent access, the type of action taken (e.g., read), and the name of the user who performed that action.

For file permissions, the pop-up lists the actual user and group names (of those who have access), along with the rights for each user and group. This list can be quite lengthy, especially in cases of complex access control lists, so the pop-up was capped at 40 characters to ensure the pop-up does not obscure most of the file manager window. Complete file permission details can be obtained by opening the sharing console, as described in the following section.

An example of a pop-up is shown in Figure 10, which shows a paper prototype created from early sketches, which was later refined through user evaluation and design activities (detailed in the following chapter). This prototype shows the three new columns, a "placeholder" set of icons, and a pop-up for the Sent column, detailing the most recent sharing event, in which the file was sent as an email attachment.

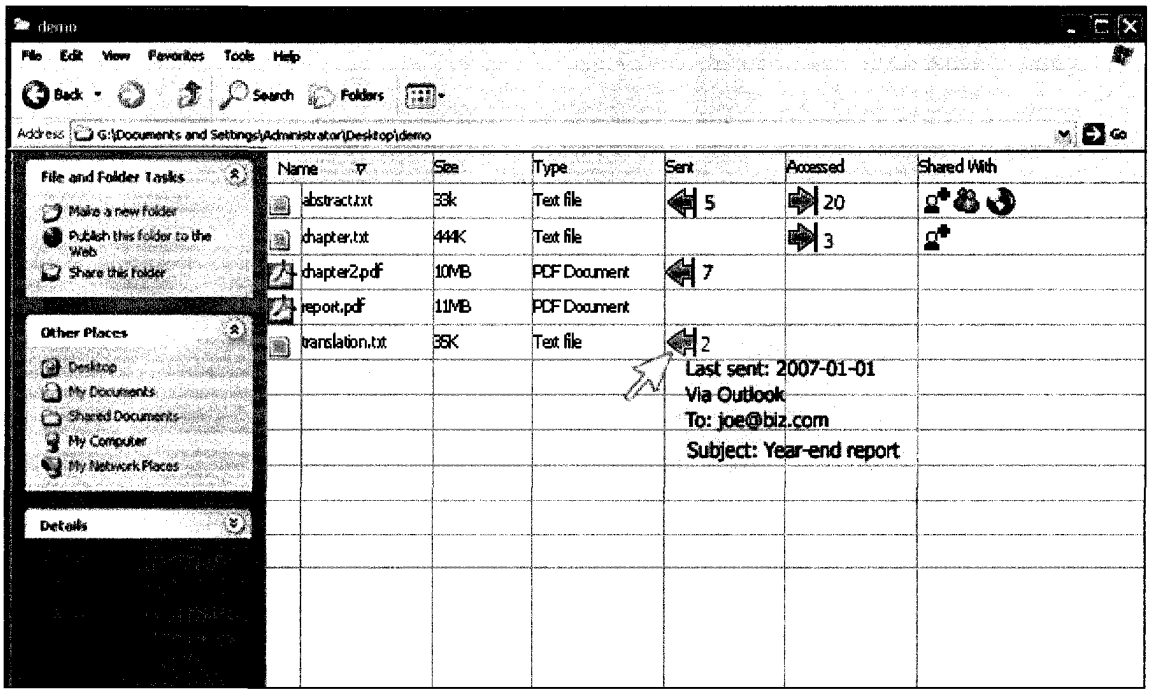


Figure 10: Paper prototype of the file manager, illustrating the basic design. Note the pop-up that provides details about the most recent "sent" event.

6.5.3 Supplementary Window: Sharing Console

Because a limited amount of file sharing information can be displayed in the file manager, and because specialized functions (such as individualized views) cannot be fully supported within the file manager, it was necessary to add a second window for these purposes. A sharing console was developed to fulfill this need, one that can be launched from the file manager when required. The sharing console includes the elements of individualized views and notifications, as well as providing more detailed information on file permissions and events. These features are detailed in the following sections.

6.5.3.1 Query Feature

As presented in Chapter 5, individualized views are supported by explicit queries. That is, if users want to generate a particular view of their files, they can do so by creating a query that identifies and displays a set of files matching the desired criteria. A sample query might be "Which files are shared with Tom?", which could consist of files stored in different folders within the file system (as well as those sent over email), brought together in a single integrated list.

A broad set of questions about file sharing was enumerated in Chapter 5, such as "How has this person received files?", which are grouped under "person-based" or "file-based" views. Thus, searching by "person/group" is included in the initial console design, as well as "file property" (e.g., permission type). Also included is "keyword", which allows for a more generic search, if desired (e.g., files containing the word "classified"). In all these cases, a set of files matching the given criteria is generated.

6.5.3.2 Detailed File Information: History and Permissions

The file-based view also requires that detailed information about specific files be made readily available, for answering questions such as "What actions have taken place on this file?" A subset of this information is visible in the file manager, but more complete information must also be available. The console design includes the full sharing history of a file, which includes "send" (*push*) events as well as file system *pull* actions (e.g., read and write). In addition, complete file permission information must be presented; for simple cases, this can be displayed in the file manager, but as described above, there are cases when there is too much data to present within the file manager. Therefore the console can be used to display these permission details, along with event

history, to give a comprehensive view of a file. As discussed in Chapter 5, this incorporates elements of the past (the file history) as well as the present (current permissions).

6.5.3.3 Alert Feature

The third (and final) element of the awareness framework—notification—is another feature included in the sharing console. Users must be able to identify events of interest on files, and configure an alert to inform them when the event happens. This configuration feature, along with a list of active alerts, is supported in the console. Note that the alert itself does not appear within the console, because users would not see it unless they were working with the console at the exact moment that the event occurred; instead, a separate pop-up window appears to let the user know that the event of interest has transpired.

These components—the query feature, the detailed file information, and the notification feature—are included within a single tool, as shown in Figure 11. The inspiration for the console design came from desktop search tools, particularly Windows Desktop Search (Microsoft Corporation, 2006) and Microsoft Research's PHLAT (Cutrell, Robbins, Dumais, & Sarin, 2006), both of which display sets of files (from various folders on disk) satisfying specific criteria. Users can interact with the console in different ways: they can bring up detailed file information in it (either via the file manager, or by using the "view another file" button), or can perform a search for specific files, or manage their alerts (view, create, or edit). When a query is performed that returns a set of matches, the user can view details for each of those matching files. This is illustrated in Figure 11: the user looked for files shared with, or sent to, `alice`, and a list of matches was generated. Clicking on the filename `quarterly_report.doc` brought up the permissions and event history for that file, in the panel beneath. In this way, users can easily bring up additional relevant sharing information within the same window.

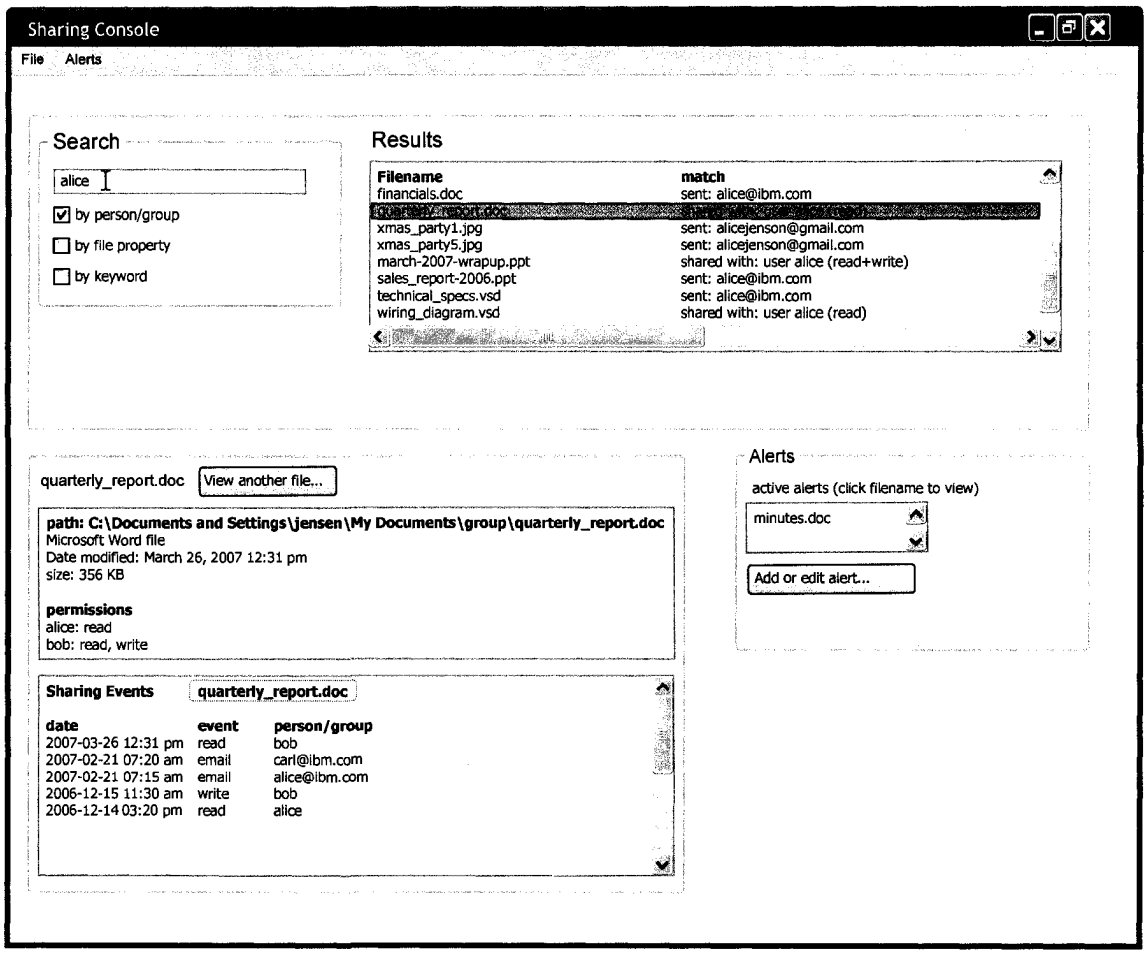


Figure 11: Preliminary design for Sharing Console

Note the detailed file information presented in Figure 11, on the lower left, for the file `quarterly_report.doc`. The filename is shown, with some basic file details, such as size. The full file permissions are presented next, for each user and group; in this example, the access rights for `alice` and `bob` are shown. The complete history is listed in the panel beneath the permissions: in this example, there are five events for this file, consisting of two sends (emails), and three accesses. For each event, a time, event type, and user or recipient name are provided.

On the right side of the console is the "Alerts" panel, which lists the active alerts and allows the user to modify current alerts and create new ones. Figure 12 shows the preliminary design for the add/edit feature, as well as the pop-up notification message.

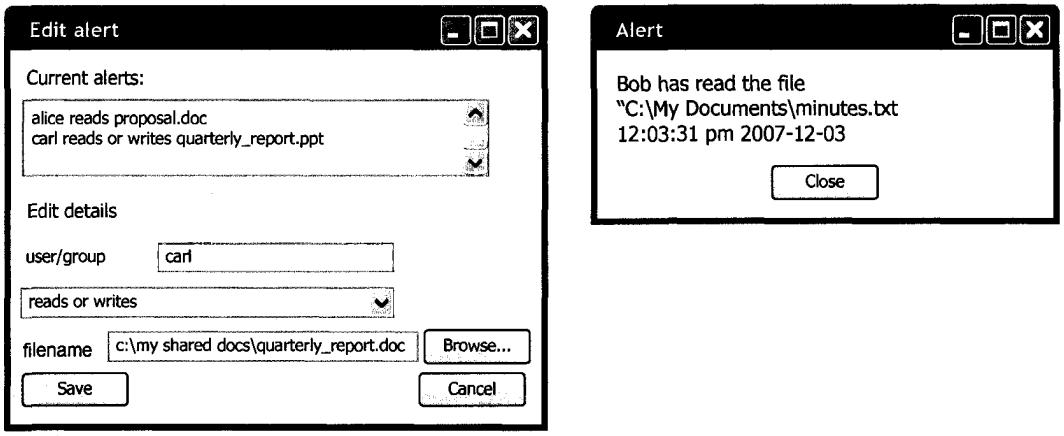


Figure 12: Notification feature: editing window (left) and notification window (right)

6.6 *Synopsis of Preliminary Design and Unresolved Design Issues*

The elements from the awareness framework were incorporated into a preliminary design, which included a file manager and a sharing console. The file manager design, which is an extension of the Details view of Windows Explorer for XP, includes three columns dedicated to file sharing information: "sent", "accessed", and "shared with". The "sent" and "accessed" contain icons indicating that sharing activity has taken place on specific files, along with the number of times that action has occurred (e.g., "sent 4 times") for each file. Additionally, pop-ups provide information about the most recent event, such as date and recipient. The "shared with" column displays graphical representations of file permissions, with icons indicating the scope of sharing (user(s), group(s), and/or everyone), along with pop-ups that provide more specific details about those permissions (e.g., user and group name, rights).

The file sharing designs presented in this chapter were preliminary versions. During the design process, a number of issues arose that required clarification before implementing the design for testing. For example, it was not clear what a good choice of icons was be for the file manager, to convey such ideas as "user" or "sent." Similarly, it was not known if the terminology used in the column headings for push-oriented and pull-oriented events would be understandable. These questions were addressed through a user-centered design study (described in the following chapter), but before this activity could be conducted, sets of potential icons and terms needed to be generated to be used in

the study. These icons and terms, and the process used to select them, is described in the following section.

6.7 *File Sharing Icons and Terminology*

6.7.1 File Sharing Icons

The file manager design included icons in the columns for "sent", "accessed", and "shared with." The icons in this design need to represent both file-oriented activities and person-oriented concepts. Seven icons needed to be selected for the final design. Two of these icons would represent push-oriented and pull-oriented sharing activities (*sent* and *accessed*), as described in the previous chapter. Five additional icons were required to indicate with whom a file is shared: *user, users, group, groups,* and *everyone*. One design goal was to ensure that there would be little confusion between file-oriented and person-oriented categories: e.g., that the icon for "sent" would be quite distinct from "user." This would help the user to distinguish between the different file manager columns when seeking information.

Because this was a new design, there was no set of pre-existing icons to draw from; indeed, the purpose of the study was to identify such a set. These icons needed to be the right size to work with Explorer for Windows XP, as this was the basis for the full design prototype that would include these icons. Therefore, the icon dimensions from XP were used: 16 by 16 pixels. In selecting the set of icons, discussions were held with designers Vicki Ha (of the University of British Columbia) and Christopher Dean (of NSCAD University). In addition, a review was conducted of icons in computer applications, such as file managers and web applications. A substantial search of high-quality, public domain 16x16 icons was carried out, and a subset of these was selected for inclusion in the study. Some of these designs were modified slightly (using Adobe Photoshop), to fulfill a designer's suggestion or to create matched pairs of icons.

The icon set was originally 30 icons. This set size was selected because the planned design activity (used to identify the best icons) included a card sorting task; in keeping with recommendations for card sorting, a maximum of 30 items should be used (Boulton, 2006). A pilot test of the card sorting task indicated that this was too large a number for achieving consensus within a reasonable time, and therefore the set was reduced to 20 icons. The icons that were removed from the set were those that the pilot

testers identified as having the most problems (e.g., considered to be confusing), as well as those judged to be redundant. The full set of 20 icons, selected for inclusion in the card sort, is shown in Figure 13.

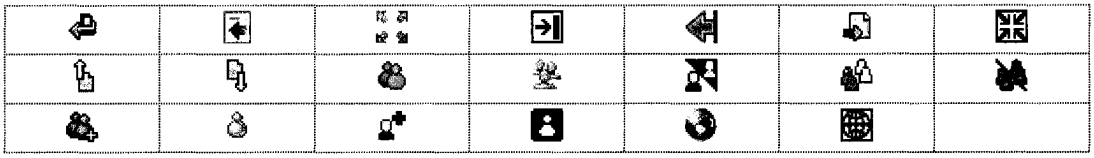


Figure 13: File sharing icons used in group card sorting activity

Note that a number of icons have a "directional" theme: these are the file-oriented icons; the concepts here are those of sending and accessing, which do not have concrete visual representations. The icons here take inspiration from such concepts as "download", and "email", which often use arrows to indicate the movement of a file. Other icons are people-oriented; the complexity here is how to distinguish between different concepts involving multiple people: "multiple users" is not the same thing as "group", from a file system perspective, and therefore this distinction must be clear to the user. Visual concepts such as borders (strokes), item placement and spacing, and colour groupings were used to try to indicate these differences.

6.7.2 File Sharing Terminology

In the file manager design, the file sharing icons are displayed under column headings, such as "sent." These column headings need to have understandable titles that indicate the type of file sharing displayed in each column.

One design challenge that arose when developing the preliminary design was what to use for the column headings that displayed push-oriented and pull-oriented file activities; the words "sent" and "accessed" were used as placeholders, but it was unclear whether these were good choices. Although the terms *push* and *pull* are used in the file-sharing literature for these concepts (specifically in (Voida, Edwards, Newman, Grinter, & Ducheneaut, 2006), it was not expected that these terms would be widely known, outside the limited confines of the research and web-application communities.

A set of terms describing file-sharing activities was created, for assessment in the design activity. These terms were in two categories: pull-oriented terms and push-oriented terms. These sets of terms were devised using thesauri (Kipfer & Chapman,

2001; Summers & Holmes, 2004), computer glossaries (Freedman, 2000; NIST, 2006) and the file-sharing literature (Voida et al., 2006).

Because the terms were to be evaluated by participants as part of a longer design activity (within a single hour-long session) the list could not be overly long, or it could not be evaluated within a reasonable time. An initial set of 19 terms was refined through pilot testing; some words were dropped and others were added, to create a final set of 18 words. The resulting sets of terms are listed in Table 27.

pull-oriented terms	push-oriented terms
accessed	delivered
changed	distributed
modified	forwarded
obtained	gave
opened	pushed
pulled	sent
received	transferred
took	
touched	
used	
viewed	

Table 27: Terms that can be used to describe pull-oriented and push-oriented file activities

6.8 *Summary: Refining the Design for Future Evaluation*

At this stage of the research, the file manager design was mostly complete. The core elements, as identified by the awareness framework, were designed, and included the file sharing information and features deemed to be most beneficial to users. This design required formal evaluation, but before that evaluation could be conducted, some components of the design needed to be finalized: namely, the specific icons and text used to express file sharing information. In addition, it was considered useful to gather feedback on the components of the overall design, to correct any flaws before the formal evaluation phase. This user feedback would provide guidance for refinement of the initial design, so that the resulting version could be implemented in a prototype for user testing.

The final design was evaluated in a laboratory setting; that study is described in a later chapter (Chapter 8). However, prior to this laboratory study, a group design activity was conducted for improving the file manager design; in particular, to determine the most

appropriate choices for terms and icons. That design activity is described in the following chapter, along with a description of the resulting prototype used for the evaluation.

Chapter 7. Validating and Refining the Preliminary Design

In the previous chapter, the preliminary design for the file manager and sharing console were presented. After creating these designs, there was a need to determine whether the design choices made were suitable, and in what ways the designs could be improved before implementation and lab testing.

For the file manager, the core design questions were:

- Are the terms and icons used to distinguish file-oriented and person-oriented information clear and distinct?
- Which terms should be used to indicate "push-oriented" versus "pull-oriented" activities?
- Which icons best represent the concepts of "push-oriented" and "pull-oriented" sharing?
- Which icons most clearly represent, and indicate the distinctions between, sharing with individuals, groups, and everyone?

For the sharing console, the core design questions were:

- Do users understand how to use the console features?
- Can users interpret and make use of the sharing information in the console?

To address these questions, a series of studies was conducted. The primary study was a design activity, which included a group card-sorting activity, an icon-labelling activity and questionnaires (performed individually), and a group discussion. The design activity was followed by two focused follow-up studies: a short icon questionnaire (to select a "groups" icon), and an online survey on terminology. The research described in this chapter was approved by the Social Sciences and Humanities Human Research Ethics Board at Dalhousie University.

7.1 *Study One: Group Design Activity*

7.1.1 Research Objectives

The group design activity (Study One) was designed to gather feedback on the preliminary designs of the file sharing tools. Specifically, there were two objectives to this study:

1. to determine the best choices for icons and terminology for use in the file manager, and
2. to discover any deficiencies in the design elements that required correction before the design was implemented

As described below, this study had four stages, and each of these stages had its own questions that addressed the two above objectives. These more focused questions will be presented in summary descriptions of each stage, presented in the following section.

7.1.2 Study Overview

There were four main stages to this study; each had a specific set of goals, expressed below as questions:

1. **Stage One**: a group card-sorting activity, designed to determine how people categorized and interpreted icons.

Questions:

- Which icons were associated with each other (i.e., most often grouped together)? Which icons were *not* associated with each other?
- What labels were associated with these groups? That is, how were icons interpreted, in terms of file sharing concepts, when not presented within the specific context of the file manager design?

2. **Stage Two**: an icon-labelling task (performed individually), designed to determine associations between terms and icons for file-sharing concepts. These concepts were in two categories: file activity-oriented, and user-oriented.

Questions:

- Which file-sharing terms were associated with which icons?
- Were the two categories (*file activity* and *user*) distinct? That is, were the icons and terms from each of these two sets associated with each other (*activity* terms with *activity* icons, and *user* terms with *user* icons)?

3. **Stage Three:** a questionnaire that gathered detailed information about the icons and labels within the context of the specific file manager design.

Questions:

- Are there any problems with interpreting the icons (e.g., were they confusing?)

- What terms are most appropriate for describing two types of file-sharing activity: *push-oriented* and *pull-oriented*?
- Using icon design, how can a distinction be made between the different types of "multiple individuals" that can be shared with (users, group, groups)?
- Is the information displayed by the file manager understandable?
- When shown in context of the file manager design: which icons are preferred for expressing file-sharing concepts (e.g., *sent*, *group*)?
- In what ways could the file manager design be improved overall?

4. **Stage Four**: a short group discussion, to gather general participant comments.

Questions:

- In what ways could the file manager design be improved overall?
- What design issues might require further exploration?

These four stages took place within a single session, which took approximately one hour to complete. A pilot study of this protocol was carried out (using a group of three participants), and some minor changes were made to the procedure and instruments before the formal study began.

7.1.3 Instruments

Four instruments were used in this study:

1. group card sort
2. labelled icon cards
3. a computer-based questionnaire, and
4. a group discussion

7.1.3.1 Instrument 1: Group Card Sort

The first stage of the study was a group card sorting activity, which was used to discover how participants categorized and interpreted file sharing information: specifically, the icons for the file manager. The instrument used in the card sort was a set of 20 icons, all representing file-sharing concepts. Participants made groupings out of these cards, placing related icons together into a grouping. Each grouping was also labelled with a word or phase that described what those icons had in common. For this study, the card sort was performed in a group (rather than individually), so that all members had to agree on the grouping.

Before the full details of this instrument are described, a brief introduction to card sorting is presented. This is followed by a description of the instrument: namely, the set of icons used for the activity.

7.1.3.1.1 Card Sorting: An Overview

Card sorting is a user-centred design method in which users sort cards depicting various pieces of information into categories. According to information architects Rosenfeld and Morville, card sorting "can provide insight into users' mental models, illuminating the way that they often tacitly group, sort and label tasks and content within their own heads." (2002, p. 235). Card sorting can be *closed* or *open*: in closed sorting, categories are provided, and the participant places cards into those categories. In open sorting, participants are free to create their own categories; this approach is often appropriate for exploratory studies (Rosenfeld & Morville, 2002), and therefore was adopted for this particular research study, because this study was intended to evaluate early design ideas. For this study, the cards consisted of icons that could be used for the new columns for the file manager (e.g., to indicate "sent").

A *group* card-sorting activity was chosen instead of an individual card-sorting activity for two reasons. First, it would generate discussion among participants, which would provide information about their interpretations of the interface elements; this advantage has been noted by card-sorting experts (Kaufman, 2006; Maurer & Warfel, 2004). The group design sessions were videotaped (with audio recording) to capture these discussions. Second, this approach would provide some consensus among interpretations: this would minimize individual preferences, revealing a more standardized interpretation of the icons.

7.1.3.1.2 Card Sorting: File Sharing Icons

The icons used in the card sorting task were intended for inclusion in the file manager, in the columns for "sent", "accessed", and "shared with." This icon set was introduced in Chapter 6, and is repeated in Table 28 below. Note that the icons were in two groups: *file activity-oriented* icons, which were intended for use in the "sent" and "accessed" columns, and *user-oriented* icons, which were intended for use in the "shared with" column.

file activity-oriented icons			user-oriented icons			
						
						
						

Table 28: File sharing icons used in group card sorting activity

As described in Chapter 6, these icons were designed to fit within a 16x16 pixel space. However, this small size was expected to make it difficult for study participants to interpret the icons on the 3x5-inch index cards during the sorting activity, especially when the cards were shared among multiple people sitting across a table from one another. To ensure that the icons were visually recognizable on an index card, the icons were presented at the "real" size (16x16 pixels, or 0.16 inches) and also enlarged (48x48 pixels, or 0.5 inches) to make them more legible on paper. An example of an index card is shown in Figure 14.

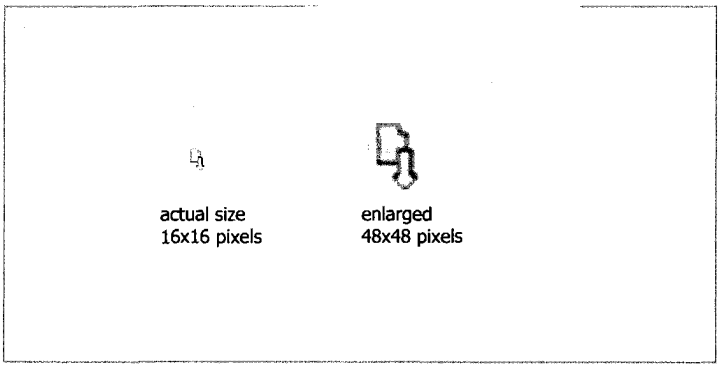


Figure 14: Example of a card used in the card sorting task. Note that the same icon is displayed twice, at two different sizes.

To summarize: at the start of the card sorting activity, the index cards were in a single stack of 20 icons, with no labels. At the end of the activity, the index cards were categorized into groupings, with a descriptive label on each grouping.

7.1.3.2 Instrument 2: Labelled Icon Cards

The same set of 20 icons (as used in the card sort) was used in a labelling task. The index cards were provided along with a set of labels; each label contained a word describing a file-sharing concept. Each icon card was labelled with the term(s) that matched that icon. An example of a labelled card is shown in Figure 15.

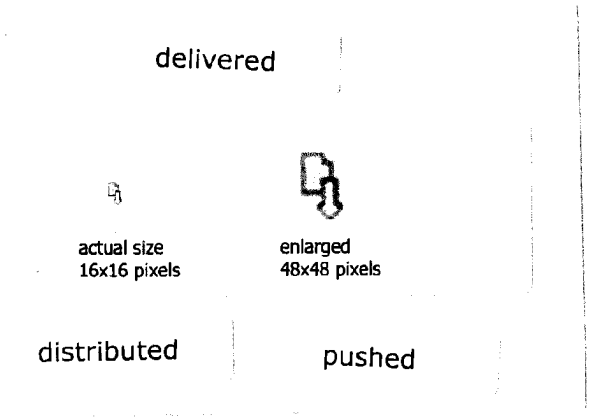


Figure 15: Example of a labelled icon card. Three words have been selected to describe this icon.

The labels consisted of 30 words related to file sharing. Specifically, these words described concepts relating to users (e.g., *user*, *group*) as well as "push-oriented" versus "pull-oriented" concepts (e.g., *forwarded*, *sent*). Eleven of the 30 terms described concepts of users and groups (e.g., "world", "people"). The remaining 19 terms referred to push-oriented and pull-oriented activity. Eighteen of these consisted of the list of terms presented in Chapter 6. One more term ("transmitted") that had been removed in pilot testing was added, simply to round the total number of terms up to 30. (This term was removed in pilot testing because the size of the word list had to be reduced, not because of any problems with the term itself.) The complete set of 30 terms is listed in Table 29.

The technique used here was a variation on *closed* card sorting: in this version, category labels are placed on cards, rather than cards being placed into categories, but the end result is similar: every icon is placed in at least one category. Note that the category names are often synonymous, rather than using distinct category names (as would be used in regular closed card sorting). In this case, the results are not used to reveal information clusters, but rather to determine appropriate labels.

push- and pull-oriented terms	user-oriented terms
accessed	all
changed	everyone
delivered	group
distributed	groups
forwarded	others
gave	people
modified	person
obtained	public
opened	user
pulled	users
pushed	world
received	
sent	
took	
touched	
transferred	
used	
viewed	

Table 29: Set of 30 terms used to label icons

Each card could have more than one label, and labels could be used on more than one card. (Multiple copies of the labels were provided to ensure that a term could be attached to as many cards as a participant wanted.) Participants were informed that additional term(s), not from the label set, could also be written onto the cards. Cards could also remain unlabelled, if no term could be found to describe it, although this was not encouraged; preferably, such an icon would have a write-in label of "unknown", rather than a blank.

This instrument was used in two ways. First, the results could be used to corroborate the findings of the group card sorting activity, in terms of showing whether there were overlaps between file-oriented and user-oriented concepts. Second, the results would indicate which icons were appropriate for use in the specific columns, particularly for the user and group information. For example, an icon that was intended to represent "user" might be rarely labelled as such, which would indicate that the icon is a poor choice for that concept.

Note that for Stage One and Stage Two, the preliminary design was not revealed; icons and terms were presented without any context other than that of the general "file

sharing" domain. The results of these stages reveal first impressions, indicating the reactions that a new user might have when encountering the file sharing tools for the first time. These findings could indicate where the designs could be modified to avoid misinterpretations, as well as to identify problematic icons and terms that should be avoided—for example, because they are unclear or confusing. (The icons and terms were revisited in Stage Three later in this study, within the context of the preliminary design, so that the participants could indicate their preferences when given more complete information.)

To summarize: at the start of the icon-labelling activity, there was a set of 20 unlabelled icon cards. At the end of the activity, each icon card had one or more terms attached to it, indicating how that icon was matched to specific file-sharing terms and concepts.

7.1.3.3 Instrument 3: Computer-Based Questionnaire

The third instrument was a computer-based questionnaire, developed using Opinio survey software (ObjectPlanet, 2007). This questionnaire had four main sections and a preamble. The questionnaire is provided in full in Appendix C, and each section is summarized below.

7.1.3.3.1 Questionnaire Preamble: Demographic Information

The preamble of the questionnaire collected demographic information and a small amount of data on computer experience and areas of training and expertise.

7.1.3.3.2 Questionnaire Section One: Icon Critique

This section of the questionnaire asked participants to comment further upon the icons, highlighting any problems and making suggestions for improvement. A sample of the checklist used is shown in Figure 16. The items used to critique the icons are based on Constantine and Lockwood's (1997) Software Usability Defects Log. For this questionnaire, a subset of the items from the Defects Log was used—only those items that were applicable to icons. In addition, the Defects Log lists only problems (i.e., negative descriptors), so for balance, a positive descriptor was added to each. For example, the Defects Log includes "confusing"; the checklist in this questionnaire includes "clear." In this way, participants can indicate which icons had positive attributes

worth preserving. Participants checked off any items they felt were applicable, and could also add their own free-form comments.

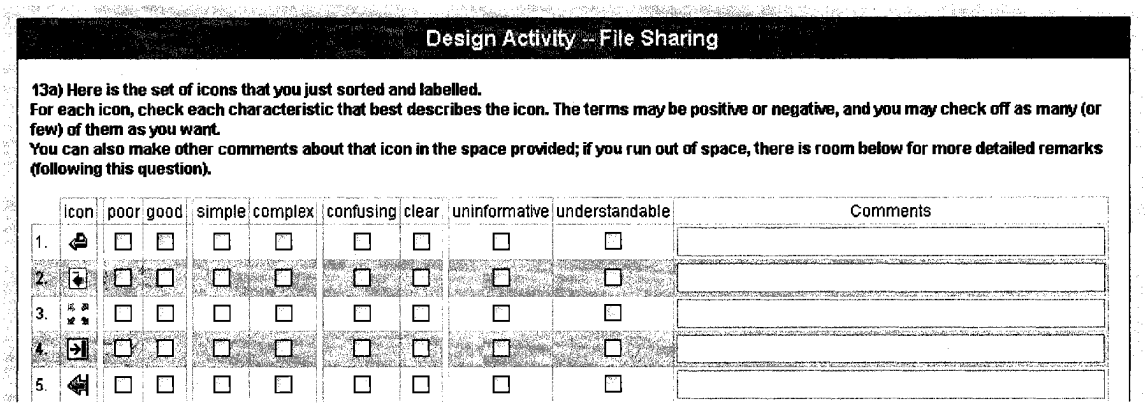


Figure 16: Sample question for analysing icons

7.1.3.3.3 Questionnaire Section Two: Push-Oriented and Pull-Oriented Terminology

This section of the questionnaire explored the question of referring to push-oriented versus pull-oriented concepts, as introduced in Chapter 6. The purpose of this section was to identify the best term to refer to each of these concepts, selected from a set of 11 pull-oriented terms and seven push-oriented terms; see Table 30 for a reminder of these sets of terms.

pull-oriented terms	push-oriented terms
accessed	delivered
changed	distributed
modified	forwarded
obtained	gave
opened	pushed
pulled	sent
received	transferred
took	
touched	
used	
viewed	

Table 30: Terms that participants could choose from to describe pull-oriented and push-oriented file activities

There were four main questions: one on pull-oriented term preference, one on push-oriented term preference, one on the best matched pair, and a freeform question about the distinction and clarity of these terms.

First, participants were given an example of pull-oriented activity (using a file in a shared folder). This was followed by a sentence that they needed to complete by selecting a pull-oriented verb to fill in the blank: "Alice _____ the file report.txt yesterday." A list of 11 words (as shown in Table 30) was presented (in random order). Participants ranked their top five choices for the best word to fill in the blank; they could also add (and rank) one word of their own to this list. In addition, participants could mark any word in the list as unsuitable, by checking a box beside it marked "I would never use this word." A freeform text box was also provided for participants to include additional words.

Following this question was the push-oriented section. This was nearly identical to the pull-oriented questions above (i.e., rank the best words for filling in the blank), except that the wordlist was different (seven words, as shown in Table 30).

The next question asked participants to choose the pair of terms that worked best together. A sentence with two blanks—one for push, one for pull—was presented, along with both sets of words in random order; participants chose a *pair* of words to use to complete the sentence.

Finally, participants were asked if the terms clearly expressed the difference between push- and pull-oriented concepts, and whether they were understandable for expressing the meaning of these concepts; these questions had freeform responses.

7.1.3.3.4 Questionnaire Section Three: File Manager Design Questions

The third section presented the preliminary design of the file manager for the first time. Note that for this section, participants were given paper printouts with full-colour screenshots of the user interface, so that they could keep the design in view at all times (while answering questions on the computer), and so they could sketch any ideas they had for improvements.

This section focused on the file listing portion of the file manager, excluding the supplementary sharing console window. There were two main subsections to this section.

First, a series of questions were presented to determine whether users could find, interpret, and comprehend the file sharing information presented in the interface. Second, the icons (from the earlier card sorting and labelling tasks) were revisited, this time within the context of the actual application they were designed for. Participants were asked to select the icons that they preferred (e.g., for *sent*), now that they could see how these icons would be used.

Using File Sharing Information

Participants were shown a series of screenshots of the file manager, all of which had some file sharing activity and file permissions displayed. There were four screenshots in total: one with no pop-ups, and three more with a pop-up for each column. Participants were asked a series of questions (e.g., "Which file is the tooltip describing?"; "Who can access this file?"), the responses to which would indicate how useful the display was for communicating file sharing information. Participants could also suggest design improvements, through comments or drawings.

File Manager Icon Preferences

Participants next revisited the icons from the earlier stages, but this time within the overall design. They were asked to choose their top picks for *sent*, *accessed*, *user*, *users*, *group*, *groups*, and *everyone*. In making their choices, participants were asked to bear in mind the design challenges of expressing pull-oriented and push-oriented concepts, as well as the distinctions between "multiple individuals" (users, group, groups).

A set of icons was given for each term (in random order), which included only those icons that were appropriate for that term. For example, for *sent*, the icons intended to represent "groups" were omitted. There were also more variations of icons in the set: every icon that was asymmetric was shown with its mirror opposite, so that, for example, both  and were shown. Similarly, every "user-oriented" icon for which there was both a singular and a plural version was shown with both variants; examples include  with , and with . Participants could pick only one icon for each, representing their top choice, and they could add their own choice (on paper) if they wanted.

In total, 16 file-oriented icons were presented to choose from, and 14 user-oriented icons. The complete sets of icons used for each question are provided in

Appendix C. More icon variants were used in the questionnaire than in the card-sorting task for two reasons. First, the number of cards to be sorted needed to be restricted; if all these icons were included, the set would have been impractically large (nearly double). Second, it was far more important to provide an exhaustive set when the icons were being analysed in context of the design, and less so when gathering initial impressions (as was done during the card sort).

7.1.3.3.5 Questionnaire Section Four: Sharing Console Design Questions

In the fourth and final section of the questionnaire, participants were shown screenshots of the sharing console, which is the supplementary window used to display detailed sharing information and individualized views, and to configure alerts. Participants were given paper printouts with a full-colour screenshots of the user interface for reference, as well as for sketching their design ideas.

As with the file manager questions, participants were asked factual questions about file sharing information shown in the screenshot; sample questions were "what is the user searching for?" and "which files have been emailed to Alice?" Participants could also suggest design improvements, through comments or drawings.

7.1.3.4 Instrument 4: Group Discussion

The final part of the design activity, completed after the questionnaire, was a short group discussion. This was a loosely moderated discussion, in which participants could discuss any outstanding issues they had, or ideas they wished to convey. To begin the discussion, participants were asked whether there were any comments that had not been captured in the earlier stages, such as in the questionnaire. This session was videotaped, so that all comments would be recorded for reference. Notes were also taken during the discussion, to underscore problems or points of interest.

7.1.4 Study Environment

The study took place in the Groupwork Lab in the Centre for Management Informatics in July 2007. This lab contained four videocameras, along with table and room microphones, for capturing group interactions and discussions. Tables were arranged for groupwork, along with individual computer stations with networked laptops for the computer-based questionnaires. A diagram of this setup is shown in Figure 17.

The videocameras were mounted on the walls, and their pan/tilt capabilities allowed them to capture activity throughout the labspace. Microphones were provided at each table, and two additional ceiling mikes captured any other ambient conversation. The audio-visual equipment was operated in an adjoining control room, by Tayze MacKenzie and Alexandra MacNutt, two research staff members at the Centre for Management Informatics. A one-way mirror connects the control room and the Groupwork Lab, to permit unobtrusive observation. As the facilitator of this activity, I remained in the room with the participants during the entire session.

The card-sorting activity stage of the study required three participants per group; up to three groups (nine participants) could be accommodated in this space at one time. This setup required three groupwork tables, and nine individual workstations, as shown in Figure 17.

7.1.5 Participants

For this study, it was desirable to collect data from people with specialized expertise. In particular, participants were recruited with backgrounds in computer science (for technical suggestions and to represent the expert computer user); in design, information science, and human-computer interaction (to give concrete suggestions for design and usability improvements); and in business and management (to represent the needs of the target user population for these tools, which is business users). The goal was to avoid a narrow participant group, to provide a multi-disciplinary perspective.

Participants were also required to have strong English language skills, due to the need to communicate one's ideas (e.g., suggestions) clearly, to understand the questionnaires, and to participate in group discussions. All participants who took part in the study were remunerated $15 for their time.

Eighteen people (13 male, 5 female) took part in this study: this number was selected mainly because of the requirements of the group card-sorting activity, for which experts recommended the inclusion of at least 15 participants (Maurer & Warfel, 2004). Participants were recruited through email invitations and posters, targeted at university departments having members with the required types of expertise.

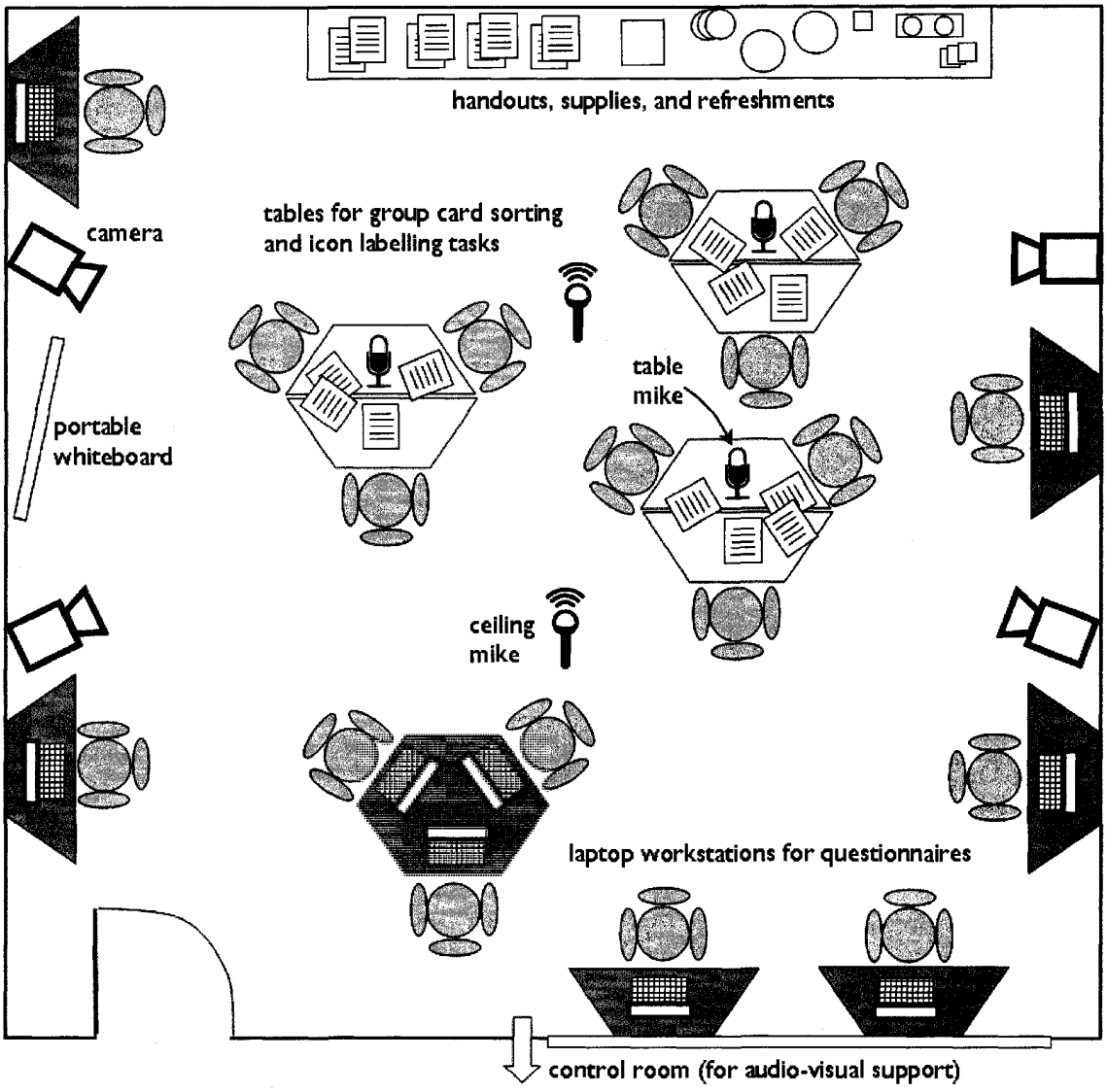


Figure 17: The design activity environment, in the Groupwork Lab. Equipment included videocameras, microphones, laptop workstations; tables were arranged for group activities and icon labelling tasks.

Participants were placed into groups of three for the card sorting part of the study, which required that each participant needed to be scheduled in a timeslot with two others. Groups were arranged to avoid homogenous clusters (e.g., all computer science students), and a timeslot was selected that accommodated the entire group's schedule. The first session included two groups of three (six participants), and all other sessions consisted of a single group of three.

Table 31 shows a summary of the participant demographics, including the areas of expertise represented in this population.

item	categories	count	comments
Sex	male	13	
	female	5	
Age	under 25:	6	
	25-35	9	
	36-45	3	
Role	graduate student	12	Note that some participants held multiple roles: for example, were both a student and a staff member
	undergraduate student	3	
	university staff	4	
	professional from outside university	5	
Areas of expertise	computer science	12	Note that some participants had expertise in multiple areas, such as computer science and human-computer interaction, or design and information science. For "other", two participants listed science degrees (biology and physics); both had computer science experience as well.
	human-computer interaction	9	
	business/management	5	
	design	4	
	information science	3	
	other	2	
Computer experience	5-10 years	3	
	11-20 years	10	
	more than 20 years	5	

Table 31: Demographics for design activity participants (n=18)

7.1.6 Procedure

When a participant arrived, he or she was greeted individually, was given a verbal introduction to the activity, and was provided with a consent form. Any questions that participants had were addressed before the consent forms were signed. Each participant was then shown to a seat at the designated groupwork table, to wait until all three group members arrived (and signed their consent forms). Refreshments were provided, and participants were encouraged to take breaks as required for food and drink.

At the start of the study, the groups were given an outline of the research topic, so that they understood the general design space in which they were working. (The script that was used in the introduction is presented in Appendix C.) However, participants were not shown any specific designs until the later stages of the study, in the computer-based questionnaire (Stage Three). This approach was taken so that their initial impressions (e.g., of the icons) approximated those of a new user, which would highlight problems with icon understandability. In addition, delaying the presentation of the full design would help to ensure that their ideas were not overly constrained, and would allow

participants to feel that the design was not "set in stone" (thus encouraging them to make suggestions).

As described above, the participants completed four stages of the study: the group card sort; the icon-labelling task; the computer-based questionnaire; and the group discussion.

1. Stage One: Group Card Sort

At the start of this activity, the groups were given a verbal overview of card sorting; see Appendix C for the full text. A handout with the instructions was also provided. For this task, the groups were asked to place unlabelled icons into related sets. Once the sets had been created, then the group should create a label for that set, one that describes what the members of that set have in common (i.e., why they were grouped together). Each set was to be paper-clipped together, and the label was to be written on a self-adhesive note and attached to the paper-clipped set. (Participants were provided with office supplies for this task, including self-adhesive notes, pens, paper clips, and binder clips.)

Because the participants were unlikely to be familiar with card sorting, they first completed a warm-up task, using ten cards with icons from a word-processing application. This warm-up gave participants a chance to get to know their group members, and to ask any questions they had about the card-sorting task. Once all the cards had been sorted, and the groups labelled, the main card-sorting task began.

Participants were given the set of 20 icons (shown in Figure 13); as before, they were asked to place these into groups and label each group. Note that the icons were unlabelled (except to indicate size), and the participants only knew that the icons represented concepts within the file sharing domain; they had not seen the proposed file manager design at this stage. Duplicate copies of icons on index cards were provided, in case the group wanted to place an icon in more than one group. Participants were told that they could also modify the icon drawings, and they were given blank cards so that they could add new icons (i.e., their own designs) to the original set if they so desired.

The group interactions (particularly discussions) were videorecorded and audio-recorded; a set of stills from the videorecordings (which could be combined in a splitscreen) is shown in Figure 18. As the facilitator, I took notes and answered questions,

although the groups were mainly left alone to interact with one another unless a problem or question arose. This activity took approximately 20 minutes, depending on how readily a group achieved consensus.



Figure 18: Video stills from group design activity, showing interaction between group members. (Faces of participants have been pixellated to preserve confidentiality.)

2. Stage Two: Icon-Labelling Task

The next activity was an icon-labelling task. Participants were informed that they were now completing an individual task, rather than a group task; for details of the script used, see Appendix C. Participants were instructed to label each card, with as many labels as they felt were applicable, and that they could write in their own labels. Participants were encouraged not to leave any cards unlabelled, although they could label such cards as "unknown" if desired; they were also informed that labels could be used multiple times for different icons.

A copy of the instructions was provided on a handout for each person. Each participant was given a stack of icon cards (30), and several sheets of adhesive labels. Each sheet contained 80 labels: two copies of the 30 terms, along with 20 blank labels for writing one's own terms (if desired). The terms were listed alphabetically, to help participants to find them on each sheet. Participants also had the same office supplies as before. Extra label sheets were provided to anyone who ran out. This activity took approximately 10 minutes to complete.

3. Stage Three: Computer-Based Questionnaire

When a participant completed the icon-labelling task, he or she was shown to a workstation to complete the questionnaire. The participant was informed that paper copies of the file manager design were available; these were face-down on the desk beside the workstation, along with pens. The participant was asked to ignore the screenshots (leaving them facedown) until they reached the portion of the questionnaire that explicitly mentioned these screenshots. In this way they could answer detailed questions about the icons "out of context"; these questions were completed in the first sections of the questionnaire. This activity took approximately 25 minutes.

4. Stage Four: Group Discussion

When a participant completed the questionnaire, he or she was brought back into the groupwork area, near a whiteboard. Participants had to wait for their group members to complete the questionnaire before beginning the discussion. When all members were ready, a short discussion was held; this was a loosely-moderated discussion intended to gather any comments that had not been recorded in the earlier stages. As with the group

card sort, participant comments were videotaped (with audio) for reference; notes were also taken during the discussion. This activity took approximately five minutes.

After the group discussion, the session ended and participants were compensated.

7.1.7 Data Analysis

7.1.7.1 Group Card Sort

The icons selected in each grouping, and the labels for each group, were transcribed from the index cards and recorded in a computerized spreadsheet. These groupings were studied to determine whether there were dominant groups for the icons (by count), and discovering what labels were associated with these groups.

Icons were also re-grouped to identify any large clusters among the categories: in particular, this analysis would indicate which icons (if any) were *never* classified together. For this analysis, icons were put into one set if they were classified together (even once) by any group. (Icons that were uncategorized or unlabelled were omitted from this clustering.) As an example, Group 3 placed icons  and  together (under "global ops"), and Group 1 placed icons , , and together (under "connections"); those two sets were therefore merged into a larger set of icons containing [, ,]. The result of this analysis was the creation of disjoint sets of icons.

7.1.7.2 Icon-Labelling

The labels for each icon were transcribed from the cards and placed in an Excel spreadsheet, such that every icon was shown with the complete set of labels that had been associated with it, and the number of times each label was used for that icon. In addition, the number of times a term was used (if at all) was counted, to indicate the popularity of each term. Finally, a comparison was performed to see whether there was any overlap between terms associated with file activity (e.g., sent, modified) and group/user concepts (e.g., public, others): how often were terms from these two sets applied to the *same* icon?

7.1.7.3 Questionnaire

Data from the questionnaire was placed in an Excel spreadsheet. The text of the freeform responses was not statistically analysed, but was reviewed for dominant impressions, opinions and concerns.

For the icon critique (i.e., the checklist of characteristics), counts were recorded: for each icon, the number of times a particular characteristic, such as "clear", or "poor" was selected for that icon.

The analysis of the push- and pull- oriented technology used calculations based on ranking of terms. The ranked items were assigned points: five points for a word selected as first choice, down to one point for the word selected as fifth choice. In addition, the number of times that a word was marked as unsuitable (e.g., would never use it) lowered its score by one point per occurrence. If, for example, 10 people chose the word "forwarded" as their first choice, but 5 marked it as "I would never use this word", the total score for that word would be 45 ((5 * 10) + (5 * -1) = 45 points total). This method of calculation included both positive and negative impressions in the overall count, but did not allow negative scores to unduly dominating the total score. These totals were used to indicate which terms were most preferred overall. For the matched-pair question, the top choices were counted (as this was not a ranking question).

For the file manager information questions (screenshots), counts were made of correct responses to the questions. Similarly, the numbers of correct responses to the sharing console questions were recorded. For the file manager icon questions (e.g., which icon would you choose for *sent*?"), count data was recorded: the number of participants choosing each icon.

7.1.7.4 Group Discussion

The videotapes from the discussion, along with the facilitator notes, were reviewed for dominant impressions, suggestions and concerns.

7.1.8 Results

7.1.8.1 Results: Group Card Sort

The categories that the icons were placed into are listed in Table 32 below. This table refers to the icons as numbers (for ease of identification). The categories created by each participant group are listed; the category description consists of the group-generated label (e.g., "add users"), followed by the set of icons in that category (identified by number for each). For example, Group 1 placed the icons  and together in a category, and called that category "add users"; in the table, using numbers to stand in for icons, this category is written as "add users: {15, 9}."

icon categories with labels from each participant group (consult key for numbers)
key: each icon is represented by a number in the lists below. The mapping is as follows:
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

Participant Group 1
add users {15, 9}; managing groups {10, 14, 13}; size (zoom) {12, 3}; undo {1}; transferring {6, 17, 16}; navigation {18, 4, 5}; connections {20, 7, 19}; status {13, 7, 2, 11, 8}
Participant Group 2
file history {"A"[new], 17, 1, 16 (modified)}; file history -> submenu: modifications – {12, 3, 18, 5}; add permissions {9, 15, 19, 7 (modified: "add blocked person", with "+"); check permissions {20, 2, 7, 8, 14, 10}; orphans{4, 11, 13, 6}
Participant Group 3
modify users/group {15, 9, 14, 10}; sizing {3, 12}; "?" {7}; global ops [operations] {19, 20}; assign users/groups {11, 2, 13, 8}; navigation {4, 18, 5}; file actions {17, 16, 1, 6}
Participant Group 4
globes {19, 20}; files w/arrows {6, 16, 17}; arrows {5, 4, 18}; sets of four arrows {3, 12}; zipping files {1}; social networks {7, 9,10, 11, 14, 13, 2, 15, 8}
Participant Group 5
replied {1}; compress/uncompress {3, 12}; networking (global vs internal) {19, 20}; file (upload, download + send) {16, 17, 6}; navigation {5, 18, 4}; online/offline {10}; sharing options {15, 9, 14, 13}; file sharing status {7, 8, 11, 2}
Participant Group 6
group/user permissions {11, 7, 9, 2, 14, 10, 15, 13, 8}; file transfer {1, 16, 17, 6}; navigation {18, 5, 4}; online status {19, 20}; view {3, 12}
Additions: Group 2 added an icon ("A", shown right), under "file history": a clock with a file and a group. Discussion indicated this icon was for displaying the list of people to whom the file was sent; to match this, icon 16 was modified to display the set of *senders*. 

Table 32: Categorization of icons, including labels, created by each participant group. Note that the elements in each category are listed as numbers (rather than graphically as icons); consult the key in the top row for the mapping between numbers and icons.

Few groups added icons, but Group 2 did provide one new one, shown in the final row; this group also modified icon 16 (as shown in Table 32, to create a pair of icons for displaying file history information (sets of files senders and recipients).

These categories were reorganized to look for disjoint sets: that is, sets of icons that were *never* categorized together. When this analysis was performed on the participant categories (from Table 32), three disjoint sets emerged, as shown in Table 33.

set of icons appearing in same group	dominant category for icons (common label)
	file actions; file transfer
	navigation
	people and group concepts (e.g., add user), file permissions

Table 33: Overall clusters of icons, based on participant groupings. The sets in Column One consist of icons that were grouped together (at least once) with at least one other icon in that set. Items in different rows were never categorized together by any participant groups.

A brief explanation of how to interpret the data in this table may be helpful: this data shows that the icons in each of these three sets were never grouped with icons in one of the others (e.g.,  was never grouped with ). For each set, the most-common labels, or the unifying theme, for the icons appearing in that set is also provided, to indicate the high-level clusters that emerged.

7.1.8.2 Results: Icon Labels

Participants also labelled these 20 icons, matching each with one or more terms. Table 34 shows the counts for each term: how often it was used for a specific icon. The rows are grouped and shaded to distinguish between different types of terms. The first set (white rows) represent *file-activity* concepts, while the second set (grey rows) represent *user-oriented* concepts. Similarly, the icons in the first nine columns (starting from the left) represent *file-activity* concepts, while the icons in the last eleven columns represent *user-oriented* concepts.

Presenting the data in this fashion demonstrates the degree to which the participants labelled the icons according to these two expected categories (file versus user); that is, it indicates whether it was clear which icons belonged to which category. This is similar to the disjoint set analysis presented above; however, in this case, the

division was based on primarily specific labels (a closed card sort rather than open). Furthermore, the results were intended to show correspondence between icons and labels, rather than similarities based simply on the icons themselves (as in the disjoint sets).

Note that if there were perfect agreement between the *expected* labelling and the *actual* labelling in these two categories, all *file-activity* counts would be in the top left quadrant and all *user-oriented* counts would be in the bottom right quadrant. Also in that case, the top right and the bottom left quadrants of Table 34 would be empty.

		icons: each cell below is count (how often term was applied specific icon)																		
	term																			
file-activity concepts	accessed	2	1		2	4	1	1			1									
	changed	1	2								4		1	2				1		
	delivered	2	1	2	1		1		2	1			1	1						
	distributed			1	2	4	1		1	1					1					1
	forwarded	3		1		1	5		10	1										
	gave			2				1	2					1						
	modified	1	4								5		1	2					1	
	obtained	1	2		7		1	2		4										
	opened	1				9		1			2									
	pulled	1	1	2	5	1	1	7		4										
	pushed			5	2	1	9			1										
	received	2	1		8			2		4				1						
	sent	2		4		1	1		8	1										
	took	2	1		3			1		1				2						
	touched										1		1							
	transferred	3		5	4		1		3	2								1		
	transmitted	1		6	2	2			3					1				1	1	
	used					2				1										
	viewed	1				3		2					1	2						
user-oriented concepts	all														4		2			2	1
	everyone														7		2		1	1	3
	group											1		1	4	14	1	2	8		
	groups														1	1	13	9	2		
	others													2	3	3	3	2			
	people		1			1								1	7	5	4	1	3	1	
	person						1	1			5	12	5								
	public													1	6		1	1	1	6	3
	user										3	13	7	2	2						
	users													3	4	9			5		
	world																	1		16	11

Table 34: Counts of how often a label was applied to each icon. The first set of rows (first 19 white rows) are intended to represent file-activity concepts, while the last set (11 grey rows) are intended to represent user-oriented concepts.

Fourteen of the 18 participants (78%) also created their own labels for icons. These labels are listed in Table 35, organized by icon.

icon	participant-created labels created for this icon
	undo, previous version, back, revert, zipped file, compress
	
	enlarge, checked out, expand, expanded
	forward, end, align right
	back
	send
	offline, unavailable, access denied, error warning, block, no access, owner, danger, block, special/flagged user
	online, available, me
	add, add user, add person, friend
	permissions, night/day, toggle, night, online/offline, hide
	
	shrink, commit changes, contrast, shrink/make smaller, smaller, closed, shrank, minimize, compressed
	contrast
	delete, permissions, remove all/everyone, split, separate, restricted, block group
	add, add users, add group, friend
	upload, uploaded
	download, downloaded
	undo, previous version, back, "?", under
	network, internet, web
	private, internet, network, map, locate

Table 35: New labels created by participants

7.1.8.3 Results: Questionnaire

The questionnaire results are presented in five parts. First is the icon critique (feedback on icons), followed by push-oriented and pull-oriented terminology results. Next are the two parts of the file manager critique (excluding the sharing console): the question-answering portion (using screenshots), and the icon preferences. The fifth and final set of results is about the sharing console, for the question-answering portion (using screenshots).

7.1.8.3.1 Icon Critique

Participants critiqued each icon by checking items from a set of descriptors (e.g., *confusing*) that they thought applied to that icon. Table 36 shows the total counts for each icon descriptor (where the maximum possible count for each descriptor was 18, because n=18). The descriptor with the highest count for an icon is highlighted in bold for that row; in some cases, multiple items tied for first.

icon	poor	good	complex	simple	confusing	clear	uninformative	understandable
	6	8	3	**13**	10	3	9	5
	11	2	3	11	**15**	1	14	1
	3	13	0	**14**	5	9	6	6
	3	10	0	**14**	8	8	8	7
	4	12	0	**15**	3	9	7	7
	1	13	1	**15**	4	12	2	13
	7	**9**	5	**9**	**9**	5	8	6
	0	**15**	1	14	2	**15**	1	14
	0	**15**	1	14	2	14	1	14
	1	12	0	**15**	6	10	6	10
	2	9	2	**14**	8	7	6	8
	9	3	11	6	**14**	1	9	2
	4	7	4	**11**	8	7	7	5
	6	7	9	6	**13**	3	8	4
	0	16	0	**17**	0	16	0	14
	1	13	0	**18**	1	11	3	11
	1	14	0	**17**	0	15	0	15
	7	8	2	**13**	11	4	12	3
	0	**17**	2	14	3	12	4	11
	9	8	8	8	**13**	3	11	4

Table 36: How participants described icons, according to the checklist. Numbers in each row represent the total number of times a descriptor was chosen for that particular icon. Columns that are shaded grey indicate negative descriptors. The bold cells indicate the descriptor with the highest count for each icon; there are some ties.

7.1.8.3.2 Push-Oriented and Pull-Oriented Terminology

The score for each of the terms in the ranking questions are in Table 37; the score was computed as described in the Data Analysis section above. Scores could be negative, if sufficient people marked a term as "I would never use this word;" each occurrence led to a one-point deduction. The total number of these occurrences is listed under the heading "negative count."

In one case, a participant added their own word to their top five choices (*altered*), ranking it third. Other terms were also added, but none of these were selected as a top-five choice (but simply as an alternative suggestion).

pull-oriented terms	total score	negative count	push-oriented terms	total score	negative count
accessed	60	0	sent	78	0
opened	41	2	forwarded	68	0
viewed	39	2	transferred	44	2
received	21	2	gave	28	2
used	19	3	distributed	19	3
modified	14	3	delivered	8	7
changed	13	3			
obtained	11	4			
touched	-1	9	pushed	-4	8
pulled	-3	9			
took	-7	9			
other: *altered* (ranked 3rd)	3				
other suggestions (not in top five): *printed, updated (x2), read (x2), downloaded*			other suggestions (not in top five): *dropped, emailed*		

Table 37: Top choices for push-oriented and pull-oriented terms. Additional suggestions also listed; any that appeared twice are marked as (x2). The *negative count* refers to the number of times that a term was marked as "I would never use this word", each occurrence of which led to a point deduction.

When asked which pair worked best together, participants most commonly selected "sent and accessed" (8 out of 14, or 57%). Note that participants did not have to select the same terms as they ranked first in the previous question, although the results were the same. The only other pairs to appear more than once were "sent and opened" and "forwarded and accessed", each of which were selected twice.

Participants also commented on how well these terms described the target file-sharing concepts, as well as expressing the distinction between the two concepts. The participants were unanimous in stating that the terms were both understandable and clearly distinct. The main concern expressed in the freeform comments was that some terms from the lists were superior to others (depending on one's personal preference for certain words), and thus not all terms were equally valid for expressing file sharing concepts.

7.1.8.3.3 File Manager Questions (Main Listing Window)

Participants answered a series of questions about file information displayed by proposed file manager (such as "how many times has file X been sent?), using screenshots. Participants answered the majority of questions correctly, with very few errors (two) occurring across questions. The most common error was misidentifying which file the tooltip was describing, concluding it was the file listed immediately below the one that the mouse was actually pointing at. (This is a common occurrence in managers with file listings: the small size of the list elements means that the mouse pointer may cover more than one element at a time.)

Suggestions for improvements were solicited, the most common of which was to present additional details in the tooltip—the types of detail provided in the sharing console, such as the full history. (At this stage of the questionnaire, participants had not seen the sharing console, so they did not know there was another way to access this information.) There were no major flaws identified, and participants described the design as appropriate and clear.

7.1.8.3.4 File Manager Icon Preferences

This section asked participants to indicate their top choices for icons for a set of seven file-sharing concepts, after they were shown the icons in the full context of the file manager. (That is, screenshots of the full design were presented, with examples of how the icons would be used.) These icons would be used for the columns in the main file listing, to indicate for file activity and permissions.

Participants could pick only one icon for each, representing their top choice, and they could add their own choice (on paper) if they wanted. The results are shown in Table 38, which shows the top choice and the second-place choice icons for each file-sharing concept. The count for this icon (i.e., how often it was selected) is listed in brackets after each icon, along with the percent of participant who chose it (n=18).

		top icon choice		second-place icon choice	
file-oriented concepts	sent		(10; 56%)		(3; 17%)
	accessed		(4; 22%)		(3; 17%)
person-oriented concepts	user		(12; 67%)		(5; 28%)
	users†		(10; 56%) [5 each]		(5; 28%)
	group		(8; 44%)		(5; 28%)
	groups		(7; 39%)		(4; 22%)
	everyone		(16; 89%)		(1; 6%)

† In this group, there were two icons that differed only by colour; each was selected 5 times, and as they are essentially interchangeable their totals were combined.

Table 38: Top choices for icons for the file manager, with counts and percentages for each (n=18).

7.1.8.3.5 Sharing Console Design Questions

As with the earlier file manager section, participants were asked factual questions about file sharing information shown in screenshots of the sharing console. Participants performed very well; they were readily able to interpret the information that was displayed in the console, with only four errors being made. (These errors consisted of overlooking elements in the file sharing history and the search results.) Participants also made accurate guesses as to where they would find information and how they would use the console features.

Participants also made some design suggestions. Predominant among these was a desire for more complex searching (as the initial design was quite rudimentary in this regard). In addition, there was a request for more information in the "results" pane, such as file size. There were also recommendations for the overall aesthetics, primarily to resize and move the console elements (e.g., the panels and windows) to create better balance, and to make the text as legible as possible.

7.1.8.4 Results: Group Discussion

Participants generally spoke briefly, perhaps because they had just completed a lengthy and involved exercise. Some participants expressed general interest in the research project, and were intrigued by the idea of designing for file sharing. One discussion point that arose was the difficulty of finding the right icons for the variations of "multiple people": *users*, *group*, and *groups*. A few participants noted they had trouble

choosing between icons for these terms when they were completing the questionnaire. (This was reflected in the questionnaire answers, where a great deal of overlap was found in these areas, particularly in the icon choices for *users* and *group*. This subject is revisited in the Discussion section below, in the Questionnaire portion.)

7.1.8.5 Discussion

7.1.8.5.1 Group Card Sort

There were several different groupings that emerged, but overall, there were three disjoint sets, as shown in Table 33: icons with only an arrow on them (generally labelled "navigation"), icons with a file plus an arrow (generally labelled "file activities"), and icons involving people and groups (generally labelled "people" or with a user-related activity). This result indicates two things. First, the abstract "arrow" label was not as successful in indicating file activity as the file-with-arrow icons were: an arrow alone was interpreted to indicate movement rather than activity. Second, the categories that were created for the file manager, in which file action concepts were separated from user and group concepts, closely matches the way in which people separated the icons. Therefore, it is unlikely that there will be confusion between the icons in these different columns, indicating that the information display is clearly organized.

7.1.8.5.2 Icon Labelling

Table 34 shows that, as with the group card-sorting activity, there was little overlap between the user-oriented concepts and the file-activity concepts. This can be seen by looking at the distribution of the counts: they are mostly clustered in the top left and bottom right quadrants, with the terms and icons from these two categories being grouped together.

Occasionally the terms and icons fell outside this strict categorization: see for instance the icon  which has a single instance of being labelled as "person," but otherwise, has all file-activity labels. This is the most common pattern: very occasional overlap between categories, and usually a single instance. There are some exceptions; however, analysis showed that in the majority of cases, this was attributable to a single participant, and did not represent the majority of responses. (That is to say, there was one individual who continually cross-categorized icons as *both* user-oriented and file-oriented, but this was quite unusual.)

One notable exception was the categorization of the icon  , which had many different types of labels matched with it (both user-oriented and file-oriented). As shown in Table 35, this was also the icon that received the most new labels, many of which indicated some type of danger or caution. It is likely that the red color was the source of this label: although participants were informed that colour was unimportant (and could be changed in the final design), red was such a strong indicator that it could not be readily ignored. Contrast the labels for this icon with those applied to the next icon in the list:  . Although both icons basically indicate a single person, the red version has many more associations, but the plain grey version is more cleanly categorized: all but one of its labels were "person" or "user." As a practical piece of testing advice, it is recommended that red be avoided (unless perhaps the entire set of icons is red), unless this colour is intended to be significant.

A second example of an icon that had many different labels was  . This icon was intended to represent multiple users, with the two-tone background meant to highlight that they were not simply member of one group. The "multiple users" concept was not conveyed to the participants; instead, it was interpreted as *duality*, as evidenced by the new labels created for it (listed in Table 35). The meaning of this icon was not at all clear out of context, and therefore it was not a strong contender for inclusion in the final design.

Despite these exceptions, the results demonstrate that overall, the two categories of file-oriented and user-oriented concepts were separate and distinct. This finding reinforces that of the group card-sorting activity: the file manager columns showing these two types of information are unlikely to be confused.

There was another issue related to the clarity of icons: there was often no clear distinction made between the concepts of users, group (singular) and groups (plural). These three concepts, when shown visually, may simply appear identical, communicating only the idea of "multiple people." This confusion is evident when looking at the results (in shown Table 34): six different icons were all matched with users, group, and groups. The design issue is that if one of those icons was chosen to convey *users*, then other people might interpret that icon to mean *group* instead.

Consider for example the set of icons in Table 39, all of which could be used to indicate *multiple people*. For each icon, a checkmark (and count) is shown in the column beneath if that icon was labelled as *users*, *group*, and/or *groups* (as listed in Table 34).

label						
users	✓9	✓3	✓4			✓5
group	✓14	✓1	✓4	✓1	✓2	✓8
groups	✓1	✓4	✓1	✓13	✓9	✓2

Table 39: Icons that participants interpreted in at least three ways: as *users*, *group*, or *groups*. A checkmark appears under any icon which was labelled with that term (e.g. *users*) at least once; the actual count is listed beside each checkmark.

Note that most of these icons were labelled with all three terms, and every icon had at least two terms applied to it. Unfortunately, the distinction between these three concepts is very important for file permissions: in the file system, these are distinct categories. Permissions given to a specific user are handled differently when that person is a member of a group, and each group's permissions are configured separately. Thus, in order to let a user know who has access to files, it is vital to communicate whether a file is shared with multiple individuals versus members of a group (or groups). Finding a suitable visual means to make a clear distinction between users and group(s) is a design challenge.

This lack of clarity was an expected result; the issue arose when selecting the original set of icons for use in the design activity. For this reason, specific questions were included in the online questionnaire to delve more deeply into the problem; this issue is revisited in the following section of this discussion (about the questionnaire).

7.1.8.5.3 Questionnaire

As in the Questionnaire Results section above, the discussion is presented in five parts: icon critique; push-oriented and pull-oriented terminology; the two parts of the file manager critique (question-answering portion and the icon preferences); and the sharing console.

Icon Critique

There results indicate some dominant impressions overall, within the wide range of opinions. Consider first the catch-all descriptors "poor" and "good," which were included to allow for sweeping judgements. The icons that were most often judged as

"poor" were , , and , with at least half of the participants making this choice. In contrast, the icons most often judged to be "good" were , , , and  . One point of note between these two results is that despite their similarity, the icons and were rated near-opposites; this is likely because was also judged as "uninformative."

Taking a look at the negative descriptor columns in Table 36, there are certain icons that were repeatedly perceived as flawed. In particular, and were judged to be complex and confusing; similarly, and were confusing and uninformative. It is important to reiterate that these icons were being analysed out of context, so these assessments reflect initial impressions. Participants' comments reflected how difficult it was to assess the icons independently from the rest of the design. For example, one participant commented that one icon "could be very effective, but difficult to understand out of context," while another remarked that "it may be probably easy to associate most of the icons once you are given what they do, but it is quite difficult to guess what they mean at first."

Of course, contextual information is very important for understanding. Therefore these icons were revisited later in the questionnaire, within the actual context of the preliminary design. Those findings (and how they compare to the initial impressions) are discussed later in this section.

Push-Oriented and Pull-Oriented Terminology

For the pull-oriented term, the top choice was "accessed," although "opened" was also a strong contender (in second place). For the push-oriented term, "sent" was the top choice, with "forwarded" taking second place. Note that top-finishing terms "sent" and "accessed" were the same as the terms used in the preliminary design, indicating that these were reasonable original choices for column headings. (Note that participants had not seen the proposed design yet, so this had no influence on their choice of terms.)

When asked which pair worked best together, participants most commonly selected *sent* and *accessed*. Thus *sent* and *accessed* were seen to work well together, as well as being rated highly as individual terms.

In the freeform comments, a few participants expressed strong preferences, judging some terms to be particularly good or poor. It is difficult to accommodate the preferences of *all* users; this concern can be at least partially addressed by choosing the

most highly-ranked terms, although this approach may not satisfy the personal preferences of every potential user.

File Manager Questions (Main Listing Window)

Participants performed very well in answering questions, with very few errors occurring across questions. Most errors could be attributed to the small size of the file manager window, rather than the information presented within it. This result indicates that the file sharing information is likely to be understandable to users. There were no major flaws identified in the overall design of the main listing window; overall, the design was judged to be suitable.

File Manager Icon Preferences

For file-oriented concepts, the results show a preference for the *sent* icon, with a majority of participants (56%) selected it as their top choice. For the *accessed* icon, only a slight preference was demonstrated, with 22% of respondents choosing that icon. (This number (four participants) was the largest count; for all other icons, the count was three or fewer.) However, as this icon was preferred (however slightly), and did not demonstrate any particular design problems (e.g., in the card sort results), it is a reasonable choice for inclusion in the design.

For the person-oriented concepts, all five terms showed a gap between the first and second-place choice, making the top choices for each icon suitable for inclusion. However, note that the same icon () was frequently chosen for both *users* and *group*, but it cannot be used (within the same design) to refer to two different things. Because this icon was chosen more often for *users* (56%) than for *group* (44%), this suggests that  should represent *users*. In that case, another icon must be chosen to represent *group*. The second-place *group* choice () seems to be the best remaining option, although it was chosen by fewer participants than any of the top choices (at 28%) and was chosen by equally many people to represent *users* instead. This outcome underlines the challenge of visually representing the distinctions between multiple-people icons; this issue was re-visited in a short follow-up study, described later in this chapter, used to resolve the choice for the *group* icon.

As a side note, one interesting finding was that the most popular *users* icon () was ranked poorly in the earlier questions, where it was seen out of context; at this early

stage, the icon was viewed as uninformative. As many participants noted in their comments, an icon's context makes a huge difference to its interpretation. In this case, the icon went from a poor rating to a very high rating; it likely helped that the icon was initially judged as "simple," meaning that it had some redeeming features that made it useful when more contextual information was added to facilitate interpretation.

Sharing Console Design Questions

As with the file manager questions, participants performed very well; they were readily able to interpret the information that was displayed in the console, and could accurately determine where to find information and how they would use the console features. Design suggestions that were made included supporting more complex search queries, providing more details in the search results pane, and to create better visual balance among the console elements.

7.1.8.5.4 Group Discussion

There were few comments in this stage, as it was short. However, one recurring issue was the problem of choosing the right icons for the variations of "multiple people," to distinguish between *users*, *group*, and *groups*. This problem was explored in a follow-up study, which was developed to gain consensus for the *group* icon, for the refined file manager design. That study is described in the following section.

7.2 Study Two: Consensus for Group Icon

7.2.1 Research Objective

As shown by the results of the group design activity—particularly the questionnaire results in Table 38—it was evident that choosing an icon for group was problematic. There was no clear consensus on the best icon; the same icons were used for both users and group, with no clear visual distinction between these concepts. To address this problem, a short follow-up study, using a questionnaire, was carried out. The objective of this study was to determine the preferred choice for the group icon.

7.2.2 Instrument: Questionnaire

As shown in Figure 19, the questionnaire consisted of a single question, with space for additional comments. In the questionnaire, participants were given a set of icons for *user*, *users*, *groups*, and *everyone*; these icons corresponded to the top-choice icons identified in the group design activity. Along with these was a set of five icons to

represent *group*; participants selected the one they thought best completed the set. (Participants could also provide icon design suggestions, if desired.)

Two of the five icons were used in the original design activity (and). Two others were new designs, based on other icons used in the design study (and). The fifth icon ( ) originated from the same public-domain icon sources that were reviewed in selecting the icons chosen for the group design activity. These icons are shown in Figure 19 within the context of the full questionnaire. This activity was intended to be quite short and informal, so that participants did not feel required to over-analyse the design, but could make a snap decision when picking an icon.

This questionnaire was provided both on paper and in an online version. For both the paper and online versions, the icons were presented in random order, to minimize bias. The only difference between the two versions was that the suggestion to draw design ideas was omitted from the online version, because freeform drawing was not possible in that environment.

7.2.3 Participants

The paper version was given (in person) to twelve students and faculty in Computer Science and Management. The 18 participants from the group design activity

This is an informal study on icon preference. I am selecting icons for a file sharing application: the icons are meant to signify who has access to a file: one user, multiple users, one group, multiple groups, and everyone. In earlier testing, the icon for **group** did not show a strong preference. Thus, I want you to choose which icon from the set below 1) best signifies "group" and 2) goes best with the other icons in the set. All icons are 16x16 pixels (very small):

user	users	group	groups	everyone
		?		

Choose from the following set: which one is best for group (singular)? Note that there is a distinction between users (just multiple individuals, like Dave, Bob, and Joe) and group (a specific set of people, like a work team). Also note that colours can be easily changed to match the other icons.

Best choice for group (put a check mark next to the icon):

_	_	_	_	_

Additional comments about your choice (e.g., suggested colour change, modifications to shape, etc.). <<paper version only:>> Feel free to draw:

Figure 19: Questionnaire for group icon followup study. The comments section was slightly modified for the online environment.

were emailed the web address for the online version; 15 responded (83%). (Note that these two participant groups were disjoint.) The participants from the group design activity were not all available for in-person meetings on campus; therefore this group was contacted by email and given the opportunity to participate online.

7.2.4 Results

The results for these two groups are shown in Table 40, with the counts for each icon for how often it was chosen.

paper version (new respondents)		online (design activity participants)	
icon	*count*	*icon*	*count*
	5		**10**
	2		2
	1		1
	1		1
	1		1
other	2 (drew own versions; all were variants of icons in set shown)		

Table 40: Results of followup group icon study (for paper version, n=12; for online version, n=15).

7.2.5 Discussion

The icon that was chosen most often by both groups (and particularly so with the design activity participants) was the icon with three individuals (). This icon most closely matched the "groups" icon that had been selected in the previous study, which may have been why it was most often selected as the best item to complete the set.

7.3 *Study Three: File Sharing Terminology Survey*

The majority of the design elements were finalized in the group design activity and the follow-up icon study. One portion of the group design activity involved terminology: determining which words would be best for describing push-oriented and pull-oriented file sharing activities. Replicating this portion of the study with a larger population could lend validity to these results. Therefore, a short survey was created, based on the design activity questions about terminology.

7.3.1 Research Objective

As in the group design activity, the objective of this survey was to identify the most appropriate terms to express the file-sharing concepts of *pushed* and *pulled.*

7.3.2 Instrument: Online Survey

A web-based survey with 12 questions was created using Opinio survey software (ObjectPlanet, 2007), and was hosted on a web server at Dalhousie University for ten days. This survey was based on questions included in the questionnaire portion of the design activity, which focused on the terms one would use in describing push-oriented activities (e.g., sending email attachments) versus pull-oriented activities (e.g., using a file in a shared folder). As in the design activity, respondents were asked to rank their top five choices from a list of terms, and to select which words worked best as a pair. The complete survey is presented in Appendix D, and is summarized below.

The first part of the survey collected demographics and information on respondents' computer backgrounds. Respondents were then presented with a list of 13 words that could be used to indicate pull-oriented file sharing (e.g., *used*, *opened*), and asked to rank their top five choices, as well as indicating which words they would never use to describe this activity; they were also given the opportunity to add their own words. This question was repeated for push-oriented sharing, with a list of seven words (e.g., *forwarded*, *sent*).

The two sets of terms were almost identical to those used in the design activity, with two additions: the words *read* and *updated* were added to the "pull-oriented" lists. These were added because both these words had been independently suggested by two different respondents (during the earlier design study). The sets of terms are listed in Table 41.

pull-oriented terms		push-oriented terms
accessed	received	delivered
changed	took	distributed
modified	touched	forwarded
obtained	updated	gave
opened	used	pushed
pulled	viewed	sent
read		transferred

Table 41: Sets of terms used in followup survey

Next, respondents selected the terms that worked best as a pair (e.g., "sent and opened"); these did not have to match their choices from the earlier questions. (Because the terms are being used together in the design, their suitability as a matched set needed

to be considered.) Finally, respondents commented about the clarity of these terms, particularly in describing these file-sharing activities and the distinction between these two approaches: push vs. pull.

7.3.3 Participants

The survey was open to all persons over the age of 18 with computer experience, and was advertised primarily within a university community, including campus-wide mailing lists and web pages. Recruitment emails were also sent to email lists for university-affiliated research groups.

Sixty-five people responded to the survey—41 male (63%) and 22 female (34%); two declined to specify sex. Approximately half of the respondents were students, and half were employees (faculty and staff). All respondents were experienced computer users: only three respondents reported fewer than five years' experience, and these three were computer science students.

7.3.4 Data Analysis

The scoring scheme was the same as for the design activity. The ranked items were assigned points: five points for a word selected as first choice, down to one point for the word selected as fifth choice. If a word was marked as unsuitable (i.e., "I would never use this"), this lowered its score by one point. If, for example, 10 people marked the word "forwarded" as their first choice, but 5 checked "never use," the total score would be 45; that is, (5 * 10) + (5 * -1) = 45 points total.

Given a participant population size of 65, the lowest possible score for a term was -65 (if all participants marked this term as unsuitable), and the highest possible score was 325 (if all participants selected this term as their top choice, at 5 points each).

7.3.5 Results

The highest-scoring and lowest-scoring terms are listed in Table 42. Total scores for these terms are shown (as calculated above) along with the "negative count": the number of times participants marked that term as "I would never use this word."

For the "pull-oriented" term, the top two choices were *opened* (with 150 points) and *accessed* (with 146 points); no other word received over 100 points. For the "push-oriented" term, the top choice was *sent*; *forwarded* was a distant second at 200, and no other word received over 200 points.

Note that the scores for push-oriented words were generally higher than for pull-oriented: the list of push-oriented terms was much shorter, increasing the likelihood that a given term was selected more frequently.

Table 42 also lists which terms had the lowest scores overall. Some terms received negative scores: *took* (at -34) was the lowest-ranked among the pull-oriented terms, followed by *pulled* and *touched* (tied at -26). The lowest-scoring push-oriented term was pushed, with a score of -48. For the number of times people marked "I would never use this term"—indicating a negative impression of that word—the highest counts were for *took* (38 [58%]) and *pulled* (37 [57%]), as well as to *pushed* (51 [78%]).

pull-oriented terms	total score	negative count	push-oriented terms	total score	negative count
top three scoring terms			top three scoring terms		
opened	150	5	sent	284	0
accessed	146	3	forwarded	200	4
viewed	90	8	gave	111	13
lowest three scoring terms			lowest three scoring terms		
took	-34	38	pushed	-48	51
pulled	-26	37	delivered	10	26
touched	-26	36	distributed	49	20

Table 42: Highest and lowest scoring terms for pull-oriented and push-oriented activities. The top three items for each category are listed, along with the three lowest-ranked term. The *negative count* refers to the number of times that a term was marked as "I would never use this word," each occurrence of which led to a one-point deduction. The highest possible total score was 325.

Respondents also indicated which terms worked best together; the vast majority (78%) chose *sent*, and 34% (the largest group) matched it with *accessed*. As in the design activity, the push-oriented term showed strong preference, with more divided opinion on the best pull-oriented term. These results differ slightly from the design activity: the same pair emerged as the top choice (sent with accessed), but the preference was much stronger in the survey, with over 60% choosing this pair.

When asked if the distinction between *push* versus *pull* activities could be clearly expressed using these terms, 53 of 59 respondents (90%) indicated that the distinction was clear. (Six respondents skipped this question.) When asked whether these terms were suitable for expressing the file-sharing activity that they were meant to describe, 50 of 57

respondents (88%) indicated that these terms were suitable. (Eight respondents skipped this question.)

7.3.6 Discussion

For pull-oriented sharing, the words *opened* and *accessed* were both ranked highly, and few people indicated they would never use these terms (five [8%] and three [5%] occurrences, respectively). Additionally, they were preferred by far over the other choices. Note that although *opened* had a slight lead over *accessed* in the survey, these ranks were reversed among the design activity participants, where *accessed* was scored at 60, and *opened* at 41. Adding the totals for both studies gives a combined total of 206 for *accessed*, and a total of 191 for *opened*. (The highest possible score for any term is 415, if it were selected as the first choice of all participants in both studies (n=83) and there were no deductions.)

The term *accessed* was chosen for the final design, due to its slight edge in scoring, and because it was most often chosen to be the best match for *sent*. (Note that both these terms were ranked highly; in other applications, *opened* might be an equally valid design choice.) For push-oriented sharing, there was only one clear favourite: *sent*, which was incorporated into the design. As it happens, this also matched the preference for the best-matching pair of terms, which indicates that the terminology choice is both clear and harmonious.

Note that the two terms that have been used in the technical sphere—*push* and *pull*—were very poorly received by the respondents, with low scores and several people indicating "I would never use this word." This was expected, although not confirmed by any studies before now. These words remain part of the technical lexicon, particularly within the domain of web-based media delivery; however, they are not widely accepted outside of that narrow field. Therefore, it is not recommended that they be used in most file sharing contexts.

7.4 Summary of Results

The goal of these studies was to select an appropriate set of icons and terms for use in the file manager, and to discover and correct design flaws before the implementation of the sharing tools. Based on the results of the design activity, and the followup studies, a final set of icons and terms was selected; these are presented in Table

43. Note that it is desirable to have a consistent set of icons, as recommended by designers (such as Horton (1994). Therefore, the colour of one icon (*accessed*) was changed, as shown in the table, to have a better colour match with the other icons; the resulting colour palette favoured greys, blues, and purples.

file manager: column headings (terms)	
concept	**column heading (term)**
push-oriented activity	sent
pull-oriented activity	accessed
file manager: icons for columns	
concept	**icon**
sent	
accessed	 colour changed: (from yellow to purple)
shared with: user	
shared with: users (plural)	
shared with: group	
shared with: groups (plural)	
shared with: everyone	

Table 43: Summary of icons and terms resulting from user study, included in the final design. Note that the colour of the "accessed" icon was changed to better match the rest of the icon set.

7.5 *Resulting Designs for File Sharing Tools: Prototypes*

The user studies identified appropriate icons and terms, and provided refinements for the original designs for the file manager and sharing console. Once the design elements were finalized, the testing prototypes could be developed. The following sections present the designs for the file manager and the sharing console, as realized in high-fidelity prototypes for Windows XP.

7.5.1 File Manager Prototype

As discussed in Chapter 6, the file manager design incorporates new columns in a "details view" of an XP file manager, such as Windows Explorer. Ideally, the prototype would match the so-called "look and feel" (appearance and behaviour) of Windows Explorer in all ways except for the new features being tested

However, creating a working prototype with all the features of this design requires modification of the file manager source code; the source code for Microsoft Windows Explorer is not available. Thus, another file manager, with functionality and appearance similar to Windows Explorer, was used as the basis of the prototype. This file manager is "Jeff's File Manager," which is written in C#; the author, Jeff Gaines, has made the source code available and has allowed others to modify it (Gaines, 2006).

The source code from Jeff's File Manager was rewritten to allow for three new columns (*sent*, *accessed*, *shared with*), with icons, text, and popups for each column. A screenshot of this file manager, with the new columns, is shown in Figure 20.

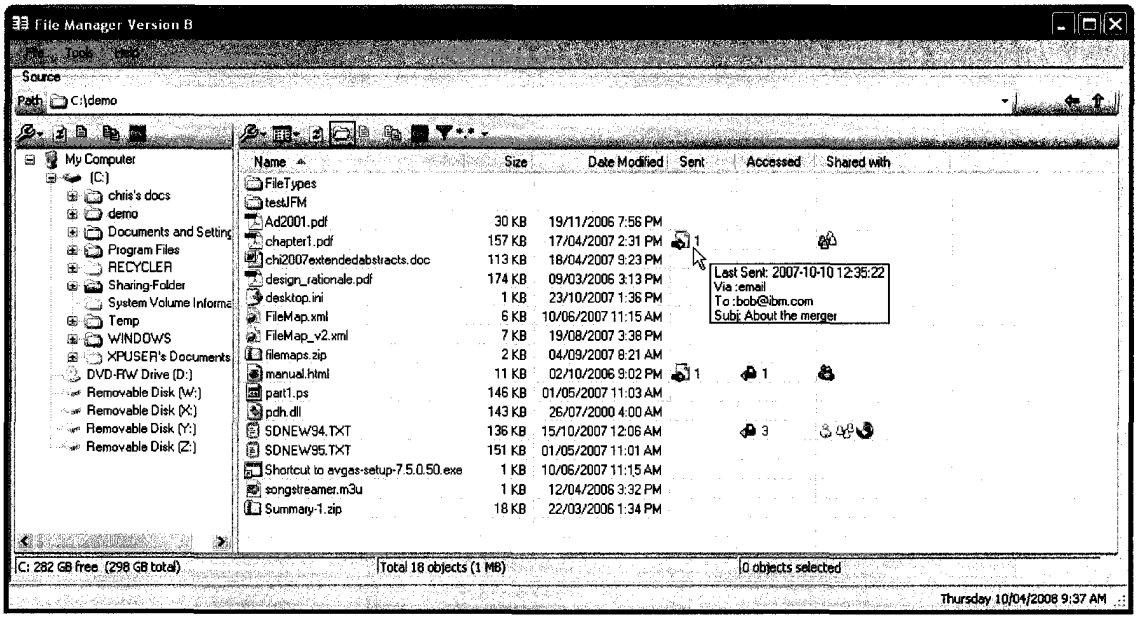


Figure 20: The file manager prototype developed for testing, based on *Jeff's File Manager.*

The file sharing information displayed by the file manager is read from a dedicated eXtensible Markup Language (XML) file. This file contains an entry for every file for which information is displayed. A sample entry is shown in Figure 21, for the file "demo_files\chapter1.pdf." In the example, the file has been sent twice (by instant messenger and email), was accessed once (read by `philip`) and has read and write permissions for one user (`philip`) and one group (`docs`). Note that reading an XML file was a technique used for prototyping purposes; for a full-fledged application, file sharing information would be read from other sources, such as file metadata.

```
<FILE id="2" path="demo_files\chapter1.pdf">
    <EVENTS fileid="2">
      <SENT date="2007-05-01 10:01:00" via="IM" to="john5"
subject="here is a copy of the draft" />
      <SENT date="2007-05-02 02:01:00" via="email"
to="alice@ibm.com" subject="Draft Chp 1" />
      <ACCESSED date="2007-05-01 01:01:00" by="philip"
action="read" />
      <PERMISSION user="philip" usertype="user"
permissiontype="read+write" />
      <PERMISSION user="docs" usertype="group"
permissiontype="read+write" />
      </EVENTS>
  </FILE>
```

Figure 21: Sample XML entry for file history and settings information, for the file "chapter1.pdf"

7.5.2 Sharing Console Prototype

Because the sharing console was a separate tool from the file manager, the prototype could be designed from scratch, as long as the two applications worked together. A context-menu item ("right-click" item) was included in the file manager for launching the console; it could also be launched as a standalone application (e.g., through a desktop shortcut). The sharing console was developed in C#, and was based primarily on the preliminary design, with some refinements suggested by the design participants. The resulting console prototype is shown in Figure 22.

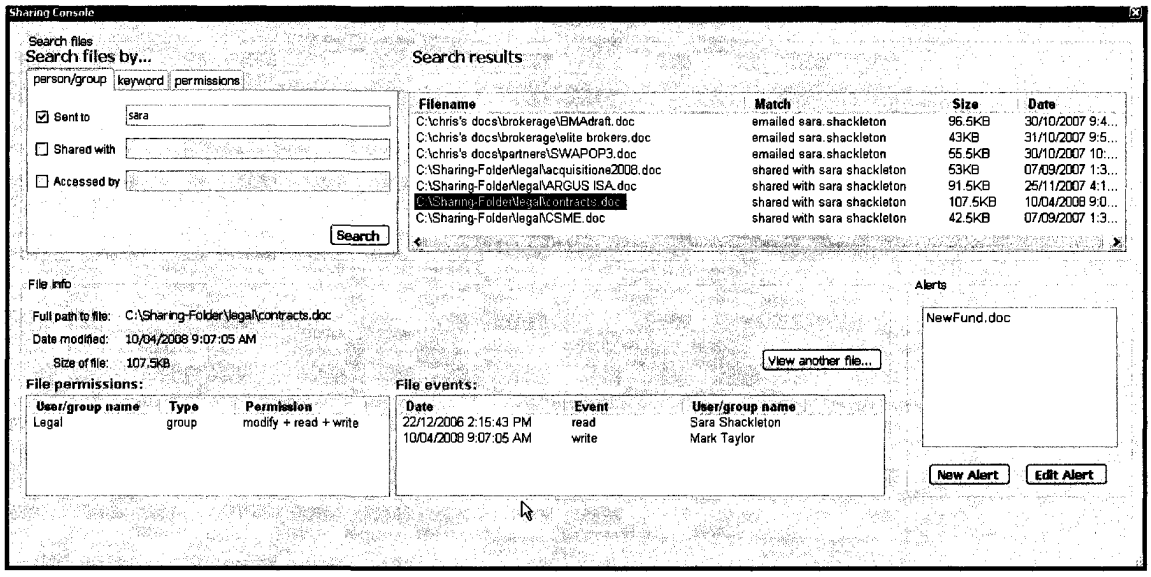


Figure 22: Sharing console prototype developed for testing

Note that the participants' design suggestions were included in this version: the "results" pane has more file information, there is a more flexible search feature (with a tabbed interface), and the panels have been rebalanced in the layout.

Because this was a prototype, the functionality was implemented to the degree required for user interaction for evaluation; this is a high-fidelity prototype rather than a fully-functional application. The functionality of the sharing console was implemented by a contract programmer as work-for-hire, according to a detailed specification that I provided; the specification is presented in Appendix E. This implementation was carried out under my supervision and guidance.

As with the file manager, some of the file sharing information was read from an XML file. The search feature, for example, returned "canned searches" from the XML file. An excerpt from a sample XML file is shown in Figure 23; this includes two sample queries (with matching files) and an alert. For the query, the "result" matches correspond to files; the numbers identify a specific file (as in Figure 23, which shows "file-id=2").

```
<queries>
    <query>
      <term field="sent-to">bob</term>
      <term field="shared-with">alice</term>
      <results>
        <result>1</result>
        <result>17</result>
        <result>8</result>
      </results>
    </query>
```

Figure 23: Excerpt from sample XML file, showing queries and alerts

7.5.3 Using the File Manager and Sharing Console

The file manager acts as a replacement for Windows Explorer; it is launched as a separate application, which can then be used just as Explorer—to browse the file system, open files, and perform other file management functions. The columns (with icons) appear automatically as extra information in the manager. When user wish to get more information about a file that is displayed in the file manager (such as sharing history), they can right-click on a file, and select "Open in sharing console," as shown in Figure 24. The information associated with that file will be automatically loaded and displayed in the console.

The console can also be launched separately from the file manager (for example, as a desktop shortcut). This approach may be followed by users who wish to perform a search, for example; this action is not related to a particular file.

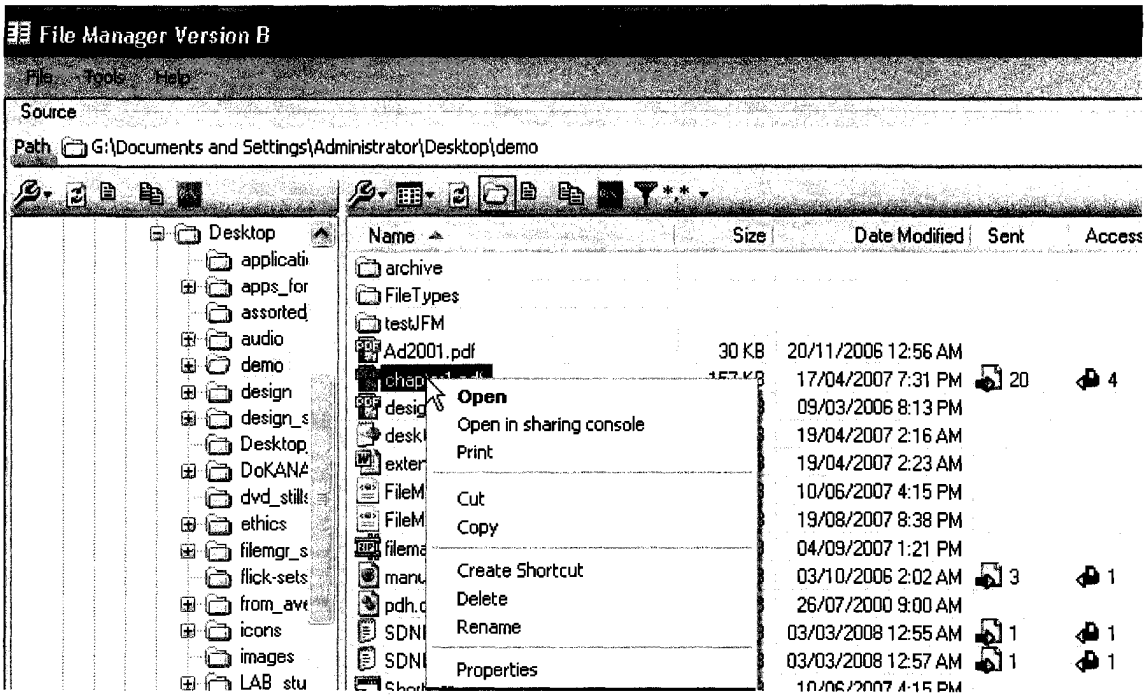


Figure 24: Right-click (context) menu for launching sharing console from file manager

7.6 *Summary*

The previous chapter presented preliminary design ideas; these ideas needed to be refined before creating prototypes for testing. This refinement involved a user-centred design process, which identified appropriate iconography and terminology for file sharing information displays. In addition, design participants demonstrated that they could effectively use the information displays, and provided design suggestions that were incorporated into the prototype designs. High fidelity prototypes were created, based on these designs, and assessed through formal evaluation. The evaluation is presented in the following chapters: Chapter 8 presents the design and methodology, and Chapter 9 contains the results.

Chapter 8. Evaluation Part One: Design and Methodology

Once the prototype for the experimental file manager was created, a formal user evaluation was carried out. This evaluation was designed to compare the experimental file manager with a control manager (similar to Windows Explorer) through a laboratory experiment. In addition, the experimental file manager was also evaluated in terms of its new features as well as intention-to-use. The research described in this chapter was approved by the Social Sciences and Humanities Human Research Ethics Board at Dalhousie University. This chapter presents the design and methodology of the laboratory study; the results are presented in the following chapter (Chapter 9).

8.1 *Study Overview*

The experiment was a within-subjects study: twenty-eight participants used two file managers (control and experimental) to complete a series of tasks, during two different sessions spaced about one week apart. The participants followed an office scenario, in which they played the role of an administrator filling in for an absent employee. The absent person asked the participant to look up information about shared files, through a set of "faxes" with questions such as "Who did I sent a copy of 'file.x' to over email?" Participants used one file manager (control or experimental) to answer questions for the first half of each session, and the other file manager for the second half. The file managers were compared on the basis of task completion time, correctness of responses to the file questions, usability score, and user preference. Participants also assessed the unique features of the experimental file manager, and provided data about their intention to use this file manager (should it become available). These components are presented in detail in the following sections.

The design and hypothesis of the research study are presented next, followed by a description of how a realistic workplace was simulated for use in the laboratory study. This is followed by a discussion of the tasks used in the study, and the study variables (including measures used to compare the control and experimental file managers). Next, the study participants are described, followed by a description of the procedures used during the evaluation. These items are summarized at the end of the chapter.

8.2 *Research Design and Hypotheses*

There were three main elements being tested in this study:

1. Whether the experimental file manager provided any advantages over existing file manager technology (used as the control condition), in terms of effectiveness, efficiency, or satisfaction;
2. Whether the unique functionality provided by the experimental file manager was effective for completing file management tasks; and
3. Whether the experimental file manager was likely to be adopted for use in the real world, should it become available.

All three elements were tested using a laboratory-based evaluation. For each of these elements, a set of research hypotheses were created, to assess whether or not the experimental file manager fulfilled the research goals. The experimental design was developed to verify these hypotheses, through data collection and analysis.

The research hypotheses are detailed in the following section; this is followed by a description of the overall research design.

8.2.1 Hypotheses and Research Questions

The testable hypotheses being tested in the experiment were as follows:

For *efficiency*:

H1: The experimental file manager will result in task completion times that are faster than those in the control file manager for at least 50% of tasks, and will not result in any slower completion times.

For *effectiveness*:

H2: The experimental file manager will result in at least 25% of question response scores that are higher than the control.

For *satisfaction*:

H3: The experimental file manager will have the majority of the higher scores for the usability questions in Session One or Session Two, or both.

H4: There will be a significant difference in at least one usability score, in Session One, between those who used the control *after* the experimental manager and those who used the control *before* the experimental manager.

H5 The experimental file manager will be selected more often than the control for the preference questions, in both Session One and Session Two.

In addition to these testable hypotheses, there were some research questions that applied to the experimental file manager:

RQ1: Do the new features of the experimental file manager help users to complete file management tasks (that cannot be completed with existing file managers)? If so, for which task types is it most effective, in terms of effectiveness, efficiency, or satisfaction?

RQ2: Is the information displayed by the experimental file manager useful, and is this file manager easy to learn to use?

RQ3: In what ways might the experimental file manager be improved?

RQ4: What is the likelihood that the experimental file manager would be adopted for use, should it become available?

8.2.2 Experimental Design

The type of evaluation chosen was a laboratory study, to allow for precise measurement of variables (including task efficiency and effectiveness). McGrath (1995) described the methodological trade offs between three factors: generalizability of evidence; precision of measurement; and realism of the setting. Measurement was a key component for demonstrating tangible benefits, and therefore a laboratory study was appropriate. (To make the testing environment more realistic, a workplace scenario was used, as detailed in Section 8.3 below; this approach was employed to mitigate the artificiality of the laboratory and allow assessment under more natural conditions.)

As described above, one component of the laboratory evaluation was the comparison of the new file manager (with sharing features) to existing file manager technology. This comparison required that measures be taken using two file managers: the experimental and a control. The control file manager used for testing was a near-replica of Windows Explorer for XP: it was a version of Jeff's File Manager (described in Chapter 7), configured to have the same features as, and similar appearance to, Windows Explorer.

The experiment was a within-subjects study, carried out over two one-hour sessions (spaced approximately one week apart). In each session, participants completed a series of tasks that required them to look up information about shared files. (These tasks are detailed in a later section.) A total of 17 tasks were completed: eight in Session One, and nine in Session Two. (These sessions are described below.) Note that each session was divided into two sub-sessions: one of the two file managers (control or experimental) was used in each sub-session; therefore, at the end of each session, participants had used both file managers. Participants were randomly assigned to these conditions (i.e., control or experimental first). The odd-numbered participants used the control first, and the numbering was determined simply by scheduling: participants were assigned a number when they arrived for Session One.

8.2.2.1 Use of Multiple Experiment Sessions

Although one session would have been sufficient for a simple comparison of the file managers, a second session was added for two reasons. First, multiple sessions allowed the participant to have more experience with the new (experimental) file manager, allowing them to gain more experience and, hopefully, provide a more accurate assessment of its features. Second, this arrangement simulated asynchronous use over time. The file sharing awareness framework was based on the asynchronous Workspace Awareness model (Tam & Greenberg, 2006), which was designed specifically for *change awareness*: people needing to keep up with any changes that occurred while they were away from the workspace. Bringing participants back to the scenario environment (after an interval) allowed for an assessment of how easily they were able to catch up with file changes in their absence. A description of both sessions is provided below.

Session One

During the first session, participants completed eight tasks: three in the control condition, and five in the experimental. As described below, six of these tasks were common to both file managers: three were completed in each of the control and experimental managers. (For example, participants looked up the "last modified" date of a file in both experimental conditions.) The remaining two tasks were used to assess the unique features of the experimental file manager only (e.g., finding the number of times a file was accessed); these tasks were conducted only in the experimental manager

condition. (Thus, there were a total of five tasks in the experimental condition: three common tasks, and two unique). The file manager condition was counterbalanced, so that half of the participants began with the control and finished with the experimental manager; the other half of the participants used the file managers in the reverse order.

Session Two

The first and second experimental sessions were similar, although not identical. The file system was changed between sessions, to simulate activity during the week. The file manager condition was counterbalanced again: those that started with the control in Session One started with the experimental file manager in Session Two.

In this session, nine tasks were completed. As in Session One, six of these tasks were completed in both file manager conditions (e.g., looking up file permissions). Also as before, there were tasks unique to the experimental file manager: for this session, there were *three* such tasks. Another task was added to permit more extensive evaluation of the new file manager features; these tasks were limited to two for the first session because at that stage, the file manager would be unfamiliar to participants, and thus they might be overwhelmed if asked to perform many tasks with it. (It was expected that by Session Two, participants would be comfortable enough with the experimental file manager to complete one additional task.) Thus, there were six tasks in the experimental condition, and three in the control.

8.2.2.2 Task Ordering

The complete task ordering, for both sessions, is shown in Table 44. (The tasks are identified by letter; these tasks are detailed in Section 8.5 below).

	Session One		Session Two	
1	ABC	DFEGH	PMNILQ	KOJ
2	HEFDG	CBA	OKJ	QLPIMN
3	GFH	BCADE	LOMJNK	IPQ
4	ECDAB	HFG	IQP	KOMNLJ
5	GAC	HDFEB	MLNJQI	OPK
6	FHBED	CGA	KPO	INMJQL
7	BFH	CEAGD	ONPMLK	IQJ
8	DEGCA	BHF	QIJ	LONKMP
9	FGC	AHBED	ILJNMK	POQ
10	HADEB	FCG	QOP	KNMLJI
11	BHA	FGDEC	MQLNPO	KIJ
12	GEDCF	HBA	JIK	OMPLNQ
13	FAC	BEDGH	JMLOQN	KIP
14	GDBEH	AFC	PKI	JNLQOM
15	HGB	ADCEF	KNPMIL	QOJ
16	EFCAD	HBG	QJO	PLIMNK
17	AHF	EGDCB	ONLMJK	QIP
18	CDEGB	FAH	PIQ	MJNLKO
19	GCB	AHDEF	LPNQIM	KJO
20	FEAHD	CGB	JOK	ILPNMQ
21	CBF	DHEAG	JNLIMK	OQP
22	GEAHD	BCF	OPQ	IKJMLN
23	GHA	CBDEF	QNPLOM	JKI
24	BEFDC	AGH	KJI	NOMQLP
25	HFA	GEBCD	NLIPKM	JOQ
26	DCGEB	AFH	OJQ	PLNKIM
27	BGC	FADHE	QLOJMN	IKP
28	EFHAD	CBG	PIK	MQJNOL

Table 44: Task ordering for Session One and Session Two. (The tasks are described in Section 8.5 below). Odd-numbered participants (grey rows) started with the control file manager, while even-numbered participants (white rows) started with the experimental manager.

Note that there are eight tasks in Session One, and nine tasks in Session Two. For Session One, tasks D and E are unique to the experimental file manager (and thus appear only in that condition); for Session Two, these unique tasks are identified as L, M, and N. The remainder of the tasks were common to both file managers, and each task was paired so that one of each was completed in each file manager. For Session One, the task pairs were: A and F (email search); B and G (check file permissions); and C and H (check file modified date). For Session Two, the task pairs were: I and O (email search); J and P (check file accesses); and K and Q (check file permissions).

Because there were unequal numbers of tasks in each condition, traditional counterbalancing methods (such as Latin Square) could not be used. Instead, tasks were randomly permuted, subject to the following conditions:

i. One of each of the paired tasks had to be performed in each condition; for example, a person had to look up a file permission in the experimental and control conditions, and never complete *both* file permission tasks in the same condition. (This means that, for instance, tasks A and F were never completed in the same sub-session.)
ii. Each paired task (i.e., not the unique tasks D-E or L-N) had to be completed in equal numbers, across the conditions. For instance, Task A needed to be completed equally often with the control and the experimental file managers, and equally in the first and second sub-sessions. This condition required that the number of participants be a multiple of four (e.g., Task A in both conditions and both sub-sessions, which is four combinations.)

These conditions led to the ordering in Table 44. The above conditions hold, and outside of those constraints, the tasks are randomly permuted (including the unique tasks in the experimental condition). The details of each task are provided in Section 8.5; note that each task involves answering a question about shared files. These question-answering tasks were designed to be realistic, to mitigate the unnaturalness of the laboratory environment. These tasks were thus completed in an office-like environment, using a realistic digital workspace; this simulated experimental environment is described in the following section.

8.3 *Simulating the Workplace: Scenario and Experimental Environment*

The experiment was designed to simulate an office environment (within the confines of a laboratory setting). An office workplace was a natural choice for this experiment: file sharing is a common activity in this environment, and confidentiality is a concern for corporations. This situation requires the careful management of shared files; therefore, finding accurate and complete information about file sharing is important.

Simulating the office workplace was done in two ways: by using an office-based scenario, and by testing in a physical environment that resembled an office. These two aspects of the experiment are described below.

8.3.1 Scenario

To add realism to the laboratory environment, a *scenario* was created for participants to follow during testing. Scenario-based design has emerged as a means for

balancing generalizability and accuracy in evaluation (Rosson & Carroll, 2002; Young & Barnard, 1987). One definition of a scenario is a "description that contains (1) actors, (2) background information on the actors and assumptions about their environment, (3) actors' goals or objectives, and (4) sequences of actions and events" (Kentaro & John, 2004, p. 46). Kentaro and John note that some elements from this list may be omitted or made implicit, depending on the testing requirements.

The specific approach that was adopted in this research was based on the theoretical methodology of Borlund (2003), who created a framework for evaluating information retrieval systems. Borlund's method involves the use of realistic work scenarios, with a "cover story" that explains why the user needs to interact with the system. For the scenario, participants gathered this information through a series of questions about files in the digital workspace.

The specifics of the office scenario were designed to coordinate with the email dataset used for the digital workspace; as described below, this dataset was the Enron corpus. Therefore, the scenario was based around an energy and telecommunications company; this is also a familiar industry within the local geographic area. According to the cover story, the participant was an administrator filling in for an absent employee. Again, substitution of employees is a common workplace occurrence, adding to the realism of the scenario.

The scenario, as presented to the participant, was as follows:

> Scenario: Session One
>
> Your name is Terry Martin. It is April 10, 2008. You are an administrator at Enercomp, a large Canadian energy and telecommunications company. You usually work in the Sackville office, but today you have been called in to Halifax to fill in for Chris Lee, who had to leave work suddenly due to illness.
>
> Chris can't get at email, so has faxed you a list of tasks to complete today. You have full access to Chris's desktop, including files and emails. Please complete each task, which involves finding the answers to some questions that Chris has.

This scenario contains the majority of the elements from Kentaro's and John's (2004) definition. The actors are the participant ("Terry Martin") and the (hypothetical) absent employee ("Chris Lee"). These names were chosen to make the scenario gender-neutral, so it could be effectively used with men and women; these forenames were

identified as being in the top four most-neutral names (Van Fleet & Atwater, 1997). Background information about the organization and the absent employee's work habits were provided on paper handouts.

The goals of the scenario were explicitly stated: to answer questions about files for the absent employee. The reason why an "absent employee" was used is because the file system and emails would be unfamiliar to the participant, and therefore it would be unrealistic to tell them that they were using, for example, "your own computer." This approach allowed the participant to use an unfamiliar file environment, while still having a plausible reason for interacting with it (i.e., to find information for another person). Finally, the actions and events were mainly implicit, except in two ways: first, an explicit order was placed on the questions, and therefore the actions. Second, participants were required to use the testing applications (Outlook and the file manager) to answer questions. However, the *exact* means of finding the information was left up to the participant.

The same basic scenario was used in the second session, with the exception that the passing of time (since the previous session) was explicitly noted:

> Scenario: Session Two
>
> Your name is Terry Martin. It is April 17, 2008. You are an administrator at Enercomp, a large Canadian energy and telecommunications company. You usually work in the Sackville office, but today you have been called back in to Halifax to fill in for Chris Lee, who had a medical emergency. Chris has been performing computer tasks since you left, such as sending and receiving email and working on documents, so the computer system is not identical to what it was when you left on April 10th.
>
> Chris can't get at email, so has faxed you a list of tasks to complete today. You have full access to Chris's desktop, including files and emails. Please complete each task, which involves finding the answers to some questions that Chris has.

8.3.2 Experimental Environment

In keeping with the computing environment and the corporate scenario, the physical space that the participants used was designed to look like an office. This environment was designed to make the setting less like a laboratory, to facilitate the role-playing aspects of the scenario. The space used for the experiment was the Usability Lab in the Centre for Management Informatics, which was designed to support ecological

validity. A photograph of the laboratory "office" is shown in Figure 25, which shows the desk that the participant used for the tasks.

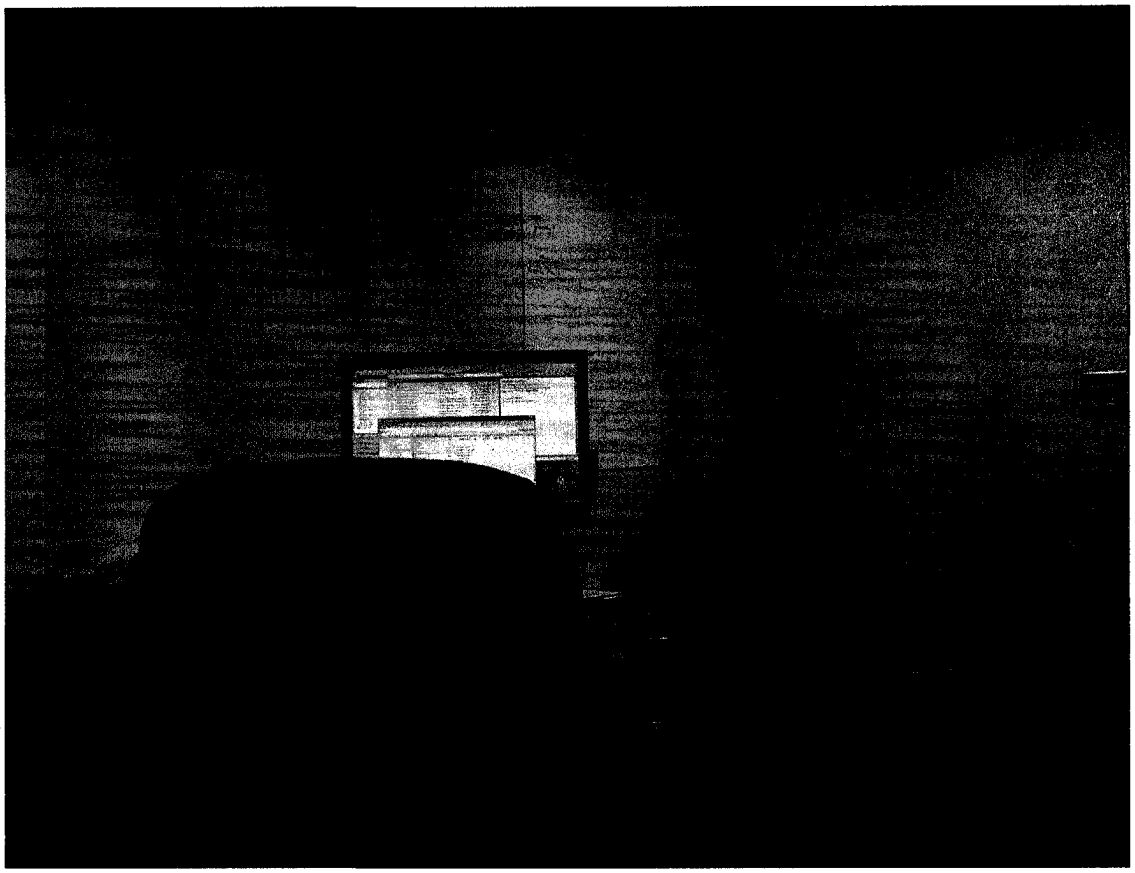


Figure 25: The experimental environment, designed to replicate an office setting

Props were used to make the environment more natural, including office supplies and books. A calendar was set to the future month of the scenario, April 2008. Also, a whiteboard listed fictional details about projects and colleagues, many of which were mentioned in the scenario tasks.

The "office" had a one-way mirror in it, through which activity could be viewed from an adjacent audio-visual control room. The participants completed the tasks while being observed from behind the glass. The facilitator (the author) was in the room with the participant for the start and end of each set of tasks (three times per session). At all other times, the facilitator was observing remotely. The experimental office and the control room were also connected with speakers and microphones, so that the participant could ask questions during the session.

The control room contained all audio-visual equipment, as well as the managers for the logging software and an eyetracker (described below). Logging was initiated and terminated remotely from this location, so that the data collection was unobtrusive to the participant. The control room equipment was managed by an assistant for all sessions, either Alexandra MacNutt or Emilie Dawe (research assistants at the Centre for Management Informatics). A photograph of the control room, showing the window overlooking the experimental environment, is shown in Figure 26 below.

While participants were completing tasks, their gaze patterns were recorded with a non-head-mounted eyetracker. This gaze data was recorded on a contingency basis, in the event that the other data sources were insufficient for explaining particular findings (e.g., anomalous user behaviours). This situation did not arise (hence, the eyetracking data were superfluous); the eyetracker is mentioned here simply for clarity and completeness.



Figure 26: Control room, used for observation and for data collection. The study participants sat at the desk that can be seen on the other side of the window.

8.4 *Digital Workspace: Email and File Collections*

Because the file sharing tools were meant to be tested in as realistic an environment as possible, a suitable computing environment needed to be created for the evaluation process. This entailed the creation of a set of files, including emails, which would realistically simulate an actual office workstation. There were two challenges in creating this environment: first, in obtaining a realistic set of emails, and second, in providing a collection of work documents that was realistic in terms of organization and size.

8.4.1 Real-World Email: The Enron Dataset

One difficulty in creating a simulated work environment was the lack of email datasets: privacy concerns typically preclude the use of real-world emails in research (Carvalho & Cohen, 2007). This problem was addressed when the Enron corporate email files were publicly released by the American Federal Energy Regulatory Commission, as a result of legal investigations in 2004; this is believed to be the first (if not the only) sizable real-world email dataset made available in the public domain (Shetty & Adibi, 2005).

Researchers cleaned the raw Enron corpus, in order to correct errors (such as integrity problems), to strip out file attachments, and to delete some emails by employee request. The resulting dataset (dated March 4, 2004), provided by William Cohen at Carnegie Mellon University, was the version used for this research (2004). This corpus consists of 517,431 emails from 151 users, in 4700 folders, collected from 1998 through 2002 (Klimt & Yang, 2004).

One user was selected from this set of 151 to represent "Chris Lee." Although any user would be a suitable choice, one was selected whose data showed a regular use of email attachments, and who had a representative number of folders for this dataset (between 3 and 20). This user's mailbox had eight folders, with a total of 23206 messages; it was converted to "Chris Lee's" mailbox (through a process described in Section 8.4.3.1 below).

8.4.2 Creating a Realistic File Environment

In addition to a realistic email dataset, testing also required a realistic workspace file environment: the study participants would need to interact with a file system, with

files and folders organized in a way that represented actual user practice. As with email, privacy concerns have precluded the collection of real-world data (Chernov, Demartini, Herder, Kopycki, & Nejdl, 2008), which required that a short study be carried out to glean how digital workspaces are organized in practice. This study was not intended to be comprehensive, but to indicate a reasonable set of parameters within which to operate. The study consisted of collecting details about workplace document collections; primarily, the numbers of files and folders, and the names of those folders.

8.4.2.1 Workspace File Study

Participants were recruited through email, and asked to share basic information about their workplace files; they could provide this information through email, or through an in-person interview. Those who were recruited were administrators and/or project managers, with responsibility for one or more projects. This group was selected because the participants closely resembled those in the proposed testing scenario (namely, an administrator at "Enercomp").

The questions they were asked are listed in Figure 27 below. These questions asked about the number of files and folders on the Desktop and in work folders: how many of each, how many file per folder, and what the folder names were (as a simple way to see how files were categorized and organized).

1. Go to your Desktop (on the computer used for most of your work). a) How many folders are on it? b) How many files?
2. For Desktop folders: What categories do your folders fall into? For example: "3 are work projects, 2 personal, 1 temporary storage..." (If you prefer, you can just list their names, with any confidential information omitted, e.g., "reports_companyX")
3. Go to the folder that has most of your work folders (or files) in it. (This could be "My Documents", or a separate drive or special folder.) a) How many folders are in it? b) How many files?
4. Continuing in the folder with your work folders and files in it: What categories do your folders fall into? For example: "12 are specific projects, 10 are for particular people, 4 are for departments", etc. (If you prefer, you can just list their names, with any confidential information omitted, e.g., "employee_x")
5. Continuing in the folder with your work folders and files in it: About how many files would you estimate, on average, are in each folder? You can check some representative folders (e.g., 5-10 folders) and calculate a "files per folder" average from these. You only need to go down one level. Suppose your work folders are in "My Documents/Work". If you have three folders here called "reports", "minutes", and "misc", then just count files in these three folders. Ignore files in folders down inside "reports", etc.

Figure 27: Questions about file organization

Ten participants provided information; four responded in email and six provided information through an interview. Those who were interviewed volunteered additional information about their organizational habits, such as the depth of their folders and their work practices, as well as the basic information requested in the set of questions. The numeric results are summarized in Table 45, grouped in ranges. A summary of the categories for folders is also presented, with the count in brackets after each category name; these category names were provided by the participants.

category	count (n=10)							
	0	1-5	6-10	11-19	20-30	31-49	50-74	75-100
desktop: folders	3	6	1	0	0	0	0	0
desktop: files	3	1	2	0	3	0	1	0
work folder: folders	0	1	1	3	2	0	2	1
work folder: files	1	2	1	1	1	2	1	0
work folder: files per folder	0	1	2	2	3	0	1	1
topics: desktop folders	project (5); temp (3); "My Documents" (2); programs (1); shortcuts (1); resumes (1); HR (1); personal (1); speakers (1); budget (1)							
topics: work folders	project (6); dates (3); organizations (2); people (2); sharing (2); default, e.g., "My Photos" (2); temporary (2); courses (1); travel (1); departments(1); suppliers(1); forms(1); personal(1); human resources (1); contracts (1); records (1); received documents (1)							

Table 45: Summary of workplace file organization. Counts are shown for the number of files and folders in the various categories, grouped into ranges.

The responses showed a wide variation in file organization habits: for example, for desktop files and folders, three participants had nothing on the desktop, while one person had sixty files. Similarly, some respondents had many folders with few files each, while others had fewer folders but many files per folder. Most commonly, people reporting having somewhere between 11 and 30 folders in their main workspace folder (e.g., "My Documents"), with six to 30 files per folder. The dominant types of folders tended to be for projects, along with departments and individuals (and date-based filing systems); there are also many unique folder types, such as "speakers" for a person whose job involved organizing invited talks.

Given the wide variations in workspace file organization, it is evident that many types of file arrangements would be suitable for a realistic laboratory scenario. For the Enercomp scenario, a simple file structure was created, using the Windows XP Desktop along with two folders—"chris's docs" and "Sharing Folder"—as the location of scenario files.

8.4.2.2 Windows Desktop files

The term "Desktop" is used here to refer to the area displayed behind the windows of the graphical user interface, such as that used in Windows XP; it commonly

contains folder and file icons and application shortcuts. This location was not a primary source of files for the scenario, but it was set up to be realistic, because it would be frequently viewed. It contained 10 items, mainly program shortcuts and folders; these included shortcuts to the file sharing tools (console and file managers), mainly for convenience of testing (i.e., in case the participants accidentally closed the application and needed to restart). There were also shortcuts to Outlook, to the "sharing folder", and to the "chris's docs" folder, in case these were helpful to the participant.

8.4.2.3 Workplace Document Folders

Two folders were created as the principal locations for storing the workplace documents used in the experiment. These folders were one level below the root of the C drive: C:\chris's docs and C:\Sharing Folder; the former primarily had the owner's personal items (some of which were sent as attachments), and the latter contained items whose permissions were set for sharing among colleagues.

The number of files and folders in these two locations was arranged in keeping with the findings of the workplace study. As shown in Table 46, "C:\chris's docs" had 19 subfolders, and "C:\Sharing Folder" had 20 subfolders. In keeping with the findings of the digital workspace study, the number of files per folders remained within the range of six to 30; the lowest number of files per folder was seven, and the highest was 28. (The details of the specific files used in this structure are described in Section 8.4.3.2. below.)

For the first session, the folder "chris's docs" contained a total of 325 files, and the "Sharing Folder" contained 340 files. An additional 20 files were added to the file collection for Session Two, to simulate activity during the week. Fifteen files were added to "chris's docs" (in three subfolders), and five files to "Sharing Folder" (in two subfolders). This information is summarized in Table 47. The vast majority of the files and folders in the system were not explicitly used in any tasks, but were included for realism.

C:\chris's docs		C:\Sharing Folder	
folder name	**#files**	**folder name**	**#files**
asia_pacific	8	Aliant	10
brokerage	22	annual_reports	12
confirmations	20	archived_projects	28
customers	11	AT&T	9
drafts	14	BP	15
eol documents	20	community	13
europe	16	finance	25
financial	17	forms	23
forms	22	government regulations	15
general_motors	21	handbook	22
guidelines	7	Human Resources	27
material for review	27	investments	24
north_america	15	legal	20
partners	10	memos	18
personal	22	PG&E	9
planning_docs	19	PR	16
presentations	8	presentations	11
proposals	24	public	24
spreadsheets	22	travel docs	12
		WHMIS	7
total	325	**total**	340

Table 46: Structure and number of workplace documents. Each subfolder (e.g., "drafts") contained only files (no subfolders). This structure was used in Session One; for Session Two, another 20 files were added (15 to "chris's docs" and five to "Sharing Folder").

	folder statistics	
	C:\chris's docs	C:\Sharing Folder
Session One		
total files	325	340
total subfolders	19	20
average files per subfolder	17	17
# of files with sharing information	45	340
Session Two		
total files	340	345
total subfolders	19	20
average files per subfolder	18	18
# of files with sharing information	50	345

Table 47: Information about files and folders used in the workspace environment for testing; numbers of files varied for each of the two sessions.

As shown in Table 47, only a subset of the files had sharing information associated with them, because not all files in the system were shared. An exception is the files in the "Sharing Folder": these all had permission information associated with them. However, in the "chris's docs" folder, only a subset of the files had sharing information:

45 files (in six subfolders) for Session One, and another five files (in two subfolders) were added for Session Two. (File sharing information includes event history, such as email sends and file accesses, as well as permissions.)

8.4.3 Implementation of the File Collection: Email and Documents

Both the Enron email data and the file structure required additional preparation before they could be used in the experiment: the mailbox required conversion, and the file structure needed to be populated with files for the tasks. These implementation steps are described below.

8.4.3.1 Creating the Mailbox

The email data that originated in the Enron dataset had to be further manipulated to be used with an email client; for example, every email message was stored in a separate file, and thus the files needed to be recombined. With the assistance of Dana Echtner (Dalhousie Faculty of Computer Science Technical Support), the dataset was converted into a format that would work with an email client (i.e., returned the data to a standard mailbox format).

To blind this data, the username of the mailbox owner was changed, to match "chris lee", the fictional user in this scenario. In addition, all references to Enron were replaced with "Enercomp", the fictional corporation. Locations were also changed from Houston to Halifax, and dates were made more recent (e.g., 2002 became 2008). All these changes were made through `perl` scripts, with some manual checking and error correction.

Email attachments also had to be re-attached to messages, because these were purged from the Enron dataset. The original files could not be used (because they were not retained), but the dataset retained the names of the files that had been stripped from each message. Files on similar topics (e.g., energy and communications), in the same file format, were downloaded from the web to replace these files; see the following section for more details. These files were renamed to match the names of each missing attachment, and manually re-attached to each message by re-drafting email messages and editing the email headers.

The resulting mailbox was imported into the Microsoft Outlook email application for the experiment. Outlook was chosen because of its widespread usage, particularly

within corporate settings (Fontana, 2007). Potential participants would likely be familiar with Outlook, and it would be natural to find this program in the simulated office setting of the experiment. (Participants' proficiency with Outlook was checked in a background questionnaire at the start of the laboratory study.)

The experiment involved two sessions, and therefore there were two different versions of the mailbox. The mailbox included messages dating from August 2005 up to the date used in the test scenario (i.e., either April 10 or April 17, 2008). The mailbox folders were reset between sessions, so that a person completing Session One would only see email dating up to April 10, and a person completing Session Two would see email dating to April 17.

The structures of these mailboxes—the Outlook folder names and sizes—are shown in Table 48, for both sessions. The number of messages increased between sessions, to simulate additional email traffic that had arrived and was created during the week; three folders were modified: "Inbox," "all_documents," and "Sent Items." In addition to the original Enron email messages, three new messages with attachments were created specifically for the testing scenario; these were placed in "Sent Items" for Session Two.

	Session One ("April 10")		**Session Two ("April 17")**	
folder	size (number of messages)	number of emails with attachments	size (number of messages)	number of emails with attachments
all_documents	7740	118	7905	123
archived_inbox	11566	115	unchanged	
Inbox	111	9	474	19
master_notice	2	0	unchanged	
misc	18	0	unchanged	
online_trading	360	0	unchanged	
Sent Items	3395	52	3431	60
trading issues	14	0	unchanged	

Table 48: Mailbox folder sizes, including numbers of email attachments, for both testing sessions.

8.4.3.2 Workplace Documents

The workplace study identified appropriate parameters for a workplace file structure; once this was determined, the structure had to be populated with actual files. To do so, files were downloaded from the web, primarily on subjects relating to the energy sector and telecommunications. Because the Enron employee's mailbox (adapted for the experiment) included a number of references to financial and legal documents, files on

these subjects were added to the collection. (Many of these were reattached to the email messages, as described above.)

These files were a mix of formats, primarily Microsoft Word (.doc), Portable Document Format (.pdf), Excel spreadsheets (.xls), Powerpoint (.ppt) and text (.txt). To find these files, Google searches were performed with keywords such as "energy", or "non-disclosure agreement" filtered to return only files matching certain file formats (such as ".xls"). As described above, 395 files were placed into the file collection, divided between two folders: "chris's docs" and "Shared Folders."

These files were used in carrying out the experimental tasks: participants answered questions about these shared files, using this realistic digital workspace. These experimental tasks are discussed in the following section.

8.5 *Tasks*

8.5.1 Mapping of Tasks to Awareness Framework

The tasks used for evaluation were designed to be realistic, as well as appropriate for formal assessment. Because the experimental file manager designed was based on a theoretical framework, that framework was consulted when developing the tasks.

The tasks in the experiment were designed primarily to assess how well people could find and interpret information about shared files, which incorporates the core aspect of awareness: "knowing what is going on" (Endsley, 1995), as discussed in Chapter 5. Therefore, the tasks in the scenario involved participants answering questions about shared files: this required them to locate information about one or more shared files, and understand this information enough to use it for answering questions.

Each of the questions relates to a real-world situation, as outlined in the testing scenario; each question was accompanied by a specific issue or need that the "absent worker" had. There were 37 questions in total: 20 in Session One and 17 in Session Two. Questions varied in complexity, from simply confirmation of a file's existence, to finding all the recipients of a certain file.

These questions can be mapped onto the questions that emerge from the awareness framework, as outlined in Chapter 6. The framework identified a number of questions that a user might need to answer about a shared file, from the perspective of both files (e.g., "Who can access this file?") and people (e.g., "What files can this person

access?). Each of the questions in the scenario comprises one or more of these framework questions; this mapping is shown in Table 49 and Table 50. The questions have been abbreviated slightly, with exact filenames and folder names omitted, to fit into the tables the full text of all questions is presented in Appendix F.

File Sharing Questions Mapped to Awareness Framework Questions (Session One)		
ID#	Framework Question	Task Question
A1	How has this file been distributed?	Was [file A] emailed to anyone?
A2	Who has a copy of this file?	If yes: to whom? (username(s) is fine; don't need to include email address)
B1	What actions can this person [or group] take on this file?	Does the "legal" group have *modify* permissions for [file B] {yes/no}
B2	What actions can this person [or group] take on this file?	If not, what permissions does legal have?
C1	Where is this file?	Is [file C] in the folder [folder A]? {yes/no}
C2	When have this file's properties been changed?	If yes: what is the "last modified" date?
D1a D1b D1c	What actions have taken place on this file? Who has accessed this file? When has this file been accessed?	For [file D] a. number of times accessed: b. person who accessed most recently: c. date of most recent access:
D2a D2b D2c	What actions have taken place on this file? Who has accessed this file? When has this file been accessed?	For [file E] a. number of times accessed: b. person who accessed most recently: c. date of most recent access:
E1	What files have been distributed to this person?	List three files that Sara Shackleton has received from me over email.
E2	What files can this person access?	List three files that Sara Shackleton currently has access to (can read or write)
F1	How has this file been distributed?	Was [file F] emailed to anyone? {yes/no}
F2	Who has a copy of this file?	If yes: to whom?
G1	What actions can this person [or group] take on this file?	Can Dave Samuels modify [file G] {yes/no}
G2	What actions can this person [or group] take on this file?	If no: what permissions does Dave have?
H1	Where is this file?	Is [file H] in [folder B]? {yes/no}
H2	When have this file's properties been changed?	If yes: was the file modified in the last 24 hours (April 9 or later)? {yes/no}

Table 49: The questions from the Session One tasks, mapped to questions from the awareness framework

File Sharing Questions Mapped to Awareness Framework Questions (Session Two)		
ID#	Framework Question	Task Question
I1	Who has a copy of this file? When was this file distributed?	List all people who received [file I] over email between April 13 and April 16
J1	When has this file been accessed?	Did anyone access [file J] between April 13th and April 17th (today)? {yes/no}
J2	What actions have taken place on this file? When was a particular change to this file made? Who has accessed this file? What actions has this person performed on this file?	If so, provide all the details you can find about any accesses (such as time), between April 13th and today:
K1	Where is this file?	Is [file K] in [folder C]? {yes/no}
K2a	What actions can be taken on this file? What actions can this person [or group] take on this file?	If yes, Do I [Chris Lee] have read and write permissions on it? {yes/no}
K2b	What actions can be taken on this file? What actions can this person [or group] take on this file?	Does Alan Aronowitz have read and write permissions on it? {yes/no}
L1	When did this person access a file?	Has Dave Samuels accessed [file L] since April 12th?
L2	What actions has this person performed on this file?	If so, what action did he perform (e.g., read, write...)?
M1	What files have been distributed to this person?	List all files emailed to Tom Moran
M2	What files has a person accessed? When did this person access a file?	List all files accessed by Tom Moran in [folder D] after April 13
N1	What actions have taken place on this file?	List the three most-accessed files in [folder E]
O1	Who has a copy of this file? When was this file distributed?	Who was emailed a copy of [file M] (April 12th or later)?
P1	When has this file been accessed?	Did anyone access [file N] on April 12th or later? {yes/no}
P2	What actions have taken place on this file? When was a particular change to this file made? Who has accessed this file? What actions has this person performed on this file?	If so, provide all the details you can find about any accesses:
Q1	Where is this file?	Is [file O] in [folder F]{yes/no}
Q2a	What actions can be taken on this file? What actions can this person [or group] take on this file?	If so, Do only Legal and Finance have access (and no one else)? {yes/no}
Q2b	What actions can be taken on this file? What actions can this person [or group] take on this file?	Do Legal and Finance have read access? {yes/no}
Q2a	What actions can be taken on this file? What actions can this person [or group] take on this file?	If so, Do only Legal and Finance have access (and no one else)? {yes/no}
Q2b	What actions can be taken on this file? What actions can this person [or group] take on this file?	Do Legal and Finance have read access? {yes/no}

Table 50: The questions from the Session Two tasks, mapped to questions from the awareness framework

8.5.2 Task Types

There were two basic categories of tasks in this study: those that could be performed using *both* file managers, and those that were feasible only with the experimental file manager. The tasks that could be performed with both file managers could be compared with each other, using such factors as task completion time. However, these tasks are restrictive: under these conditions, the additional functionality of the experimental file manager could not be explored. Therefore, additional tasks were included that were *only* carried out in the experimental file manager condition. These were tasks that, using the control file manager, were either impossible to complete, or prohibitively difficult to complete within the one-hour timeframe of each testing session.

8.5.2.1 Tasks Completed with Both File Managers

Three tasks were completed using both the control file manager and the new file manager, in both sessions. As described in the Design section of this chapter, each of these tasks was in a pair, with one version of the equivalent task being performed with each file manager. For example, if the user looked up a file permission using the control file manager, then he or she would also look up a file permission using the experimental new file manager. The tasks are identified by letter: tasks A-H were performed in Session One, and tasks I-O in Session Two. (These are the same letters used in the question identification numbers, in the previous section.)

The paired tasks are those identified by the letters [A,B,C] and [F,G,H] in Session One, and by [I, J, K] and [O,P,Q] in Session Two. (Tasks [D,E] and [L,M.N] were performed only with the experimental file manager.) The pairs are listed in order (e.g., A is mapped to F, B is mapped to G). This mapping is shown in Table 51: each row contains the pair of equivalent tasks, including a short description for clarity.

Tasks for Both File Managers: Session One	
A file: clickpaper-mar-05.xls 1. Was it emailed to anyone? 2. If so, to whom?	**F** file: 2009_plans.doc 1. Was it emailed to anyone? 2. If so, to whom?
B 1.Does the "legal" group have modify permissions for the file acquisitions2008.doc? 2. If not, what permissions does legal have?	**G** 1. Can Dave Samuels modify template.doc? 2. If no: what permissions does Dave have?
C 1. Is the file energy_2008_q3.doc in the folder C:\chris's docs\drafts? 2. If yes: what is the "last modified" date?	**H** 1. Is the file contracts.doc in the folder C:\Sharing-Folder\legal? 2. If yes: was the file modified in the last 24 hours (April 9 or later)?
Tasks for Both File Managers: Session Two	
I 1. List all people who received the file PR_apr13_08.pdf over email between April 13 and April 16 (provide email addresses)	**O** 1. Who was emailed a copy of the file clickpaper-mar-05.xls (April 12th or later)?
J 1. Did anyone access the file nda.pdf between April 13th and April 17th (today)? 2. If so, provide all the details you can find about any accesses (such as time), between April 13th and today:	**P** 1. Did anyone access the file merger.doc on April 12th or later? 2. If so, provide all the details you can find about any accesses (such as time):
K 1. Is GE_prospect.doc in the folder C:\Sharing-Folder\legal? 2. If yes, a. Do I [Chris Lee] have read and write permissions on it? b.Does Alan Aronowitz have read and write permissions on it?	**Q** 1. Is options.doc in the folder C:\Sharing-Folder\BP? 2. If so, a. Do only Legal and Finance have access (and no one else)? b. Do Legal and Finance have read access?

Table 51: Paired tasks performed in laboratory study. Each row contains a pair of equivalent tasks (such as finding a list of email recipients), such as A and F, or J and P.

These tasks were objectively compared, in terms of how quickly they were completed and how accurately the questions were answered. (The analysis is in the following chapter).

Note that both sessions included one task pair for which the experimental file manager was not expected to demonstrate any improvements in efficiency: instead, one simple task was included to demonstrate whether or not the experimental file manager caused any *problems*, as compared to the control. (That is, for certain simple file tasks, would the experimental file manager be detrimental for users?) In Session One, it was task pair C and H (checking the "file modified" date); for Session Two, it was task pair J

and P (checking details of recent file accesses). For these tasks, the experimental file manager was unlikely to provide a significant efficiency improvement, but it was hoped that it would not make the tasks *more* difficult; the two file managers should be equivalent. (This information would be used to prove hypothesis H1: that the experimental file manager would *not* lead to any *slower* completion times.) The remainder of the tasks were to be used to determine whether improvements in effectiveness and efficiency were provided by the experimental file manager.

8.5.2.2 Tasks Completed Only with the Experimental File Manager

Some tasks could only be carried out using the functionality of the experimental file manager. Two of these tasks were included in Session One, and three in Session Two. (As described in the Design action, one more task was included in Session Two because the participant would be more familiar with the features; having three such tasks in Session One was likely to be too overwhelming for the initial session.) The sets of tasks are listed in Table 52.

The two experimental manager-only tasks in Session One involved looking up file access history (which is not available in the control file manager). The participants also had to find out which files had been emailed to, and were accessible to, a specific person; the latter subtask was intractable with the control file manager, as it would require checking every folder in the entire file system. Session Two included three tasks; all three tasks involved checking the file access history, either for specific actions or to find the most-accessed files in a folder. Because the file access history is not available in the control file manager, this task requires the use of the new functionality. Task "M" (in Session Two) could be partly completed using Outlook, but the second part required the experimental file manager features.

Experimental Manager-Only Tasks: Session One	
D 1. file merger_guide.doc a. number of times accessed: b. person who accessed most recently: c. date of most recent access: 2. file Leave Guidelines.doc a. number of times accessed: b. person who accessed most recently: c. date of most recent access:	**E** 1. List three files that Sara Shackleton has received from me over email. 2. List three files that Sara Shackleton currently has access to (can read or write)
Experimental Manager-Only Tasks: Session Two	
L 1. Has Dave Samuels accessed the file template.doc since April 12th? 2. If so, what action did he perform (e.g., read, write...)?	**M** 1. List all files emailed to Tom Moran 2. List all files accessed by Tom Moran in folder C:\Sharing-Folder\legal after April 13
N 1. List the three most-accessed files in Sharing Folder\public\:	

Table 52: Tasks performed only in the Helium file manager condition for both sessions

In addition to the explicit tasks, there was an "alert" event added to Session Two, so that the notification feature could be evaluated. During Session Two, a pop-up alert appeared when the participant was using the experimental file manager, three minutes after the tasks began. The alert message informed the participant that "Sheri Stack has read the file NewFund.doc." The participant did not need to do anything with this alert, or with the information displayed; instead, this was a simple test of the alert feature, to determine whether the alert were noticeable, understandable, or intrusive. The participant was asked about the alert during the talk-after session at the end of the study, to find out if they noticed it and how they interpreted it.

8.5.2.3 Addendum: Tasks Completed with Microsoft Outlook

Note that the tasks that involved email attachments could often be completed using Microsoft Outlook; indeed, for the control condition, this was the only tool available for finding file attachment information. Participants were free to use Outlook whenever they wanted, even when the experimental file manager would allow them to find the information more quickly; they were not directed to choose a specific method of answering the questions, but could choose their own approach.

8.6 *Variables*

8.6.1 Independent Variables

The independent variables were:

i) file manager: control ("Carbon") or experimental ("Helium")

ii) order of file manager use: whether the control or the experimental file manager was used first or second in a session

8.6.2 Dependent Variables (Measures)

The control and the experimental file managers were compared on four variables:

- task completion time
- correctness of question responses (accuracy and completeness)
- usability scores
- user preference

8.6.2.1 Task Completion Time

Completion times were recorded for all tasks; comparisons were made for the 12 tasks completed using *both* environments (i.e., control and experimental), as described above. Times were compared to determine whether the experimental file manager led to faster task completion times than the control, for any tasks in either session.

The completion times were determined by analysing screen captures: participants' interaction with the computer was logged using Morae Recorder (Techsmith Corporation, 2007). This software logging tool captured all activity on the screen, including mouse clicks; all events included timestamps. This recording was used to determine how much time was taken to complete tasks. The start time was marked when the participant first moved the mouse to begin work on the task. The end time was marked when the participants stopped moving the mouse after completing the task. The total time includes periods when the participant may not have been working directly with the computer: this includes thinking, or reading the tasks and instructional handouts, as well as writing down responses on the paper handouts. The start and end times were marked by Alexandra MacNutt, a research assistant at the Centre for Management Informatics, who manually reviewed all the Morae files. I checked a random sample of 42 tasks (out of 336) to validate accuracy of the times.

8.6.2.2 Correctness of Question Responses

The two file managers were also compared in terms of how well they assisted participants in answering the questions that formed the basis of the tasks. Each question was given a maximum score, and the response would be scored on how accurate and complete it was. A complete scoring guide was developed for all questions. Some short questions, such as yes/no questions, had simple scores (usually "1": correct or incorrect). Those yes/no questions for which a correct response was critical for confidentiality were given a higher maximum score—of four, rather than one—to reflect the relative importance of this item. (These questions are in Tasks A and F, and I and O, as listed in Table 50 above.). Other questions were more complex, with lengthier answers, and could receive partial scores depending on completeness of response (e.g., if respondent omitted one item from a list of three files). These questions had higher maximum scores (e.g., between five and nine), in order to distinguish between the various types of incomplete or incorrect answers.

Most tasks had multiple questions; there were twelve questions in each of the two sessions (24 in total). Each question was individually scored for comparison, rather than being totalled on a per-task basis. Scores were compared for those questions that participants answered using both file managers. (As described below, some additional responses were recorded for tasks performed using the experimental file manager only.)

8.6.2.3 Usability Scores

After each set of tasks, participants completed a usability questionnaire, which consisted of 19 questions: 17 of these were from the Computer System Usability Questionnaire (CSUQ)(Lewis, 1995), with two more questions from the After-Scenario Questionnaire (Lewis, 1991). The CSUQ is designed to gather subjective measures of efficiency, effectiveness, and satisfaction. The ASQ is used for scenario-based testing, to subjectively assess satisfaction with the scenario tasks. (Four usability questionnaires were completed in total.)

Each question score consisted of a number, representing the degree to which the respondent agreed with a statement about the usability of the file manager. The score came from a seven-point Likert scale, ranging from "strongly disagree" (1) to "strongly agree" (7). The full questionnaire is presented in Table 53. Note that the items in the

	strongly disagree 1	2	3	4	5	6	strongly agree 7
CSUQ							
SYSUSE (questions 1-8)							
1. Overall, I am satisfied with how easy it is to use this file manager	o	o	o	o	o	o	o
2. It was simple to use this file manager	o	o	o	o	o	o	o
3. I could effectively complete my work using this file manager	o	o	o	o	o	o	o
4. I was able to complete my work quickly using this file manager	o	o	o	o	o	o	o
5. I was able to efficiently complete my work using this file manager	o	o	o	o	o	o	o
6. I felt comfortable using this file manager	o	o	o	o	o	o	o
7. It was easy to learn to use this file manager	o	o	o	o	o	o	o
8. I believe I became productive quickly using this file manager	o	o	o	o	o	o	o
INFOQUAL (questions 9-13)							
9. Whenever I made a mistake using the file manager, I could recover easily and quickly	o	o	o	o	o	o	o
10. It was easy to find the information I needed	o	o	o	o	o	o	o
11. The information provided by the file manager was easy to understand	o	o	o	o	o	o	o
12. The information was effective in helping me complete the tasks and scenarios	o	o	o	o	o	o	o
13. The organization of information in the file manager was clear	o	o	o	o	o	o	o
INTERQUAL (questions 14-16)							
14. The interface of this file manager is pleasant	o	o	o	o	o	o	o
15. I liked using the interface of this file manager	o	o	o	o	o	o	o
16. This file manager has all the functions and capabilities I expect it to have	o	o	o	o	o	o	o
OVERALL (questions 1-17)							
17. Overall, I am satisfied with this file manager	o	o	o	o	o	o	o
ASQ (questions 18-19)							
18. Overall, I am satisfied with the ease of completing the tasks in this scenario	o	o	o	o	o	o	o
19. Overall, I am satisfied with the amount of time it took to complete the tasks in this scenario.	o	o	o	o	o	o	o

Table 53:Usability questionnaire used to generate usability scores: CSUQ and ASQ

CSUQ are combined into four factors: system usefulness (SYSUSE), information quality (INFOQUAL), interface quality (INTERQUAL), and an overall score (OVERALL). The mapping between questionnaire items and these factors is also shown in Table 53; ASQ is added to these four factors to make a total of five. These scores were used to assess the degree of user satisfaction with the two file managers (control and experimental). The comparison between the usability of the two file managers was based on the mean scores for these five factors.

8.6.2.4 User Preference

The final variable used to compare the two file managers was subjective preference. This was assessed by questionnaires: participants indicated which file manager they preferred, on a number of factors. For Session One, the questionnaire had 11 questions in all. The first three questions centered on the three *tasks* performed using both file managers. Respondents indicated which of the two managers most helped them complete each task; they could also choose "both the same." The remaining eight questions were forced-choice, and included questions about awareness and perceived usefulness. The full text of all questions for Session One is listed in Table 54.

The questionnaire for Session Two was similar, but with 14 questions: three task questions, and additional questions about catching up with changes between sessions. (In the table, the control is referred to as "Carbon" and the experimental as "Helium"; recall that this is how the file managers were named for the participants.) The full text of all questions for Session Two is listed in Table 55.

Note that the order of the choices was based on the order in which the participant had used the two tools. For instance, for Session One, Part One, odd-numbered participants chose between control and experimental (not experimental and control). Therefore there was no bias introduced in the questions, in terms of which option was presented first.

Session One: Task Questions	
Which file manager best helped you complete the task:	
Q1: Checking file permissions	_Carbon _Helium _both the same
Q2: Determining who a file was sent to, if anyone	_Carbon _Helium _both the same
Q3: Checking the "last modified" date of a file	_Carbon _Helium _both the same
Session One: Preference Questions	
Q4: Overall, it was simpler to use this file manager	_Carbon _Helium
Q5: I could more effectively complete the tasks using this file manager	_Carbon _Helium
Q6: I was able to complete the tasks quicker using this file manager	_Carbon _Helium
Q7: It was easier to find the information I needed using this file manager	_Carbon _Helium
Q8: I had greater awareness of which files were shared with whom using this file manager	_Carbon _Helium
Q9: I had more complete information about shared files using this file manager	_Carbon _Helium
Q10: I believe I could manage my shared files better using this file manager	_Carbon _Helium
Q11: Overall, I prefer:	_Carbon _Helium

Table 54: User preference questions, between control "Carbon" and experimental "Helium" file managers, for Session One.

8.6.3 Additional Measures: Experimental File Manager Only

The four variables above were used for comparison between the control and experimental file managers. In addition to these, the experimental file manager had additional assessment: first, of its unique features and functionality (e.g., based on tasks that could not be completed using the control file manager), and second, for participant intention-to-use.

8.6.3.1 Evaluation of New Features and Functionality

The unique features of the experimental file manager were evaluated in five ways. The first evaluation was using task completion times, for those tasks performed *only* in the experimental file manager condition; this analysis would show which tasks were performed more quickly than others. (Task completion time was computed in the same way as described above, using the log files.)

The second evaluation method was through question response scores; again, these were computed in the same way as described above, for those tasks performed only in the experimental file manager. This analysis would show which questions were effectively answered using the experimental manager.

Session Two: Task Questions	
Which file manager best helped you complete the task:	
Q1: Checking file permissions	_Carbon _Helium _both the same
Q2: Determining who a file was sent to, if anyone	_Carbon _Helium _both the same
Q3: Finding information about past accesses on a file	_Carbon _Helium _both the same
Session Two: Preference Questions	
Q4: Overall, it was simpler to use this file manager	_Carbon _Helium
Q5: I could more effectively complete the tasks using this file manager	_Carbon _Helium
Q6: I was able to complete the tasks quicker using this file manager	_Carbon _Helium
Q7: It was easier to find the information I needed using this file manager	_Carbon _Helium
Q8: I was better able to find out how the workspace had changed during my absence (e.g., the "week" since the last session) using this file manager	_Carbon _Helium
Q9: I was better able to keep track of events on shared files using this file manager	_Carbon _Helium
Q10:I was better able to find out who had made changes to shared files using this file manager	_Carbon _Helium
Q11: I had greater awareness of which files were shared with whom using this file manager	_Carbon _Helium
Q12: I had more complete information about shared files using this file manager	_Carbon _Helium
Q13: I believe I could manage my shared files better using this file manager	_Carbon _Helium
Q14: Overall, I prefer:	_Carbon _Helium

Table 55: User preference questions, between control "Carbon" and experimental "Helium" file managers, for Session Two.

The third evaluation focused specifically on the notification feature, which is unique to the experimental file manager. As described above, this feature displayed an alert during Session Two, in the sub-session that used the experimental file manager. The evaluation of this feature was based on a talk-after session, conducted at the end of Session Two. In the talk-after session, the screen captures were replayed to participants, who described what they did when completing tasks (e.g., how they approached the task, what they were thinking). When the notification window appeared during the talk-after replay, participants were asked whether it was intrusive, and whether it was meaningful to them (e.g., did they know what the alert information was referring to?).

The fourth evaluation was based on subjective ratings, focusing on the quality of the information displayed by the experimental file manager, and whether or not the

manager was easy to learn to use. This evaluation came from a questionnaire with seven questions, which was completed at the end of Session Two. Each question consisted of a seven-point Likert scale, ranging from "strongly agree" (1) to "strongly disagree" (7). The questions were reverse-worded, so that every odd-numbered question described a negative characteristic, and every even-numbered question described a positive characteristic. The reverse-wording approach was taken to prevent participants from automatically checking the response boxes without reading the questions. The full text of the questions is provided below, in Table 56.

		1	2	3	4	5	6	7	
1. I found it difficult to understand the information about shared files provided by the file manager.	strongly agree	O	O	O	O	O	O	O	strongly disagree
2. I was able to easily find the information I needed, about shared files, using this file manager.	strongly agree	O	O	O	O	O	O	O	strongly disagree
3. I found it difficult to learn to use this file manager.	strongly agree	O	O	O	O	O	O	O	strongly disagree
4. I was able to perceive information about shared files "at a glance" using this file manager	strongly agree	O	O	O	O	O	O	O	strongly disagree
5. The information provided by the file manager was ***not*** useful for completing the tasks.	strongly agree	O	O	O	O	O	O	O	strongly disagree
6. The information displayed by this file manager is presented in a non-intrusive manner.	strongly agree	O	O	O	O	O	O	O	strongly disagree
7. There is too much information displayed in this file manager.	strongly agree	O	O	O	O	O	O	O	strongly disagree
Comments:									

Table 56: Evaluation of information display and ease of learning, for experimental file manager. Note that the odd-numbered questions are reverse-worded.

The fifth and final way that the experimental file manager's features were assessed was by gathering suggestions for improvement, using a questionnaire that was completed in Session Two. Participants were asked to provide suggestions for improvement for the file manager (including the sharing console). They were presented with a list of elements that they could comment on; this list included the new file manager columns and the three components of the sharing console (search, file details,

and alert). Participants could also list any features they felt should be added to the manager or console. These questions were presented on a paper questionnaire, giving participants the opportunity to draw their ideas. This questionnaire is presented in full in Appendix F.

8.6.3.2 Intention to Use

Because the experimental file manager was a new technology, it was beneficial to assess "intention to use": that is, would participants adopt this file manager, if it were made available? Using a new technology requires additional effort (e.g., to learn to use it), and therefore there must be sufficient perceived benefit to make it worth this effort.

The intention-to-use factor for the experimental file manager was assessed in two ways: through a Technology Acceptance Model questionnaire, and through a follow-up email, sent two weeks after the laboratory study.

8.6.3.2.1 Technology Acceptance Model Questionnaire

Near the end of the second session, participants completed a Technology Acceptance Model (TAM) questionnaire (Davis, 1989). TAM theory posits that two factors influence the adoption of new technology: perceived ease-of-use and perceived effectiveness. The Technology Acceptance Model has been repeatedly validated in studies; see, for example, the overview of 88 TAM studies provided by King and He (2006).

The associated questionnaire (developed by Davis) is designed to measure the two driving factors for adoption, and thus indicate the likelihood that respondents would employ a new technology (1989). For this study, the questions asked about the new file manager, in terms of whether it would help participants in their job. (For participants who were not currently working, they were asked to consider how well the technology would have worked in their former job.)

The specific TAM questionnaire used in this study included the twelve items from Davis's scale (1989), with one additional item from a similar scale by Laitenberger and Dreyer (1998), making 13 questions in all. These items consist of seven-point Likert scales, ranging from "extremely likely" to "extremely unlikely", for statements such as "Using [the experimental] file manager would improve my job performance." (The full

questionnaire is provided in Appendix F.) The new question from Laitenberger and Dreyer is "Assuming that [the experimental] file manager would be available on my job, I predict that I would use it on a regular basis" (1998). The items in Davis's questionnaire have been validated by several studies (e.g., Lin, Shih, & Sher, 2007; Roca, Chiu, & Martínez, 2006; Sivo & Pan, 2005)). These studies demonstrate that the questionnaire is a reliable testing instrument. The additional question from Laitenberger and Dreyer is not part of the original TAM scale, but is supplementary; it provides a self-rating of expected usage (1998).

The results of the TAM questionnaire can be used to indicate the likelihood that a certain technology will be adopted: positive ratings for perceived usefulness and perceived ease of use are correlated with higher intention to use (i.e., more likely to be adopted).

8.6.3.2.2 Follow-Up Email

Participants were given the choice to opt-in to a follow-up email communication, to take place two weeks after the second laboratory session. Those who opted in were emailed screenshots to remind them what the file manager and console looked like, along with the following question:

> Since the research study, have you encountered any situations in which you thought these sharing tools might be helpful, or recalled a time in the past when you could have used these tools to help you? If so, please provide me with a few details, so I can understand how the tool might be used in actual practice.

The goal of this communication was to get feedback after the participants had been exposed to the new features and after they had time to think about the potential uses of the technology. The results would be anecdotal, but nevertheless valuable indicators of the real-world usefulness of the new file manager functionality.

8.7 *Participants*

8.7.1 Eligibility

To be eligible for the experiment, participants had to be at least 18 years of age. Because the testing scenario involved an office worker, participants needed at least one year's experience in an office or similar environment; this background would permit the participant to more easily adopt the role during testing. In addition to office experience,

participants also needed to be experienced with Windows XP; a minimum of one year's experience was required. Participants also needed to be comfortable using the Microsoft Outlook email program. (Participants self-rated their comfort level with XP and Outlook at the start of the experiment, and indicated how long they had been using XP.) These requirements were included in the recruitment notices.

8.7.2 Recruitment

Recruitment was carried out through campus mailing lists and posters. The recruitment notice read as follows:

> As part of my doctoral research, I am conducting a study on the management of shared files. You are eligible to participate if you are at least 18 years of age, have used Windows XP for at least one year, are comfortable using XP and Microsoft Outlook, and have at least one year's experience working in any type of office environment (including work terms and summer jobs). Experienced office workers are especially encouraged to participate.
>
> The time commitment would be two 1-hour sessions, spaced approximately one week apart, to take place in the Usability Lab of the Centre for Management Informatics. Scheduling is at your convenience -- evening and weekend appointments are available. An honorarium of $7 will be provided for the first session and $13 for the second, for a total of $20.
>
> To participate, please contact Tara Whalen {*contact email and phone provided*}.

Because of the office scenario, special encouragement was included for office workers, as they represent the ideal testing participant. Note that as stated above, participants were compensated twenty dollars if they attended both sessions: seven dollars for the first session and thirteen dollars for the second session. The second session was more lucrative to encourage participants to return, completing both sessions.

8.7.3 Sample Size

Twenty-eight participants took part in this study, along with two pilot participants. The experimental design required that the number of participants be a multiple of four (to balance the tasks across both conditions, in both sub-sessions). A minimum of ten participants per condition is a common usability testing guideline, including scenario-based testing (see, for example, Rosson & Carroll, 2002); this experiment used two conditions (the two file managers). Similarly, a sample size of twenty participants in a usability study has been shown to be a highly effective number for uncovering defects (Faulker, 2003). Therefore, a minimum of twenty participants was

recommended; the actual number of participants who took part in this experiment was higher than this (n=28).

8.7.4 Participant Characteristics: Demographics and Experience

The 28 participants provided basic demographic data as well as information about their professional and computer backgrounds. These data are summarized in Table 57. The participant group was predominantly female (71%), with most participants coming from the 25-34 age group (54%). The population was a mix of students and employees, from a variety of faculties and careers. The majority of participants (68%) had at least three years of work experience. In addition, participants reported being technically proficient: no one had less than five years' computer experience, and the majority were very comfortable (or extremely comfortable) using Windows XP (68%) and Outlook (54%).

8.8 *Procedures*

Prior to the formal laboratory evaluation, the instruments and procedures were pilot tested. A set of preliminary pilot tests were run with four participants; a refined pilot test was conducted with two more participants. Feedback from these tests was used to refine the instruments and the procedures. The laboratory evaluation was conducted from December 2007 to February 2008. As described in the Design section, the study took place over two sessions. The procedures used for each session are described below.

8.8.1 Session One

For the first session, the following activities took place, in the order presented:

1-1. Participants arrived at the Usability Lab and received brief verbal overview of study

1-2. Participants read and signed consent form; any questions they had about the study were answered

1-3. Demographic information was collected through a background questionnaire (see Appendix F).

Participant demographics and experience (n=28)				
sex	male: **8**	female: **20**		
age	under 25: **8**	25-34: **15**	35-44: **5**	
students (n= 19) [1]	undergraduate: **6**	graduate: **13**		
	programs: library/information studies (5); computer science (3); environmental studies/management (2); electronic commerce (1); health informatics (1); political science (1); biochemistry (1); marketing/commerce (1); international development (1); MBA (1); psychology (1); unspecified (1)			
employees (n= 18)[1]	*job titles*[2]: senior administrator (5); administrative secretary (3); campus security (1); IT sales (1); community planner (1); freelance journalist (1); teaching/research assistant (1); senior consultant (1); IT technician (1); architecture consultant (1); retail employee (1); associate professor (1)			
years of work experience	about one year: **2**	5-9 years: **5**		
	1-2 years: **7**	10-20 years: **4**		
	3-4 years: **8**	over 20 years: **2**		
years of computer experience	5-10 years: **11**	11-20 years: **13**	over 20 years: **4**	
years of experience with Windows XP	1-2 years: **7**	3-4 years: **8**	5-6 years: **9**	over 6 years[3]: **4**
comfort level with Windows XP (7-point scale)[4]	extremely comfortable [7]	**11**		
	very comfortable [6]	**8**		
	comfortable [5]	**7**		
	neither comfortable nor uncomfortable [4][5]	**2**		
comfort level with Microsoft Outlook email program (7-point scale)[4]	extremely comfortable [7]	**7**		
	very comfortable [6]	**8**		
	comfortable [5]	**9**		
	neither comfortable nor uncomfortable [4]	**4**		

Comments

[1] A participant could be both a student and an employee.
[2] Some job titles have been made less specific to hide potentially-identifying details.
[3] At the time of testing, Windows XP had been available for just over seven years.
[4] Scale endpoints were "extremely uncomfortable" (1) and "extremely comfortable" (7).
[5] One participant who chose "4" had 1-2 years' experience; the other had 5-6 years.

Table 57: Information about study participants

Sub-session One starts

1-4. Participant received verbal introduction to file manager (either control or experimental), and completed warm-up task (navigating in file system; using file manager features). Printed handout with instructions for file manager was provided.

1-5. Participant received verbal introduction to Outlook search features and XP file permission tool. Printed handout with instructions for these tools was provided.

1-6. Morae Recorder was launched to record user interactions for remainder of session.

1-7. Scenario was presented; as a verbal introduction with accompanying handout (see Appendix F).

1-8. Participant received the set of tasks ("faxes"), stapled in the order that they were to be completed.

1-9. Facilitator left the room; observation was made from adjoining control room. Participant could ask questions at any time; communication with facilitator was via microphone and speakers.

1-10. Participant completed tasks (either three or five, depending on condition), consisting of reading fax questions, finding answers in file system, and writing them on the fax sheet provided. Participant told facilitator when all tasks were done.

1-11. Participant directed to complete computer-based usability questionnaire (CSUQ/ASQ)

Sub-session One ends

Sub-session Two starts

1-12. Facilitator returned to Usability Lab with second set of tasks.

1-13. Participant received verbal introduction to file manager (either control or experimental), and completed warm-up task (navigating in file system; using file manager features). Printed handout with instructions for file manager was provided.

1-14. Participant received the set of tasks ("faxes"), stapled in the order that they were to be completed.

1-15. Facilitator left the room; observation was made from adjoining control room. Participant could ask questions at any time; communication with facilitator was via microphone and speakers.

1-16. Participant completed tasks (either three or five, depending on condition), consisting of reading fax questions, finding answers in file system, and writing them on the fax sheet provided. Participant told facilitator when all tasks were done.

1-17. Participant directed to complete computer-based usability questionnaire (CSUQ/ASQ)

Sub-session Two ends

Final Phase

1-18. Facilitator instructed participant (over speakers) to complete final paper-based questionnaire (preference data). Screen capture (logging) terminated and saved.

1-19. Participant completed questionnaire, told facilitator when done

1-20. Facilitator returned to Usability Lab; participant was paid for session ($7) and was scheduled for Session Two (four to ten days in the future).

End of Session One

8.8.2 Session Two

For the second session, the following activities took place, in the order presented:

2-1. Participants arrived at the Usability Lab and received brief verbal overview of study (reminder from previous session). Note that participant completed tasks using file manager condition in opposite order from previous session; if they started with control last time, they started with the experimental this time (and vice versa).

Sub-session One starts

2-2. Participant received verbal introduction to file manager (either control or experimental), and completed warm-up task (navigating in file system). Printed handout with instructions for file manager was provided.

- If this was the experimental condition: during the introduction, the notification feature was mentioned, showing that the absent employee had already configured an alert for an event of interest. (This alert appeared during the experimental session.)

2-3. Participant received verbal introduction to Outlook search features and XP file permission tool. Printed handout with instructions for these tools was provided.

2-4. Morae Recorder was launched to record user interactions for remainder of session.

2-5. Scenario was presented; as a verbal introduction with accompanying handout (see Appendix F). Included in the scenario was the fact that "one week" had passed since participant was here and that the computer files had changed during that time.

2-6. Participant received the set of tasks ("faxes"), stapled in the order that they were to be completed.

2-7. Facilitator left the room; observation was made from adjoining control room. Participant could ask questions at any time; communication with facilitator was via microphone and speakers.

2-8. Participant completed tasks (either three or five, depending on condition), consisting of reading fax questions, finding answers in file system, and writing them on the fax sheet provided. Participant told facilitator when all tasks were done.

- If the condition was the experimental file manager, a notification message appeared after three minutes (this was the alert message that was mentioned during the introduction to the experimental file manager).

2-9. Participant directed to complete computer-based usability questionnaire (CSUQ/ASQ)

Sub-session One ends

Sub-session Two starts

2-10. Facilitator returned to Usability Lab with second set of tasks.

2-11. Participant received verbal introduction to file manager (either control or experimental), and completed warm-up task (navigating in file system). Printed handout with instructions for file manager was provided.

- If this was the experimental condition: during the introduction, the notification feature was mentioned, showing that the absent employee had already configured an alert for an event of interest. (This alert appeared during the experimental session.)

2-12. Participant received verbal introduction to Outlook search features and XP file permission tool. Printed handout with instructions for these tools was provided.

2-13. Participant received the second set of tasks ("faxes"), stapled in the order that they were to be completed.

2-14. Facilitator left the room; observation was made from adjoining control room. Participant could ask questions at any time; communication with facilitator was via microphone and speakers.

2-15. Participant completed tasks (either three or five, depending on condition), consisting of reading fax questions, finding answers in file system, and writing them on the fax sheet provided. Participant told facilitator when all tasks were done.

- If the condition was the experimental file manager, a notification message appeared after three minutes (this was the alert message that was mentioned during the introduction to the experimental file manager).

2-16. Participant directed to complete computer-based usability questionnaire (CSUQ/ASQ)

Final Phase

2-17. Facilitator instructed participant (over speakers) to complete paper-based questionnaire (preference data, suggestions for improvement). Screen capture (logging) terminated and saved.

2-18. Participant completed questionnaire, told facilitator when done.

2-19. Participant instructed to complete Technology Acceptance Model questionnaire; told facilitator when done.

2-20. Facilitator returned to Usability Lab for talk-after; audio and screen recordings made of this session. Screen captures replayed for participant, of the experimental condition tasks. Participant asked to describe what they did as they completed tasks. When the alert message (notification) appeared during replay, participant asked whether they noticed it, if it were meaningful and whether it was intrusive.

2-21. Participant was paid for session ($13) and asked whether they would consent to a follow-up email in two weeks' time.

End of Session Two

8.9 *Summary*

This chapter presented the experimental design and methodology used to evaluate the experimental file manager. A realistic workspace environment was created, including a file collection, and a series of tasks was developed from the awareness framework. Data was collected from these tasks (task completion time, correctness of question response) to determine whether the experimental file manager provided improvements in effectiveness or efficiency. Questionnaire data also indicated user satisfaction and intention-to-use; intention-to-use was also assessed through a post-experiment email. The data was collected and analysed, and used to address the experimental hypothesis and research questions. These results are presented in the following chapter.

Chapter 9. Evaluation Part Two: Results

The results of the experiment are presented in detail in this chapter. The results are presented in three parts, corresponding to the goals of the evaluation:

1. to compare the experimental file manager to a control file manager, on the basis of efficiency, effectiveness, and satisfaction
2. to evaluate the new functionality and features unique to the experimental file manager, and
3. to assess the *intention to use* factor for the experimental file manager

9.1 *Overview of the Presentation of Results*

As described above, the results of the laboratory evaluation are presented in three parts in this chapter.

Part One: this part compares the performance of the experimental file manager to the control file manager to determine whether the experimental file manager provided any benefits for efficiency, effectiveness, or satisfaction. The two file managers (control and experimental) were compared using four variables:

- task completion time
- correctness of question responses (accuracy and completeness)
- usability scores
- user preference

The relationship of these four variables to three usability factors—efficiency, effectiveness, and satisfaction—is shown in Table 58. Each of these four variables was compared between both file managers, using inferential statistics; these results are discussed in the first part of this chapter.

	usability measures mapped to testing variables			
usability measure	task completion time	correctness of question response	usability scores	user preference
efficiency	✓			
effectiveness		✓		
satisfaction			✓	✓

Table 58: Mapping of the testing variables to usability measures, showing how each was assessed in evaluation.

Part Two: this part presents results that are concerned *only* with the experimental file manager, evaluating its innovative features and design. (Because these new components are not available in the control file manager, they cannot be compared with it; hence, these findings are presented separately from the comparative results.)

The new features and functionality were assessed using both objective and subjective data. The information used in this analysis came from the following sources:

- for tasks performed *only* in the experimental file manager condition:
    - task completion times (descriptive statistics)
    - correctness of question responses (descriptive statistics)
- participant ratings and comments on:
    - effectiveness of the notification feature
    - quality of information display and ease of learning
    - suggestions for improvements of the new features

Part Three: the third and final section presents results about intention to use. This factor indicates how likely people are to adopt a new technology, such as the experimental file manager. Intention to use was evaluated using two sources:

- Technology Acceptance Model (TAM) questionnaire scores
- participant comments: real-world examples solicited in a post-study email

All statistical analysis was performed using SPSS 14.0 for Windows; statistical advice was provided by Vaneeta Grover (Dalhousie University Department of Mathematics and Statistics) and Holly Etchegary (IWK Health Centre Interdisciplinary Research).

9.2 *Comparison between File Managers*

As stated in the introduction to this chapter, the experimental file manager was compared against a control file manager. Four variables were recorded in both conditions: task completion times, correctness of question response, usability scores, and user preference. These four variables were statistically analysed; an alpha level of .05 was used for all statistical tests. The results are presented in the following sections.

9.2.1 Task Completion Times

As discussed in Chapter 8, one of the research hypotheses for this evaluation was that the experimental file manager would lead to faster completion times for at least one task, and would not be slower than the control for any tasks.

H1: The experimental file manager will result in task completion times that are faster than those in the control file manager for at least 50% of tasks, and will not result in any slower completion times.

9.2.1.1 Task Completion Times: Data Analysis

The completion times for each of twelve tasks, completed using both the control and the new file manager, were compared using the Mann-Whitney U test, which is a non-parametric statistical test. A non-parametric test was chosen instead of a parametric test, such as a t-test, because it does not rely on an assumption of normality within the data (Field, 2005). (The task completion times were tested for normality, using the Shapiro-Wilk test; results showed that not all task times were normally distributed.)

9.2.1.2 Task Completion Times: Results

The results of the Mann-Whitney U test, comparing the completion times for each task in the two conditions (control and experimental) for both study sessions, are shown in Table 59. The means and standard deviations for all task times are also provided.

As shown in Table 59, in Session One, two tasks demonstrated significant improvement for the experimental file manager: A and F (finding email recipients of a file). For Session Two, four tasks showed significant improvements: I and O (finding email recipients of a file) and K and Q (checking file permissions).

9.2.2 Correctness of Question Responses

The two file managers were also compared in terms of how well they assisted participants in answering the questions that formed the basis of the tasks. These questions had a variety of scores, depending on the complexity of the question; there were 24 questions in all.

The research hypothesis was that the experimental file manager would lead to more accurate and complete answers to questions about shared files.

H2: The experimental file manager will result in at least 25% of question response scores that are higher than the control.

Task ID	task description	M (sec), SD **control**	M (sec), SD **experimental**	Mann-Whitney U test U statistic, p-value
Session One				
A	finding email recipients of file (if any)	M=309.5 SD=208.3	M=102.1 SD=82.1	U=37 , p=.004**
B	checking file permissions	M=61.1 SD=29.2	M=57.6 SD=62.4	U=64, p=.125
C	checking "last modified" date of file	M=34.8 SD=25.3	M=44.7 SD=28.5	U= 84, p=.541
F	finding email recipients of file (if any)	M=196.1 SD=81.6	M=94.9 SD=131.3	U=33, p=.002**
G	checking file permissions	M=54.6 SD=45.5	M=61.2 SD=25.7	U=64, p=.125
H	checking "last modified" date of file	M=39.1 SD=22.0	M=52.1 SD=52.0	U=92, p=.804
Session Two				
I	finding all email recipients of a file	M=194.3 SD=100.8	M=79.0 SD=52.0	U=32, p=.002**
J	checking recent file access information	M=71.3 SD=49.8	M=52.8 SD=18.9	U=73, p=.265
K	checking file permissions	M=85.7 SD=47.7	M=48.9 SD=21.1	U=43, p=.011*
O	finding all email recipients of a file	M=199.6 SD=64.2	M=64.2 SD=41.1	U=17, p<.01**
P	checking recent file access information	M=104.5 SD=71.0	M=60.8 SD=32.3	U=71, p=.227
Q	checking file permissions	M=114.6 SD=43.6	M=68.3 SD=40.2	U=39, p=.006**

Table 59: Results of task completion times; mean times in seconds are presented with standard deviation. Mann-Whitney U test results are shown; single asterisk indicates p<.05; double asterisk indicates p<.01.

9.2.2.1 Question Responses: Data Analysis

The scores for each of the 24 questions (answered using both file managers) were compared between conditions, using an independent samples t-test.

9.2.2.2 Question Responses: Results

The results are provided in Table 60. Each question is listed (by ID); a key at the bottom of the table gives a description of the question itself, for reference. The means and

question ID	mean score M (max), SD **control**	mean score M (max), SD **experimental**	**t-test results** t statistic, p-value
Session One			
A1	*M*=1.1 (4), *SD*=1.9	*M*=4.0 (4), *SD*=0.0	t(26)=-5.70 p<.01**
A2	*M*=1.4 (5), *SD*=2.3	*M*=4.6 (5), *SD*=1.2	t(26)=-4.50, p<.01**
B1	*M*=1.0 (1), *SD*=0.0	*M*=1.0 (1), *SD*=0.0	SD=0: no output
B2	*M*=1.0 (1), *SD*=0.0	*M*=1.0 (1), *SD*=0.0	SD=0: no output
C1	*M*=1.0 (1), *SD*=0.0	*M*=1.0 (1), *SD*=0.0	SD=0: no output
C2	*M*=0.9 (1), *SD*=0.4	*M*=1.0 (1), *SD*=0.0	t(26)=-1.47, p=.165
F1	*M*=1.7 (4), *SD*=2.1	*M*=4.0 (4), *SD*=0.0	t(26)=-4.16, p=.001**
F2	*M*=1.7 (4), *SD*=2.1	*M*=4.0 (4), *SD*=0.0	t(26)=-4.16, p=.001**
G1	*M*=0.9 (1), *SD*=0.3	*M*=1.0 (1), *SD*=0.0	t(26)=-1.00, p=.336
G2	*M*=1.9 (2), *SD*=0.5	*M*=2.0 (2), *SD*=0.0	t(26)= -1.00, p=.336
H1	*M*=1.0 (1), *SD*=0.0	*M*=1.0 (1), *SD*=0.0	SD=0: no output
H2	*M*=1.0 (1), *SD*=0.0	*M*=1.0 (1), *SD*=0.0	SD=0: no output
Session Two			
I1	*M*=5.6 (9), *SD*=3.9	*M*=9.0 (9), *SD*=0.0	t(26)=-3.31, p=.006**
J1	*M*=1.0 (1), *SD*=0.0	*M*=1.0 (1), *SD*=0.0	SD=0: no output
J2	*M*=1.9 (9), *SD*=0.3	*M*=7.7 (9), *SD*=0.7	t(26)=-27.97, p<.01**
K1	*M*=1.0 (1), *SD*=0.0	*M*=1.0 (1), *SD*=0.0	SD=0: no output
K2a	*M*=0.9 (1), *SD*=0.3	*M*=1.0 (1), *SD*=0.0	t(26)=-1.00, p=.336
K2b	*M*=0.9 (1), *SD*=0.4	*M*=1.0 (1), *SD*=0.0	t(26)=-1.47, p=.165
O1	*M*=3.4 (6), *SD*=2.9	*M*=6.0 (6), *SD*=0.0	t(26)=-3.35, p=.005**
P1	*M*=1.0 (1), *SD*=0.0	*M*=1.0 (1), *SD*=0.0	SD=0: no output
P2	*M*=1.4 (4), *SD*=0.9	*M*=3.5 (4), *SD*=0.7	t(26)=-7.23, p<.01**
Q1	*M*=1.0 (1), *SD*=0.0	*M*=1.0 (1), *SD*=0.0	SD=0: no output
Q2a	*M*=0.9 (1), *SD*=0.3	*M*=1.0 (1), *SD*=0.0	t(26)=-1.00, p=.336
Q2b	*M*=1.0 (1), *SD*=0.0	*M*=1.0 (1), *SD*=0.0	SD=0: no output
Key: description of questions matched to IDs (in first column)			
A1: Was [file A] emailed to anyone? A2: If yes: to whom?			
B1: Does the "legal" group have *modify* permission for [file B]? B2: If not, what permissions does legal have?			
C1: Is [file C] in the folder [folder A]? C2: If yes: what is the "last modified" date?			
F1: Was [file F] emailed to anyone? F2: If yes: to whom?			
G1: Can Dave Samuels modify [file G] G2: If no: what permissions does Dave have?			
H1: Is [file H] in [folder B]? H2: If yes: was the file modified in the last 24 hours?			
I1: List all who received [file I] over email between [dates]			
J1: Did anyone access [file J] between [dates]? J2: If so, provide all the details you can find about any accesses (such as time), between [dates]			
K1: Is [file K] in [folder C]? K2a: If yes, does [person X] have read & write permission on it? K2b: If yes, does [person Y] have read & write permission on it?			
O1: Who was emailed a copy of [file M] [since date]?			
P1: Did anyone access [file N] on [date] or later? P2: If so, provide all the details you can find about any accesses (such as time), between [dates]:			
Q1: Is [file O] in [folder F]? Q2a: If so, do only [groups A&B] have access, and no one else? Q2b: Do [groups A & B] have read access?			

Table 60: Results for question response scores for both control and experimental file managers. The mean score is shown with the maximum possible score for that question, and the standard deviation. Independent samples t-test results are shown; single asterisk indicates p<.05; double asterisk indicates p<.01. (n=28) When both results have an SD of zero, the t-test cannot be run; these items are marked as SD=0. A description of each question (with its ID) is provided for reference.

standard deviations for each question are provided, along with the t-test result. Note that for some questions, particularly those whose maximum score was "1", the results for both conditions had a standard deviation of zero. This indicates that all the values in the data are identical; there is no t-test output in these cases. These cases are marked in Table 60 as "SD=0: no output." (For interpretation, note that a higher score indicates a more correct response to the question about files.)

As shown in Table 60, for Session One, there were significant differences found in response scores for four questions: A1, A2, F1, and F2 (listing email recipients of a file). For Session Two, there were significant differences found for four questions as well: I1 and O1 (listing email recipients of a file) and J2 and P2 (listing recent file access information).

9.2.3 Usability Scores

After each set of tasks (twice per session), participants completed a usability questionnaire, consisting of 19 questions: 17 from the Computer System Usability Questionnaire (CSUQ) and two questions from the After-Scenario Questionnaire (ASQ). (Lewis, 1991). Questions scores are based on seven-point Likert scale, ranging from "strongly disagree" (1) to "strongly agree" (7).

The research hypothesis was that the experimental file manager would score higher than the control on the usability questionnaire. However, it was expected that this result might not occur until the second session, when the participants had more experience with using the experimental file manager.

H3: The experimental file manager will have the majority of the higher scores for the usability questions in Session One or Session Two, or both.

It was also expected that order mattered, particularly in Session One, in terms of relative ranking of the two file managers. For example, at the end of Session One, some participants who began with the control file manager mentioned that they would have ranked it lower if they had already used the experimental file manager and its new features. (In other words, they *preferred* the experimental file manager, but they could not retroactively lower their scores for the earlier control to reflect this change in their perception.)

H4: There will be a significant difference in at least one usability score, in Session One, between those who used the control *after* the experimental manager and those who used the control *before* the experimental manager.

9.2.3.1 Usability Scores: Data Analysis

The CSUQ questions are divided into four factors: system usefulness (SYSUSE), information quality (INFOQUAL), interface quality (INTERQUAL) and an overall score (OVERALL) (Lewis, 1995). (The mapping between the items in the questionnaire and these factors was presented in Chapter 8.) The mean scores for the items (question) for each of these four factors were computed, along with the mean ASQ score. This resulted in a set of five usability scores (SYSUSE, INFOQUAL, INTERQUAL, OVERALL, ASQ) for both file managers in both sessions.

These usability scores were compared in two ways, using an independent samples t-test for each. First, scores were compared between the control and the experimental file managers for each of Session One and Session Two, to determine which file manager scored better (to respond to H3). Scores were compared using an independent samples t-test. Second, the scores were compared to check for order effects (to address H4). Comparisons were made for each file manager between those who had used it *first*, versus those who had used it *second*; this comparison was made for both sessions, for a total of four comparisons. In other words, this result would show whether participants rated a particular file manager differently (higher or lower) depending on the order in which they used it. Scores were compared using an independent samples t-test.

9.2.3.2 Usability Scores: Results

The results of the first analysis—comparing the two file managers, irrespective of order—are presented in Table 61. Results are presented for both sessions; scores are based on a seven-point scale. The results of the second comparison—order effects for the two managers—are shown in Table 62, for each of the four combinations of order and file manager condition.

factor	mean score M (max), SD **control**	mean score M (max), SD **experimental**	**t-test results** t statistic, p-value
Session One			
SYSUSE	*M*=5.4, *SD*=1.3	*M*=6.1, *SD*=0.8	t(54)= -2.64, p=.011*
INFOQUAL	*M*=5.2, *SD*=1.3	*M*=6.1, *SD*=0.7	t(54)= -3.54, p=.001**
INTERQUAL	*M*=5.1, *SD*=1.2	*M*=5.9, *SD*=0.9	t(54)= -2.79, p=.007**
OVERALL	*M*=5.2, *SD*=1.2	*M*=6.1, *SD*=0.7	t(54)= -3.22, p=.002**
ASQ	*M*=4.9, *SD*=1.8	*M*=6.3, *SD*=0.8	t(54)= -3.85, p<.01**
Session Two			
SYSUSE	*M*=4.6, *SD*=1.2	*M*=6.3, *SD*=0.5	t(54)= -6.98, p<.01**
INFOQUAL	*M*=4.5, *SD*=1.3	*M*=6.1, *SD*=0.7	t(54)= -5.67, p<.01**
INTERQUAL	*M*=4.2, *SD*=1.3	*M*=6.1, *SD*=0.7	t(54)= -6.82, p<.01**
OVERALL	*M*=4.5, *SD*=1.2	*M*=6.2, *SD*=0.5	t(54)= -7.15, p<.01**
ASQ	*M*=4.2, *SD*=1.6	*M*=6.5, *SD*=0.7	t(54)= -7.02, p<.01**

Table 61: CSUQ and ASQ scores for the control and experimental file managers, for each session. The mean score and standard deviation are listed for each factor. Independent samples t-test results are shown; single asterisk indicates p<.05; double asterisk indicates p<.01. (n=28)

factor	mean score M (max), SD **control mgr. first**	mean score M (max), SD **experimental mgr. first**	**t-test results** t statistic, p-value
Session One			
	control file manager		
SYSUSE	*M*=6.4, *SD*=0.5	*M*=5.0, *SD*=1.2	t(26)= -4.30, p<.01**
INFOQUAL	*M*=6.5, *SD*=0.5	*M*=4.8, *SD*=1.2	t(26)= -4.53, p<.01**
INTERQUAL	*M*=6.1, *SD*=0.9	*M*=4.7, *SD*=1.1	t(26)= -3.72, p=.001**
OVERALL	*M*=6.4, *SD*=0.5	*M*=4.8, *SD*=1.1	t(26)= -4.60, p<.01**
ASQ	*M*=6.8, *SD*=0.5	*M*=4.5, *SD*=2.1	t(26)= -3.96, p=.001**
	experimental file manager		
SYSUSE	*M*=6.4, *SD*=0.5	*M*=5.8, *SD*=0.9	t(26)= -2.34, p=.027*
INFOQUAL	*M*=6.5, *SD*=0.5	*M*=5.8, *SD*=0.8	t(26)= -2.70, p=.013*
INTERQUAL	*M*=6.1, *SD*=0.9	*M*=5.7, *SD*=0.8	t(26)= -1.27, p=.217
OVERALL	*M*=6.4, *SD*=0.5	*M*=5.8, *SD*=0.7	t(26)= -2.46, p=.021*
ASQ	*M*=6.8, *SD*=0.5	*M*=5.9, *SD*=0.9	t(26)= -2.98, p=.008**
Session Two			
	control file manager		
SYSUSE	*M*=5.0, *SD*=0.9	*M*=4.1, *SD*=1.3	t(26)=2.28, p=.039*
INFOQUAL	*M*=5.0, *SD*=1.0	*M*=4.0, *SD*=1.3	t(26)=2.21, p=.036*
INTERQUAL	*M*=4.5, *SD*=1.2	*M*=3.9, *SD*=1.3	t(26)=1.20, p=.242
OVERALL	*M*=4.9, *SD*=1.0	*M*=4.0, *SD*=1.3	t(26)=2.08, p=.048*
ASQ	*M*=4.9, *SD*=1.1	*M*=3.5, *SD*=1.6	t(26)=2.62, p=.015*
	experimental file manager		
SYSUSE	*M*=6.3, *SD*=0.5	*M*=6.4, *SD*=0.6	t(26)= -.523, p=.606
INFOQUAL	*M*=6.1, *SD*=0.5	*M*=6.1, *SD*=0.8	t(26)=.055, p=.956
INTERQUAL	*M*=6.1, *SD*=0.8	*M*=6.1, *SD*=0.7	t(26)= -.003, p=.998
OVERALL	*M*=6.2, *SD*=0.5	*M*=6.2, *SD*=0.6	t(26)= -.261, p=.796
ASQ	*M*=6.5, *SD*=0.5	*M*=6.4, *SD*=0.9	t(26)=.138, p=.891

Table 62: CSUQ and ASQ scores, compared for each file manager on the basis of order effect Independent samples t-test results are shown; single asterisk indicates p<.05 and double asterisk indicates p<.01. (n=28)

In terms of the usability scores for each session (irrespective of order), significant differences were found for all five usability factors, in both sessions, as shown in Table 61. When order effects are taken into account, as shown in Table 62, significant differences are again shown for usability. This is a more complex case to describe (and is discussed in more detail in the Discussion section below), but in brief: the control manager was rated significantly worse for usability by participants who used the experimental file manager beforehand (in both sessions, for at least four of the five usability factors). Also, participants who used the control manager first rated the experimental manager significantly higher in comparison, for Session One.

As an addendum, note that as shown in Table 62, mean scores for the control file manager decreased between sessions. These scores cannot be directly compared in formal analysis (for example, as repeated measures), because the environment and tasks differed in each session. However, the trend suggests that as participants gained more experience with both file managers, the control was perceived as less usable.

9.2.4 Preference Data

The final variable used to compare the two file managers was subjective preference. This was assessed by questionnaires: participants indicated which file manager they preferred, on a number of factors. For Session One, the questionnaire had 11 questions in all. The first three questions centered on the three *tasks* performed using both file managers. Respondents indicated which of the two managers most helped them complete each task; they could also choose "both the same." The remaining eight questions were forced-choice, and included questions about awareness and perceived usefulness. (The full text of all questions is listed below in the Results tables.) The questionnaire for Session Two was similar, but with 14 questions: the same 11 questions as in Session One, and three additional questions about catching up with changes to files between sessions.

It was expected that the experimental file manager would perform at least as well as the control, in terms of overall preference. That is, if significant differences were found for any questions, the majority of these differences should favour the experimental manager. Moreover, this condition was expected to hold true for both sessions, despite the unfamiliarity of the software.

H5 The experimental file manager will be selected more often than the control for the preference questions, in both Session One and Session Two.

9.2.4.1 Preference Data: Data Analysis

The preference questions were all analysed using chi-square tests. For the task-based questions (the first three questions of both questionnaires), there were three options, including "both were the same." In this instance, the expected frequency for responses, when there is no difference, is not equally divided among all three possibilities. To normalize the data, whenever a participant selected "both the same", this count was halved: 0.5 for the control, and 0.5 for the experimental. In this way, the counts for each could be directly compared, using 50% expected frequency for each response in the chi-square test.

For the other questions, which used a forced-choice, the basic chi-square test was applied, with expected frequency of 50% for each cell. This test was performed for the eight questions in Session One, and the 11 questions in Session Two.

9.2.4.2 Preference Data: Results

The results of the chi-square tests for Session One, for the task questions, are shown in Table 63.

Session One: Task Questions		
Which file manager best helped you complete the task:	**descriptive statistics (n=28)**	**chi-square**
Q1	exp.=16 (57%); control =4 (14%), same=8 (29%)	$\chi2(1, n=28)= 5.14, p=.023$*
Q2	exp.=28 (100%); control=0, same=0	$\chi2(1, n=28)= 28.00, p<.001$**
Q3	same=24 (86%); exp.=4 (14%) control=0	$\chi2(1, n=28)= .571, p=.450$
Questions about tasks:		
Q1: Checking file permissions		
Q2: Determining who a file was sent to, if anyone		
Q3: Checking the "last modified" date of a file		

Table 63: Descriptive statistics and chi-square result for preference data on tasks (Session One). "Exp." refers to the experimental file manager. A single asterisk indicates p<.05; double asterisk indicates p<.01.

The results for the remaining Session One questions are presented in Table 64, including the question text, with descriptive statistics and chi-square results. (The descriptive statistics for only the experimental file manager statistics are explicitly

presented in Table 64, for simplicity. (Because there were only two choices, the control results can be computed as the complement.) One participant did not respond to questions Q4 or Q8, thus n=27 for those two questions.

Session One: Preference Questions (choice between control and experimental)		
Which file manager best helped you complete the task:	**descriptive statistics (n=28, except Q4 & Q8: n=27)**	**chi-square**
Q4	*Experimental*= 22 (81%)	$\chi 2(1, n=27)=10.704, p=.001^{**}$
Q5	*Experimental*= 25 (89%)	$\chi 2(1, n=26)=18.615, p<.01^{**}$
Q6	*Experimental*= 25 (89%)	$\chi 2(1, n=28)=17.286, p<.01^{**}$
Q7	*Experimental*= 25 (89%)	$\chi 2(1, n=28)=17.286, p<.01^{**}$
Q8	*Experimental*= 27 (100%)	$\chi 2(1, n=27)= 27.000, p<.01^{**}$
Q9	*Experimental*= 28 (100%)	$\chi 2(1, n=28)=28.000, p<.01^{**}$
Q10	*Experimental*= 24 (86%)	$\chi 2(1, n=26)=18.615, p<.01^{**}$
Q11	*Experimental*= 25 (89%)	$\chi 2(1, n=28)=17.286, p<.01^{**}$
Preference questions:		
Q4: Overall, it was simpler to use this file manager		
Q5: I could more effectively complete the tasks using this file manager		
Q6: I was able to complete the tasks quicker using this file manager		
Q7: It was easier to find the information I needed using this file manager		
Q8: I had greater awareness of which files were shared with whom using this file manager		
Q9: I had more complete information about shared files using this file manager		
Q10: I believe I could manage my shared files better using this file manager		
Q11: Overall, I prefer:		

Table 64: Preference data for forced-choice questions (Session One), expressed as descriptive statistics and chi-square results. Descriptive results shown only for experimental files manager (the control represents the remainder of the total). A single asterisk indicates p<.05; double asterisk indicates p<.01.

The results of the task questions for Session Two are shown in Table 65, again using descriptive statistics along with chi-square.

Session Two: Task Questions		
Which file manager best helped you complete the task:	**descriptive statistics (n=28)**	**chi-square**
Q1	exp.=21 (75%); same=7 (25%); control=0	$\chi 2(1, n=28)=15.75, p<.001^{**}$
Q2	exp.=28 (100%); same=0; control=0	$\chi 2(1, n=28)= 28.00, p<.001^{**}$
Q3	exp.=27 (96%); same=1 (4%); control=0	$\chi 2(1, n=28)= 26.035, p<.001^{**}$
questions about task:		
Q1: Checking file permissions		
Q2: Determining who a file was sent to, if anyone		
Q3: Finding information about past accesses on a file		

Table 65: Descriptive statistics and chi-square result for preference data on tasks (Session Two). A single asterisk indicates p<.05; double asterisk indicates p<.01.

The results for the remaining Session Two questions are presented in Table 66, including the question text.

Session Two: Preference Questions (choice between control and experimental)		
Which file manager best helped you complete the task:	**descriptive statistics (n=28, except Q4 & Q8: n=27)**	**chi-square**
Q4	Experimental= 24 (86%)	$\chi 2(1, n=28)=14.286, p<.01^{**}$
Q5	Experimental= 28 (100%)	$\chi 2(1, n=28)=28.000, p<.01^{**}$
Q6	Experimental= 27 (96%)	$\chi 2(1, n=28)=24.143, p<.01^{**}$
Q7	Experimental= 27 (96%)	$\chi 2(1, n=28)=24.143, p<.01^{**}$
Q8	Experimental= 28 (100%)	$\chi 2(1, n=28)= 28.000, p<.01^{**}$
Q9	Experimental= 28 (100%)	$\chi 2(1, n=28)=28.000, p<.01^{**}$
Q10	Experimental= 28 (100%)	$\chi 2(1, n=28)=28.000, p<.01^{**}$
Q11	Experimental= 28 (100%)	$\chi 2(1, n=28)= 28.000, p<.01^{**}$
Q12	Experimental= 28 (100%)	$\chi 2(1, n=28)= 28.000, p<.01^{**}$
Q13	Experimental= 27 (96%)	$\chi 2(1, n=28)=24.143, p<.01^{**}$
Q14	Experimental= 27 (96%)	$\chi 2(1, n=28)=24.143, p<.01^{**}$
Preference questions		
Q4: Overall, it was simpler to use this file manager		
Q5: I could more effectively complete the tasks using this file manager		
Q6: I was able to complete the tasks quicker using this file manager		
Q7: It was easier to find the information I needed using this file manager		
Q8: I was better able to find out how the workspace had changed during my absence (e.g., the "week" since the last session) using this file manager		
Q9: I was better able to keep track of events on shared files using this file manager		
Q10:I was better able to find out who had made changes to shared files using this file manager		
Q11: I had greater awareness of which files were shared with whom using this file manager		
Q12: I had more complete information about shared files using this file manager		
Q13: I believe I could manage my shared files better using this file manager		
Q14: Overall, I prefer:		

Table 66: Preference data for forced-choice questions (Session Two), for either control or experimental file manager, presented as descriptive statistics and chi-square results. A single asterisk indicates p<.05; double asterisk indicates p<.01.

In summary: for both sessions, the experimental file manager was found to be the strongly-preferred choice, for all of the preference questions except for Q3 in Session

One. (Q3 asked which file manager best helped in finding file modification times: most participants found the two managers equivalent for this one task.)

9.3 *Evaluation of New Features and Functionality*

Assessing the usability of the experimental file manager, as well as its potential utility for real-world environments, involved the use of multiple data sources. The objective task data will be presented first: completion times and question response scores. This will be followed by results derived from participant ratings and comments, on three subjects. The first is effectiveness of the notification feature (pop-up alert). The second is the participant ratings of quality of the file sharing information display and the ease of learning the experimental file manager. The third item presented is the participant suggestions for improvements of the experimental file manager.

9.3.1 Task Completion Times: Experimental File Manager-Only Tasks

For each session, there were some tasks completed only with the experimental file manager, used to evaluate the unique features of that file manager. In Session One, two additional tasks were performed (D and E): one to check file access history, and the other to create a "person-centric" list of shared files (i.e., items that person can access and has copies of). In Session Two, three additional tasks (L, M, and N) were performed. The first two tasks were similar to those in Session One: check an element of file access history, and create a list of files emailed to, and accessed by, a specific person (within a date range). The third task was to find the three most-accessed files in a specified folder.

9.3.1.1 Results

The task completion times (mean and standard deviation) for Tasks D and E (Session One) and Tasks L, M, and N (Session Two) are presented in Table 67.

Mean completion time for tasks completed only in experimental file manager (n=28)		
task	task description	Time to complete in seconds: M, SD
Session One		
D	find details about recent accesses on two different files	M=127.0, SD= 77.3
E	find three files emailed to a specified person, and three files that person has read or write access to	M=246.3, SD=185.1
Session Two		
L	check if a certain person accessed one file since date dd/mm; if so, find what action was performed	M=192.46, SD= 47.2
M	list all files emailed to a specified person; list all files accessed by that person after date dd/mm	M=404.6, SD= 80.7
N	list the three most-accessed files in a specified folder	M=159.9, SD= 41.4

Table 67: Task completion times (in seconds) for experimental file manager-only tasks, presented as means. Standard deviation is included for each measure.

9.3.2 Correctness of Question Responses: Experimental File Manager-Only Tasks

For each of the tasks above, a question was responded to (on the scenario fax), and that response was scored. Just as with the responses analysed previously (in Section 9.2.2), each question had a maximum score, and the response received a score based on how accurate and complete it was. Each task had at least one question; the text for each question, and its maximum possible score, is presented in Table 68 below.

These questions were grouped into categories with counts: how many times the participants got the questions completely right (correct), partially correct, or completely wrong (incorrect). These categories were used to highlight which types of questions were effectively (or ineffectively) answered using the experimental file manager.

9.3.2.1 Results

The response scores are presented in Table 68, including the text for each question and the maximum score for that question (in parentheses). The response scores are presented as count data in three categories: correct, partial (partially correct) and incorrect; these counts are also presented as percentages (where n=28).

Session One				
Ques. ID	**Question text summary (max score)**	**correct**	**partial**	**incorrect**
D1a	For [file D], list number of times accessed: (1)	28 (100%)	0	0
D1b	person who accessed most recently (1)	28 (100%)	0	0
D1c	date of most recent access (1)	26 (93%)	0	2 (7%)
D2a	For [file E], list number of times accessed (1)	28 (100%)	0	0
D2b	person who accessed most recently (1)	28 (100%)	0	0
D2c	date of most recent access (1)	28 (100%)	0	0
E1	List three files that [person A] has received from me over email. (3)	27 (96%)	0	1 (4%)
E2	List three files that [person A] currently has access to (can read or write) (3)	28 (100%)	0	0
Session Two				
Ques. ID	**Question text summary (max score)**	**correct**	**partial**	**incorrect**
L1	Has [person B] accessed [file L] since [date]? (1)	28 (100%)	0	0
L2	If so, what action did he perform (e.g., read, write...)? (2)	26 (93%)	1 (3.5%)	1 (3.5%)
M1	List all files emailed to [person C] (7)	16 (57%)	8 (29%)	4 (14%)
M2	List all files accessed by [person C] in [folder D] after [date] (1)	22 (79%)	5 (17.6%)	1 (3.5%)
N1	List the three most-accessed files in [folder E] (4)	26 (93%)	2 (7%)	0

Table 68: Results for question correctness (experimental file manager-only tasks). Response results presented in categories (correct, partially correct and incorrect), with counts for each category; percentages are based on n=28.

To summarize these results, nearly all questions in Session One were answered correctly. For Session Two, three of the five questions were answered correctly by at least 90% of participants, but questions M1 and M2 demonstrated far more errors.

9.3.3 Event Notification: Alert Message

In addition to the tasks, participants encountered a pop-up message when they used the experimental file manager in Session Two. Participants did not need to explicitly act upon this message, but they were asked about it during the talk-after session at the end of the study. Specifically, they were asked whether they noticed it, whether the information was meaningful, and whether the alert window was obtrusive.

9.3.3.1 Results

In some cases, the pop-up window was buried under other windows (such as the file manager). Because the alert window did not "grab focus" when it appeared, it would not appear on top of any window being actively used. For example, if the participant was typing in the console, at the exact time when the alert appeared, the alert window would

appear underneath the console window. In some of these cases, the alert went unnoticed. Therefore, the alert window was observed by 22 out of 28 participants.

In the talk-after activity (when the screen-capture video was replayed for the participant for discussion), one participant mentioned closing the window by accident. Two others closed it almost immediately without reading it, due to a strong dislike of pop-up messages; they did not want to be interrupted and thus wanted to get rid of the window as soon as possible. (One person mentioned the annoyance of web browser pop-up windows, and how these pop-ups had conditioned him to immediately close any window that suddenly appeared.)

Of the 19 people who actually read the message, 18 people were able to associate it with the alert that had been set in the sharing console. They knew that the alert was telling them that the event that Chris Lee was watching for (i.e., Sheri Stack reading the file NewFund.doc) had occurred. One person did not recall this information from the session introduction, and thus did not know what the message meant. In addition, none of the 19 participants found the message window to be overly intrusive. The two participants who immediately closed it stated that they might not have done so if they themselves had set the alert, because they would be expecting the interruption.

9.3.4 Subjective Ratings: Information Display and Ease of Learning

At the end of Session Two, participants were given a questionnaire that asked them to rate various aspects of the experimental file manager. This questionnaire included seven questions about the information presented in the experimental file manager, and the perceived ease of learning to use this manager. Each question consisted of a seven-point Likert scale, ranging from "strongly agree"(1) to "strongly disagree"(7). The full text of the questions is provided below, in Table 69 (along with the results).

9.3.4.1 Results

The responses for each question were entered into a spreadsheet, and the means and standard deviations for all questions were computed; these are presented in Table 69. Note that reverse-wording was used in the questionnaire, so the questions are grouped according to the reverse-wording: positively-worded questions are presented first, followed by negatively-worded questions. For the first three statements, a *lower* score indicates a better rating; for the last four statements, a *higher* score is a better rating.

statement	mean, SD 7-point scale
positively-worded statements	
Q2: I was able to easily find the information I needed, about shared files, using this file manager.	M=1.57, SD=0.69
Q4: I was able to perceive information about shared files "at a glance" using this file manager	M=1.64, SD=0.78
Q6: The information displayed by this file manager is presented in a non-intrusive manner.	M=2.25, SD=1.55
negatively-worded statements	
Q1: I found it difficult to understand the information about shared files provided by the file manager.	M=6.32, SD=1.12
Q3: I found it difficult to learn to use this file manager.	M=6.00, SD=1.51
Q5: The information provided by the file manager was not useful for completing the tasks.	M=6.86, SD=0.36
Q7: There is too much information displayed in this file manager.	M=6.29, SD=1.08

Table 69: Responses to questions about the file sharing information and the ease of learning to use the experimental file manager. The responses are means, based on a 7-point Likert scale (where "1" means "strongly agree"). For the first three statements, a lower score is better; for the last four statements, a higher score is better (because of the negative wording).

9.3.5 Suggestions for Improvement

Participants were asked to provide suggestions for improvement for the file manager (including the sharing console), as part of a questionnaire conducted at the end of the study. They were presented with a list of elements that they could comment on; this list included the new file manager columns and the three components of the sharing console (search, file details, and alert). Participants could also list any features they felt should be added to the manager or console.

9.3.5.1 Results

Twenty-two participants provided suggestions for improvement. These suggestions are described below, presented in terms of the feature being critiqued: the file manager columns, the search feature (in the console), the file details (in the console), the alert, and any additional features.

9.3.5.1.1 File Manager Columns

There were two suggestions that were repeatedly requested for this feature. The first was to make the icons clickable, so that the sharing console could be opened by clicking on the icon. (In the implementation of the prototype used for evaluation, user had to click on the filename in order for the sharing console to work correctly.)

The second suggestion was to change the amount of information displayed in the columns. For some participants, less information was desired, particularly file permission information. Others wanted to have more information, particularly in the file event history (e.g., showing more than just the most recent event in the info-tip).

9.3.5.1.2 Search Feature (Console)

The most common suggestion for the search feature was to make it easier to access. One participant suggested that it become a top-level item in the right-click menu, so that you could directly select "Search" as well as the "Open in Console" option.

9.3.5.1.3 File Details (Console)

Three suggestions were made to improve this feature. Two participants indicated that the filename itself was not prominent enough, making it hard to tell which file the details were connected with. Another participant had trouble resizing the column widths, and suggested that the resize handles be made more visible. Finally, one participant suggested that the filetype icon (such as those used in the file manager) be used in the file detail window as well, to help identify the file type (e.g., Microsoft Word, Excel, etc.).

9.3.5.1.4 Alert

Two suggestions were proposed for the alert. The first was to provide a "disable" option; no reason was given, although it may have been to prevent interruptions. The second suggestion was to add an audio component, so the alert could be heard.

9.3.5.1.5 Additional features

Participants could also propose features that they would like to see added to the file manager; there were a few suggestions. One person suggested integrating the console with the built-in XP permission-configuration tool, to facilitate permission updates. Another suggested that the information be extended to include information about files sent *to* an individual (as opposed to just *outgoing* files, as in the current version).

9.4 *Intention to Use*

The final set of measures gauged intention to use: the likelihood that the new file manager would be adopted, should it become available. This was measured in two ways: using a Technology Acceptance Model (TAM) questionnaire, and through a post-study email.

9.4.1 Technology Acceptance Model

At the end of the second session, participants completed a Technology Acceptance Model questionnaire, in which they indicated how well they felt that the experimental file manager would help them in their job. This questionnaire measured the factors of perceived usefulness and perceived ease-of-use, using six items for each factor; one supplementary question (added to the TAM questionnaire) measured self-rated expected usage.

9.4.1.1 Technology Acceptance Model Results

The questionnaire items are listed in Table 70 ; perceived usefulness items are labelled U1-U6, perceived ease-of-use items are labelled E1-E6, and the expected usage question is labelled X1. The mean and standard deviation for each question response are also provided in Table 70.

Note that these responses were based on a seven-point scale, and thus a score of four represents the neutral point. As shown in Table 70, all items scored at least 5.29 on the scale, indicating positive ratings for perceived usefulness, perceived ease-of-use, and expected usage.

	Question	**mean, SD 7-point scale**
U1	Using this file manager in my job would enable me to accomplish tasks more quickly	*M*=5.82, *SD*=1.06
U2	Using this file manager would improve my job performance	*M*=5.39, *SD*=1.13
U3	Using this file manager in my job would increase my productivity	*M*=5.29, *SD*=1.21
U4	Using this file manager would enhance my effectiveness on the job	*M*=5.50, *SD*=1.26
U5	Using this file manager would make it easier to do my job	*M*=5.57, *SD*=1.37
U6	I would find this file manager useful in my job	*M*=6.00, *SD*=1.18
E1	Learning to operate this file manager would be easy for me.	*M*=6.39, *SD*=0.92
E2	I would find it easy to get this file manager to do what I want it to do	*M*=6.29, *SD*=0.81
E3	My interaction with this file manager would be clear and understandable	*M*=6.07, *SD*=0.86
E4	I would find this file manager to be flexible to interact with	*M*=6.07, *SD*=0.94
E5	It would be easy for me to become skilful at using this file manager	*M*=6.50, *SD*=0.64
E6	I would find this file manager easy to use	*M*=6.50, *SD*=0.64
X1	Assuming this file manager would be available on my job, I predict that I will use it on a regular basis	*M*=6.00, *SD*=1.31

Table 70: Responses to TAM questionnaire items. U1-U6 and E1-E6 are from Davis (1989); X1 is from Laitenberger and Dreyer (1998). Each item used a 7-point Likert scale. Mean (M) and standard deviation (SD) are listed in the third column.

The consistency of items from the original TAM questionnaire (U1-U6 and E1-E6) were measured with Cronbach's alpha. This result was used to check for any

unexpected inconsistencies in the responses. If the items U1-U6 had high inter-item correlation (i.e., all measuring the same construct, "perceived usefulness"), the value of Cronbach's alpha should be above 0.7 (Field, 2005). The same holds for items for E1-E6. The Cronbach's alpha measure for both factors is provided in Table 71.

factor	Cronbach's alpha (α)
perceived usefulness (six items: U1-U6)	.931
perceived ease-of-use (six items: E1-E6)	.917

Table 71: Cronbach's alpha measure of consistency, for the twelve items in the TAM questionnaire. Values of α above 0.7 are considered to have high inter-item correlation.

Because the alpha measure for both factors is above 0.7, there are no inconsistencies in these questionnaire results.

9.4.2 Follow-up Email

Participants were given the choice to opt-in to a follow-up email communication, to take place two weeks after the second laboratory session. Those who opted in were contacted by email and asked whether they had encountered or recalled a situation (since the study) when the experimental file manager might have been useful to them.

9.4.2.1 Follow-up Email Results

All 28 participants agreed to be contacted with the follow-up email question, and 18 (64%) responded. Fifteen of these (54% of total respondents) provided examples of how the file manager could have helped them; three others (11% of total respondents) stated that they had no instances to report.

The comments from the 15 positive responses were coded for dominant themes, in terms of how the experimental file manager was seen as helpful or positive. Five categories emerged in terms of the benefits, such as ease of search. A sixth category was added to record instances where participants explicitly solicited the file manager: that is, expressed a desire for the file manager, such as "Can't wait till it is mainstream!" These categories are presented in Table 72 along with the counts of comments for each category. Most participants made comments in more than one category.

category of positive attribute	count
reduced effort	6
ease of search	6
keeping track of files	6
keeping track of activity	6
interface simplicity	1
expressions of strong desire for file manager	3

Table 72: Categories of positive attributes of the experimental file manager, derived from post-study emails about real world use.

Of the three participants who had no instances to report, two stated that their current job did not involve situations in which the experimental file manager would be helpful. The third stated that no instances were encountered or recalled since the study, but did not specify more information than that.

9.5 *Discussion*

The discussion of the results is presented in three parts, one for each of the evaluation goals: comparison between experimental and control file managers, evaluation of unique functionality of the experimental file manager, and assessment of the intention to use.

9.5.1 Comparison Between Experimental and Control File Managers

The purpose of this part of the evaluation was to determine whether the experimental file manager provided any improvements over a control file manager, in terms of efficiency, effectiveness, or satisfaction. Efficiency was measured using task completion times, effectiveness through question response scores (correctness), and satisfaction through usability scores and preference data. Each of these results is discussed below.

9.5.1.1 Task Completion Times

The testable hypothesis for this variable was as follows:

H1: The experimental file manager will result in task completion times that are faster than those in the control file manager for at least 50% of tasks, and will not result in any slower completion times.

A comparison of the times showed that significant differences were found for tasks in both sessions. In Session One, two of the six tasks (email search tasks) were significantly faster with the experimental file manager. In Session Two, four of the six tasks (email search and checking file permissions) were significantly faster with the experimental file manager. Note that significant differences were detected for the email-based tasks (A and F, and I and O) in both sessions. Participants were able to locate information more quickly using the experimental file manager, using a simple search and without having to switch back to the email application.

Some tasks did not demonstrate a marked improvement. Note that each session contained a pair of uncomplicated file tasks, intended to detect whether the experimental file manager caused any delays. That is, for tasks that were already reasonably efficient in the control, it was necessary to determine whether the experimental file manager made things *slower* for users (for example, by presenting superfluous details that interfered with simple tasks).

In Session One, the uncomplicated task type was checking the "last modified date" (tasks C and H), and in Session Two, it was checking recent file access information (tasks J and P). No significant difference was found for task completion time in any of these cases: neither an increase nor a decrease. This suggests that the experimental file manager did not cause any significant delays in completing the file tasks.

The other tasks that did not demonstrate an improvement were the file permission tasks in Session One (tasks B and G). Note, however, that for similar tasks in Session Two (K and Q), a significant difference *was* found. This is likely due to the experience gained by participants, allowing them to make better use of the tool. Reviewing the screen captures for Session One suggests that some participants forgot about the new functionality of the experimental file manager, and therefore reverted to completing tasks in the familiar way, without using the time-saving features. By the second session, participants were more fluid and comfortable with the new features, and were much more likely to take advantage of them.

In summary, the experimental file manager showed improvements in task completion time for six of the twelve tasks (email search and file permission checks). Thus, the experimental file manager was faster for at least 50% of tasks, when counting

across both sessions. In addition, no significant decreases in completion time were found for any task. This result confirms hypothesis H1.

9.5.1.2 Question Response Scores

The testable hypothesis for this variable was as follows:

H2: The experimental file manager will result in at least 25% of question response scores that are higher than the control.

A subset of the responses demonstrated significance in scoring: four of the twelve questions in both Sessions One and Two (33%). These responses were in two question categories: questions that involved finding email recipients, and questions that requested details about past accesses. The email questions had confidentiality repercussions, and thus a correct response was important. In Task F, for example, the participant was supposed to detect whether a leak had happened (i.e., whether a document had been mistakenly mailed out). This question was answered correctly by all the participants who used the experimental file manager, but by fewer than half (43%) of those who used the control (and thus had to rely on only Outlook for searching.) If the Outlook search was performed incorrectly, or was not exhaustive, this information was missed. Therefore, the experimental file manager was a key tool for avoiding a security-critical mistake.

For the responses that detailed file-access information (Tasks J and P), it was simply not possible for the same amount of detail to be discovered using the control file manager, because it maintains little information to begin with: it stores data on only the most recent access, including the time and whether or not the file was modified. Therefore, those who used the experimental file manager were able to yield a much richer data set (such as usernames) from a longer history. This explains why the experimental condition scores were much higher for these questions.

The remainder of the questions did not show significant differences. Note that for the simpler questions, such as the "yes/no" questions, a difference was unlikely to be found unless the participants got the answer completely wrong. (This is due to the smaller possible distribution of scores: either zero or one.) Because these questions were simple, they were usually answered correctly. Note also that no questions showed a *decrease* in score using the experimental file manager.

In summary, approximately 33% of the questions in each session were answered with a higher level of correctness (accuracy and/or completeness) using the experimental file manager; some of these questions were security-critical, requiring accurate answers. An improvement in 25% of the questions was hypothesized, and the file manager performed better than this at 33%. These results confirm hypothesis H2.

9.5.1.3 Usability Scores

The two testable hypotheses for this variable were as follows:

H3: The experimental file manager will have the majority of the higher scores for the usability questions in Session One or Session Two, or both.

H4: There will be a significant difference in at least one usability score, in Session One, between those who used the control *after* the experimental manager and those who used the control *before* the experimental manager.

The results relating to both these hypotheses are discussed together in this section. In addition to the tests performed to address these two hypotheses, usability scores were also compared between sessions, to test if the ratings of each file manager changed, although this was not one of the central research questions. In all cases, the comparisons were performed on scores from five factors: four CSUQ factors (system use, information quality, interface quality, overall) and ASQ (the After-Scenario Questionnaire)..

The first test compared usability scores between the two file managers, irrespective of order. For this test, significant differences were found for all five usability factors (for both sessions) in favour of the experimental file manager.

The second test compared scores based on the order of the conditions (control versus experimental); significant differences were found, but the results are not straightforward to interpret. What they reveal are significantly lower control scores, for both sessions, when the experimental file manager was used first; this occurred for all five factors in Session One, and four of the five factors in Session Two (all but interface quality). Similarly, there were significantly higher scores for the experimental file manager in those cases where the control was used first, for four of the five factors (all but interface quality). There was no difference in scores for the experimental manager in the remaining case.

These results indicate that the control was rated more poorly after the participant had encountered the experimental file manager. Similarly, those who used the control first tended to rank the experimental manager higher in comparison. The core finding of these results: after people had exposure to the experimental file manager, they found the control file manager less satisfying.

In summary, the experimental file manager had higher usability scores (indicating greater satisfaction) for both sessions, confirming hypothesis H3. Hypothesis H4 was also confirmed: order effects were demonstrated that showed a decrease in satisfaction for the control file manager scores after exposure to the experimental file manager. A related finding (although not a strictly provable hypothesis) was that satisfaction with the control had a tendency to decrease between sessions.

9.5.1.4 User Preference

The testable hypothesis for this variable was as follows:

H5 The experimental file manager will be selected more often than the control for the preference questions, in both Session One and Session Two.

This data was collected through sets of questions asking participants to choose one file manager over the other, based on different attributes. For both sessions, the experimental file manager was the preferred choice, for all questions. In no case was the control preferred by the majority, and the two file managers were rarely shown to be equivalent.

Note that for Session One, the two file managers were rated as equivalent for checking "file modified" dates. Note that no significant difference was expected (or found) for task completion time for this uncomplicated task (as discussed earlier), and therefore the participants' perceptions accurately reflected the equivalency of the two file managers for this task.

For many questions, the experimental file manager was the near-unanimous choice: it was chosen by 100% of participants for three of the 11 questions in Session One, and for seven of the 14 questions in Session Two. For all but two question, the experimental file manager was preferred by at least 75% of participants. (The only exceptions were Q1 and Q3 in Session One.)

The final preference question asked participants to select one file manager as the overall choice. In Session One, three people chose the control; in Session Two, one person chose the control. Participants were asked to specify the reason for their choice. For Session One, in the cases where the control was preferred, participants made the following comments (recall that "Carbon" is the name used for the control file manager, and "Helium" for the experimental):

> Although it [Carbon] provided more work to find info such as...accessed, I found it easier to use because it had less information up-front, but still provided all that I needed within a few clicks of the mouse. Perhaps, also more familiar with this method and could learn to use the other better with time and use.

> Because it [Carbon] is less confusing and more intuitive. Possibly I found Carbon easier because it was the second option, so I warmed up on Helium and was more confident with file sharing once I got to the Carbon part of the test...

> It feels more familiar, but I also think I just got more comfortable and it was the one I used second.

There are two common themes in these comments. First, there is the issue of familiarity: for some, having to learn a new piece of software was seen as detrimental, and thus they preferred to use what they already knew. Second, having time to warm up to the tasks made a difference; for those who needed some time to get used to the computing tasks and the environment, they had a harder time with the software that was used first: in this case, the experimental file manager. The latter cannot be avoided (and was a factor in the experimental design), although the former can be analysed further by looking at the outcome of Session Two.

After Session Two, only one person preferred the control, down from three in the previous session. This suggests that those who felt they would get accustomed to the experimental file manager after more practice were correct. The individual who preferred the control at this stage had also preferred it in Session One, and stated the following (in the Session Two questionnaire):

> More steps to perform [in the control] to access info, but still simpler and therefore, sometimes faster to use.

For this participant, there was a perceived advantage for efficiency with the control file manager (although this was not demonstrated by the task completion time

results). However, this subjective data does indicate that the participant did not believe that there was enough of an advantage provided by the experimental file manager. This was definitely a minority opinion, but it demonstrates the degree to which *perceived usefulness* (in this case, effectiveness) influences technology adoption. This issue will be discussed later in this chapter, when the Technology Acceptance Model questionnaire results are presented.

9.5.1.5 Summary of Comparative Results

The five testable hypotheses, summarized in Table 73, were all confirmed by the experimental results. Significant results were found for all four variables (task times, response scores, usability scores, user preference) demonstrating that the experimental file manager provided improvements in efficiency, effectiveness, and satisfaction.

hypothesis		factor	confirmed?
H1	The experimental file manager will result in task completion times that are faster than those in the control file manager for at least 50% of tasks, and will not result in any slower completion times.	efficiency	**yes**
H2	The experimental file manager will result in at least 25% of question response scores that are higher than the control.	effectiveness	**yes**
H3	The experimental file manager will have the majority of the higher scores for the usability questions in Session One or Session Two, or both.	satisfaction	**yes**
H4	There will be a significant difference in at least one usability score, in Session One, between those who used the control after the experimental manager and those who used the control before the experimental manager.	satisfaction	**yes**
H5	The experimental file manager will be selected more often than the control for the preference questions, in both Session One and Session Two.	satisfaction	**yes**

Table 73: Summary of hypothesis results, for comparisons between experimental and control file managers

9.5.2 Evaluation of New Features and Functionality of Experimental File Manager

In addition to comparisons between the two file managers, additional evaluation was performed for features and functions available only in the experimental file manager. The results were presented above, in five areas: task completion times, question response scores, evaluation of the notification feature, subjective ratings (information display and

ease of learning), and suggestions for improvement. Each of these results is discussed in the sections below.

9.5.2.1 Task Completion Times: Experimental File Manager-Only Tasks

In the absence of comparators, the timing data cannot be formally analysed, but there are a few findings of note. Most of the tasks performed with the experimental file manager were complex, particularly those involving individualized views (e.g., tasks E and M), leading in some instance to long task completion times. Task M was unexpectedly high; a review of the screen capture files indicated that some participants neglected to use the search feature to create the email list. Those who used this approach could still answer the question, but it would take much longer; these infrequent (but large) elapsed times skewed the mean towards the higher end. Similarly, the completion times for tasks D, E, and M were highly variable (as evidenced by the standard deviation); this variability is attributable to the difference in times between those who exploited the new features immediately (e.g., recalled they were available and used them) versus those who took longer to use the features. This problem was more evident in Session One, when the experimental file manager was completely new to the participants. Reviews of the screen captures, along with participant comments, indicate that the features became more familiar over time, and thus more likely to be exploited, which suggests that the tasks completion times would decrease after more experience with the experimental file manager.

9.5.2.2 Question Response Scores: Experimental File Manager-Only Tasks

The majority of questions were answered correctly, for all questions in both sessions. There are two questions for which the correct response rate is substantially lower, and both are in task M: finding all files emailed to a specific person, as well as checking for files accessed by that person after a certain date. As discussed in the task completion time analysis, this task was not always performed efficiently, and the new file manager features were not always used. In such a case, the task time suffered, but so did the accuracy of the task result.

The screen capture files and the talk-after recordings were analysed for the four cases in which the participants gave an incorrect answer. In three of these cases, the search feature in the sharing console was not used at all. Instead, a search was performed

using the Outlook email application, and the correct result was not found. In the fourth case, a search was performed, and the list of matching files with the correct information was generated. However, in this instance, the participant ignored the information and returned to Outlook for searching. When asked about this action in the talk-after, the participant stated that "I guess I wasn't used to using something like this, so I went back to Outlook, and thought I could figure it out that way." This statement could indicate that the person lacked confidence that she was using the tool correctly, or that she did not believe the tool was reliable. The result was that the question was answered incorrectly, despite the availability of correct information. It is likely that with more experience with the new file manager features, these problems would become less prevalent.

There were some other minor errors that occurred when answering these same two questions (M1 and M2): the list of files was either incomplete by one item, or included files from the wrong category: emailed but not accessed, or vice versa. This latter problem was likely due to a misreading of the search results: the results show whether or not the file sharing event was a send or an access, but these differences in distribution method might be overlooked. Making the difference more visually distinct (such as listing the "sent" file matches in one colour and the "accessed" file matches in another) could mitigate this problem.

9.5.2.3 Evaluation of the Notification Feature

The notification feature, implemented as a pop-up alert message was found to be unobtrusive, and could be readily understood by participants. Note that there is a trade-off between the need for interruption (when an important event occurs), and the annoyance of interruption when it is unwanted. The default setting for the prototype window was for it to not grab focus, as this was seen to be overly intrusive. In fact, two participants noted that they were glad that the window did not require them to stop what they were doing to deal with it; thus, this approach was seen as an advantage. However, if an event is critical, then the alert may need to be more obvious, to avoid being lost underneath other windows. This requirement could be accommodated by adding an option to the alert: when the alert is set, the user could decide whether that event was sufficiently important to warrant a "pop-over" window that grabs focus. In non-urgent

cases, the alert could be set to not grab focus (as in the experiment), ensuring it did not interrupt more important tasks.

9.5.2.4 Subjective Ratings: Information Display and Ease of Learning

The scores indicate strong agreement with the positive attributes of the experimental file manager, including ease of locating information, seeing information "at a glance", and non-obtrusiveness of the display. Additionally, participants did not find the information hard to understand, irrelevant, or overwhelming, and did not think that the file manager was difficult to learn to use. This last element—ease of learning—was evidenced by the screen capture reviews; as described above, people were more likely to take advantage of the new features in Session Two, demonstrating that they could more easily recall how to use them (and that these features were available).

Most scores were on the extreme ends of the scale: between one and two, or between six and seven. The only exception was the obtrusiveness score, at 2.25; this is still on the low end of the (seven-point) scale, but indicates that this attribute could perhaps be improved. Participants could add comments to these scores; unfortunately, none of those who felt that the display could be less intrusive gave suggestions.

One avenue for improvement could be providing a more customizable interface. A standardized interface, which displayed the same types of data to all participants, was necessary for the formal testing environment. However, other types of displays can be readily provided, which would allow users to tailor the amount of data shown within the file manager columns. Some people might not wish to view the permission information in certain folders, for instance, or might only want to view the "sent" column. This approach could make the information display less intrusive, for those who want to see fewer details.

9.5.2.5 Suggestions for Improvement

Participants provided some suggestions for improvement, for a variety of features. These changes are all enhancements that could be readily incorporated into a working prototype, and provide useful starting points for a refined version of the design.

One suggestion, for the file manager columns, was to make the icons clickable, so that the sharing console could be opened by clicking on the icon. This is a simple feature to provide, which would make the file manager more convenient to use. A second

suggestion was to change the amount of information displayed in the columns. As discussed earlier, making the columns customizable is the most promising approach. This is a requirement for accommodating diverse preferences; even within this test population, there was no consensus on what the "right" amount of information was for display.

Some participants also wanted the search feature to be more readily accessible. One participant suggested that it become a top-level item in the right-click menu, so that you could directly select "Search" as well as the "Open in Console" option. Allowing the console window to open by clicking on icons (as suggested above) would also alleviate this search feature problem as well, making it quicker to access.

The console could also be improved by incorporating participant suggestions; namely, making the filename, as displayed in the file details panel, more prominent. Making the filename bolder, larger, or in a different colour, could address this problem. Other minor improvements that could be included in the console are making the resize handles more visible (for file details), and to include filetype icons in the file detail window as well. For the alert, some possible design enhancements are adding a disable function or an audible alarm.

A few new features were proposed: one participant suggested integrating the console with the built-in XP permission tools, to facilitate changes. Another suggested including information about incoming files. Both of these features were already in consideration for future, extended versions of the file manager; having other people request these features indicates this is a fruitful area for further exploration.

9.5.3 Intention to Use Experimental File Manager

The assessment of participants' intention to use the experimental file manager was based on the TAM questionnaire results as well as the post-study email comments.

The Technology Acceptance Model proposes that if technology is perceived to be useful and easy to use, then it is more likely to be adopted. Both of these factors had positive scores on the 7-point Likert scale: the overall mean for perceived usefulness was 5.6 (out of 7), and the overall mean for perceived ease-of-use was 6.30. These factors indicate a positive tendency towards adopting the technology. Note that this finding correlates with the self-rated result in X1 (i.e., participants' estimate of whether they would use the technology), which had a mean score of 6.0 (out of 7). In other words,

participants believed they would use the technology, which is in keeping with their positive ratings for usefulness and ease of use. The TAM results suggest that if the new file manager technology were to be introduced, it would be adopted by users who were similar to the study participants.

In the post-study email, the majority of those who responded (15/18) described concrete instances in their own lives for which the new file manager functionality would be helpful. These 15 respondents represent over half the total study population; even under the assumption that the non-respondents had no examples to report (which may not be true), more than half of the participants were able to concretely demonstrate the value of this tool. This result reinforces the Technology Acceptance Model questionnaire findings: participants were attitudinally-inclined to use the technology, and could also show that there were occasions where the technology would be helpful. Therefore, the potential benefits of this new file manager have been demonstrated, for practical application in real-world situations.

9.6 *Summary*

The laboratory evaluation demonstrated that the experimental file manager, with specialized file sharing features, provides tangible benefits. It offers an improvement over the standard file managers provided for Windows and provides new functionality for supporting shared file management. In comparison to a control file manager, the experimental file manager allowed participants to complete a number of file sharing information tasks faster, and with fewer errors. Participants also scored the experimental file manager higher than the control on usability scores, and for overall preference.

The new functionality provided by the experimental file manager was also shown to offer benefits, by allowing users to successfully complete file sharing information tasks that they could not conduct using existing file managers. Participants provided high subjective ratings for the new file manager, finding the information display clear and useful, and considering the tool to be easy to learn.

The results from the Technology Acceptance Model questionnaires indicate that participants would adopt the new file manager technology if it were introduced into their workplace, giving it high ratings for perceived usefulness and perceived ease of use. In

addition, a majority of participants were able to identify a concrete, real-world example of how the experimental file manager could be beneficial.

The evaluation identified areas of strength and weakness in the experimental file manager, and outlined some future directions for research. These future directions are explored in the following chapter, which concludes this thesis. That concluding chapter also includes a summary of this research project, which highlights the contributions and limitations of the work, along with a reflection on the overall thesis research.

Chapter 10. Conclusion

This chapter summarizes the research contributions of this thesis, discusses the limitations of the research, and presents areas for future research.

10.1 *Research Contributions*

This thesis has made contributions in the field of human-computer interaction, particularly within the sub-domain of information management. These contributions have been made in both theoretical and applied areas; these two categories are presented below.

10.1.1 Theoretical Contributions

10.1.1.1 Characterizing File Sharing Practice

At the outset of this research, little information existed about the practice of file sharing, such as the prevalent tools used, the types of file shared, the ways of managing limits, and the groups with whom people share files. The surveys conducted to characterize file sharing filled this gap, providing information about user behaviour and concerns. This data is a resource for other researchers in HCI and information science, who can use this information to guide designs and solutions that address user practice. In particular, this research has been used in designing solutions for improved security, such as usable access control policies (Bauer, Cranor, Reeder, Reiter, & Vaniea, 2008) as well as for collaborative technologies (e.g., Nelson, Smetters, & Churchill, 2008).

10.1.1.2 Creation of a File Sharing Awareness Framework

The Workspace Awareness models (both synchronous and asynchronous) of Gutwin and Greenberg (2002), and Tam and Greenberg (2006) formed the basis for the file sharing awareness framework. These frameworks were adapted for the file sharing domain, which demonstrated how the theoretical model could be applied to other aspects of awareness in computer-supported collaboration.

The awareness framework used to design the file manager provides a theoretical model for other applications in the domain of file sharing. It identifies elements that are key to providing awareness for file sharing: *feedthrough*, *individualized views*, and *notifications*. These elements should be supported in any application that uses file sharing, such as collaboration tools. The framework can be used for new designs, as well

as to identify ways in which current designs could be modified or augmented for awareness support.

10.1.1.3 Creation of Iconography and Terminology for File Sharing Concepts

The design process for the file manager led to a set of icons and terminology for describing file sharing concepts, which were developed, refined, and validated through a user-centred design process. These elements were evaluated in the context of a specific file manager design, and were shown to provide understandable information to users. These terms and icons allowed users to quickly glean information (as evidenced by the improved task completion times), and led to greater comprehension (as evidenced by improved scores on information-finding questions).

Because these items were shown to be understandable to users, other applications that use file sharing (such as online collaboration tools) can include them (or adapt them) in their own interface designs, and have confidence that these elements will clearly communicate file sharing concepts to their users.

10.1.2 Applied Contributions

10.1.2.1 Design of File Manager to Support File Sharing Awareness

The file sharing awareness framework identified elements that were necessary for supporting awareness; these elements needed to be instantiated into a specific design. This design involved the creation of a file manager listing window (with specialized text and icons), as well as a supplementary console window with search features, a file details panel, and an alert feature. All these elements were developed using a meticulous design and evaluation process. This research demonstrated that the resulting design was successful for file management; these elements can be adopted for use in other file sharing applications.

10.1.2.2 Demonstration of the Value of Awareness for File Sharing

The core goal of this research was to facilitate the management of shared files, by supporting user awareness. The file-sharing awareness framework elements (feedthrough, individualized views, and notifications) were instantiated in a specific file manager design. Evaluation of the file manager showed that presenting file sharing information (supporting awareness) led to more effective, efficient, and satisfactory file management. Therefore, the approach of supporting user awareness was shown to be beneficial for

shared file management. This finding can be used by other researchers who wish to improve file sharing technologies (and thus enhance security and collaboration). This research can also be utilized by those working in other areas of awareness research, who may wish to see how awareness theories can be applied in specific computing domains, to provide concrete benefits for users.

10.2 *Research Limitations*

10.2.1 Limitations of File Sharing Surveys

The surveys conducted at the outset of this research were carried out within specific organizations: a research institute and a university. The majority of the survey respondents were very experienced with computers, and many had experience with complex technical tasks. Therefore, these results may not be generalizable to a population with less sophisticated computer experience. However, note that even the technically-adept had difficulties carrying out file sharing correctly, which indicates that this problem are likely to be worse for those with lesser technical skills.

10.2.2 Limitations of Laboratory Evaluation

The experimental file manager was evaluated in a laboratory setting; while this approach allowed for precise measurement of variables, it also had limitations. One of these is that the file manager was assessed over a short period using an artificial file system; participants did not get to use the manager longitudinally, in their own dynamic file systems. There are two issues of concern. The first issue is that file sharing information would grow over time (for example, as more files were created and emailed to others). There may be too much information to display in these circumstances and the interface might become cluttered; the design might need to be adopted to present more streamlined data or to present prioritized information based on user preference. The second issue is that participants did not get a chance to use the file manager in their own personal setting; it was not possible to observe how they might interact with this tool in a natural work environment (rather than the laboratory). Participants might have used the file manager differently in their own digital workspace; they would likely carry out more types of tasks than were used in the evaluation, and would have their own unique file management concerns (e.g., nature of workplace collaboration; organization of files;

priorities for certain types of file sharing information). These elements could not be observed without testing the file manager in a longitudinal field study.

10.2.3 Limitations of Prototype Functionality

The experimental file manager that was evaluated was a high-fidelity prototype, rather than a fully-functional application that was ready for deployment. The scope of the prototype was limited, excluding certain features, to accommodate the practicalities of testing. One feature that was not included was the incorporation of any file sharing tools other than email; information about files that had been sent using other applications (e.g., instant messenger) was not recorded or displayed in the file manager.

In addition, the prototype did not allow for customization, owing primarily to the constraints of the experimental requirements (i.e., so that all participants used an identical tool). This limitation did not allow for the testing of individual user preference; for example, to see whether users wanted to view only certain types of information, or whether they might add or remove file sharing information depending on task.

10.2.4 Limitations of Integration

The experimental file manager functionality was limited to file information tasks (e.g., presenting event history of, and settings on, shared files). Integrated support for making changes was not included: for example, if users looked at the file permission icons, they might notice that the permissions were set incorrectly. The current version of the experimental file manager is not designed to assist users in making the actual update; they need to launch the operating system tools to perform the actual corrections. Thus, there was no assessment of how well the file manager helped people to update the file system (i.e., make changes), because the file manager did not include explicit support for making changes; it was limited to *presenting* file information only.

10.3 *Future Work*

There are a number of avenues for further exploration in this research topic. Some of these address the limitations described above, while others take the research into new areas.

10.3.1 Addressing Limitations: Extending the File Manager

As discussed above, evaluating the file manager in a longitudinal field study would be beneficial. In order for this to occur, the prototype would have to be made fully functional. This would require that information be gathered from various file-sharing tools (including email, instant messenger, web browsers, etc.). This is not a trivial technical challenge, as it requires a series of application plug-ins for data collection, as well as the logging of operating system actions (e.g., reading and writing of files). However, this has been shown to be a feasible approach: a similar technique was applied for information retrieval, using plug-ins to collect document information from multiple applications (Gonçalves & Jorge, 2006). This suggests that converting the prototype to a full application is a practical possibility, which would allow for further evaluation of the file manager in an actual work setting.

Additional features could also be added to the file manager, to address the limitation described above. One such enhancement would be to allow customization of the interface; this was also identified by the study participants as a desirable feature. As well, the file manager needs to be better integrated with other applications, so that users can complete their file management tasks in a more seamless fashion. One key aspect of this integration would be the incorporation of file permission tools, for updating sharing settings. There has been some research on making Windows XP file permission tools more usable; this has included the creation of new interfaces for authoring rules (Maxion & Reeder, 2005) as well as for visualizing access control policies (Reeder et al., 2008). Integrating these usable configuration tools with the file manager would provide advantages for users: they could use the file manager to understand their settings, and exploit the usable file permission tools to configure any necessary updates.

Finally, there were some improvements that were suggested by the study participants. In addition to customization (mentioned above), participants suggested interface enhancements, such as making the file icons clickable and modifying the file details display in the sharing console window (e.g., making the filename more distinctive and adding filetype icons). These improvements could be easily added to the functional file manager application.

10.3.2 Integrating the File Manager with Search Tools

The search features of the file manager are currently standalone; that is, searching using shared file information is not available outside of the sharing console window. It could be beneficial to integrate the manager with other search tools, such as Windows Desktop Search (Microsoft Corporation, 2006) or Google Desktop Search (Google Inc., 2007). This would allow users to conduct searches that included file sharing criteria, when using other types of search tools. For example, if a user were already using Google Desktop Search to look for a file with a certain keyword, which was also shared with Bob, then both these criteria could be included in the Google search. Currently, such file sharing metadata is not supported as a part of desktop search.

10.3.3 Additional Sharing Information for the File Manager

Another way in which the file manager could be extended is by adding information about "incoming" files. The current prototype shows only outgoing files (i.e., items sent to others). It might be helpful to also show files that were sent to a specific user, which could assist them in visualizing their relationship with colleagues. For example, users might wish to identify their most frequent collaborators; this would require the inclusion of both incoming and outgoing activity. File sharing information is a part of more general social relationship information, which is used in many ways, including for personal recall (Viegas, Boyd, Nguyen, Potter, & Donath, 2004) and for visualizing asynchronous collaboration (Heer, Viégas, & Wattenberg, 2007).

Information about received files could also be used to create collections of files from particular users or groups (e.g., "everything sent to me by Human Resources"). This technique could be used for simple information retrieval purposes (e.g., "where did this document come from?"), or to identify related items that, for instance, need to be stored together in an archive, or need to be distributed to another colleague (e.g., "all the documents sent to me by Client X").

10.3.4 Tailoring the File Manager for Specialized Domains

The file manager was designed to incorporate broad sharing practices (as identified in the studies conducted at research and educational institutes). Although this design was shown to be appropriate for the test user population, which included people from a variety of professions, there may be organizational contexts for which the design

would need to be changed to better serve the needs of a specific population. For example, if the file manager was used in a high-security facility (such as a government office that handled classified documents), then different types and amounts of file sharing information might be more appropriate. For example, users might need a specialized icon for a specific user or group (e.g., "contractor") to show who had access, or who a file was sent to. Similarly, a domain in which the focus was on document availability, such as a public archive, might wish to display more detailed information about each of the downloads (such as displaying a long detailed list, including the amount of activity over time), to show how the community was using the public files.

10.4 *Conclusion*

File sharing is a common task within organizations, requiring careful management of access to ensure that a balance is reached between availability of files for teamwork, and the protection of confidential information. Prior to the start of this research, there were outstanding research questions about file sharing, and how this task might be made more effective, particularly in terms of enhancing security and collaboration. The initial phase of the research characterized file sharing practices, revealing the complexity of this task, such as the multiple sharing methods used and the requirements to keep track of shared files in dynamic environments. These factors were shown to impair security and collaboration; one cause of this was a lack of awareness about the status of shared files.

This issue was addressed through the creation of an awareness framework for file sharing, which was used to design a file manager that presented information about shared files, through the methods of feedthrough, individualized views, and notifications. These elements were shown to improve the management of shared files. Overall, the awareness information in the file manager led to greater effectiveness, efficiency, and satisfaction for file sharing tasks: users made fewer security-critical errors, and completed their collaborative tasks faster and with greater satisfaction.

From this research has emerged a better understanding of the complexity of file sharing practices within organizations, and a framework for driving application design for file sharing. In addition, this research resulted in a standard set of icons and terminology to express file sharing concepts and the design for a new file manager informed by the awareness framework. Finally, this research found that presenting clear, readily-

accessible information about shared files will enable collaboration while at the same time offering a better level of security.

References

37signals LLC. (2005). Project collaboration, management, and task software: Basecamp. Retrieved April 30, 2007, from http://www.basecamphq.com/

Adams, A., & Sasse, M. A. (1999). Users are not the enemy. *Communications of the ACM, 42*(12), 40-46.

Adams, M. J., Tenney, Y. J., & Pew, R. W. (1995). Situation awareness and the cognitive management of complex systems. *Human factors, 37*(1), 85-104.

Allison, B., Hawley, R., Borr, A., Muhlestein, M., & Hitz, D. (1998). *File System Security: Secure Network Data Sharing for NT and Unix.* Paper presented at the Large Installation System Administration of Windows NT Conference, Seattle, WA.

Appelt, W. (1999). WWW Based Collaboration with the BSCW System. *Proceedings of SOFSEM, 99*, 66-78.

Baker, K., Greenberg, S., & Gutwin, C. (2001). Heuristic Evaluation of Groupware Based on the Mechanics of Collaboration. *Proceedings of the 8th IFIP International Conference on Engineering for Human-Computer Interaction*, 123-140.

Bauer, L., Cranor, L. F., Reeder, R. W., Reiter, M. K., & Vaniea, K. (2008). A user study of policy creation in a flexible access-control system. *Conference on Human Factors in Computing Systems*, 543-552.

Bellotti, V. (1996). What You Don't Know Can Hurt You: Privacy in Collaborative Computing (pp. 241-261): Springer-Verlag London, UK.

Bentley, R., Appelt, W., Busbach, U., Hinrichs, E., Kerr, D., Sikkel, K., et al. (1997). Basic support for cooperative work on the World Wide Web. *International Journal of Human Computer Studies, 46*(6), 827-846.

Berlage, T., & Sohlenkamp, M. (1999). Visualizing Common Artefacts to Support Awareness in Computer-Mediated Cooperation. *Computer Supported Cooperative Work (CSCW), 8*(3), 207-238.

Borlund, P. (2003). The IIR evaluation model: a framework for evaluation of interactive information retrieval systems. *Information Research, 8*(3), 1–38.

Boulton, M. (2006). Card Sorting - Part One. Retrieved June 16, 2007, from http://www.markboulton.co.uk/articles/detail/card_sorting_part_1/

Cao, X., & Iverson, L. (2006). Intentional access management: making access control usable for end-users. *Proceedings of the second symposium on Usable privacy and security*, 20-31.

Carvalho, V. R., & Cohen, W. W. (2007). Preventing Information Leaks in Email. *Proceedings of SIAM International Conference on Data Mining (SDM-07), Minneapolis, MN.*

Chernov, S., Demartini, G., Herder, E., Kopycki, M., & Nejdl, W. (2008). *Evaluating Personal Information Management Using an Activity Logs Enriched Desktop Dataset.* Paper presented at the Personal Information Management Workshop, Florence, Italy.

Cohen, W. W. (2004). Enron Email Dataset. Retrieved September 6, 2007, from http://www.cs.cmu.edu/~enron/

Constantine, L., & Lockwood, L. (1997). Software Usability Defects Log, Version 3.1. Retrieved June 20, 2007, from http://www.foruse.com/publications/templates/defectsoft.pdf

Cranor, L. F., & Garfinkel, S. (Eds.). (2005). *Security And Usability: Designing Secure Systems That People Can Use.* Sebastopol, CA: O'Reilly.

Cutrell, E., Robbins, D., Dumais, S., & Sarin, R. (2006). *Fast, flexible filtering with PHLAT.* Paper presented at the Conference on Human Factors in Computing Systems (CHI).

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319-340.

de Paula, R., Ding, X., Dourish, P., Nies, K., Pillet, B., Redmiles, D., et al. (2005). *Two experiences designing for effective security.* Paper presented at the Symposium on Usable Privacy and Security.

DiGioia, P., & Dourish, P. (2005). Social navigation as a model for usable security. *Symposium on Usable Privacy and Security*, 101-108.

Dix, A., Finlay, J., Abowd, G., & Beale, R. (2003). *Human-Computer Interaction* (3rd ed.): Prentice Hall.

Dourish, P., & Chalmers, M. (1994). Running Out of Space: Models of Information Navigation. *Adjunct Proceedings (Short Papers) of HCI: People and Computers XI*

Duchenaut, N., & Bellotti, V. (2001). Email as habitat. *Interactions, 8*(5), 30-38.

Element Software Pty Ltd. (2003). Copper Project - Collaboration, Projects, Tasks, Contacts and File Management Software. Retrieved June 4, 2007, from http://www.copperproject.com/

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems: Situation awareness. *Human factors, 37*(1), 32-64.

Erickson, T. (2006). From PIM to GIM: Personal Information Management in group contexts. *Communications of the ACM, 49*(1), 74-75.

Erickson, T., & Kellogg, W. A. (2000). Social Translucence: An Approach to Designing Systems that Support Social Processes. *Transactions on Computer Human Interface, 7*(1), 59-83.

Faulker, L. (2003). Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. *Behavior Research Methods, Instruments, & Computers, 35*(3), 379-383.

Field, A. P. (2005). *Discovering Statistics Using SPSS*: Sage Publications Inc.

file management. (2006). *High Definition: A-Z Guide to Personal Technology.*

Fitzpatrick, G. (2003). *The Locales Framework: Understanding and Designing for Wicked Problems*: Kluwer Academic Publishers.

Fontana, J. (2007). Microsoft adding more users than IBM in e-mail battle report shows. *Network World, 24*(5), 12-12.

Freedman, A. (2000). *The Computer Glossary: The Complete Illustrated Dictionary* (Ninth ed.). New York: AMACOM.

Fuchs, L., Pankoke-Babatz, U., & Prinz, W. (1995). Supporting Cooperative Awareness with Local Event Mechanisms: The GroupDesk System. *Proceedings of the 1995 European Conference on Computer-Supported Cooperative Work (ECSCW'95)*, 247-263.

Gaines, J. (2006). Jeff's File Manager. Retrieved May 2, 2007, from http://www.jgaines.eu/html/jeff_s_file_manager.html

Gal, T., Singer, H. M., & Popkin, L. (2003). The IP war: apocalypse or revolution? *Proceedings of the 2003 ACM workshop on Digital rights management*, 39-46.

Garfinkel, S. L., & Miller, R. C. (2005). *Johnny 2: a user test of key continuity management with S/MIME and Outlook Express*. Paper presented at the Symposium on Usable Security and Privacy (SOUPS).

Gaw, S., Felten, E. W., & Fernandez-Kelly, P. (2006). *Secrecy, flagging, and paranoia: adoption criteria in encrypted email.* Paper presented at the SIGCHI conference on Human Factors in computing systems.

Gonçalves, D., & Jorge, J. A. (2006). Quill: a narrative-based interface for personal document retrieval. *Conference on Human Factors in Computing Systems*, 327-332.

Good, N. S., & Krekelberg, A. (2003). *Usability and privacy: a study of Kazaa P2P file-sharing.* Paper presented at the Conference on Human Factors in Computing Systems (CHI).

Google Inc. (2007). Google Desktop Search. from http://desktop.google.com/

Greenberg, S., Fitzpatrick, G., Gutwin, C., & Kaplan, S. (1999). Adapting the Locales Framework for Heuristic Evaluation of Groupware.

Grinter, R. E., & Smetters, D. K. (2003). Three Challenges for Embedding Security into Applications. *Proceedings of CHI 2003 Workshop on HCI and Security Systems* Retrieved June 14, 2007, from http://www.andrewpatrick.ca/CHI2003/HCISEC/hcisec-workshop-grinter.pdf

Groove Networks. (2005, March 30, 2005). User's Guide for Groove Virtual Office 3.1. Retrieved September 5, 2006, 2006

Gutwin, C., & Greenberg, S. (2002). A Descriptive Framework of Workspace Awareness for Real-Time Groupware. *Computer Supported Cooperative Work (CSCW), 11*(3), 411-446.

Gutwin, C., & Greenberg, S. (2004). The Importance of Awareness for Team Cognition in Distributed Collaboration. *Team Cognition: Understanding the Factors that Drive Process and Performance, 201.*

Heer, J., Viégas, F. B., & Wattenberg, M. (2007). Voyagers and voyeurs: supporting asynchronous collaborative information visualization. *Proceedings of the SIGCHI conference on Human factors in computing systems*, 1029-1038.

Horton, W. (1994). *The Icon Book: Visual Symbols for Computer Systems and Documentation*. New York: Wiley.

Hourizi, R., & Johnson, P. (2004). Designing to support awareness: a predictive, composite model. *Proceedings of the 2004 conference on Human factors in computing systems*, 159-166.

information management. Retrieved August 15, 2006, from http://pim.ischool.washington.edu/final%20PIM%20report.pdf

Jones, W. (2005). Personal Information Management. IS Technical Report IS-TR-2005-11-01. Retrieved June 19, 2007, from http://hdl.handle.net/1773/2155

Jones, W., & Bruce, H. (2005). A Report on the NSF-Sponsored Workshop on Personal

Kaufman, J. (2006, November). Card Sorting: An Inexpensive and Practical Usability Technique. *infocom: The Magazine of the Society for Technical Communication, 53,* 17-19.

Kentaro, G., & John, M. C. (2004). The blind men and the elephant: views of scenario-based system design. *interactions, 11*(6), 44-53.

King, W. R., & He, J. (2006). A meta-analysis of the technology acceptance model. *Information & Management, 43*(6), 740-755.

Kipfer, B. A., & Chapman, R. L. (Eds.). (2001). *Roget's International Thesaurus* (Sixth ed.). Toronto: Harper Collins Canada Limited.

Klimt, B., & Yang, Y. (2004). Introducing the Enron corpus. *First Conference on Email and Anti-Spam (CEAS).*

Laitenberger, O., & Dreyer, H. M. (1998). Evaluating the Usefulness and the Ease of Use of a Web-based Inspection Data Collection Tool. *Proceedings of the Fifth International Software Metrics Symposium*, 122-132.

Legout, A., Liogkas, N., Kohler, E., & Zhang, L. (2007). Clustering and sharing incentives in BitTorrent systems. *Proceedings of the 2007 ACM SIGMETRICS international conference on Measurement and modeling of computer systems*, 301-312.

Lewis, J. R. (1991). Psychometric evaluation of an after-scenario questionnaire for computer usability studies: the ASQ. *ACM SIGCHI Bulletin, 23*(1), 78-81.

Lewis, J. R. (1995). IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction, 7*(1), 57-78.

Lin, C. H., Shih, H. Y., & Sher, P. J. (2007). Integrating Technology Readiness into Technology Acceptance: The TRAM Model. *Psychology and Marketing, 24*(7), 641-657.

Ma, R. T. B., Lee, S. C. M., Lui, J. C. S., & Yau, D. K. Y. (2004). A game theoretic approach to provide incentive and service differentiation in P2P networks. *Proceedings of the joint international conference on Measurement and modeling of computer systems*, 189-198.

Marchionini, G. (2006). Exploratory search: from finding to understanding. *Communications of the ACM, 49*(4), 41-46.

Maurer, D., & Warfel, T. (2004). Card sorting: a definitive guide. Retrieved June 15, 2007, from http://www.boxesandarrows.com/view/card_sorting_a_definitive_guide

Maxion, R. A., & Reeder, R. W. (2005). Improving user-interface dependability through mitigation of human error. *International Journal of Human-Computer Studies, 63*(1-2), 25-50.

McGrath, J. E. (1995). Methodology matters: doing research in the behavioral and social sciences. In (pp. 152-169): Morgan Kaufmann Publishers Inc. San Francisco, CA, USA.

McWilliams, B. (2002). Microsoft Spills Customer Data [Electronic Version]. *Wired News*. Retrieved September 1, 2006, from http://www.wired.com/news/infostructure/0,1377,56481,00.html

Microsoft Corporation. (2006). Windows Desktop Search. Retrieved July 30, 2006, 2006, from http://desktop.msn.com/

Microsoft Corporation. (2008). Tooltips and Infotips. Retrieved March 5, 2008, from http://msdn2.microsoft.com/en-us/library/aa511495.aspx

Montemayor, J., Freeman, A., Gersh, J., Llanso, T., & Patrone, D. (2006). Information Visualization for Rule-based Resource Access Control. *Poster: Symposium on Usable Privacy and Security (SOUPS)* Retrieved May 21, 2007, from http://cups.cs.cmu.edu/soups/2006/posters/montemayor-poster_abstract.pdf

Neisser, U. (1976). *Cognition and reality: principles and implications of cognitive psychology*: WH Freeman.

Nelson, L., Smetters, D., & Churchill, E., F. (2008). Keyholes: selective sharing in close collaboration. *CHI '08 extended abstracts on Human factors in computing systems*, 2443-2452.

Net Applications. (2007). Market share for browsers, operating systems and search engines: Operating system market share for May 2007. Retrieved May 15, 2007, from http://marketshare.hitslink.com/report.aspx?qprid=10&qpmr=24&qpdt=1&qpct=3&qpcal=1&qptimeframe=M&qpsp=100

Neustaedter, C., Brush, A. J. B., & Smith, M. A. (2005). *Beyond "from" and "received": exploring the dynamics of email triage.* Paper presented at the Conference on Human Factors in Computing Systems.

Neustaedter, C., Brush, A. J. B., Smith, M. A., & Fisher, D. (2005). The Social Network and Relationship Finder: Social Sorting for Email Triage. *Proceedings of the 2005 Conference on Email and Anti-Spam (CEAS).*

NextPage Incorporated. (2006). NextPage 2 Document Collaboration End User Guide. Retrieved September 5, 2006, 2006, from http://www.nextpage.com/support/collab/doc/index.htm

NIST. (2006). NIST Glossary of Key Information Security Terms. from http://csrc.nist.gov/publications/nistir/NISTIR-7298_Glossary_Key_Infor_Security_Terms.pdf

Norman, D. A. (1993). *Things that make us smart: defending human attributes in the age of the machine*: Addison Wesley, Reading Mass.

ObjectPlanet. (2007). Opinio Survey Software. Retrieved June 10, 2007, from http://www.objectplanet.com/opinio/

Olsen, D. S. (1999). Interacting in Chaos. *interactions*(September/October), 43-54.

Pew Internet & American Life Project. (2006, January 22, 2006). Generations Online. Retrieved September 22, 2006, 2006, from http://www.pewinternet.org/PPF/r/170/report_display.asp

Plaue, C., Miller, T., & Stasko, J. (2004). Is a picture worth a thousand words?: an evaluation of information awareness displays. *Proceedings of the 2004 conference on Graphics interface*, 117-126.

Reeder, R. W., Bauer, L., Cranor, L. F., Reiter, M. K., Bacon, K., How, K., et al. (2008). Expandable grids for visualizing and authoring computer security policies.

Renaud, K. (2005). Evaluating Authentication Mechanisms. In L. F. Cranor & S. Garfinkel (Eds.), *Security and Usability: Designing Secure Systems That People Can Use* (pp. 103–128). Sebastopol, CA: O'Reilly.

Robertson, T. (2002). The Public Availability of Actions and Artefacts. *Computer Supported Cooperative Work (CSCW), 11*(3), 299-316.

Roca, J. C., Chiu, C. M., & Martínez, F. J. (2006). Understanding e-learning continuance intention: An extension of the Technology Acceptance Model. *International Journal of Human-Computer Studies, 64*(8), 683-696.

Rode, J., Johansson, C., DiGioia, P., Silva Filho, R., Nies, K., Nguyen, D. H., et al. (2006). Seeing further: extending visualization as a basis for usable security. *Symposium on Usable Privacy and Security*, 145-155.

Rosenfeld, L., & Morville, P. (2002). *Information Architecture for the World Wide Web* (Second ed.): O'Reilly.

Rosson, M. B., & Carroll, J. M. (2002). *Usability Engineering: Scenario-Based Development of Human-Computer Interaction*: Morgan Kaufmann.

Sandhu, R. S., & Samarati, P. (1994). Access control: principle and practice. *Communications Magazine, IEEE, 32*(9), 40-48.

Savage, C. (2004). Infiltration of files seen as extensive: Senate panel's GOP staff pried on Democrats [Electronic Version]. *Boston Globe*. Retrieved September 1, 2006, from http://www.boston.com/news/nation/articles/2004/01/22/infiltration_of_files_seen_as_extensive/

Schechter, S. E., Dhamija, R., Ozment, A., & Fischer, I. (2007). *The Emperor's New Security Indicators*. Paper presented at the 2007 IEEE Symposium on Security and Privacy.

Schmidt, K. (2002). The Problem with "Awareness": Introductory Remarks on "Awareness in CSCW" *Computer Supported Cooperative Work 11* (3), 285-298

Segal, L. D. (1994). *Effects of Checklist Interface on Non-verbal Crew Communications*: National Aeronautics and Space Administration; National Technical Information Serviceo. Document Number)

Shetty, J., & Adibi, J. (2005). Discovering important nodes through graph entropy: the case of Enron email database. *Proceedings of the 3rd international workshop on Link discovery*, 74-81.

Sivo, S., & Pan, C. (2005). Undergraduate engineering and psychology students' use of a course management system: A factorial invariance study of user characteristics and attitudes. *The Journal of Technology Studies, 31*(2), 94–103.

Smetters, D. K., & Grinter, R. E. (2002). *Moving from the design of usable security technologies to the design of useful secure applications.* Paper presented at the 2002 New Security Paradigms Workshop, Virginia Beach, Virginia.

Stallings, W. (2006). *Cryptography and Network Security: Principles and Practice* (4th ed.): New Jersey: Prentice-Hall.

Summers, E., & Holmes, A. (Eds.). (2004). *Collins Canadian English Dictionary & Thesaurus*. Toronto: Harper Collins Canada Limited.

Tam, J. (2002). *Supporting Change Awareness in Visual Workspaces.* Unpublished MSc Thesis, University of Calgary, Alberta.

Tam, J., & Greenberg, S. (2006). A framework for asynchronous change awareness in collaborative documents and workspaces. *International Journal of Human-Computer Studies, 64*(7), 583-598.

Tang, J. C., Lin, J., Pierce, J., Whittaker, S., & Drews, C. (2007). *Recent shortcuts: using recent interactions to support shared activities.* Paper presented at the SIGCHI conference on Human factors in computing systems.

Techsmith Corporation. (2007). Morae Usability Testing Solution for Web Sites and Software. Retrieved October 11, 2007, from http://www.techsmith.com/morae.asp

Van Fleet, D. D., & Atwater, L. (1997). Gender Neutral Names: Don't Be So Sure! *Sex Roles, 37*(1), 111-123.

Viegas, F. B., Boyd, D., Nguyen, D. H., Potter, J., & Donath, J. (2004). Digital artifacts for remembering and storytelling: posthistory and social network fragments. *System Sciences, 2004. Proceedings of the 37th Annual Hawaii International Conference on*, 109-118.

Voida, S., Edwards, W. K., Newman, M. W., Grinter, R. E., & Ducheneaut, N. (2006). *Share and share alike: exploring the user interface affordances of file sharing.* Paper presented at the Conference on Human Factors in Computing Systems.

Webreakstuff. (2007). Goplan - Online project management and collaboration tool. Retrieved June 27, 2007, from http://goplan.info/

Whalen, T., & Inkpen, K. (2004). Techniques for Visual Feedback of Security State. *DIMACS Workshop on Usable Privacy and Security Software* Retrieved June 4, 2007, from dimacs.rutgers.edu/Workshops/Tools/abstract-whalen.pdf

Whalen, T., & Inkpen, K. M. (2005). *Gathering evidence: use of visual security cues in web browsers.* Paper presented at the 2005 Graphics Interface conference.

Whalen, T., Smetters, D., & Churchill, E. F. (2006). *User experiences with sharing and access control.* Paper presented at the CHI '06 extended abstracts on Human factors in computing systems.

Whitten, A., & Tygar, J. D. (1999). *Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0.* Paper presented at the Usenix Security Symposium.

Wisneski, C., Ishii, H., Dahley, A., Gorbet, M., Brave, S., Ullmer, B., et al. (1998). Ambient Displays: Turning Architectural Space into an Interface between People and Digital Information. *Proceedings of the First International Workshop on Cooperative Buildings, Integrating Information, Organization, and Architecture*, 22-32.

Wolff, F., De Angeli, A., & Romary, L. (1998). Acting on a Visual World: The Role of Perception in Multimodal HCI. *Proceedings of AAAI Workshop on Multimodal Representation.*

Wu, M., Miller, R. C., & Garfinkel, S. L. (2006). Do security toolbars actually prevent phishing attacks? *CHI*, 601-610.

Wu, M., Miller, R. C., & Little, G. (2006). *Web wallet: preventing phishing attacks by revealing user intentions.* Paper presented at the second Symposium on Usable Privacy and Security.

Yan, J., Blackwell, A., Anderson, R., & Grant, A. (2005). The Memorability and Security of Passwords. In L. F. Cranor & S. Garfinkel (Eds.), *Security and Usability: Designing Secure Systems that People Can Use* (pp. 129-142). Sebastopol, CA: O'Reilly.

Young, R. M., & Barnard, P. (1987). *The use of scenarios in human-computer interaction research: turbocharging the tortoise of cumulative science.* Paper presented at the Proceedings of the SIGCHI/GI conference on Human factors in computing systems and graphics interface.

Appendix A: Questions, Survey One

The following questions appeared in the first survey conducted for this thesis research, as described in Chapter 3. This was a wide-ranging, exploratory survey, which included many questions on general security concerns and security management, and the use of computers in a mixed home/work context. As this research was refined within the file-sharing domain, only a *subset* of these questions remained relevant. The questions whose responses were included in this thesis data are marked with an asterisk.

Survey Questions

Your Background

*What is your gender?

- O Female
- O Male

What is your age? Select from a range:

O under 25 O 25-35 O 36-45 O 46-55 O 56-65 O over 65

*What is your current job? Choose from the list below, or add your title under "Other".

- O researcher
- O manager
- O administrative support
- O technical support
- O intern or student

Other job title/category: _____

How long have you been in your current career?

O Less than 5 years O 5-10 years O 11-20 years O More than 20 years

*How long have you been using computers?

O Less than 5 years O 5-10 years O 11-20 years O More than 20 years

Your Computer Tasks

*Select the computer activities that you do within a typical week. These include activities you may do for work (either at the office or at home), or for personal purposes. For each activity that you do, check off as many categories as apply. Remember that you can tell us in confidence about personal activities that you do at work; your responses are private and we will not report your activities.

	personal: at home	personal: at office	work: at home	work: at office
reading and composing email	☐	☐	☐	☐
writing code (programming)	☐	☐	☐	☐
word processing	☐	☐	☐	☐
playing games	☐	☐	☐	☐
creating web pages	☐	☐	☐	☐
reading research papers	☐	☐	☐	☐
using mathematical software (e.g., for research data)	☐	☐	☐	☐
using spreadsheets (e.g., for budgets)	☐	☐	☐	☐
using financial software (e.g., Quicken)	☐	☐	☐	☐
managing personal records (e.g., tax returns, personnel files, health insurance data)	☐	☐	☐	☐
installing software	☐	☐	☐	☐
desktop videoconferencing	☐	☐	☐	☐
instant messaging	☐	☐	☐	☐
system maintenance (e.g., patching, configuring software/hardware)	☐	☐	☐	☐
editing/managing photos	☐	☐	☐	☐
listening to/managing music (downloading, making CDs, etc)	☐	☐	☐	☐
Internet banking	☐	☐	☐	☐
Internet shopping	☐	☐	☐	☐
reading news, sports, or financial info on web	☐	☐	☐	☐

(Optional) Other computer activities you do weekly (specify task as well as the context, such as work/home):

Computers at Home and Work

How many desktop computers do you regularly use at work?

O 0 O 1 O 2 O 3 O 4-6 O 7-10 O more than 10

How many desktop computers do you regularly use at home?

O 0 O 1 O 2 O 3 O 4-6 O 7-10 O more than 10

How many laptops do you use, that you move from one location to another?

O 0 O 1 O 2 O 3 O 4-6 O 7-10 O more than 10

Do you move any of your computers between home and the office?

OYes O No

(Optional) If you regularly use computers at another location, please specify this location (e.g., library):

Are any of your computers used for both personal and work tasks?

OYes O No

Where are the computers that you use for both personal and work tasks? Check as many as apply:

☐ multiple locations (laptop) ☐ home ☐ work (office)

If you share any of your computers regularly, with whom do you share them (e.g., spouse, co-worker, children)? Please specify in the blanks below, for each location in which you have a computer:

Home _____

Office _____

Other (also specify location, or "laptop") : _____

Who takes the main responsibility for administering your computer(s) (e.g., yourself, computer support, spouse)? Please specify in the blanks below, for each location in which you have computer(s):

Home _____

Office _____

Other (also specify location, or "laptop") : _____

Files You Share

*What types of files do you share with people (e.g., photos, documents, music)? Check as many as apply. You may add types under "other", below:

- ☐ photos or graphics
- ☐ written documents (e.g., Word files, PDF)
- ☐ audio (music)
- ☐ video
- ☐ spreadsheets/numeric data
- ☐ source code
- ☐ slides (PowerPoint)
- ☐ none

*(Optional) Are there any other types of files that you share? Please specify:

File Sharing Methods

In the following questions, we will ask about methods you use to share files, using a list. Here is a more detailed list for reference, with examples to make the questions clearer:

- Portable hard drives/tokens (e.g., USB flash drive, iPod)
- Writable disks (CD-RW, DVD-RW, floppy)
- Direct connection between machines (Firewire, Ethernet, USB cable)
- Email attachments
- Web transfer (over http)
- Other network transfer (ftp, scp)
- Network file sharing (shared folders, Appleshare, Windows shared folders, NFS)
- Infrared
- Synchronization software (rsync, ghost, ActiveSync, unison)
- Peer-to-peer networks (Grokster, Bittorrent, KaZaa)
- Docushare

In this section we are focusing on digital documents. If you regularly make paper copies of electronic documents and send them by mail or fax, you can list these sharing methods in the boxes marked "other", below.

*How do you share files with work colleagues? For each of the situations below, select all the methods that apply. (For examples, refer to list above.)

	portable drive	disks	over a cable	attach to email	web (over http)	other network transfer (ftp)	network file share	infrared	sync software	peer to peer network	Docushare
With work colleagues in the same location	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
With work colleagues in a remote location	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
With work associates (clients, partners) who work in a different organization	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

*(Optional) Other ways that you share work files with colleagues:

* How do you share personal (i.e., non-work) files? For each of the recipients below, select all the methods that apply:

	portable drive	disks	over a cable	attach to email	web (over http)	other network transfer (ftp)	network file share	infrared	sync software	peer to peer network	Docushare
spouse/partner	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
children	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
family	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
close friends	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Acquaintances (e.g., online interest group)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Anybody online (may not know them)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

* (Optional) Other methods you use to share personal files, and the people/groups with whom you share them:

* When sharing work files with colleagues, what are the 3 ways you most often use to share them? (Choose a maximum of 3, including "other"; you may choose fewer than 3.)

- ☐ Portable hard drives/tokens (e.g., USB flash drive, iPod)
- ☐ Writable disks (CD-RW, DVD-RW, floppy)
- ☐ Direct connection between machines (Firewire, Ethernet, USB cable)
- ☐ Email attachments
- ☐ Web transfer (over http)
- ☐ Other network transfer (ftp, scp)
- ☐ Network file sharing (shared folders, Appleshare, Windows shared folders, NFS)
- ☐ Infrared
- ☐ Synchronization software (rsync, ghost, ActiveSync, unison)
- ☐ Peer-to-peer networks (Grokster, Bittorrent, KaZaa)
- ☐ Docushare
- ☐ Other:

* When sharing personal files with others, what are the 3 ways you most often use to share them? (Choose a maximum of 3, including "other"; you may choose fewer than 3.)

- ☐ Portable hard drives/tokens (e.g., USB flash drive, iPod)
- ☐ Writable disks (CD-RW, DVD-RW, floppy)
- ☐ Direct connection between machines (Firewire, Ethernet, USB cable)
- ☐ Email attachments
- ☐ Web transfer (over http)
- ☐ Other network transfer (ftp, scp)
- ☐ Network file sharing (shared folders, Appleshare, Windows shared folders, NFS)
- ☐ Infrared
- ☐ Synchronization software (rsync, ghost, ActiveSync, unison)
- ☐ Peer-to-peer networks (Grokster, Bittorrent, KaZaa)
- ☐ Docushare
- ☐ Other:

When sharing files between your own computers--such as your laptop and desktop-- what are the 3 ways you most often use to move them? (Choose a maximum of 3, including "other"; you may choose fewer than 3.)

- ☐ Portable hard drives/tokens (e.g., USB flash drive, iPod)
- ☐ Writable disks (CD-RW, DVD-RW, floppy)
- ☐ Direct connection between machines (Firewire, Ethernet, USB cable)
- ☐ Email attachments
- ☐ Web transfer (over http)
- ☐ Other network transfer (ftp, scp)
- ☐ Network file sharing (shared folders, Appleshare, Windows shared folders, NFS)
- ☐ Infrared
- ☐ Synchronization software (rsync, ghost, ActiveSync, unison)
- ☐ Peer-to-peer networks (Grokster, Bittorrent, KaZaa)
- ☐ Docushare
- ☐ Other:

File Sharing Management

Do you share entertainment files (music, video) with other people over peer-to-peer networks, such as Grokster, Bit Torrent or KaZaa? This includes both uploading and downloading files.

○Yes ○ No

If yes, do you

- ○ upload only
- ○ download only
- ○ both upload and download

When you store files to be shared over a network, where do you store them? Under each column for "work" and "personal", check all the locations that apply.

Remember that your responses are private; if you break or bend corporate security policy to make it easier to manage files and do your job, you can tell us this in confidence.

	work files	personal files (non-work)
On my own computers that I manage myself	☐	☐
On an internal corporate-run network server (i.e., can only be accessed locally within the organization)	☐	☐
On an external corporate-run network server (i.e., can be accessed remotely outside the organization)	☐	☐
On another external site (e.g., commercial/ISP web space)	☐	☐

Other locations where you store shared files; please indicate whether these are work or non-work files:

* Have you ever removed your files from a shared space? For example: taken them off KaZaa, or deleted files from shared folders.

O Yes O No

*If yes, what were the circumstances?

*Have you ever reset permissions on your shared files to disallow sharing?

O Yes O No

*If yes, what were the circumstances?

Do you have both work and personal files on the same computer?

- O No - never
- O Yes - occasionally
- O Yes - frequently
- O I don't make a distinction between work and personal files

If you answered "yes" above: how do you segregate work and personal files? (You may respond "I don't" if you do not keep them separate.)

Managing Sharing at Work

Do other people edit files that you share with them (such as text documents)?

OYes O No

If so, which of the following editing methods do they use? Check as many as apply:

- ☐ make changes to their own copy of the document, then re-share it (e.g., under a new filename or resend as attachment)
- ☐ edit your copy (modify original)

Do you use any of the following methods to handle conflicts when multiple people make changes to a document? Check as many as apply:

- ☐ arrange turn taking (through email or conversation)
- ☐ automated change tracking system (e.g., as in Word)
- ☐ automated version control system (e.g., CVS)

(Optional) Other method(s) that you use to manage editing with multiple people:

Do you use a remote desktop (e.g., Windows Remote Desktop, VNC, other)?

OYes O No

If so, what machines do you access via remote desktop? (For example: your own computers, others' computers when you are helping them with technical problems, shared machines, etc). Please specify:

Do you maintain separate email addresses for work and personal matters?

OYes O No

If so, what are your reasons?

If not, how do you segregate work and personal email (if at all)?

Do you use a non-corporate email address for work email in order to make your job easier in some way (e.g., easier to get access to this email when travelling)?

OYes O No

If yes, please describe why and how you use this additional email address:

Sharing in Everyday Life

We are now going to ask some questions outside of the computer domain: this next section deals with personal possessions.

What was the last time you shared your car key with someone? Please describe the circumstances (the person, the reason, and whether or not you asked for the key back at the end):

What was the last time you shared your house key with someone? Please describe the circumstances (the person, the reason, and whether or not you asked for the key back at the end):

Who are the people who have keys to your house?

Can you lock your desk at work?

○Yes ○ No ○ I don't know

If so, do you lock it, and why or why not?

Do you use a shredder?

○Yes ○ No

If so, do you use it at one or both of the following:
☐ home ☐ office

What sorts of documents do you shred, if any?

Do you have paper files in your office at work that you control access to? (This includes simple actions such as keeping certain files out of sight, and stronger methods such as locking drawers.) If so, what kinds of documents, and how do you control access?

Do you have paper files in your home that you control access to? (This includes simple actions such as keeping certain files out of sight, and stronger methods such as locking drawers.) If so, are they work or personal? What type are they (e.g., tax return), and how do you control access?

Passwords and Privacy

Do you share your computer passwords with other people?

OYes O No

If yes, please describe the circumstances and reasons:

Do you share your ATM PIN number with anyone? If so, with whom and why?

Do you have a cell phone or PDA?

OYes O No

Is there a password (or PIN) on your cell phone or PDA?

- O Yes - it was put there already (e.g., automatically or by the vendor)
- O Yes - I created one or changed the default
- O No - there is no password
- O I do not know

If there is a password: What sort of information do you keep on your PDA or cell phone that you wish to control access to (if any)?

If there is no password: Why have you not added a password?

Consider the people who have legitimate access to your computer and/or files, such as colleagues with whom you share folders or family members who share your laptop.

* Do you have items on your computers that you do not want these people to have access to? If so, please list, including both work and personal items:

* What do you do to ensure that people do not have access to these items?

If you accidentally left your laptop in a restaurant, what would you be most worried about people having access to?

Permissions and Access Control

* Do you ever set permissions on files (such as read-only, or limit to a certain group)?

OYes O No

* If so, what are the circumstances in which you do this (e.g., what sort of files, groups, etc)?

* If you have change/set permissions, which of the following have you done? Check all that apply:

- ☐ give more people access
- ☐ give fewer people access
- ☐ allow people to do more actions (such as adding write permission)
- ☐ allow people to do fewer actions (such as removing write permission)

* Have you ever needed to set temporary permissions?
(For example, setting the permissions on a file or resource to allow someone to access it, intending to reset them after they had obtained a copy of the file or used the resource?)

OYes O No

*If yes, did you ever go back and reset these permissions?

OYes O No

* If you have any examples of when you managed file permissions, please provide details in the space below. This could include information about when you had to reset permissions, how you dealt with temporary file access, different groups you need to set permissions for, and/or problems you had when setting permissions.

* Have you ever been unable to get work done due to problems with permissions (such as being denied access to a file or server)? If so, please describe:

* When using work resources remotely (such as over a VPN), have you ever had problems with security access? For example: being refused access after you entered a password you thought was correct; being unable to figure out how to use the security software correctly and therefore being unable to connect to work.
If so, please describe:

Sharing with Groups and the Public

Do you maintain any content on the web (photo gallery, journal/blog, home page, etc)?

○Yes ○ No

If yes, is your material open to the entire public, or have you controlled access?

- ○ all material open to the entire public
- ○ some material protected by controlled access
- ○ all material protected by controlled access

If you have controlled access, please describe the sort of material that you limit access to:

If you limit access, what method(s) do you use? Check all that apply:

- ☐ login name/password shared within a group of people
- ☐ different login name/password for each individual person
- ☐ "hidden" links (i.e., tell people where to find the site but don't link to it directly)

Do you have any materials that you share, but in different ways with different groups? For example, if you share photos with friends, do you have some photos that only a small number of friends are allowed to see, and others that everyone can see?
If so, please describe the sort of material and how you manage the sharing.
(If not, just write "n/a")

Your Security Concerns

What do you feel are the major threats to your personal digital information? Choose your top 3 concerns (including any "Other" threats you may add below):

- ☐ Data loss (e.g., important files deleted)
- ☐ Private data exposed (e.g., email, health information)
- ☐ Identity theft
- ☐ Financial loss (e.g, credit card or bank theft)
- ☐ Other threats: _____

What do you feel are the major threats to your digital work information? Choose your top 3 concerns (including any "Other" threats you may add below):

- ☐ Data loss (e.g., important files deleted)
- ☐ Private data exposed (e.g., contract information)
- ☐ Financial loss
- ☐ Legal difficulties arising from a security breach
- ☐ Damage to corporate reputation
- ☐ Other threats: _____

Security Software and Home Networks
Which of the following types of software are running on your computers? Check all that apply.

For each type (such as anti-virus), your answer depends on whether you know that the software is running, for home and office desktops, as well as laptop computers:
- If you are fairly certain that the software is running, then choose "for sure."
- If you are not very certain, or have no idea, then choose "unsure/don't know"
- If you are fairly certain that the software is not running, then choose "not there"
- If the question does not apply (e.g., you don't have a laptop), leave it blank

There should be no more than 3 boxes checked per row - at most one each for home, office and laptop.

	home: for sure	home: unsure/don't know	home: not there	office: for sure	office: unsure/don't know	office: not there	laptop: for sure	laptop: unsure/ don't know	laptop: not there
Spyware detection	☐	☐	☐	☐	☐	☐	☐	☐	☐
Anti-virus	☐	☐	☐	☐	☐	☐	☐	☐	☐
firewall	☐	☐	☐	☐	☐	☐	☐	☐	☐
Spam filters	☐	☐	☐	☐	☐	☐	☐	☐	☐
Automated patches/updates	☐	☐	☐	☐	☐	☐	☐	☐	☐
VPN software	☐	☐	☐	☐	☐	☐	☐	☐	☐

Are there any other types of security software that are running on your computers? If so, please specify, along with where this software runs (home, work, laptop):

Who manages your security software (i.e., who takes responsibility for installation, configuration, updates and general maintenance)?
Check all that apply; sometimes more than one person may be responsible for a single piece of software.

	myself	tech support (work)	other
Spyware detection	☐	☐	☐
Anti-virus	☐	☐	☐
firewall	☐	☐	☐
Spam filters	☐	☐	☐
Automated patches/updates	☐	☐	☐
VPN software	☐	☐	☐

If you selected "other" above, please specify who this is, and what software they manage:

Do you have a home network connected to the Internet?

OYes O No

If so, what steps have you personally taken to make it secure, if any?

Has someone else taken responsibility for managing your home computer security? If so, please explain:

Do you have a wireless network at home (connected to the Internet)?

OYes O No

If yes: Have there been limits set on who can have access to your wireless network? (For example, by using authentication, MAC filtering, etc.)

OYes O No

(Optional) If you have worked out a policy for who can and cannot use your wireless network, please tell us about it. (For example: you keep it entirely open so that houseguests and neighbors can easily use it. Or, you have closed it down so that no strangers can access it.)

Do you ever make backups of files at work? (If this is done automatically, without your intervention, please mention that.) If so, what sorts of files?

Do you ever make backups of files at home? If so, what sorts of files?

*(Optional) Is there anything we didn't ask in this survey (impossible as that may seem) about sharing, permissions, or access control that you want to discuss? If so, enter it below:

Appendix B: Questions, Survey Two

Information About You

What is your gender?
O Female O Male

What is your age? Select from a range:
Ounder 25 O25-35 O36-45 O46-55 O55-65 Oover 65

What is your current role at Dalhousie? Check as many as apply:

- ☐ undergraduate student
- ☐ graduate student
- ☐ faculty
- ☐ staff
- ☐ Other (please specify)_____

What program of study are you in, or what is your job title? _____

Your Technical Background

How many years of computer experience do you have?
O less than 5 years O 5-10 years O 11-20 years Omore than 20 years

Which computer operating systems do you currently use regularly? Check as many as apply:

Windows:
☐ DOS ☐ 3.x☐ 98☐ NT/2000 ☐ XP

Apple:
☐ Mac OS ☐ OS X
for OS X users: also check here if you use the terminal window part (Darwin/Unix) ☐
Unix:
☐ any type of Unix (Linux, Solaris, BSD, etc)
check here if you use the command line/terminal window in Unix ☐
(Optional) If you wish to specify which particular type(s) of Unix you use, do so here:

Any other operating systems that you use regularly:

Which operating system do you use most frequently? Select one, or write your answer in the space below:

Which other operating system(s) are you familiar with?

Your Computer Skills and Experience

Which of the following computer tasks have you performed, and are comfortable doing? Check as many as apply:

- ☐ installing software
- ☐ maintaining web content (e.g., creating web pages/blog, running a website)
- ☐ installing hardware (e.g., internal hard drive, memory)
- ☐ creating documents (e.g., using Word)
- ☐ formatting a hard drive
- ☐ programming (e.g., Java, Visual Basic)
- ☐ installing an operating system (e.g., XP, OS X, Linux)

Which of the following tasks do you currently perform on a computer in an average week? Check as many as apply:

- ☐ programming (writing code)
- ☐ developing web material (home page, corporate web site, blog)
- ☐ writing and composing email
- ☐ word processing
- ☐ installing software
- ☐ system maintenance (e.g., configuring hardware/software, installing patches)
- ☐ using/managing media files (photos, music, videos)

What would you say your level of computer proficiency is (skills and ability)? Choose one:

- ☐ very low: I use computers infrequently and/or do a limited number of tasks
- ☐ low: I use computers regularly, but am not confident in my skills and ability
- ☐ medium: I am somewhat confident in my computer skills and ability
- ☐ high: I am fairly proficient using computers and can do many different tasks
- ☐ very high: I am very comfortable using computers and can successfully perform complex tasks

Do you have training or specialized experience with computers, such as studying computer science or working in technical support? If so, please specify:

Your Experiences with Files and Sharing

This survey asks questions about file sharing. When we use the term "sharing", we generally mean any method you use to give other people access to your file(s). This could be "active" (e.g., you send a file as an email attachment), or "passive" (e.g., you place a file in a shared folder for the recipient to access).

How often do you share files with others as part of your professional life (academic or job)?

O daily O weekly O few times a month O monthly
O few times a year O rarely O never

(Optional) Comments:

How often do you share files with others for personal reasons?

O daily O weekly O few times a month O monthly
O few times a year O rarely O never

(Optional) Comments:

How often is security and/or privacy a concern for you when sharing files?

O very frequently O often O sometimes O rarely O never

(Optional) Comments:

The following section consists of several statements about sharing files. Choose the option that best describes how you feel about each statement. If you wish to elaborate on any items, please feel free to use the space provided below each statement.

I am generally aware of all of the files that I am sharing, and with whom I am sharing them.

O strongly disagree O disagree O neither agree nor disagree O agree O strongly agree
(Optional) Comments:

When I share my files, it is important to me to know exactly who has access to the files.

O strongly disagree O disagree O neither agree nor disagree O agree O strongly agree
(Optional) Comments:

When I share my files, it is important to me to know when others are using the shared files.

O strongly disagree O disagree O neither agree nor disagree O agree O strongly agree

(Optional) Comments:

When I share my files, I generally allow others to access them for a limited amount of time.

O strongly disagree O disagree O neither agree nor disagree O agree O strongly agree

(Optional) Comments:

I would like to be able to know who was using each file.

O strongly disagree O disagree O neither agree nor disagree O agree O strongly agree

(Optional) Comments:

I would like to be able to specify precisely how files could be used by others.

O strongly disagree O disagree O neither agree nor disagree O agree O strongly agree

(Optional) Comments:

When I work with sensitive/private files, I make an effort to protect them.

O strongly disagree O disagree O neither agree nor disagree O agree O strongly agree

(Optional) Comments:

File Sharing Methods

There are multiple ways to share documents, files, music, and other data. Please select the methods that you have used to provide others with access to your file(s). Check as many as apply:

- ☐ email attachments
- ☐ instant messenger (e.g., MSN, Yahoo, AIM, ICQ, etc)
- ☐ physical devices (e.g., USB token, portable drive, disc, CD)
- ☐ network file share (e.g., Windows file sharing, NFS)
- ☐ web server (e.g., webpage, blog, wiki)
- ☐ file copy protocol (e.g., scp, ftp)
- ☐ iTunes music sharing
- ☐ peer-to-peer (P2P) program (e.g., torrent client, KaZaa, Limewire)
- ☐ proprietary/commercial sharing systems (e.g., Groove, Glide, Lotus)
- ☐ Other (please specify). Examples: synchronization software (ActiveSync); direct connection (cable); infrared _____

Of the methods listed above, which is the one that you use the most frequently, or the most heavily? Choose one, or specify your own method in the space below:

Why do you use this method the most? (What are its positive features?)

Have you encountered any drawbacks to this sharing method?

Although email attachments are a common way to share files, there are many alternatives. Consider the times in which you did not use email to share a file, but instead picked a different method. What are the reasons why you would use an alternative to email attachments? Choose as many as apply from the list below, and/or add your own reason:

- ☐ another method was just as easy to use (e.g., existing file sharing system, USB token)
- ☐ the convention of the person/group dictated that we use something other than email (e.g., they already used a web server for this project)
- ☐ requirement of the person, group or organization (e.g., policy said ""no email attachments"")
- ☐ wanted to avoid sending out multiple versions of a file/document
- ☐ needed to support version control/change history (e.g., on a software project)
- ☐ wanted to keep all related documents together (e.g., within shared folder)
- ☐ wanted be able to track files easily, rather than locating them within email
- ☐ needed additional security
- ☐ attachments wouldn't work (e.g., email system stripped off attachments, attachment was too big)
- ☐ network was not available (needed to share using USB token, disc)
- ☐ wanted to avoid problems with email errors (e.g., forgetting to include people in cc: list)
- ☐ needed to allow new collaborators to have access to earlier documents (i.e., if you are new to a group, you miss all the earlier email exchanges with the older files)
- ☐ needed to have data backup
- ☐ needed auditing/logging of who accessed specific file(s)
- ☐ too many files to easily share by email
- ☐ email sharing was too slow
- ☐ Other (please specify)

Among the reasons listed in the previous question, what is the most common reason why you choose an alternative to email attachments? Choose from the drop-down menu (or add your reason in the blank space below):

<<dropdown menu with same options as previous question – respondent selects only one, including "other">>

If you use peer-to-peer (P2P) software (e.g., Bittorrent, Azareus, Kazaa, Limewire) or iTunes to share files:

Do you allow other people to download your files? Choose one:

- O I don't use this software
- O No - I don't allow others to download my files
- O Yes - I allow others to download my files
- O I don't know whether other people can download my files
- O Other (please specify)

(optional) Any further comments or clarifications to your answers on this page?

The People You Share Files With

Consider all the distinct groups that you share files with.
Approximately how many different groups would you say you are currently regularly sharing files with?

Example:
friends/family: share personal photos
planning committee: share meeting minutes
project team members: share research reports
general public: share your music files

total: 4

Number of groups: _____

(optional) If you would like to list these groups, please do so below. You don't need to list more than 10.:

Think about the last time you can remember sharing a file with someone (either sending it or otherwise making it available). Feel free to stop and check your own computer files/email (if convenient) to refresh your memory.
With these details in mind, please answer the following 2 questions about this occasion:

A. What were the details of your file sharing?
Material being shared (report, music, etc):_____
Person/group being shared with (boss, spouse) :_____
Method of sharing (email, shared folder) :_____

B. What was the major reason (or reasons) why you chose this mechanism for sharing? (For example, why choose email instead of a shared folder?) Select your main reason(s) from the following list:

- ☐ it was the only way to share the file (e.g., music over iTunes)
- ☐ it was the best way to reach all intended recipients (e.g., across different organizations and systems)
- ☐ it did not require any additional set-up to work (is already configured and ready to go)
- ☐ the sharing system needed to be set up to save time/effort in the future, and I decided to do the work on this occasion
- ☐ it provided the required level of security
- ☐ it provided the required level of accessibility (e.g., needed to get outside a firewall)
- ☐ it provided the required level of reliability (e.g., provided backups, version history)
- ☐ it allowed for both read and write access, rather than read-only (i.e., wanted people to be able to make changes)
- ☐ it allowed for only read access, rather than read-write (i.e., wanted to prevent people from making changes)
- ☐ it alerted me to updates (e.g., an email attachment informs you about a new draft)
- ☐ Other (please specify)

(optional) Do you have any additional comments or clarifications that you wish to make about your responses above? If your example above was an uncommon activity for you (i.e., it was not a representative example), you can describe a more common sharing example here:

Managing File Sharing

The following questions focus on sharing methods other than email: for example, sharing with people on the same computer/file system, over a server, networked file system, etc.

Have you ever enabled or disabled sharing of a particular file or folder, set permissions on a file or folder, or restricted access to resources (e.g. set up password on website)?

○ yes ○ no ○ don't know/don't remember

If yes:

Which of the following types of control have you tried to set up? Check off all that apply (including those you tried and failed to set up):

☐ enable sharing of an item that was not previously shared (e.g., via Windows shared folders)
☐ disable sharing of an item that was currently shared
☐ mark an item as private: readable and writable only by me
☐ mark an item as read-only: readable, but not writeable by some set of others
☐ mark an item as public: specifically accessible (readable and writable) by everyone, or everyone who can log into the system
☐ allow item to be edited, but not deleted
☐ mark an item as accessible (readable and writable) to a specific individual
☐ mark an item as accessible (readable and writable) to a specific group (e.g., list of people, or named group within system)
☐ mark an item as accessible to multiple named groups
☐ Other (please specify)

If yes:
Think about the last time you remember trying to manage access to a file/files. Which of the actions in the question above were you taking? Choose from the menu, or add your own response in the box below:

<<dropdown menu with same set of items as previous question, along with "other" and a space to elaborate>>

If yes:
Do you tend to set restrictions on individual items (e.g. files, documents) one by one, or do you collect items together (e.g., into same directory or folder) and restrict/allow access to the entire collection? Choose one:

○ mainly individual files ○ mainly grouped collection ○ both methods equally

Do you ever share the same file with multiple different groups of people? Examples: email file to different groups; put copies in multiple shared folders; enable access through permissions for multiple groups; make it public so that all the various specific groups can access it easily.

O very frequently O often O sometimes O rarely O never

Have you ever wanted to place specific limits on how, or with whom, a particular item would be shared, and were unable to do so? This might be because the technology did not allow it, or because you were unable to figure out how to do so.

O yes O no O don't know/don't remember

If yes:
Please describe what you were intending to do.

(optional) Any other comments?

Changing Sharing Settings

Have you ever changed permissions (sharing settings), that were previously set by you or somebody else? That is, not modifying the default/built-in settings, but changed what someone deliberately configured in order to allow or restrict sharing.

O yes Ono

If yes:
Why? Check all that apply, and/or add your own reason(s):

☐ an error was made, and I had to correct it
☐ the previous settings were for temporary sharing, and I needed to restrict access again
☐ the previous settings were temporary restrictions (deny sharing), and I had to permit sharing again
☐ an item that was not being shared needed to be shared
☐ an item that was being shared shouldn't be shared anymore
☐ the set of people with whom the items should be shared had changed
☐ the set of shared items changed (e.g., added new files)
☐ the manner in which the shared items could be used needed to change (e.g., needed to change read-only access to allow for document editing)
☐ Other (please specify)

If yes:
How often have you had to make such changes, on average?

O daily O weekly O few times a month O monthly

O few times a year O yearly O less than yearly

If you have corrected an error in sharing settings, consider the last time that you remember when such an error occurred.

What type of error was it? Choose one:

- O shared something you didn't want to share
- O failed to share something that you wanted to share
- O shared something with the wrong people (more people, or different people) than you meant to -- too open
- O prevented someone you wanted to share with from having access -- too restricted
- O Other (please specify)_____

Was there a delay between the time that the error was made and when it was detected? Choose from the list below:

O unknown delay O error detected immediately O delay: minutes O delay: hours

O delay: days Odelay: weeks Odelay: months Odelay: years

Experiences with Sensitive Information

Do you currently work with any sensitive documents on a computer, either at work or at home? (That is, any file containing information that you wish to keep private from some person, group, and/or the general public.)

O yes Ono O don't know

What sorts of sensitive documents do you currently work with, and share, on a computer (either at home or at work)? Check: the types that you have on your computer; and the types you share with other people (e.g., over a network, through shared folders, by email attachments, etc.). This could be one person (your spouse) or a large group (the public).

For example, if you both use and share personal financial data, check both boxes.

	Use myself	Share with others
personal items containing socially-awkward information that you wish to keep private (e.g., personal email, photos)		
personal financial or medical information (e.g., credit card numbers)		
professional data or documents that are sensitive to your work organization (e.g., proprietary documents, trade secrets, grades)		
professional data or documents that are governed by law/regulation (e.g., medical records, personnel files)		

(optional) If you would like to provide more specific details about the items above, please do so:

Are there any other sensitive files that you handle? If so, please specify, along with whether or not they are shared with others.

Have you taken any actions to ensure that others do not have access to these sensitive materials? These may be technical (examples: setting passwords, moving files to obscure folder) or social (don't let people access laptop). Please provide details:

Do you currently use any computer systems that other people have accounts on? This includes accounts on the same physical machine (e.g., shared Windows desktop computer with separate accounts), as well as on the same server (e.g., user account on a Unix server with multiple users).

O yes Ono O don't know

If yes:
Do you have any personal/private material in your files on this machine?

O yes Ono O don't know

If yes (i.e., yes, you have private information):
Have you taken any steps to keep this information private from the other users, and/or to check that the material is inaccessible? If so, please specify:

File Permissions

Please select which statement best matches your experience:

- O I do not know what file permissions are, or know very little about them
- O I know what file permissions are, but I have never used them (neither checked them nor set them)
- O I know what file permissions are, but I have only looked at them (never set them)
- O I have set/changed file permissions

[* skip logic was used for this questionn: if respondent chose Option 1, the section "Your Experience Using File Permissions" was skipped entirely; after completing the following two questions (on read-only), the respondent was presented with the section entitled "Document Control"]

Have you ever set a file to "read-only"? Choose the answer that best matches your experience:

- O No - I do not understand what 'read-only' means
- O No - I have needed to do this, but do not know how
- O No - I have never needed to do this
- O I have tried, but did not succeed
- O Yes - I have set a file/files to read-only

If you have set a file to read-only:
What was your reason (or your most common reason) for doing so?

Your Experience Using File Permissions

How often do you change/manage file permissions?

- O daily
- O few times a week
- O weekly
- O few times a month
- O monthly
- O few times a year
- O once per year or less
- O never
- O Other (please specify)

What operating system(s) do you manage permissions on?

Do you ever encounter difficulties when setting permissions? If so, please describe:

Please select which statement best matches your experience:

- O I do not know what Access Control Lists are, or know very little about them
- O I know what Access Control Lists are, but I have never used them
- O I have used Access Control Lists

If you have used an Access Control List, please provide details.
- Did you configure the Access Control List, or just check it?
- What sort of operating system did you use?
- What was your reason for using an ACL?
- Did you encounter any difficulties?

Do you have any more permissions-related experiences or thoughts that you wish to describe? If so, do so below:

Document Control

Have you ever used "document locking", or similar security features, in document processing software (such as Microsoft Word or Adobe Acrobat)? Choose the answer that best matches your experience:

- O No - I never use this kind of software to create documents
- O No - I use this software (such as Word), but did not know that these features were available
- O No - I knew about these features, but have never had a need to use them
- O Yes - I have used these features

If yes: consider the last time you can remember using these features. Answer the following questions:

What software were you using?

What type of document did you protect? examples: committee report, financial data.

With whom were you sharing your protected document?

Why did you use locking?

How often would you say that you manage (create or modify) document security (such as Word password protection)?

O daily O weekly O few times a month O monthly

O few times a year O rarely O never

(optional) Any clarifications about your answers above, or further comments?

Controlling Shared Files

Have you ever taken technical steps to prevent someone from modifying a file that you have shared with them? (For example: converting a file to PDF; adding password protection; setting read-only permissions)

O yes O no O don't know/don't remember

If yes, consider the last time this happened. Please provide details below. (If you have already described this occurrence in your survey responses, just state "already described.")

What sort of file was it (document, spreadsheet, etc), and with whom were you sharing it?

How did you share the file (shared folder, email attachment, etc)?

What method(s) did you use to prevent modification? (password, file permission settings, etc)

How often do you share a file with someone that you don't want them to modify? Choose one response:

O daily O weekly O few times a month O monthly

O few times a year O rarely O never

(optional) Is there anything we didn't ask about that you want to add? Please feel free to provide your comments:

Appendix C: Group Design Activity: Script, Handouts, and Questionnaires

Script

> As each person arrives, perform introduction and handle questions and consent form (individually).

Introduction
My name is Tara Whalen, and I am a doctoral student in CS here at Dalhousie. I am creating a user interface as part of my dissertation on file sharing. I am looking for input on various aspects of the interface design. As part of this process, I am running a set of simple exercises to help me to understand how people think about different concepts around file sharing. This will consist of a card-sorting activity in groups of three, followed by two tasks done individually: a labelling task and an online questionnaire. The final activity is a moderated large group discussion.

Before this activity gets started, I need to ensure that you are willing to take part in this exercise. If you have any questions, please ask me. [Take questions] If your questions have been answered to your satisfaction, please fill out this consent form. I have a short note about confidentiality: because this is a group activity, your participation and comments will be observed by other participants. I ask that you respect the confidentiality of the session, by not discussing other people's participation, actions, or comments to anyone outside this room.

I will be using subject IDs to identify you during this session. You will need to use this number when filling out questionnaires. Your subject ID is _____

<<subject ID includes group ID as well>>

> Place people in groups of three depending on their expertise/background – as much as possible, avoid homogeneous groups. Seat each group at a table.

Introduction to subject (file sharing)
Welcome to the small group design activity session. Thank you all for coming.
I'm going to give you a quick overview of my research area, so that you have some idea of the concepts and tasks before we get started.

My research is looking at file sharing – specifically, how we can present information about shared files, and activity on shared files. My goal is to help people manage their shared files by providing them with a clear picture of file sharing, including past activities and current settings (such as file permissions). The idea is that by making this information more visible, people will be able to tell whether the right people have access to the right files, which is necessary both for collaboration and for security. One way in which I plan to do this is by presenting an overview of shared files and sharing activity in the XP File Manager, as well as providing a specialized console that displays a more

detailed, customizable view. Today, I'd like your input to help me finalize a design for user testing.

I've given you all handouts with a detailed scenario on it. These are the types of tasks that I am trying to support, by displaying relevant sharing information. Take a moment to read the handout, and then we can start the design activities.

Handout for participants: Scenario

You work for a telecommunications company, in the accounting department. You are responsible for three different internal divisions: Sales, Acquisitions, and Human Resources. The people you interact with in each of these departments are

Sales: Alice, Ben, Carl
Acquisitions: Deborah, Eliza, Frank
Human Resources: George, Herbert, Ivy

You need to perform the following tasks:
Task 1: George tells you that he is concerned that confidential human resources files have been sufficiently protected: they are only supposed to be accessible to members of Accounts and of Human Resources. He wants you to check your computer to see if the files in the "Human Resources" shared network folder are restricted so that only those two groups have access. If they are not shared with the right groups, correct this problem.

Task 2: You have sent files to Ben over the last month, using both email and instant messenger. Identify which files have been sent to him.

Task 3: You have recently made a new version of the document "Annual Report.doc", which is located on the Desktop. You have sent this file to other people in the last three months, and you need to know who they are so that they can get the new copy. Find out who received a copy of this file in the last three months.

Task 4: There is a shared folder in "My Documents" called "Memos". This folder is shared with all three departments: Sales, Acquisitions, and Human Resources. You would like to know whether these files are useful to anyone, by seeing how much activity there is on them. Find out which file has been read by the most people, and which file has been changed by the most people.

Task 5: There is a shared folder in "My Documents" called "Reports". Identify who has access to this file: who can read it, and who can change the files inside it, or both.

Group Tasks

Card sorting

Setup: each table has a stack of icon cards, blank cards, Post-it notes, pens, paperclips, instruction sheet, sheet of all the icons (for reference).

Warm-up task: the set of index cards with word processing icons

Instructions (also printed on handout):
You have been given a stack of index cards with icons (pictures) on them. What I would like you to do is to work together to sort the cards into groups that make sense. Note that

these should be conceptual groups, rather than, for example, icons that have the same graphical "style", or colour. You can put a card into more than one group; I have extras at the front of the room if you run out. You can also leave a card "unsorted" if you can't find a group that it fits into, and mark it "unsorted."

Once you have a set of groups, I'd like you to label each group: take a Post-It note, write a label on it that describes the group, and stick it to the group of cards. This would be what you think all the items have in common—why you put them together in a group. If you simply cannot come up with a label for a group, you can try to describe (in a phrase) why you grouped those items together.

3A. Test task:
Read instructions (as on sheet) to introduce the task.
We'll do a test task first, using a few cards with some icons from a word processing application. First make a set of groups, then create a label for each group.

[complete test task; answer questions that participants have]

3B. Main card-sorting task.
Now we'll move to the main card-sorting task. On these cards are pictures that represent concepts that relate to files and sharing. For the icons, note that exact colour and orientation are not deliberately fixed at this stage: these can be easily changed in the final design (for example, to make the set of icons consistent with each other). As before, sort the cards into groups that make sense, and then label each group. Remember that you can put a card in more than one group, or leave any card "unsorted." Extra cards are on this table. You can clip together your stack when you're done labelling it, with a paperclip.

[ensure everyone has cards; answer questions; take observational notes]
[collect sets of cards; ensure they don't get separated.]

3C Scenario
The icons you have just seen are being evaluated to see how understandable they are for a file-sharing tool. I provided you with a sample scenario earlier. Read the scenario and consider whether there are other icons (or information) that you would need to add, beyond those you just saw. (There is a sheet with all the icons on it on the table.) Are there any icons you would have added to the stack, or those you would have re-designed, which would help you in this scenario? If so, draw them on a card with a few notes.

Individual tasks
Labelling task (similar to closed card sorting)

Now the group task is over, and we have an activity that each person will complete alone.
[arrange items for task at each person's work space – 9 places: label sheets, icon cards, sheet of icons, word list, pens]

I have a set of words that can be associated with file sharing concepts. I'd like you to associate each of the icons with a word. You can use more than one word for each icon. There are sheets of labels with the words on them – just stick the label on the icon card if you think that word describes the icon. If you want to, you can write your own label and stick that onto the card as well. A list of all the terms is also provided on your table, for reference.

List of terms (labels), randomized.
Person, people, user, users, group, groups, world, public, team, others, open, available, accessible, sent, pushed, distributed, transferred, dispatched, delivered, forwarded, gave, transmitted, accessed, pulled, opened, viewed, changed, modified, received, touched, used, obtained, took

[make sure everyone has labels, sheets and cards; answer questions; take observational notes]
[Collect cards.]

Questionnaire
I would like to know what types of words and phrases, relating to file sharing, are most useful to you. I have a short computer-based questionnaire for each of you to complete. It is available on the computer at your workspace.

Small Group Discussion: Design

Now we will all come back together for a semi-structured discussion, which will give you a chance to talk more about issues that might have come up during the task, and more refined feedback about design.

Free-form discussion; to get the ball rolling, ask

- Any problems you had, or interesting ideas, that you want to share?

Group Design Activity: Computer-Based Questionnaire
Preamble: demographic information

Design Activity -- File Sharing

Background Information

What is your subject ID?

[ ]

1. What is your gender?

- ◯ female
- ◯ male
- ◯ prefer not to respond

2.What is your age? Select from a range:

- ◯ under 25
- ◯ 25-35
- ◯ 36-45
- ◯ 46-55
- ◯ 56-65
- ◯ over 65

3. What is your role at, or affiliation with, a university, if any? Check as many as apply:

- ☐ Undergraduate student
- ☐ Graduate student
- ☐ Faculty
- ☐ Staff
- ☐ Not affiliated with a university
- ☐ Other (please specify): [ ]

4.a) What program of study are you in, and/ or what is your job title?

[ ]

4b) Which of the following subjects do you have specialized training (education) or experience with? Check as many as apply:

- ☐ business/ management
- ☐ information science
- ☐ computer science
- ☐ human-computer interaction/ human factors
- ☐ design
- ☐ Other (please specify): [ ]

5. Which high school/college/university diplomas/degrees have you been awarded?

	Degree/Major	Field
☐	none	
☐	high school	
☐	college diploma	
☐	undergraduate	
☐	graduate: masters or equivalent	
☐	graduate: doctorate, PhD or equivalent	
☐	professional degree (medicine, law)	
☐	other	

6. Which high school/college/university diplomas/degrees are you in the process of completing?

	Degree/Major	Field
☐	none	
☐	high school	
☐	college diploma	
☐	undergraduate	
☐	graduate: masters or equivalent	
☐	graduate: doctorate, PhD or equivalent	
☐	professional degree (medicine, law)	
☐	other	

7. How many years of computer experience do you have?

◯ less than 5 years
◯ 5-10 years
◯ 11-20 years
◯ more than 20 years

8. What operating systems do you use regularly, or are quite familiar with? Check as many as apply :

☐ Windows NT/2000/XP/Vista
☐ Apple: Mac OS
☐ Apple: OS X
☐ Unix: any (Linux, Solaris, BSD, etc)
☐ other (please specify) ☐

9. What operating systems have you used for file sharing, using shared folders or directories? Check as many as apply:

- ☐ Windows NT/2000/XP/Vista
- ☐ Apple: Mac OS
- ☐ Apple: OS X
- ☐ Unix: any (Linux, Solaris, BSD, etc)
- ☐ I do not understand this question
- ☐ Other (please specify): [ ]

10. Which of the following tasks have you performed, and are quite comfortable doing? Check as many as apply:

- ☐ installing software
- ☐ maintaining web content (e.g., creating web pages/blog, running a website)
- ☐ installing hardware (e.g., internal hard drive, memory)
- ☐ creating documents (e.g., using Word)
- ☐ formatting a hard drive
- ☐ programming (e.g., Java, Visual Basic)
- ☐ installing an operating system (e.g., Windows XP, OS X, Linux)

11. Which of the following tasks do you currently perform on a computer in an average week? Check as many as apply:

- ☐ programming (writing code)
- ☐ developing web material (home page, corporate web site, blog)
- ☐ writing and composing email
- ☐ word processing
- ☐ installing software
- ☐ system maintenance (e.g., configuring hardware/software, installing patches)
- ☐ using/managing media files (photos, music, videos)

12. Briefly describe any file sharing activities that you do regularly, in your personal and work life. Include the frequency, type of file, and the person or group that you share with.
Example: email photos to family monthly; put documents in shared folder for work colleagues every week.

[ ]

Next

Section One: Icon Critique

Design Activity -- File Sharing

13a) Here is the set of icons that you just sorted and labelled.
For each icon, check each characteristic that best describes the icon. The terms may be positive or negative, and you may check off as many (or few) of them as you want.
You can also make other comments about that icon in the space provided; if you run out of space, there is room below for more detailed remarks (following this question).

	icon	poor	good	simple	complex	confusing	clear	uninformative	understandable	Comments
1.		☐	☐	☐	☐	☐	☐	☐	☐	
2.		☐	☐	☐	☐	☐	☐	☐	☐	
3.		☐	☐	☐	☐	☐	☐	☐	☐	
4.		☐	☐	☐	☐	☐	☐	☐	☐	
5.		☐	☐	☐	☐	☐	☐	☐	☐	
6.		☐	☐	☐	☐	☐	☐	☐	☐	
7.		☐	☐	☐	☐	☐	☐	☐	☐	
8.		☐	☐	☐	☐	☐	☐	☐	☐	
9.		☐	☐	☐	☐	☐	☐	☐	☐	
10.		☐	☐	☐	☐	☐	☐	☐	☐	
11.		☐	☐	☐	☐	☐	☐	☐	☐	
12.		☐	☐	☐	☐	☐	☐	☐	☐	
13.		☐	☐	☐	☐	☐	☐	☐	☐	
14.		☐	☐	☐	☐	☐	☐	☐	☐	
15.		☐	☐	☐	☐	☐	☐	☐	☐	
16.		☐	☐	☐	☐	☐	☐	☐	☐	
17.		☐	☐	☐	☐	☐	☐	☐	☐	
18.		☐	☐	☐	☐	☐	☐	☐	☐	
19.		☐	☐	☐	☐	☐	☐	☐	☐	
20.		☐	☐	☐	☐	☐	☐	☐	☐	

13b) Additional comments, if any, on icons from previous question.
Please use the number next to each icon (in table above) to refer to it (1-20), so it is clear which icon you are talking about. You can describe any problems with the above icons, if any, and how you might fix these problems, or just comment on general impressions.

Next

Section Two: Push-Oriented and Pull-Oriented Terminology

Design Activity -- File Sharing

We are trying to identify appropriate terms to describe file sharing activity. This activity includes:

- "outgoing" activity (files you send)
Example: attaching a file to an email message

- "incoming" activity (files placed in a location for others to use)
Example: a document moved into a shared folder for a colleague to read

The following questions ask about the words used to describe these different types of sharing.
(Do not worry about remembering the information on this page: each question will have instructions.)

Next

Design Activity -- File Sharing

People sometimes put computer files in a location (such as a folder or directory) for other people to use.

What is the best word to describe what other people do when they use the file, including both reading it and/or changing it?

14a). Consider this example: "Alice _____ the file report.txt yesterday."
Which word would you use to fill in that blank?

From the list below, choose the five best words and rank them: 1 is the best choice, 5 is the worst choice. Write the number in the space. If you want to add your own word, use the "other" option, and rank that word in your Top 5.

For all other words in the list: if there is a word that you would never use to fill in the blank, check the box for that word.

term	Top 5	Never use
viewed	☐	☐
opened	☐	☐
used	☐	☐
changed	☐	☐
touched	☐	☐
accessed	☐	☐
took	☐	☐
received	☐	☐
modified	☐	☐
obtained	☐	☐
pulled	☐	☐
other (specify)	☐	☐

14b). What other words, if any, would you use to fill in the blank?

Next

Design Activity -- File Sharing

People sometimes share a file by sending it to a person or group, such as using an email attachment or sending over instant messenger.
What is the best word to describe this means of sharing files?

15a). Consider this example: "" I _____ the file minutes.txt to Bob and Carl yesterday."
Which word would you use to fill in that blank?

From the list below, choose the five best words and rank them: 1 is the best choice, 5 is the worst choice. Write the number in the space. If you want to add your own word, use the "other" option, and rank that word in your Top 5.

For all other words in the list: if there is a word that you would never use to fill in the blank, check the box for that word.

term	Top 5	Never use
forwarded		☐
delivered		☐
distributed		☐
pushed		☐
sent		☐
transferred		☐
gave		☐
other (specify)		

15b). What other words, if any, would you use to fill in the blank?

Next

Design Activity -- File Sharing

16. Now, identify the best-matching pair of words from each of the lists below.
Pick only one from each list. These selections do not have to be the same as your number one choices in the previous questions.

" I ____ the report file to Alice over email,

- ◯ gave
- ◯ forwarded
- ◯ sent
- ◯ transferred
- ◯ pushed
- ◯ delivered
- ◯ distributed
- ◯ other [ ]

...and she _____ the salary file in the shared folder this morning."

- ◯ modified
- ◯ opened
- ◯ changed
- ◯ obtained
- ◯ took
- ◯ viewed
- ◯ used
- ◯ accessed
- ◯ pulled
- ◯ touched
- ◯ received
- ◯ other [ ]

Next

Design Activity -- File Sharing

These questions were designed to find words to describe two different ways of sharing:

Case One: the sender provides the file to the receiver (e.g., email attachments).
Case Two: the sender makes sure that the receiver can get access to the file at their convenience (e.g., adding a file or changing permissions in a shared directory).

Lists of words were provided for each case:
Case One: forwarded, delivered, distributed, pushed, sent, transferred, gave
Case Two: viewed, opened, used, changed, touched, accessed, took, received, modified, obtained, pulled

17. Do the words in these lists clearly show the distinction between these two cases?

18. Are the words in this list understandable, in terms of describing file sharing concepts?

Next

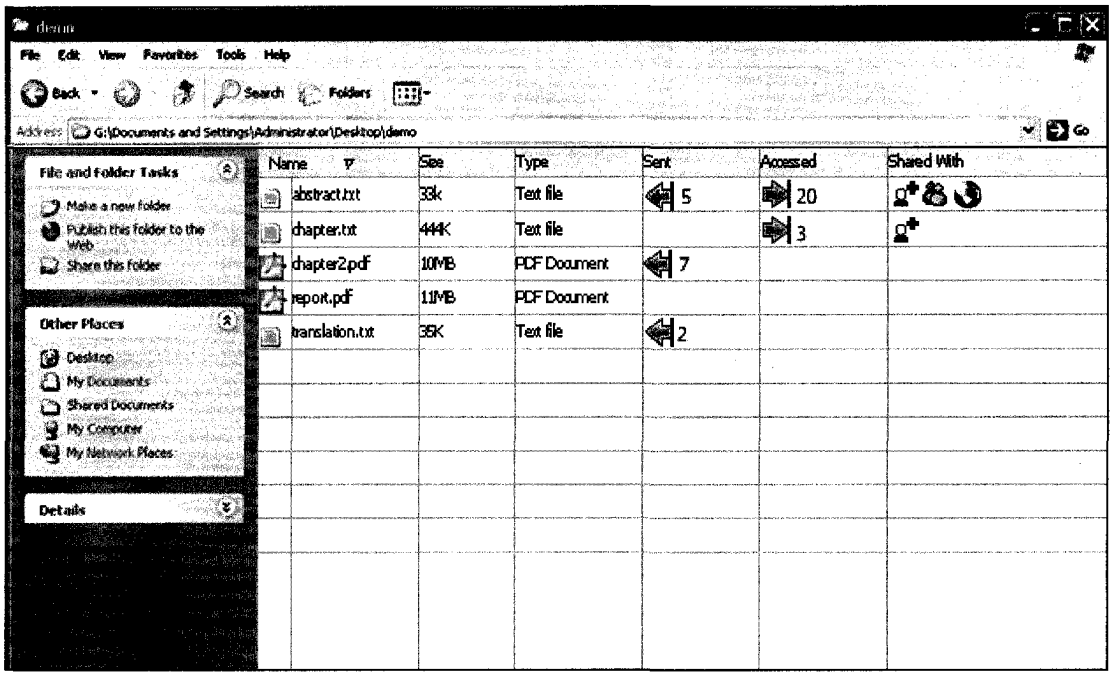
Section Three: File Manager Design Questions

Design Activity -- File Sharing

Interface Design Questions

The following questions ask about specific interface designs.

Here is a picture of the proposed interface for the modified File Manager. Answer the questions that follow.



19.a)How many times has the file chapter.txt been accessed?

19.b) How many times has the file chapter2.pdf been sent?

19.c) How many times has the file report.pdf been sent?

19.d) Is the file abstract.txt being shared with anyone?

If so, whom do you think it might be shared with (in general)?

19.e) Is the file chapter.txt being shared with anyone?

If so, whom do you think it might be shared with (in general)?

19.f) Is the file report.pdf being shared with anyone?

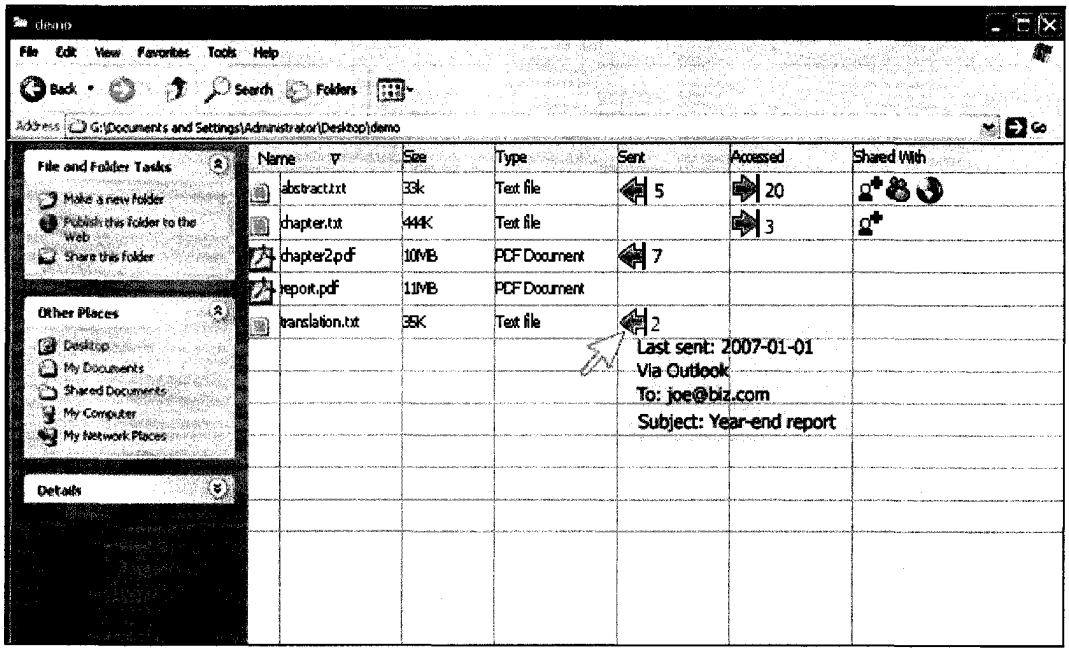
If so, whom do you think it might be shared with (in general)?



Design Activity -- File Sharing

Interface Design Questions

Here is the file manager again, but this time with tooltips ("infotips") that give some additional details. These tips appear when you hover over the icon with the mouse.



20a) Which file is the tooltip describing?

20b) What do you think the information in the tooltip describes?

20 c) To whom was the file sent?

20 d) How was the file sent?

21. Is there any other information that you think is necessary, or helpful, to include in this tooltip? If so, describe:

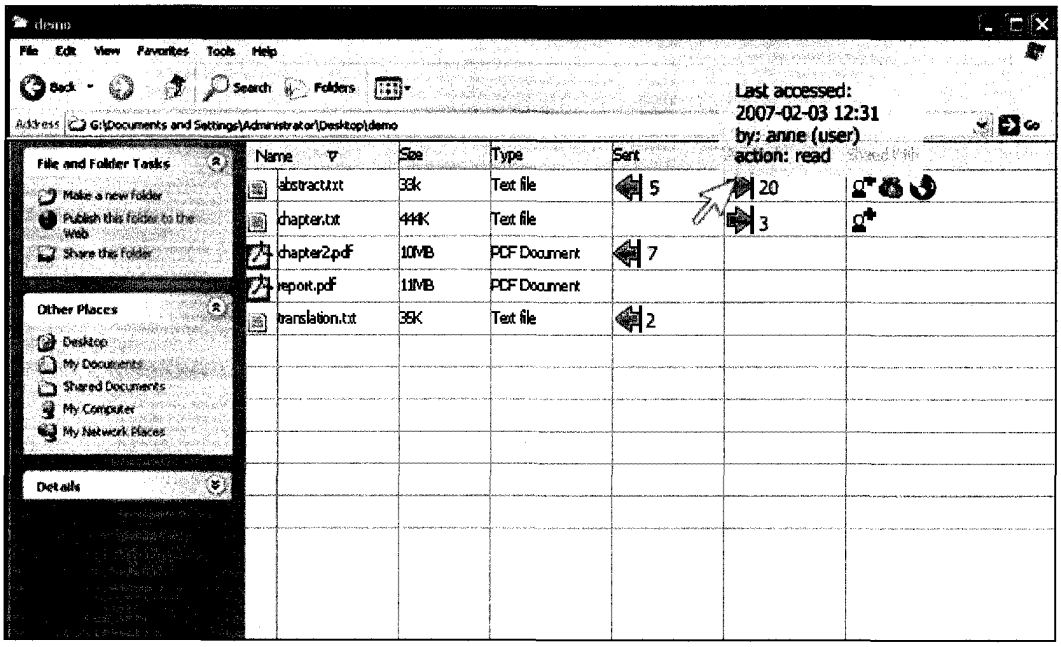
22. Are there any problems with this tooltip? If so, describe, along with any potential solutions:

Next

Design Activity -- File Sharing

Interface Design Questions

Here is the file manager again, but with a different tooltip:



23 a) Which file is the tooltip describing?

23 b) What do you think the information in the tooltip describes?

23 c) Who last used the file, and what did they do?

24. Is there any other information that you think is necessary, or helpful, to include in this tooltip? If so, describe:

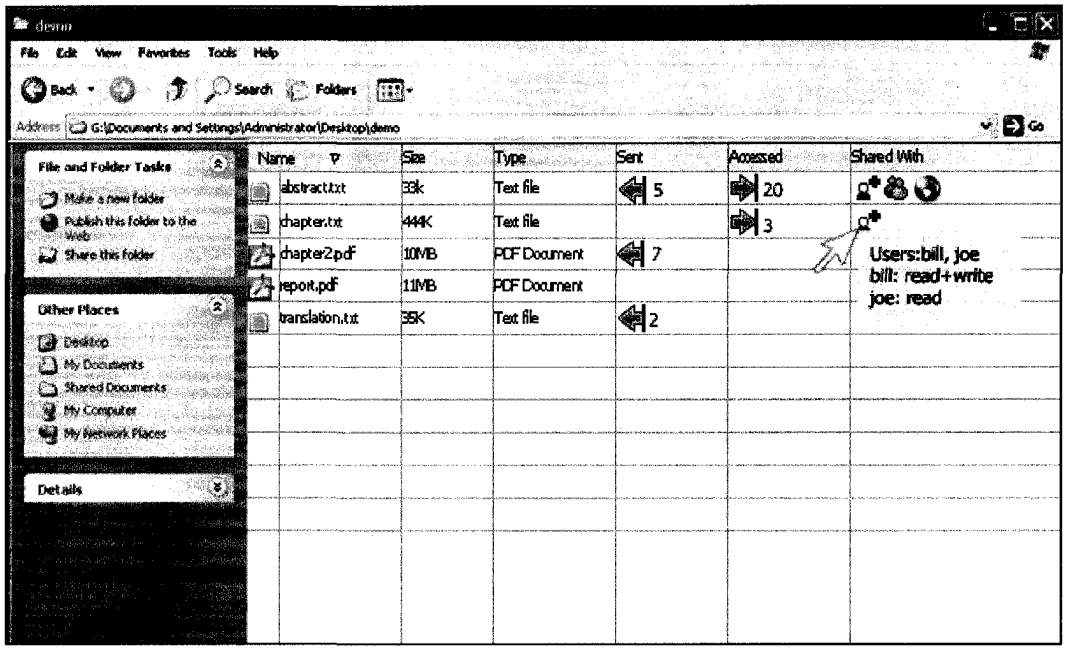
25. Are there any problems with this tooltip? If so, describe, along with any potential solutions:

Next

Design Activity -- File Sharing

Interface Design Questions

Here is the file manager again, but with another (final) tooltip:



26 a) Which file is the tooltip describing?

26 b) What do you think the information in the tooltip describes?

26 c) Who can access this file?

26 d) Who can read this file?

26 e) Who can change this file?

27. Is there any other information that you think is necessary, or helpful, to include in this tooltip? If so, describe:

28. Are there any problems with this tooltip? If so, describe, along with any potential solutions:

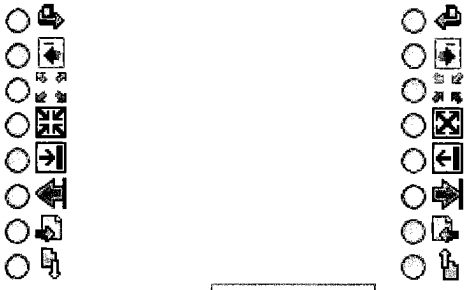
Next

Design Activity -- File Sharing

Icons, Revisited

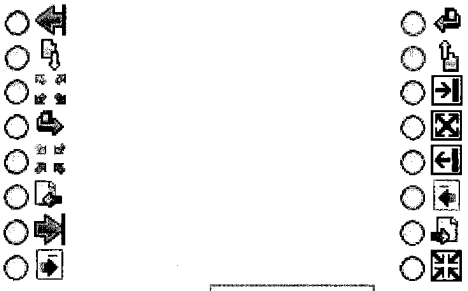
Now that you have seen the icons in context, within a file manager design, choose which ones you think are most suitable. (Note that there are a few new icons which are slight variations on what you have already seen, e.g., flipped to mirror-image.)

29. Which icon would you choose to represent "sent" (i.e., 'outgoing' activity: could be called forwarded, gave, etc.)? Choose one item from the list below, or draw your own icon on the paper provided:



○ other (on paper) _____

30. Which icon would you choose to represent "accessed" (i.e., 'incoming' activity: could be called used, opened, etc.)? Choose one item from the list below, or draw your own icon on the paper provided:

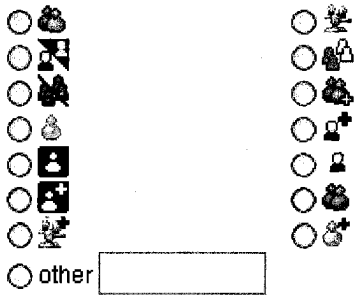


○ other (on paper) _____

There are also icons to describe those with whom you share: people (individual users), specific groups, and everyone.

First, consider the "user" and "users" icons. (One user would be Alice; multiple users might be Alice and Bob and Carl.)

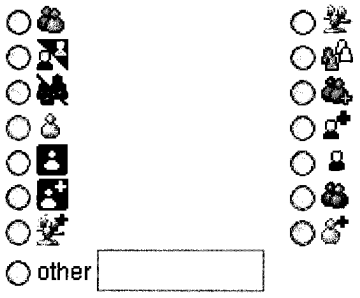
31. From the set below, choose one icon that best represents a single user, or draw your own icon on the paper provided:



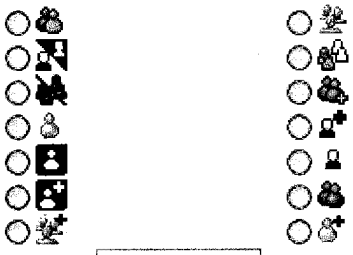
Next, consider what would be a good icon for users (plural).
Note that there may be a distinction between users and a group: "users" is just multiple individuals (such as Alice and Bob), but a "group" is a specific set of people (e.g., Payroll, carpool, Monday_Lunch_Group). This difference is important in the file system, where the permissions are set.

If possible, it would be helpful to find the best icon for users, so that it is not confused with "group."

32. Which of the following icons is your top choice for "users" (plural)? Choose only one, or draw your own icon ("other") on the paper provided.



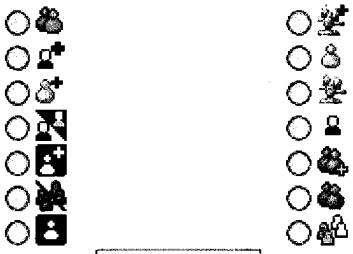
33. Which of the following icons is your top choice for "group? Choose only one, or draw your own icon ("other") on the paper provided.



○ other

Now select the best icon for groups (plural). (It may be difficult to choose one that does not also look like a single group.)

34. Which of the following icons is your top choice for "groups" (plural)? Choose only one, or draw your own icon ("other") on the paper provided.



○ other

35. Finally, which is your choice for "everybody" (e.g., sharing with all users)?
Choose one icon from the list below, or draw your own icon ("other") on the paper provided.



○ other

36. (optional) Do you have any additional comments about the overall design of the File Manager? Feel free to discuss here, and/or making drawings on the paper provided.

Next

Section Four: Sharing Console Design Questions

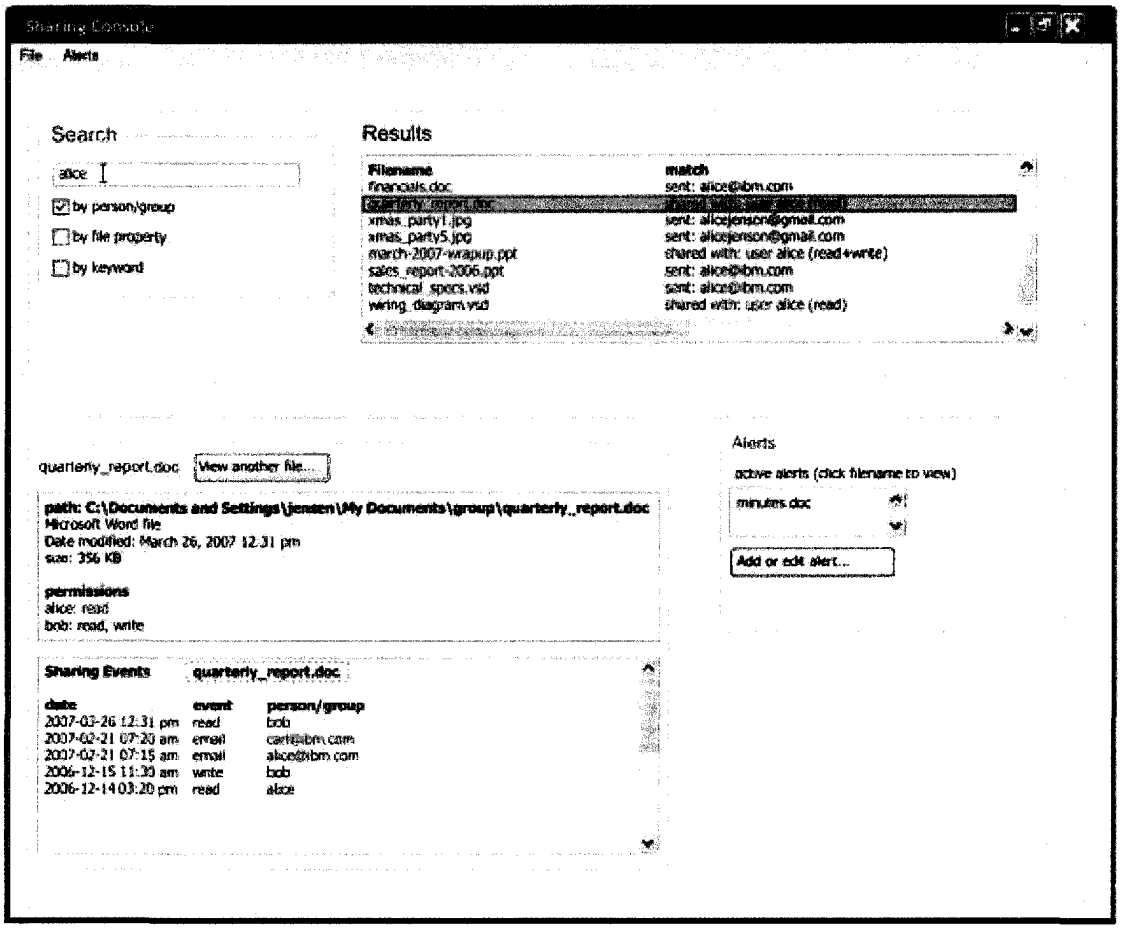
Design Activity -- File Sharing

Last Section: Sharing Console

This is the final set of questions.

Here is a console for viewing more detailed sharing information. The preliminary design is shown below. It includes the following components:

- a search feature, for finding files matching certain criteria
- a detailed view of a specific file, including permissions and activity history
- alerts, to pop-up when a certain event occurs (e.g., a person accesses it)



37 a) What is the user searching for in this console?

37 b) Which files have been emailed to Alice? Check as many as apply:

- [ ] financials.doc
- [ ] quarterly_report.doc
- [ ] xmas_party1.jpg
- [ ] xmas_party5.jpg
- [ ] march-2007-wrapup.ppt
- [ ] sales_report-2006.ppt
- [ ] technical_specs.vsd
- [ ] wiring_diagram.vsd

37 c) Who can read the file "quarterly_report.doc"?

37 d) Who has received email copies of the file "quarterly_report.doc"?

37 e) Is there an active alert on the file "quarterly_report.doc"?

37 f) How many different people have interacted with "quarterly_report.doc"?

37 g) What would you do if you wanted to set an alert?

37 h) What would you do if you wanted to view the details of the file "minutes.doc"?

37 i) What would you do if you wanted to know which files were shared with Bob?

38. (general):Do you have any comments on the design of the console? Please describe. If you identify any problems, please note, along with your proposed solution. You can use the paper provided to draw any items you wish.

Finish

Appendix D: Terminology Survey

This appendix contains the questions that composed the online terminology survey conducted during the design phase for the file sharing tools. The survey was created using Opinio (Version 5), and was hosted on an internal Dalhousie web server.

Survey: File Sharing Terminology

DALHOUSIE UNIVERSITY
Inspiring Minds

Survey: File Sharing Terminology

Tara Whalen (hci@dal.ca; phone: 494-6978)
Dr. Elaine G. Toms (hci@dal.ca; phone: 494-6978)
Dr. James Blustein (hci@dal.ca; phone: 494-6978)

Introduction	Thank you for your interest this research study. This research is part of Tara Whalen's doctoral research for a Computer Science PhD. Your participation is voluntary. You may withdraw from participation at any time, and you may decline to answer any question. You will not be asked for any information that identifies you.
Eligibility	You are eligible to participate if you are 18 years of age or older and a computer user, and have **not** already taken part in Tara Whalen's group design activity (research study) that occurred this July.
What you will be asked to do	In this study we will ask you to complete 20 questions which will take a total of 5 minutes to complete. We will ask you about your experiences with computers and file sharing, as well as your preferences among different words and phrases that relate to file sharing.
Purpose	The purpose of this study is to discover how you understand terminology that relates to file sharing. These results will be used in designing user interfaces for file sharing tools.
What you will be asked to do	In this study we will ask you to complete 20 questions which will take a total of 5 minutes to complete. We will ask you about your experiences with computers and file sharing, as well as your preferences among different words and phrases that relate to file sharing.
Risks/Benefits	The risks are comparable to those of browsing the web. Participating in the study may not benefit you directly, but we might learn about computer file sharing approaches that would benefit others.
Compensation	There is no formal compensation for participating.
Confidentiality and anonymity	Your survey responses are anonymous and thus results will be reported with no reference to you specifically. All information that you provide will be treated confidentially.
Data retention	Once the study has been completed, all data files will be retained in a locked cabinet in the research office for five years, post-publication as per Dalhousie University's policy.
In the event that you have any difficulties with, or wish to voice concern about, any aspect of your participation in this study, you may contact Patricia Lindley, Director of Dalhousie University's Office of Human Research Ethics Administration: (902) 494-1462.	

I consent to participating in this study ☐

Next

Powered by Opinio

Survey: File Sharing Terminology

1. What is your gender?

- ◯ female
- ◯ male
- ◯ prefer not to respond

2.What is your age? Select from a range:

- ◯ under 25
- ◯ 25-35
- ◯ 36-45
- ◯ 46-55
- ◯ 56-65
- ◯ over 65

3. What is your role at, or affiliation with, a university, if any? Check as many as apply:

- ☐ Undergraduate student
- ☐ Graduate student
- ☐ Faculty
- ☐ Staff
- ☐ Not affiliated with a university
- ☐ Other (please specify): _____

4. What program of study are you in, and/ or what is your job title?

_____

5. Which high school/college/university diplomas/degrees have you been awarded?

	Degree/Major	Field
☐	none	
☐	high school	
☐	college diploma	
☐	undergraduate	
☐	graduate: masters or equivalent	
☐	graduate: doctorate, PhD or equivalent	
☐	professional degree (medicine, law)	
☐	other	

6. Which high school/college/university diplomas/degrees are you in the process of completing?

	Degree/Major	Field
☐	none	
☐	high school	
☐	college diploma	
☐	undergraduate	
☐	graduate: masters or equivalent	
☐	graduate: doctorate, PhD or equivalent	
☐	professional degree (medicine, law)	
☐	other	

7. How many years of computer experience do you have?

◯ less than 5 years
◯ 5-10 years
◯ 11-20 years
◯ more than 20 years

8. What operating systems do you use regularly, or are quite familiar with? Check as many as apply:

☐ Windows NT/2000/XP/Vista
☐ Apple: Mac OS
☐ Apple: OS X
☐ Unix: any (Linux, Solaris, BSD, etc)
☐ other (please specify)

9. What operating systems have you used for file sharing, using shared folders or directories? Check as many as apply:

☐ Windows NT/2000/XP/Vista
☐ Apple: Mac OS
☐ Apple: OS X
☐ Unix: any (Linux, Solaris, BSD, etc)
☐ I do not understand this question
☐ Other (please specify):

10. Which of the following tasks have you performed, and are quite comfortable doing? Check as many as apply:

☐ installing software
☐ maintaining web content (e.g., creating web pages/blog, running a website)
☐ installing hardware (e.g., internal hard drive, memory)
☐ creating documents (e.g., using Word)
☐ formatting a hard drive
☐ programming (e.g., Java, Visual Basic)
☐ installing an operating system (e.g., Windows XP, OS X, Linux)

11. Which of the following tasks do you currently perform on a computer in an average week? Check as many as apply:

- ☐ programming (writing code)
- ☐ developing web material (home page, corporate web site, blog)
- ☐ writing and composing email
- ☐ word processing
- ☐ installing software
- ☐ system maintenance (e.g., configuring hardware/software, installing patches)
- ☐ using/managing media files (photos, music, videos)

12. Briefly describe any file sharing activities that you do regularly, in your personal and work life. Include the frequency, type of file, and the person or group that you share with.
Example: email photos to family monthly; put documents in shared folder for work colleagues every week.

Next

Survey: File Sharing Terminology

We are trying to identify appropriate terms to describe file sharing activity. This activity includes:

- "outgoing" activity (files you send)
  Example: attaching a file to an email message

- "incoming" activity (files placed in a location for others to use)
  Example: a document moved into a shared folder for a colleague to read

The following questions ask about the words used to describe these different types of sharing.
(Do not worry about remembering the information on this page: each question will have instructions.)

Next

Survey: File Sharing Terminology

People sometimes put computer files in a location (such as a folder or directory) for other people to use.

What is the best word to describe what other people do when they use the file, including both reading it and/or changing it?

13. Consider this example: "Alice _____ the file report.txt yesterday."
Which word would you use to fill in that blank?

From the list below, choose the five best words and rank them: 1 is the best choice, 5 is the worst choice. Write the number in the space. If you want to add your own word, use the "other" option, and rank that word in your Top 5.

For all other words in the list: if there is a word that you would never use to fill in the blank, check the box for that word.

term	Top 5	Never use
viewed		☐
opened		☐
used		☐
changed		☐
touched		☐
read		☐
updated		☐
accessed		☐
took		☐
received		☐
modified		☐
obtained		☐
pulled		☐
other (specify)		

14. What other words, if any, would you use to fill in the blank?

Next

Survey: File Sharing Terminology

People sometimes share a file by sending it to a person or group, such as using an email attachment or sending over instant messenger.
What is the best word to describe this means of sharing files?

15. Consider this example: "" I _____ the file minutes.txt to Bob and Carl yesterday."
Which word would you use to fill in that blank?

From the list below, choose the five best words and rank them: 1 is the best choice, 5 is the worst choice. Write the number in the space. If you want to add your own word, use the "other" option, and rank that word in your Top 5.

For all other words in the list: if there is a word that you would never use to fill in the blank, check the box for that word.

term	Top 5	Never use
forwarded		☐
delivered		☐
distributed		☐
pushed		☐
sent		☐
transferred		☐
gave		☐
other (specify)		

16. What other words, if any, would you use to fill in the blank?

Next

Survey: File Sharing Terminology

17. Now, identify the best-matching pair of words from each of the lists below.
Pick only one from each list. These selections do not have to be the same as your number one choices in the previous questions.

" I ____ the report file to Alice over email,

- ◯ sent
- ◯ distributed
- ◯ forwarded
- ◯ transferred
- ◯ pushed
- ◯ delivered
- ◯ gave
- ◯ other [ ]

...and she _____ the salary file in the shared folder this morning."

- ◯ took
- ◯ touched
- ◯ opened
- ◯ changed
- ◯ pulled
- ◯ accessed
- ◯ received
- ◯ obtained
- ◯ modified
- ◯ used
- ◯ viewed
- ◯ other [ ]

Next

Appendix E: Specification for Sharing Console

The following specification was provided to guide the implementation of the sharing console prototype.

Detailed Specification: Sharing Console

Overview:

The sharing console is designed to support three main functions:

1. Provide a detailed view about settings (permissions) and sharing activity (read, writes, etc.) of a specific file
2. Provide search functions to identify a set of files, mainly in terms of their sharing properties (e.g., shared with Alice, sent to Bob, all read-only files, etc)
3. Provide a means to set alerts on files – to pop-up an alert box when a specific event (file access of a certain type) occurs

Preliminary design

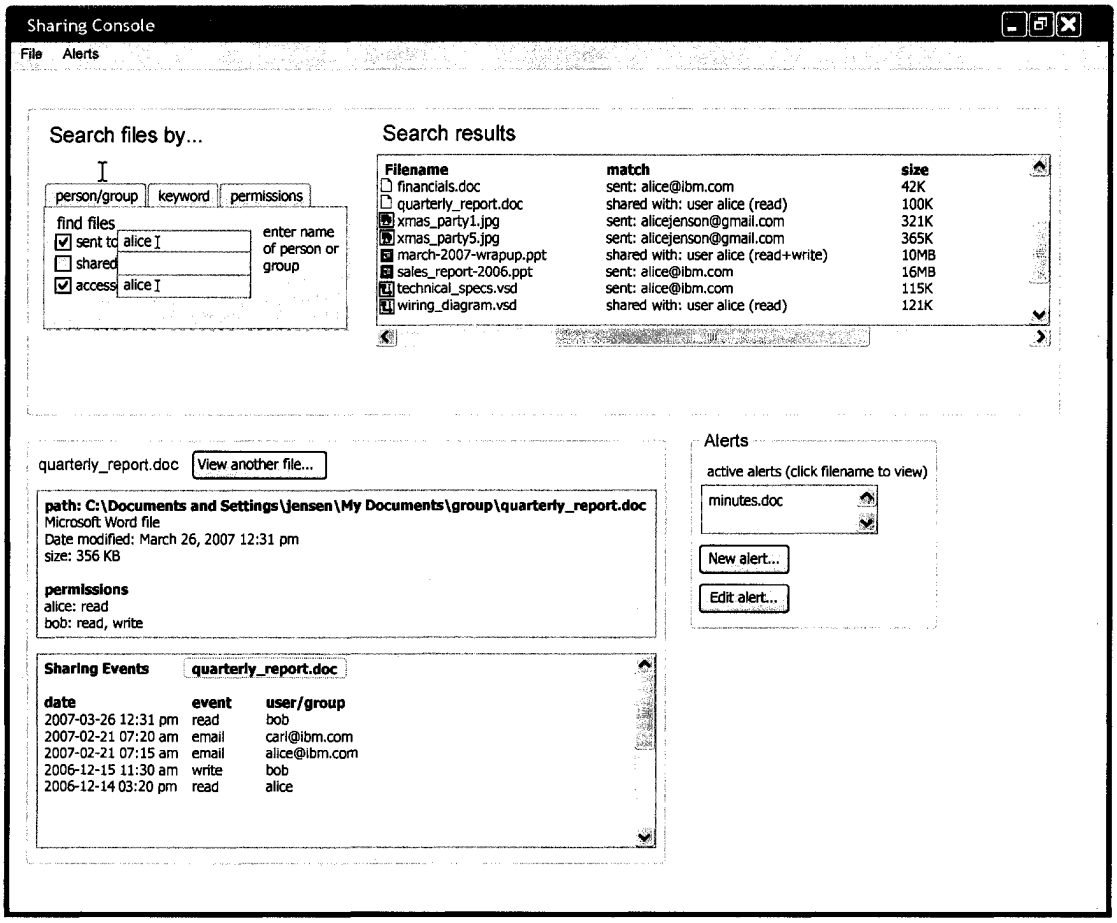


Figure 1: preliminary design for prototype console

Note that the design was drawn in Visio, so it is merely an approximation of what the final version will be; the C# controls and elements will look slightly different. The layout

of the items may also change, although the search box and results must always be adjacent to each other. The exact layout will be discussed between us, and finalized as the prototype is developed.

Interaction with other elements

The console must be able to be launched as an independent application (preferably with a desktop icon provided as well). It must also integrate with the JFile Manager, through the right-click menu: if a user right-clicks on a file, that file must open in the "details" panel in the Sharing Console. (The existing code for the JFile Manager, with some modifications for sharing, will be provided to you.)

Brief overview of JFile Manager

The File Manager that the console is launched from looks like the following:

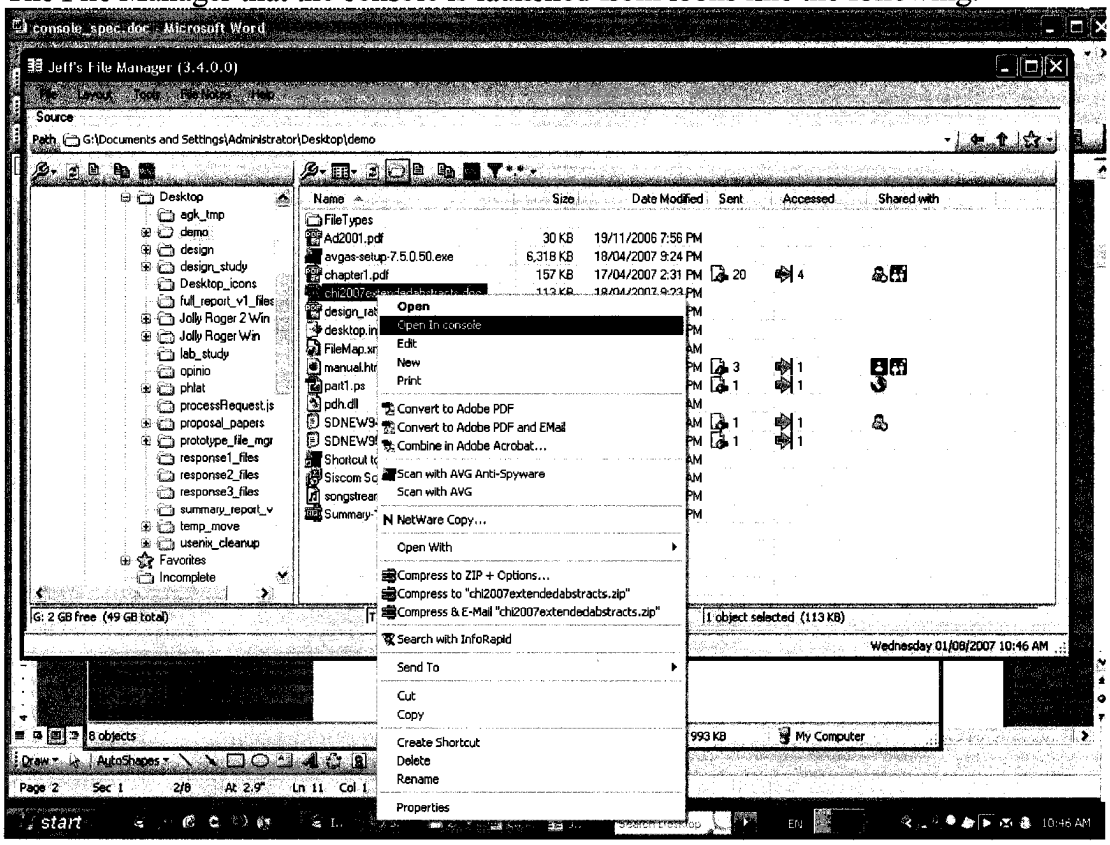


Figure 2: JFile Manager, with right-click menu

Note that this is a basic XP Explorer-type of application, with some extra details. Specifically, extra columns that describe file sharing settings and details (number of times sent and accessed, plus permissions), as shown with the icons and numbers above. This information is read from an XML file.

Right now, the right-click menu item is a stub; you can see an item marked "open in console". When that is selected, the information from the XML file (related to that specific shared file) is displayed in a very simple (scrolling) window (titled "sharing console", below):

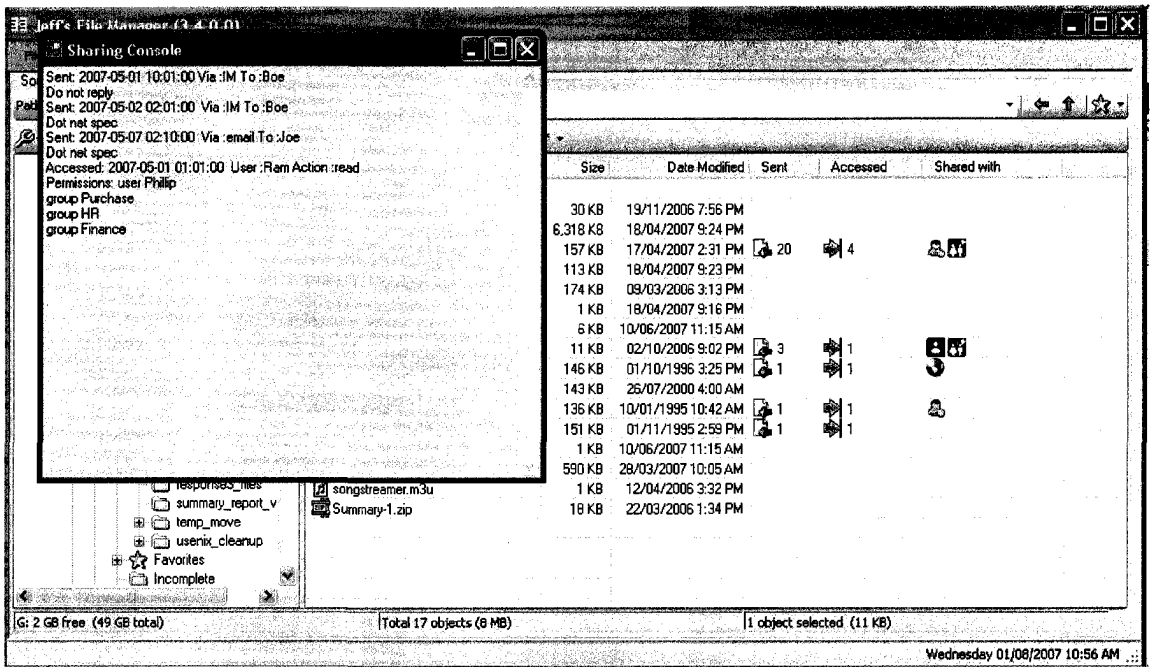


Figure 3: the stub (placeholder) version of the console

The console displays the history of interaction, for each "sent" event (time/date, method (email, IM, etc), recipient, text string (could be the subject line of email), followed by accesses (date, by whom, action taken), with permissions at the end. **This** is the basic information that needs to be displayed in the new console you will implement, although with a much improved interface, and some additional features. More information about the file manager is presented in the appendix.

Implementation Language
C#, with shell extensions as necessary.

Colour and Design
There is a some flexibility with the colour. It should look like other XP applications, in terms of the colours chosen. (For example, blue and white are common.) It needs to be clear what the various sections/regions are (e.g., the search box, the search results…); currently we have used panels to do this.

Elements of console
1. Search and Search Results

The "search" and "results" are related: the user formulates a search query (using the elements in the search box) and the files that match that query are listed in the "results" box.

Search
There are three types of search supported:

- By person/group
- By keyword

- By permissions

Each type of search is slightly different, although they all use textboxes to formulate queries.

Examples of things that a user might want to find:
Person/group: find all files sent to alice; find all files shared with bob
Keyword: find all files containing the word "bob" (including the sharing data associated with the file)
Permissions: find all files that alice has write access to; find all files that everyone can read.

Because these queries require slightly different features (e.g., different checkboxes, numbers of textboxes, etc), it is best to have tabbed interface, with one tab for each type of search. The default tab (opened when the application launches) will be "person/group."

Here is a picture of the proposed control:

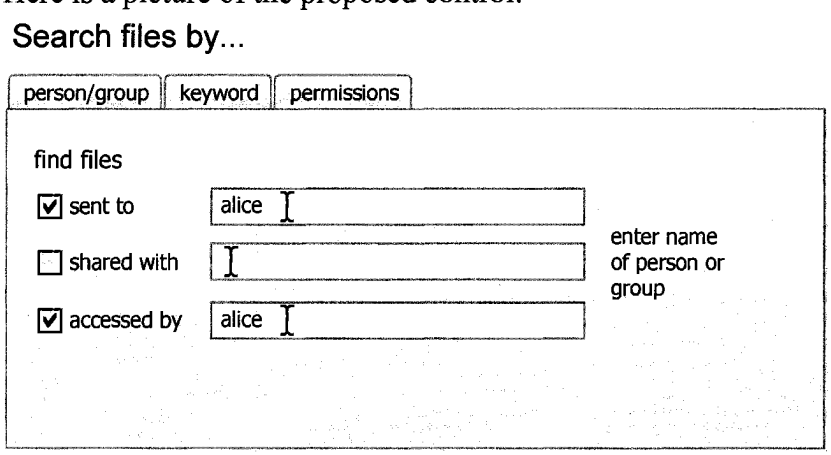


Figure 4: tabbed control for file search queries

By person/group:

This requires three pairs of checkboxes and textboxes. The textboxes are used to write the search queries in. The checkboxes refer to three different types of sharing settings and activity: sent to, shared with, and accessed by. (This is the tab shown in Figure 4 above).

It would be helpful to support Boolean queries ("Alice and not Bob"), but this feature is not mandatory. If no Boolean search is supported, all queries are assumed to be "and": if two terms are typed in the search box, *both* must appear in the file data to create a match. Also, a person can check off one or more boxes: "sent to", "shared with", and/or "accessed by". If more than one is checked, then all items must appear in order for there to be a match.

The matches are made to the data in the xml file (which is used by the file manager to display the "sharing" information). If the term(s) match a file, then that file is listed in the "results" box. "Sent to" and "Shared with" are explicitly listed in the xml file; "accessed by" is in the event history for that file (also in the xml data).

By keyword:
Very similar to person/group, except there is only one textbox (and no checkbox). As with the person/group search, this would match any files with that word in the file itself (e.g., a Word document) or in the metadata (the information in the xml file.) As with "person/group", it would be helpful (although not essential) to support Boolean queries.

Rather than implement a full text search, it is better to simplify the implementation of this feature by extending the xml data for the file manager. Add a new field (along with "shared with", etc) called "keyword", which allows for a set of words, and match using this group of words. The "keyword" information does not appear in either the file manager or the detailed view of the file information: it is only included for this search feature (and can appear in matches).

The design is shown in Figure 5:

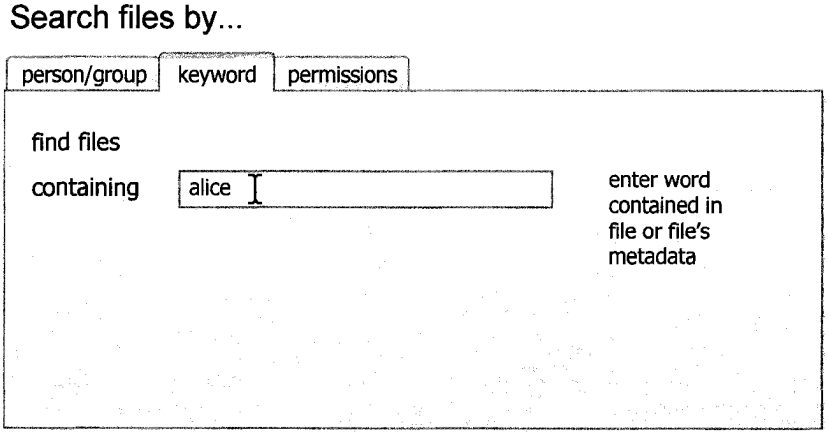


Figure 5: searching by keyword

By permissions:

This is the most complex of the three searches, because the queries are complicated. Permissions involve rights (read, write, etc) as well as a set of people and/or groups with specific rights, and user may wish to make a combination of these. The proposed design is shown in Figure 6:

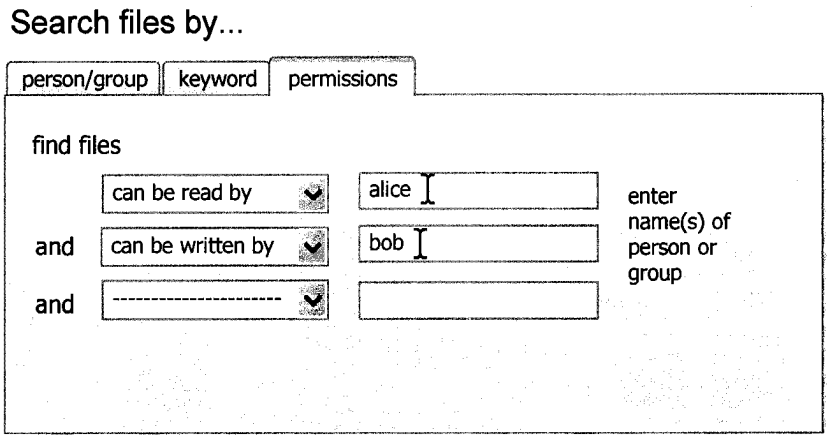


Figure 6: searching by file permissions

The basic design should have three pulldowns with text boxes. Each pulldown should have the same set of options:

can be read by
can be written by
can be read + executed by
can be modified by
full control is held by

This covers the full set of NTFS file permissions. These permissions will be included in the file metadata in the xml file, to make a match. This type of matching might prove too complicated, given that this is only a simple prototype at this stage. (Note, for example, that NTFS permissions are cumulative: a person with read + executed also has read, as does a person with write, modify, or full control rights. Thus, a search for "can be read by" must include all of these.)

Important note: If the search is too complex, some hand-created matches could be created: that is, allow the tester to identify a set of files that will be matched when a specific query occurs (rather than actually performing the search). This could be done similarly to the way the xml file is being used currently (although this data might be stored in a separate file to keep things simple.) This feature likely needs to be discussed in more detail as the prototype develops, and I will provide further guidance.

Search Results

The results of the above queries need to be presented in a list. The list should be very similar to that of the file manager detail view, but with fewer columns. One new column would be "match", which shows what the matching term was (i.e., sent to alice@ibm.com). The columns should be sortable.

The desired columns:

Filename (with file icon); match; file type; date modified; size. (All these, except for "match", are the default XP items for the file view).

The ability to right-click on the file, to get the context menu, is important: this should be very much like the file manager. In particular, the user may need to reset permissions (which is under the "Properties" item of the built-in right-click menu in XP.

Also, if a person clicks on a file in the list of results, its information (full permissions, history, etc) should appear in the "sharing details" box (described in the following section).

The results box will look similar to that shown in Figure 7 below:

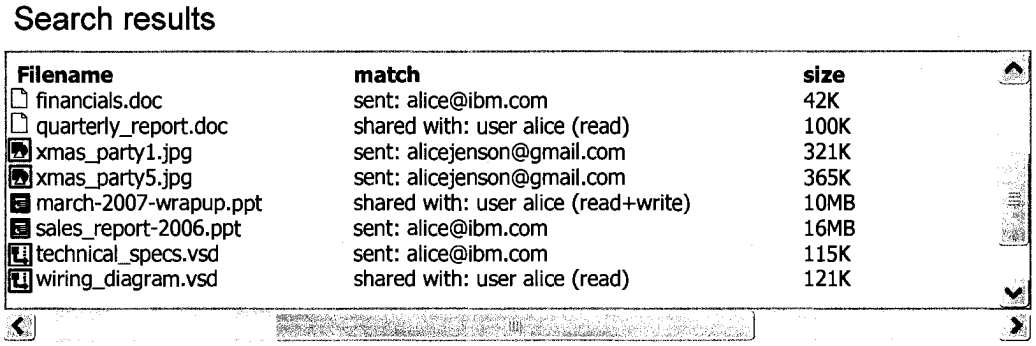


Figure 7: sharing results (list of matching files)

The file list requires vertical and horizontal scrolling, so the user can see all the details. (The full width and height of the box is dependent upon the exact measures and layout of the final design.) Filename, the matching element (e.g., string), and size of file must appear.

Sharing Details

The "sharing details" control presents the full sharing information (permissions) and history of activity on that file. This information is available in the .xml file that has been already created (for the file manager). The proposed design is shown in Figure 8:

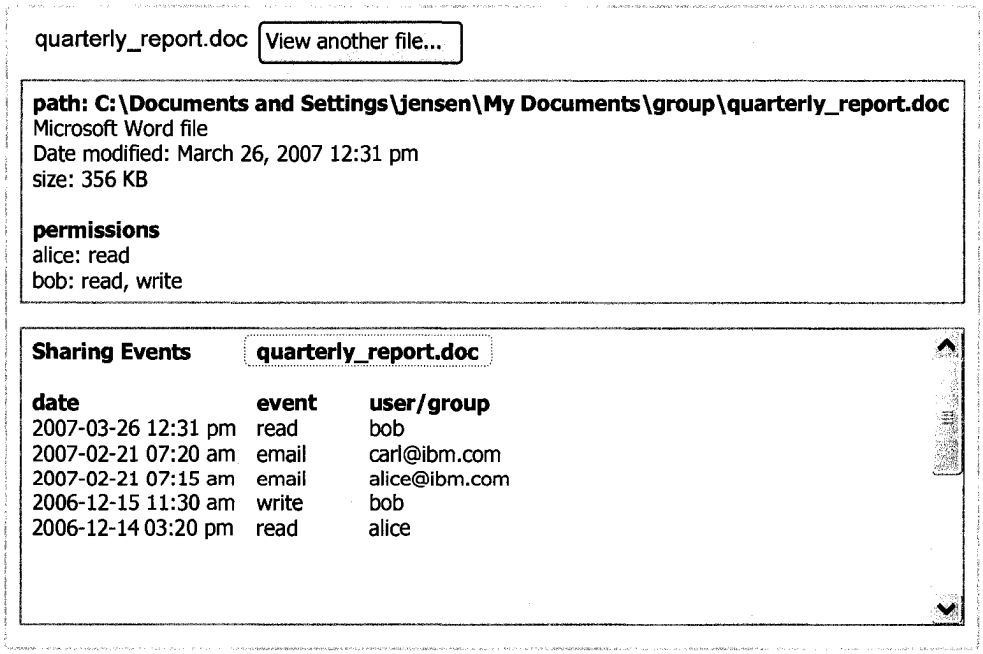
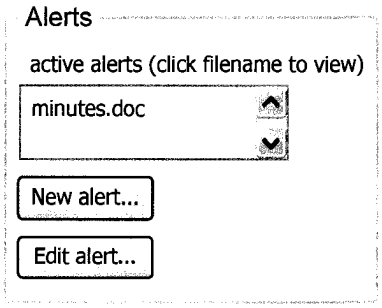


Figure 8: file sharing details

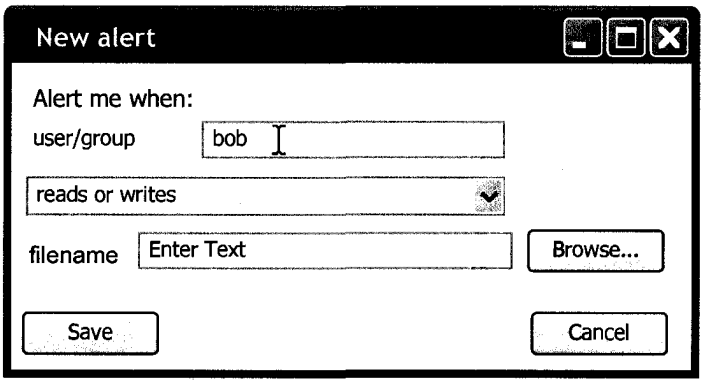
The filename must appear at the top of the box, in order to make it clear which file is being described. There is also a button for loading another file; this must launch the file manager (with the sharing interface), from the same folder as the current file being viewed. (If there is no file currently being viewed, then it should default to the desktop.) The next set of details display the file settings/information: path; file type; date modified; size; permissions (from xml file). The final set of information is the set of events, which are gathered from the .xml file; again, this box is labelled with the filename (not a button, just a simple label), so it's clear which file is being described. The events are listed in a scrolling box. The events must include: date (full date and time), the event (read, write, etc…), and the person/group/application performing the event (e.g., username, groupname…). If the event is email or IM, the "person/group" is the recipient (e.g., bob@ibm.com). All this should be read from the .xml file.

Alerts

The final component is the "alert" box, which allows a user to set, edit, and view alerts on files. The idea is to have a pop-up appear at some interval during the test, to let the user know that some action has occurred (e.g., "Bob read the file"). This feature does not genuinely need to monitor the file for that action. Instead, it should read a time (from a file, for example), and a message to display. This box should have a close button.
The box, as part of the console, is shown below:



It consists of a list of files that have alerts on them – this can be read from a file, likely the same one that has the time and message information in it. (It could have a filename, plus a time for the alert, plus a string to display). This list can scroll if there are several files.
The "new alert" brings up a window for the user to set an alert:



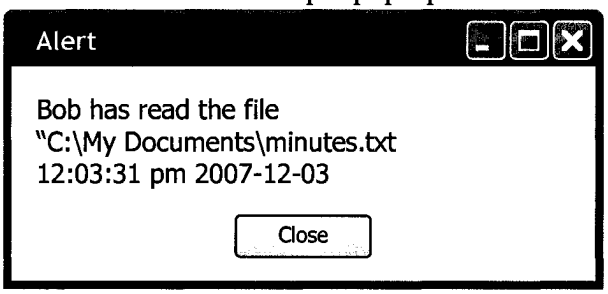
The user/group is a textbox for the user to fill in. The dropdown has a list of actions:

- reads
- writes
- executes
- read or writes
- reads, writes, or executes
- takes full control of

The "filename" box can be typed into, but it also must be able to be filled by the "browse" feature (which brings up a file manager). If possible, it would be best if that were the new file manager; but it id all right if it simply uses the XP default one. The "save" will make the box close, but no actual data is stored. (It's just to test the interface...) Cancel also closes the box.

The "edit alert" is similar to "add alert", but includes a list of the current alerts. (This list can be read from a file.) When one alert is clicked on, the dropdowns and text match the details from that alert. The user can then update them, just as above (type, and use pulldown, and use the browse button). However, *no real data* is saved to the file (it is ignored, as above.)

The *actual* alert is a simple pop-up box:



The string to be displayed ("Bob has read the file...") can be read from a file. The filename (with path) should be on a new line, and so should the date and time. The "close" button should close the alertbox.
The alert should pop up at a specified time, which is read from a file. This file contains the info to display (e.g., "Bob has read the file file.doc, 12:03 pm") along with a time at which this box needs to appear on screen.

General notes
The proposed layout above (in Figure 1) is subject to change. The placement of the items (e.g., sharing details box, alert box) may change. The main idea is to lay these out in a pleasing and usable form; there is a lot of flexibility on how this is achieved. This may take a small amount of experimentation.

List of required elements:

1. a sharing console prototype that is launched from the right-click menu item in the JFileManager. This console must have the following features:
    a. Provide a detailed view about settings (permissions) and sharing activity (read, writes, etc) of a specific file
    b. Provide search functions to identify a set of files, mainly in terms of their sharing properties (e.g., shared with Alice, sent to Bob, all read-only files, etc). If necessary, this function can return a set of "canned" (simulated) search results to match a set of specific keywords.
    c. Provide a means to set alerts on files – to pop-up an alert box when a specific event (file access of a certain type) occurs. This is simulated (in that it happens at a specific time).

Appendix: The File Manager

Here are a few more details about the file manager: a bit more context and information about the input data.

This is part of an interface study; as part of the study, we need to be able to manipulate the input files to match what actions our test user has taken. Right now, the implementation is as an xml file.

In order for this to work, there are a couple of elements you need to create and modify. The file "FileMap.xml" is the one with all the file details (events, etc) that are displayed in the file manager columns. Here is an excerpt from my own demo (the results are shown in Figures 2 and 3, for the file "manual.html" – you can see the details plus the info in the placeholder "sharing console").

```
<FILE id="2" path="G:\Documents and
Settings\Administrator\Desktop\demo\manual.html">
    <EVENTS fileid="2">
      <SENT date="2007-05-01 10:01:00" via="IM" to="Bob" subject="Do
not reply" />
      <SENT date="2007-05-02 02:01:00 " via="IM" to="Bob" subject="Dot
net spec" />
      <SENT date="2007-05-07 02:10:00 " via="email" to="Joe"
subject="Dot net spec" />
      <ACCESSED date="2007-05-01 01:01:00 " by="Ron" action="read" />
      <PERMISSION user="Phillip" usertype="user"
permissiontype="read+write" />
      <PERMISSION user="Purchase" usertype="group"
permissiontype="read+write" />
      <PERMISSION user="HR" usertype="group"
permissiontype="read+write" />
      <PERMISSION user="Finance" usertype="group"
permissiontype="read+write" />
    </EVENTS>
  </FILE>
```

You may need to do this yourself in order to understand the existing code and for debugging. There is one entry for each file that needs to be "displayed" for the test. (Obviously we will create this data for our own tests.)

You also need to modify "JFileManager.exe.config":

1. add the location of your FileMap.xml file. e.g.,

<add key ="fileMapPath" value ="C: FileMap.xml"/>

2. provide the location of the icon images files, e.g.:

<add key ="sentImage" value ="C:\Documents and Settings\avenkate\My Documents\My Completed Downloads\customfv\EXListView_demo\EXListView_demo\comedy.png"/>

<add key ="accesedImage" value ="C:\Documents and Settings\avenkate\My Documents\My Completed Downloads\customfv\EXListView_demo\EXListView_demo\down.png"/>

In my own filemap.xml file, I used:

```
  <appSettings>
    <add key ="fileMapPath" value ="G:\Documents and
Settings\Administrator\Desktop\demo\FileMap.xml"/>
    <add key ="sentImage" value ="G:\Documents and
Settings\Administrator\Desktop\design\icons_for_cardsort\original_icons
\left-purple.gif"/>
    <add key ="accesedImage" value ="G:\Documents and
Settings\Administrator\Desktop\design\icons_for_cardsort\original_icons
\go-last.png"/>
    <add key ="userImage" value ="G:\Documents and
Settings\Administrator\Desktop\design\icons_for_cardsort\original_icons
\user.gif"/>
    <add key ="usersImage" value ="G:\Documents and
Settings\Administrator\Desktop\design\icons_for_cardsort\original_icons
\user_add.png"/>
    <add key ="groupImage" value ="G:\Documents and
Settings\Administrator\Desktop\design\icons_for_cardsort\original_icons
\user-group3.gif"/>
    <add key ="groupsImage" value ="G:\Documents and
Settings\Administrator\Desktop\design\icons_for_cardsort\original_icons
\group-plus.png"/>
    <add key ="everyoneImage" value ="G:\Documents and
Settings\Administrator\Desktop\design\icons_for_cardsort\original_icons
\wi0064-16.gif"/>
    <add key ="popupmessage" value ="true" />
  </appSettings>
```

Appendix F: Materials from Laboratory Study
Script, Handouts and Questionnaires

Script, Session One

Preliminaries

[meet participant at door; introduce self]

Thank you for participating in this study.
We are studying how to effectively manage shared files, and part of this research involves the design of file management tools. Today you will be performing a few tasks with computer files: you will be playing the role of an administrator who is filling in for an absent employee. You will be completing two sets of tasks with different tools. We will be recording your computer session (for example, mouse clicks) as well as your eye movements using this eyetracker.
First we'd like you to read the details of this study and we need your signed consent that you have willingly agreed to participate.

[participant reads and sign consent form; give them a second copy to keep]

Before we start, do you have any additional questions?

First, we need some basic demographic information from you as well as some information about your professional and computer experience. The questionnaire is online, on this laptop. You will need to enter a participant ID; yours is [#], which you will use whenever you answer questionnaires today.

[participant completes background questionnaire]

Introduction to the tools

You will be using two applications today: Microsoft Outlook for email, as well as a file manager. Are you familiar with the term file manager? {If not: "It is the program you use to view your files and folders, and to navigate through your files."} The file manager you will use now is shown here on the screen. You should keep both the file manager and Outlook open throughout the tasks, rather than opening and closing them. You can minimize them if you need to, though, using this button on the top right.

If they are using Carbon:

This is the file manager you will be using. It is similar to other XP managers, such as Windows Explorer. Note that the "up" and "back" arrows are here on the top right of the window. Feel free to navigate through the file system now to get comfortable with the tool....Now, I'll get you to navigate to the folder "C:\demo". {If they can do that without problems, you can stop the warm-up, otherwise allow them more practice.} Do you have any questions? Here is a handout with information if you need it for reference.

[give participant the handout]

If they are using Helium:

This is the file manager you will be using. It is similar to other XP managers, such as Windows Explorer. Note that the "up" and "back" arrows are here on the top right of the window. Feel free to navigate through the file system now to get comfortable with the tool....Now, I'll get you to navigate to the folder "C:\demo". {If they can do that without problems, you can stop the warm-up, otherwise allow them more practice.}

You may have noticed that there are some unfamiliar columns in this manager: "sent"; "accessed"; and "shared with". These columns describe actions that have taken place on shared files, as well as file permissions.

"Sent", with this icon {indicate on screen}, indicates that the file has been sent to someone, for example, over email; the number indicates how many times the file has been sent. If you roll your mouse over the icon {wait for them to do this}, you will get details about the most recent "send" event: the date, recipient, method of sending, and extra details such as the subject line of the email that the file was attached to.

"Accessed" is similar; it's this icon here {indicate}, with a number showing how many times the file was used by someone. This could be a file in a shared folder, for instance, that other people can read or modify. If you roll your mouse over the icon {wait for them to do this}, you will get details about the most recent "access" event: the date, person who accessed the file, and the action they performed (such as read).

"Shared with" displays the permissions for a shared file. Are you familiar with file permissions? {If not: "permissions describe what can be done to the file, and by which people, such as "only Pat can read this file, and Pat cannot write (that is, modify) the file"} There are icons in that column that indicate whom a file is shared with: one user, multiple users, one group, multiple groups, or everyone. {point out each of these icons on the screen}. If you roll your mouse over the icon {wait for them to do this}, you will get more details about the specific permissions that each of these users and groups has, such as read plus write.

You can also look up more information about a file, such as detailed file permissions and the full history of sends and accesses. If you right click on the filename, you will see a menu item called "Open in sharing console". Try this with "Chapter1.pdf".
{wait for window to pop up}
You'll see the details in this pane at the bottom": here are some file details (such as the name), along with the full file permissions. There is also a list of all the sharing events that happened on that file, such as emailing and accesses. The date of the event, plus the type of event and the user or recipient, are all listed here. You can also look up details about other file by clicking on this button marked "View another file."

In addition, this console has two other features. One is the search feature. You can create a list of files sent to a specific person, or accessed by a person, for example, and the results will be displayed for you here. {indicate pane in upper right}. To perform a search, check the box next to the type of search (e.g., "sent to"), and write a name.

Try this out. What would you enter to find the files that were emailed to marie?
You can also click on the filenames that appear as search results, and the file details will appear in the pane below for each file. {let them try this}

The final feature is the "alert" feature, which you can set in order to be informed when a specific action takes place on a file, such as a specific user reading it. Clicking on the name of the alert will show you the details, and you can edit them if you wish. {let the person view the alert that is pre-loaded}

Any questions about this tool? Here is a handout with information if you need it for reference.

[give participant the reference handout]

Other Tools: Outlook + File Permissions

One task you will need to perform today is looking up file permissions; are you familiar with file permissions? {If not: "permissions describe what can be done to the file, and by which people, such as "only Pat can read this file, and Pat cannot write (that is, modify) the file"}
Are you familiar with how to find permissions in XP? If not, right-click on the file, choose "permissions"; select the "security" tab. Highlighting the name of the person will tell you what permissions they have. The checkboxes tell you which rights the person has; even if the box is greyed-out, if the checkbox is there, then they have that permission (such as "read").

In Outlook, you will need to search through emails. Are you familiar with the Advanced Find feature? It's under Tools->Find->Advanced Find, or under the "Find" pulldown on the far right of the search toolbar {show the pulldown}
It allows you to perform complex searches. More options are available under the tabs; feel free to explore these options.

Any questions about these tasks or tools? Here is a handout with information if you need it for reference.

[give participant the reference handout]

CALIBRATION

{this is done by the assistant}
I'm working with {Emile or Alexandra} in the next room; she will be setting up the eyetracker today. We have a mike and speakers so we can communicate.

Before you begin the tasks, we need to set up and calibrate the eyetracker. First I'll minimize the two programs; can you maximize them again without any trouble? {make sure this is okay}. These numbers on the screen will be used for calibrating.

[perform calibration process; start gazetracker recording]
[use Morae observer to launch recorder on remote computer]

TASKS

Okay, now we're ready to start.
Here is the scenario you will be using today; I'll read it to you and give you a copy.

Scenario: Session One

Your name is Terry Martin. It is April 10, 2008. You are an administrator at Enercomp, a large Canadian energy and telecommunications company. You usually work in the Sackville office, but today you have been called in to Halifax to fill in for Chris Lee, who had to leave work suddenly due to illness.

Chris can't get at email, so has faxed you a list of tasks to complete today. You have full access to Chris's desktop, including files and emails. Please complete each task, which involves finding the answers to some questions that Chris has.

[note that Chris and Terry are both gender-neutral names; don't assume that "Chris" must be a male or female; participant must be allowed to assume whichever makes the scenario most realistic for him or her.]
[give a copy of scenario to participant]

I'll give you the set of tasks in a moment. Please do them in the order they appear. Try your best, but if you can't find the answer to a question, just note that on the sheet and move to the next task. I realize this is a lot to take in all at once; please ask any questions you have at any time.

TASK SET #1 BEGIN

Okay, Terry, I have a set of questions from Chris, who had to fax us today. Chris left a description of this computer setup, to help you find the relevant files. Here is that description along with the tasks; Chris wants you to answer some questions and write your responses on the paper provided. I'll be next door if you have any concerns. When you're finished, just tell us – we'll hear you in the next room and we'll give you further instructions.

[give participant the stapled set of tasks. Leave]

TASK SET #1 COMPLETE

[return to the room]
I'd like you to fill out the Questionnaire marked "Post-Task Questionnaire, Part One", on the laptop. I'll just need to use your computer for a moment while you do that.

[While participant fills out questionnaire, reset environment for Task Set #2: the other file manager, Outlook back to Inbox]

- file away the paperwork from the first session
- put the yellow folder on the desk (far side)

WARM-UP – for other file manager

If they are using Carbon:

This is the file manager you will be using. It is similar to other XP managers, such as Windows Explorer. Note that the "up" and "back" arrows are here on the top right of the window. Feel free to navigate through the file system now to get comfortable with the tool....Now, I'll get you to navigate to the folder "C:\demo". {If they can do that without problems, you can stop the warm-up, otherwise allow them more practice.} Do you have any questions? Here is a handout with information if you need it for reference.

[give them the handout]

If they are using Helium:

This is the file manager you will be using. It is similar to other XP managers, such as Windows Explorer. Note that the "up" and "back" arrows are here on the top right of the window. Feel free to navigate through the file system now to get comfortable with the tool....Now, I'll get you to navigate to the folder "C:\demo". {If they can do that without problems, you can stop the warm-up, otherwise allow them more practice.}

You may have noticed that there are some unfamiliar columns in this manager: "sent"; "accessed"; and "shared with". These columns describe actions that have taken place on shared files, as well as file permissions.

"Sent", with this icon {indicate on screen}, indicates that the file has been sent to someone, for example, over email; the number indicates how many times the file has been sent. If you roll your mouse over the icon {wait for them to do this}, you will get details about the most recent "send" event: the date, recipient, method of sending, and extra details such as the subject line of the email that the file was attached to.

"Accessed" is similar; it's this icon here {indicate}, with a number showing how many times the file was used by someone. This could be a file in a shared folder, for instance, that other people can read or modify. If you roll your mouse over the icon {wait for them to do this}, you will get details about the most recent "access" event: the date, person who accessed the file, and the action they performed (such as read).

"Shared with" displays the permissions for a shared file. Are you familiar with file permissions? {If not: "permissions describe what can be done to the file, and by which people, such as "only Pat can read this file, and Pat cannot write (that is, modify) the file"} There are icons in that column that indicate whom a file is shared with: one user, multiple users, one group, multiple groups, or everyone. {point out each of these icons on the screen}. If you roll your mouse over the icon {wait for them to do this}, you will get more details about the specific permissions that each of these users and groups has, such as read plus write.

You can also look up more information about a file, such as detailed file permissions and the full history of sends and accesses. If you right click on the filename, you will see a menu item called "Open in console". Try this with "Chapter1.pdf".

{wait for window to pop up}

You'll see the details in this pane at the bottom": here are some file details (such as the name), along with the full file permissions. There is also a list of all the sharing events that happened on that file, such as emailing and accesses. The date of the event, plus the type of event and the user or recipient, are all listed here. You can also look up details about other file by clicking on this button marked "View another file."

In addition, this console has two other features. One is the search feature. You can create a list of files sent to a specific person, or accessed by a person, for example, and the results will be displayed for you here. {indicate pane in upper right}. To perform a search, check the box next to the type of search (e.g., "sent to"), and write a name.

Try this out. What would you enter to find the files that were emailed to marie?

The final feature is the "alert" feature, which you can set in order to be informed when a specific action takes place on a file, such as a specific user reading it. Clicking on the name of the alert will show you the details, and you can edit them if you wish. {let the person view the alert that is pre-loaded}

Any questions about this tool? Here is a handout with information if you need it for reference.

[give them the reference handout]

RE-CALIBRATION

Before you start the second half, we need to recalibrate the eyetracker. As before, I'll minimize the two programs, and {Emile or Alexandra} next door will talk you through the process, as before.
[perform calibration process]

I'll give you the set of tasks in a moment. Please do them in the order they appear. Try your best, but if you can't find the answer to a question, just note that on the sheet and move to the next task. Again, please do ask any questions you have.

TASK SET #2 BEGIN

Okay, Terry, I have some more faxed questions from Chris. As before, Chris wants you to answer some questions and write your responses on the paper provided. I'll be next door if you have any concerns. Tell me when you're done and I'll give you further instructions.

[give participant the stapled set of tasks. Leave room]

TASK SET #2 COMPLETE

Now that you're finished the tasks, I'd like you to complete a questionnaire on the laptop behind you. It's the link called "Post-Task Questionnaire #2" Let me know when you're done.

FINAL QUESTIONNAIRE
I have one final questionnaire for you to complete today – this one is on paper. You'll find it in the yellow folder on your desk. Again, just let me know when you're done.

END – go back to the room.
Thank you very much for coming in today. You are scheduled to return for the second session on [....]. Here is your honorarium.
[get participant to complete receipt; pay them]

**Save the Morae file when they leave **

Script, Session Two
Preliminaries

Thank you for coming back for the second half of the study.

As I told you last time, we are studying how to effectively manage shared files, and part of this research involves the design of file management tools. As before, you will be performing a few tasks with computer files: you will be playing the role of an administrator who is filling in for an absent employee. This time, one week has passed since you were here last, and the computer files have changed in your absence.

You will be completing two sets of tasks with different tools; these are the same tools that you used when you were here last. At the very end of the session we'll have a discussion about your experience with the tools. We will be recording your computer session (for example, mouse clicks) as well as your eye movements using this eyetracker.

Before we start, do you have any additional questions?

Introduction to the tools

Just as a reminder from last time…
You will be using two applications today: Microsoft Outlook for email, as well as a file manager. Are you familiar with the term file manager? {If not: "It is the program you use to view your files and folders, and to navigate through your files."} The file manager you will use now is shown here on the screen. You should keep both the file manager and Outlook open throughout the tasks, rather than opening and closing them. You can minimize them if you need to, though, using this button on the top right.

If they are using Carbon:
This is the file manager you will be using. It is similar to other XP managers, such as Windows Explorer. Note that the "up" and "back" arrows are here on the top right of the window. Feel free to navigate through the file system now to get comfortable with the tool….Now, I'll get you to navigate to the folder "C:\demo". {If they can do that without problems, you can stop the warm-up, otherwise allow them more practice.} Do you have any questions? Here is a handout with information if you need it for reference.

[give them the handout]

If they are using Helium:
This is the file manager you will be using. It is similar to other XP managers, such as Windows Explorer. Note that the "up" and "back" arrows are here on the top right of the window. Feel free to navigate through the file system now to get comfortable with the tool….Now, I'll get you to navigate to the folder "C:\demo". {If they can do that without problems, you can stop the warm-up, otherwise allow them more practice.}
You may have noticed that there are some unfamiliar columns in this manager: "sent"; "accessed"; and "shared with". These columns describe actions that have taken place on shared files, as well as file permissions.
"Sent", with this icon {indicate on screen}, indicates that the file has been sent to someone, for example, over email; the number indicates how many times the file has been sent. If you roll your mouse over the icon {wait for them to do this}, you will get details about the most recent "send" event: the date, recipient, method of sending, and extra details such as the subject line of the email that the file was attached to.
"Accessed" is similar; it's this icon here {indicate}, with a number showing how many times the file was used by someone. This could be a file in a shared folder, for instance, that other people can read or modify. If you roll your mouse over the icon {wait for them to do this}, you will get details about the most recent "access" event: the date, person who accessed the file, and the action they performed (such as read).

"Shared with" displays the permissions for a shared file. Are you familiar with file permissions? {If not: "permissions describe what can be done to the file, and by which people, such as "only Pat can read this file, and Pat cannot write (that is, modify) the file"} There are icons in that column that indicate whom a file is shared with: one user, multiple users, one group, multiple groups, or everyone. {point out each of these icons on the screen}. If you roll your mouse over the icon {wait for them to do this}, you will get more details about the specific permissions that each of these users and groups has, such as read plus write.

You can also look up more information about a file, such as detailed file permissions and the full history of sends and accesses. If you right click on the filename, you will see a menu item called "Open in sharing console". Try this with "Chapter1.pdf".
{wait for window to pop up}
You'll see the details in this pane at the bottom": here are some file details (such as the name), along with the full file permissions. There is also a list of all the sharing events that happened on that file, such as emailing and accesses. The date of the event, plus the type of event and the user or recipient, are all listed here. You can also look up details about other file by clicking on this button marked "View another file."

In addition, this console has two other features. One is the search feature. You can create a list of files sent to a specific person, or accessed by a person, for example, and the results will be displayed for you here. {indicate pane in upper right}. To perform a search, check the box next to the type of search (e.g., "sent to"), and write a name.

Try this out. What would you enter to find the files that were emailed to marie?
You can also click on the filenames that appear as search results, and the file details will appear in the pane below for each file. {let them try this}

The final feature is the "alert" feature, which you can set in order to be informed when a specific action takes place on a file, such as a specific user reading it. Clicking on the name of the alert will show you the details, and you can edit them if you wish. {let the person view the alert that is pre-loaded; this is used during the session } I'll have you open the NewFund.doc alert: see that Chris wants to know when Sheri Stack reads the NewFund.doc file.

Any questions about this tool? Here is a handout with information if you need it for reference.
[give them the reference handout]

Other Tools: Outlook + File Permissions
One task you will need to perform today is looking up file permissions; are you familiar with file permissions? {If not: "permissions describe what can be done to the file, and by which people, such as "only Pat can read this file, and Pat cannot write (that is, modify) the file"}
Are you familiar with how to find permissions in XP? If not, right-click on the file, choose "permissions"; select the "security" tab. Highlighting the name of the person will tell you what permissions they have. The checkboxes tell you which rights the person has; even if the box is greyed-out, if the checkbox is there, then they have that permission (such as "read").

In Outlook, you will need to search through emails. Are you familiar with the Advanced Find feature? It's under Tools->Find->Advanced Find, or under the "Find" pulldown on the far right of the search toolbar {show the pulldown}
It allows you to perform complex searches. More options are available under the tabs; feel free to explore these options.

Any questions about these tasks or tools? Here is a handout with information if you need it for reference.
[give participant the reference handout]

CALIBRATION
{this is done by the assistant}
I'm working with {Emile or Alexandra} in the next room; she will be setting up the eyetracker today. We have a mike and speakers so we can communicate.

Before you begin the tasks, we need to set up and calibrate the eyetracker. First I'll minimize the two programs; can you maximize them again without any trouble? {make sure this is okay}. These numbers on the screen will be used for calibrating.

- perform calibration process
- start gazetracker recording
- use Morae observer to launch recorder on remote computer

TASKS

Okay, now we're ready to start.
Here is the scenario you will be using today; I'll read it to you and give you a copy.

Scenario: Session Two

Your name is Terry Martin. It is April 17, 2008. You are an administrator at Enercomp, a large Canadian energy and telecommunications company. You usually work in the Sackville office, but today you have been called back in to Halifax to fill in for Chris Lee, who had a medical emergency. Chris has been performing computer tasks since you left, such as sending and receiving email and working on documents, so the computer system is not identical to what it was when you left on April 10th.

Chris can't get at email, so has faxed you a list of tasks to complete today. You have full access to Chris's desktop, including files and emails. Please complete each task, which involves finding the answers to some questions that Chris has.

- give a copy of scenario to participant

I'll give you the set of tasks in a moment. Please do them in the order they appear. Try your best, but if you can't find the answer to a question, just note that on the sheet and move to the next task. I realize this is a lot to take in all at once; please ask any questions you have at any time.

- Turn on the alert NOW (3 minutes) BUT only if this is Helium

TASK SET #1 BEGIN

Okay, Terry, thanks for coming back. I have another set of questions from Chris, who had to fax us today. Chris left a description of this computer setup, to help you find the relevant files. Here is that description along with the tasks; Chris wants you to answer some questions and write your responses on the paper provided. I'll be next door if you have any concerns. When you're finished, just tell us – we'll hear you in the next room and we'll give you further instructions.

[give participant the stapled set of tasks. Leave]

TASK SET #1 COMPLETE

[return to the room]
I'd like you to fill out the Questionnaire marked "Post-Task Questionnaire, Part One", on the laptop. I'll just need to use your computer for a moment while you do that.
[While participant fills out questionnaire, reset environment for Task Set #2: the other file manager, Outlook back to Inbox]

- if they are starting with Helium *only*: set up the Alert timer (3 min)
- file away the paperwork from the first session
- put the yellow folder on the desk (far side)

WARM-UP – for other file manager
If they are using Carbon:

This is the file manager you will be using. It is similar to other XP managers, such as Windows Explorer. Note that the "up" and "back" arrows are here on the top right of the window. Feel free to navigate through the file system now to get comfortable with the tool....Now, I'll get you to navigate to the folder "C:\demo". {If they can do that without problems, you can stop the warm-up, otherwise allow them more practice.} Do you have any questions? Here is a handout with information if you need it for reference.

[give them the handout]

If they are using Helium:
This is the file manager you will be using. It is similar to other XP managers, such as Windows Explorer. Note that the "up" and "back" arrows are here on the top right of the window. Feel free to navigate through the file system now to get comfortable with the tool....Now, I'll get you to navigate to the folder "C:\demo". {If they can do that without problems, you can stop the warm-up, otherwise allow them more practice.}

You may have noticed that there are some unfamiliar columns in this manager: "sent"; "accessed"; and "shared with". These columns describe actions that have taken place on shared files, as well as file permissions.

"Sent", with this icon {indicate on screen}, indicates that the file has been sent to someone, for example, over email; the number indicates how many times the file has been sent. If you roll your mouse over the icon {wait for them to do this}, you will get details about the most recent "send" event: the date, recipient, method of sending, and extra details such as the subject line of the email that the file was attached to.

"Accessed" is similar; it's this icon here {indicate}, with a number showing how many times the file was used by someone. This could be a file in a shared folder, for instance, that other people can read or modify. If you roll your mouse over the icon {wait for them to do this}, you will get details about the most recent "access" event: the date, person who accessed the file, and the action they performed (such as read).

"Shared with" displays the permissions for a shared file. Are you familiar with file permissions? {If not: "permissions describe what can be done to the file, and by which people, such as "only Pat can read this file, and Pat cannot write (that is, modify) the file"} There are icons in that column that indicate whom a file is shared with: one user, multiple users, one group, multiple groups, or everyone. {point out each of these icons on the screen}. If you roll your mouse over the icon {wait for them to do this}, you will get more details about the specific permissions that each of these users and groups has, such as read plus write.

You can also look up more information about a file, such as detailed file permissions and the full history of sends and accesses. If you right click on the filename, you will see a menu item called "Open in sharing console". Try this with "Chapter1.pdf".
{wait for window to pop up}
You'll see the details in this pane at the bottom": here are some file details (such as the name), along with the full file permissions. There is also a list of all the sharing events that happened on that file, such as emailing and accesses. The date of the event, plus the type of event and the user or recipient, are all listed here. You can also look up details about other file by clicking on this button marked "View another file."

In addition, this console has two other features. One is the search feature. You can create a list of files sent to a specific person, or accessed by a person, for example, and the results will be displayed for you here. {indicate pane in upper right}. To perform a search, check the box next to the type of search (e.g., "sent to"), and write a name.

Try this out. What would you enter to find the files that were emailed to marie?
You can also click on the filenames that appear as search results, and the file details will appear in the pane below for each file. {let them try this}

The final feature is the "alert" feature, which you can set in order to be informed when a specific action takes place on a file, such as a specific user reading it. Clicking on the name of the alert will show you the details, and you can edit them if you wish. {let the person view the alert that is pre-loaded; this is used during the session } I'll have you open the NewFund.doc alert: see that Chris wants to know when Sheri Stack reads the NewFund.doc file.

Any questions about this tool? Here is a handout with information if you need it for reference.
[give them the reference handout]

RE-CALIBRATION
Before you start the second half, we need to recalibrate the eyetracker. As before, I'll minimize the two programs, and {Emile or Alexandra} next door will talk you through the process, as before.
[perform calibration process]

I'll give you the set of tasks in a moment. Please do them in the order they appear. Try your best, but if you can't find the answer to a question, just note that on the sheet and move to the next task. Again, please do ask any questions you have.

- Turn on the alert NOW (3 minutes) BUT only if this is Helium

TASK SET #2 BEGIN

Okay, Terry, I have some more faxed questions from Chris. As before, Chris wants you to answer some questions and write your responses on the paper provided. I'll be next door if you have any concerns. Tell me when you're done and I'll give you further instructions.

[give participant the stapled set of tasks.]

TASK SET #2 COMPLETE

Now that you're finished the tasks, I'd like you to complete a questionnaire on the laptop.
It's the link called "Post-Task Questionnaire #2" Let me know when you're done...

- stop Gazetracker
- stop Morae. Save wmv file to USB keychain drive

PAPER QUESTIONNAIRE
I have another questionnaire for you to complete today – this one is on paper. You'll find it in the yellow folder on your desk. Again, just let me know when you're done.

TAM QUESTIONNAIRE
I have one final questionnaire for you to complete today – it is on the laptop, under the link marked "TAM questionnaire". Again, just let me know when you're done.

TALK-AFTER
- start the A/V recording (online + the three DVD recorders)

[Go back to the room.[
I'll come back now and replay your session so we can discuss your experience with the file managers.

- Put USB key in the drive; bring completed task set with Helium (to follow along with)

Talk-after questions:
Very open-ended. "Talk me through what you did here"; If there are problems: "What was your approach?""What are you looking for here?"

END
Thank you very much for coming in today. Here is your honorarium.

- get participant to complete receipt; pay them
- ask if they want to be contacted via email for one followup question in two weeks time.

Task Faxes

The following are the texts that were given to each participant, describing the information they were meant to find as part of the scenario. Some whitespace has been removed for space considerations. Session One had eight faxes; Session Two had nine. The introductions are also provided, in which "Chris Martin" describes the work environment (file system organization as well as professional colleagues).

Session One

Introduction

Fax ID: S11876111OVW
From: Chris Lee
Attn: Terry Martin
Date: April 10, 2008

Terry:

Thanks so much for filling in on short notice!

I'll give you a quick overview of my desktop setup, to help you get started:

All my work-related documents are stored in "C:\chris's docs", with folders for various projects. Files that I share with my colleagues are stored under "C:\Sharing-Folder".

My colleagues here at Enercomp Legal are Alan Aronowitz, Sara Shackleton, Mark Taylor, and Dave Samuels. You'll find a lot of email correspondence between us -- feel free to search through it today if you need to.

Most of my incoming email is filed away in folders; my Inbox has recent email in it. Most of my outgoing correspondence is in the "Sent Items" folder. However, I file a lot of my incoming and outgoing email in "all_documents": I keep important messages here, and this is where a lot of my email with attachments is. So you'll find attachments in both "all_documents" and "Sent Items".

I have a few tasks that I need you to do for me today. Most of them require you to look up information on my computer. If you can't find the answer, just do your best; you can write down "I couldn't answer this question" if you need to.

You don't need to fix anything: just write down what you find on these sheets, and I'll read them when I get back and make any required changes. Thanks!

Task A
Fax ID: S1P12008CPAB0021A
From: Chris Lee
Attn: Terry Martin
Date: April 10, 2008

Terry:

I've been revising a spreadsheet for the ClickPaper project: the file is clickpaper-mar-05.xls. It can be found in C:/chris's docs/forms/.

I don't remember if I emailed a copy of this to anyone: could you check for me? If I did, could you tell me all the people I emailed it to, so I can send them an updated version?

Your notes

a. Was it emailed?

b. If so, to whom? (You can just tell me the name(s); don't bother with the address part.)

Task B
Fax ID: S1P12008AQAB0020B
From: Chris Lee
Attn: Terry Martin
Date: April 10, 2008

Terry:

I'm not sure if I set up the file permissions correctly for one of my recent documents:
C:\Sharing-Folder\legal\acquisitions2008.doc

I want the "legal" group to be able to modify this file. Could you check whether this group has modify permission? If not, what are the actual permissions on that file (acquisitions2008.doc)?

Your notes

Is "modify" allowed for *legal*?

If not, the actual permissions for *legal* are:

Task C

Fax ID: S1P12008NRGAB0024C
From: Chris Lee
Attn: Terry Martin
Date: April 10, 2008

Terry:

I'm creating an "industry forecast" on the Canadian oil and gas sector. I was working on it at home and I don't remember if I copied that file back to my office computer; could you check? The filename is "energy_2008_q3.doc"; if I remembered to copy it over, it will be in C:\chris's docs\drafts. Let me know if it's there or not. If it is, could you check the "last modified" date for me, so I'll know when it was updated?

Your notes

Is the file there?

If yes, last modified date:

Task D

Fax ID: S1P12008HB0B0024*D
From: Chris Lee
Attn: Terry Martin
Date: April 10, 2008

Terry:

Our group recently provided some policy guidelines for other internal Enercomp groups; they're in the "handbook" folder in C:/Sharing-Folder. I've been wondering whether these guidelines are useful to any employees. In particular, could you check the files "merger_guide.doc" and "Leave Guidelines.doc" to see if they have been accessed by anyone?

If they have, could you please write down

i) how many times the file was accessed;
ii) the name of the person who accessed the file most recently; and
iii) the date of most recent access:

Your notes
merger_guide.doc:

* number of times accessed:
* person who accessed most recently:
* date of most recent access:

Leave Guidelines.doc:

* number of times accessed:
* person who accessed most recently:
* date of most recent access:

Task E

Fax ID: S1P*2008SA20025E
From: Chris Lee
Attn: Terry Martin
Date: April 10, 2008

Terry:

We're hiring an intern to work with Sara Shackleton this summer, and I want to make sure that person can "hit the ground running." I need to provide copies of all files that I've made available to Sara, so I can ensure they are available to the intern: this includes all emailed documents as well as items in shared folders. Could you please make a list of all files that:

a) Sara has received from me over email, and
b) all files that Sara currently can access (either read or write)?

I'm going to start small; you can just write down the first three files you find in each category. Also, don't worry about writing down the location (i.e., the folder name(s)) - just the filename.

Your notes
Emailed files:
1.
2.
3.

Files in shared folders:
1.
2.
3.

Task F

Fax ID: S1P2x2008PLA032F
From: Chris Lee
Attn: Terry Martin
Date: April 10, 2008

Terry:

Head office wants to make sure we didn't leak any copies of our strategy document (C:\chris's docs\planning-docs\2009_plans.doc). I am reasonably confident that I didn't send this out...but I would really like to make sure. Could you check that I didn't email this as an attachment to anyone?

Your notes

Did I send it out?

If so, to whom?

Task G

Fax ID: S1P2x2008PLA053G
From: Chris Lee
Attn: Terry Martin
Date: April 10, 2008

Terry -

Dave Samuels is the newest member of our team, and we're still fixing technical problems for him. He told me he needs to edit C:\Sharing-Folder\forms\template.doc. Could you check whether Dave has the correct permissions for modifying that file? If he doesn't, tell me what his permissions are currently set to:

Your notes

Are permissions correct? __ yes __ no

Actual permissions:

Task H

Fax ID: S1P2x2008LE021H
From: Chris Lee
Attn: Terry Martin
Date: April 10, 2008

Terry -

I meant to put a copy of the file "contracts.doc" in the legal group shared folder ("C:\Sharing-Folder\legal"): could you confirm whether that file is there or not? If it is, could you tell me if it was modified in the last 24 hours (April 9 or later), while I've been away?

Your notes

Is file there? __ yes __ no

Was it modified in the last 24 hours? __ yes __ no

Session Two

Introduction

Fax ID: S218761110VW
From: Chris Lee
Attn: Terry Martin
Date: April 17, 2008

Terry -

Hello again, one week later! Thanks for your help last time and this time, too; I really appreciate your coming back to help.

Like last time, I have some unfinished items I need you to deal with. As a reminder, all my work-related documents are stored in "C:\chris's docs", with folders for various projects. Files that I share with my colleagues are stored under "C:\Sharing-Folder".

My colleagues here at Enercomp Legal are Alan Aronowitz, Sara Shackleton, Mark Taylor, and Dave Samuels. You'll find a lot of email correspondence between us -- feel free to search through it today if you need to.

Most of my incoming email is filed away in folders; my Inbox has recent email in it. Most of my outgoing correspondence is in the "Sent Items" folder. However, I file a lot of my incoming and outgoing email in "all_documents": I keep important messages here, and this is where a lot of my email with attachments is. So you'll find attachments in both "all_documents" and "Sent Items".

I have a few tasks that I need you to do for me today. Most of them require you to look up information on my computer. If you can't find the answer, just do your best; you can write down "I couldn't answer this question" if you need to.

You don't need to fix anything: just write down what you find on these sheets, and I'll read them when I get back and make any required changes. Thanks!

Task I

Fax ID: S2Bx2008PLA01I
From: Chris Lee
Attn: Terry Martin
Date: April 17, 2008

Terry -

I just got a call from Communications & Marketing: apparently there were some major errors in the April 13th press release; it's C:\chris's docs\Communications\PR_apr13_08.pdf

I was asked to provide updated versions of the press release to anyone who was sent a copy with the errors in it. First I need to find out who got a copy of that file.

Please make me a list of all the email recipients of this file (i.e., everyone who I emailed this attachment to). This could have gone out anytime between April 13th and yesterday (April 16th). I just need their email addresses.

Your notes

Email addresses of all who got a copy of PR_apr13_08.pdf via email:

Task J

Fax ID: S2Bx2008ND051J
From: Chris Lee
Attn: Terry Martin
Date: April 17, 2008

Terry -

One of our non-disclosure agreements was accidentally put in the public shared folder: C:\Sharing-Folder\public. The filename is nda.pdf. That's a sensitive document, and I'm worried that it might have leaked. I don't know if anybody accessed it, though, so we might be safe. If the file was accessed between April 13th and today, then we may have a problem.

If there were any accesses, then give me as many details as you can (such as the time of access, or anything else you can find). That will help with damage control!

Your notes

1) Did anyone access the file nda.pdf between April 13th and April 17th (today)?
2) If so, give me all the details you can find about any accesses (such as time), between April 13th and today:

Task K

Fax ID: S2Bx2008GE083K
From: Chris Lee
Attn: Terry Martin
Date: April 17, 2008

Terry -

Sara wants me and Alan to review and edit a draft of the General Electric prospectus; she told me it was called "GE_prospect.doc", and it should be in C:\Sharing-Folder\legal. Is there a copy of that document in that folder? If so, can you check if Alan and I have both read and write permission on that file?

Your notes

1. Is GE_prospect.doc in the legal folder?

2. If so,

i) Do I have read and write permissions on it?

ii) Does Alan Aronowitz have read and write permissions on it?

Task L

Fax ID: S2P*x2008PLA04L
From: Chris Lee
Attn: Terry Martin
Date: April 17, 2008

Terry -

I fixed some more file permissions for Dave - thanks for giving me the information last week. I think it's all correct now, but just as a quick check: could you see whether Dave has accessed the file C:\Sharing-Folder\forms\template.doc since April 12th? If he did, what action did he perform (e.g., read, write...)?

Your notes

a) Has Dave accessed template.doc since April 12th?

b) If so, what action did he perform?

Task M

Fax ID: S2P*x2008TM014M
From: Chris Lee
Attn: Terry Martin
Date: April 17, 2008

Terry -

On April 12th, one of the Finance employees left Enercomp to take a job at another company. His name is Tom Moran; he used to work with us in Legal, a few months ago. I need to create a list of documents that he might have had access to, in case there are any intellectual property concerns in future.

Could you please create a list of all files that a) Tom has received from me over email, and b) whether Tom accessed any files in our shared legal folder since he left on April 13th?

Your notes

Files emailed to Tom (you can just write the filename; I don't need the path/folder):

Tom's accesses in C:\Sharing-Folder\legal after April 13th:

Task N

Fax ID: S2P*x2008ACK035N
From: Chris Lee
Attn: Terry Martin
Date: April 17, 2008

Terry -

Communications & Marketing are creating a summary of how our reports are being used by other departments and outside the firm. Legal put some materials in the C:\Sharing Folder\public folder, so that customers and others could get useful information. We're supposed to report on our successes: could you tell me which three files have been accessed the most? We'll probably move those to a more prominent spot on the Enercomp external web site.

Your notes

Three most-accessed files in Sharing Folder\public\:
1.
2.
3.

Task O

Fax ID: S2Bx2008CP031O
From: Chris Lee
Attn: Terry Martin
Date: April 17, 2008

Terry -

I'm still re-drafting that ClickPaper spreadsheet - if I recall, it's in C:\chris's docs\forms\clickpaper-mar-05.xls. I vaguely recall emailing this out to someone for fact checking, but I have no idea who that was; it would have been since you were here, so April 11 or later.

So could you check who it was I sent that file to?

Your notes

Who did I email clickpaper-mar-05.xls to (April 11th or later)?

Task P

Fax ID: S2Bx2008MER042P
From: Chris Lee
Attn: Terry Martin
Date: April 17, 2008

Terry -

I put a copy of the file "merger.doc" in the "forms" shared folder ("C:\Sharing-Folder\forms").

I'd like to know if the document is useful to the team or not. Find out any details you can about accesses to file since you were here last, which would be April 11th or later. (Details would include time of last access, for example, or any other information about accesses that you can find out.) That will help me to see how the file is being used by my colleagues.

Your notes:

Did anyone access C:\Sharing-Folder\forms\merger.doc on April 11th or later?

If so, provide all the details you can find about any accesses (such as time):

Task Q

Fax ID: S2Bx2008BP062Q
From: Chris Lee
Attn: Terry Martin
Date: April 17, 2008

Terry -

We've created a joint shared folder for the British Petroleum project, for Legal and Finance. It's C:\Sharing-Folder\BP. I believe there is a file in there called options.doc: everybody in Legal as well as Finance should have read access to this, but nobody else should. Could you check whether the file is there, and if the permissions are set properly?

If there's an error, write it down below and I'll get Tech Support to fix it ASAP:

Your notes

1. Is options.doc in the BP folder?

2. If so: check file permissions for C:\Sharing-Folder\BP\options.doc

i) Do only Legal and Finance have access (and no one else)?

ii) Do Legal and Finance have read access?

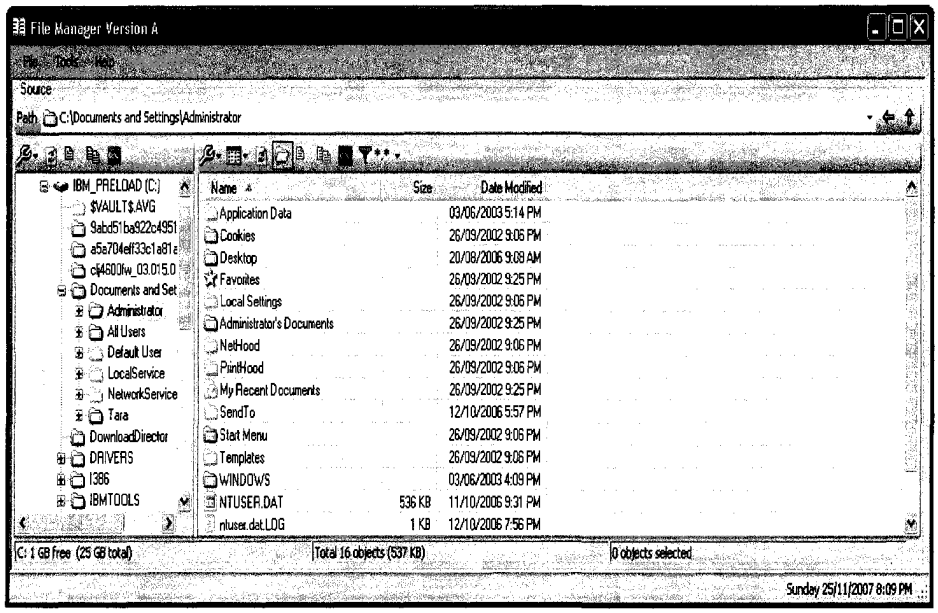
Usability Questionnaire

This questionnaire consists of 17 items (Q1-17) from the Computer System Usability Questionnaire (CSUQ) and 2 items (Q18-19) from the After-Scenario Questionnaire (ASQ).

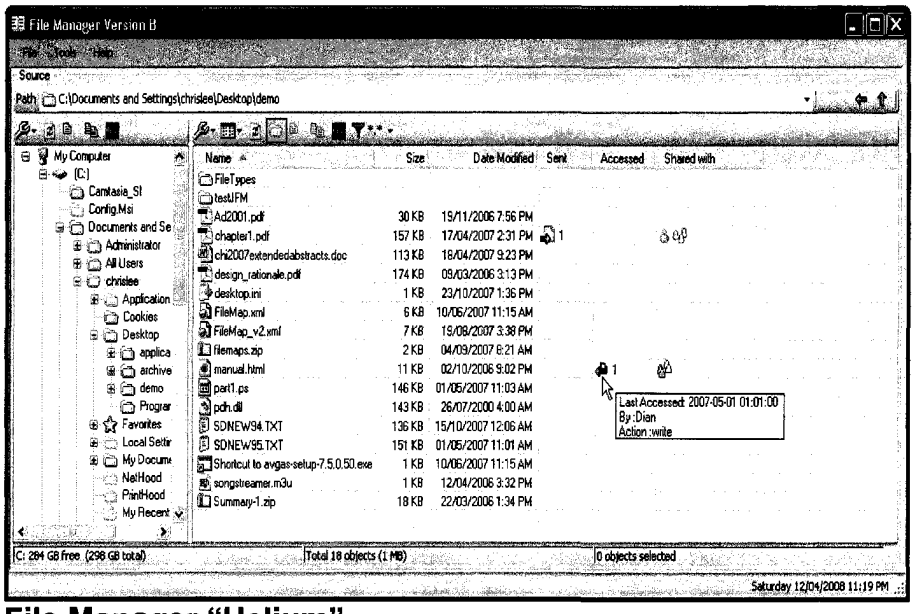
	strongly disagree 1	2	3	4	5	6	strongly agree 7
CSUQ							
1. Overall, I am satisfied with how easy it is to use this file manager	o	o	o	o	o	o	o
2. It was simple to use this file manager	o	o	o	o	o	o	o
3. I could effectively complete my work using this file manager	o	o	o	o	o	o	o
4. I was able to complete my work quickly using this file manager	o	o	o	o	o	o	o
5. I was able to efficiently complete my work using this file manager	o	o	o	o	o	o	o
6. I felt comfortable using this file manager	o	o	o	o	o	o	o
7. It was easy to learn to use this file manager	o	o	o	o	o	o	o
8. I believe I became productive quickly using this file manager	o	o	o	o	o	o	o
9. Whenever I made a mistake using the file manager, I could recover easily and quickly	o	o	o	o	o	o	o
10. It was easy to find the information I needed	o	o	o	o	o	o	o
11. The information provided by the file manager was easy to understand	o	o	o	o	o	o	o
12. The information was effective in helping me complete the tasks and scenarios	o	o	o	o	o	o	o
13. The organization of information in the file manager was clear	o	o	o	o	o	o	o
14. The interface of this file manager is pleasant	o	o	o	o	o	o	o
15. I liked using the interface of this file manager	o	o	o	o	o	o	o
16. This file manager has all the functions and capabilities I expect it to have	o	o	o	o	o	o	o
17. Overall, I am satisfied with this file manager	o	o	o	o	o	o	o
ASQ							
18. Overall, I am satisfied with the ease of completing the tasks in this scenario	o	o	o	o	o	o	o
19. Overall, I am satisfied with the amount of time it took to complete the tasks in this scenario.	o	o	o	o	o	o	o

Questionnaire: Post-Session Comparison (Session One)

Note that the order in which the two managers were presented varied depending on the order actually used in that session. The example below is for an odd-numbered participant, who used the control (Carbon) first. In addition, the questionnaire was printed in landscape orientation, so the managers were shown side by side; they are re-arranged in the example below due to space constraints.



File Manager "Carbon"



File Manager "Helium"

The file managers are shown above, in the order you used them. For each of the tasks you performed today, which file manager do you think best helped you complete that task? Check the appropriate box for each task below:

1. Task: Checking file permissions Carbon__
   Helium__ both were the same ___

2. Task: Determining who a file was sent to, if anyone Carbon__ Helium__
   both were the same ___

3. Task: Checking the "last modified" date of a file Carbon__
   Helium__ both were the same ___

{screenshots reprinted here in original questionnaire, for convenience; omitted for space}

Please answer these questions about the two file managers you used today. They are shown above in the order you used them.

1. Overall, it was simpler to use this file manager

   Carbon__ Helium__

2. I could more effectively complete the tasks using this file manager

   Carbon__ Helium__

3. I was able to complete the tasks quicker using this file manager

   Carbon__ Helium__

4. It was easier to find the information I needed using this file manager

   Carbon__ Helium__

5. I had greater awareness of which files were shared with whom using this file manager

   Carbon__ Helium__

6. I had more complete information about shared files using this file manager

   Carbon__ Helium__

7. I believe I could manage my shared files better using this file manager

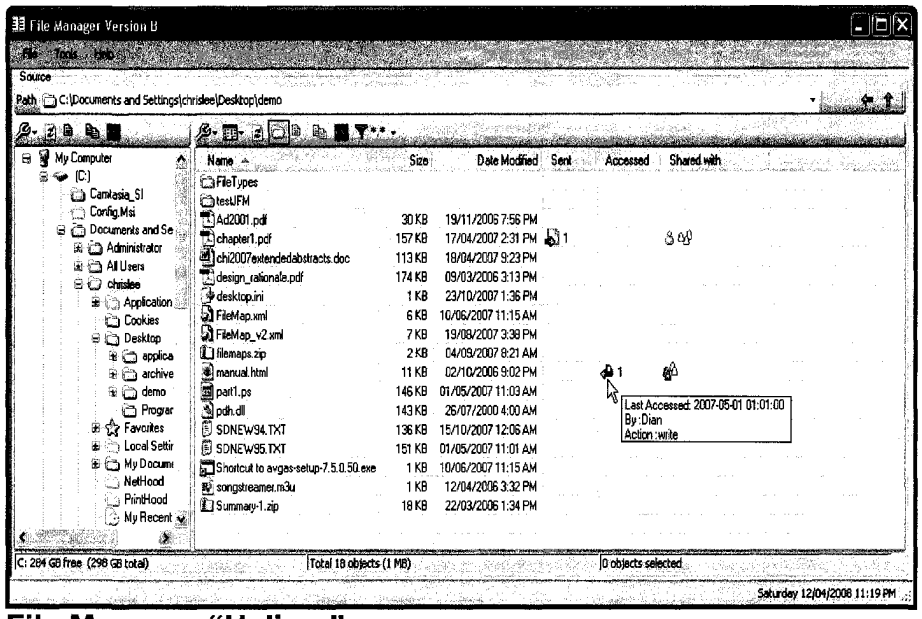
   Carbon__ Helium__

8. Overall, I prefer: Carbon__ Helium__

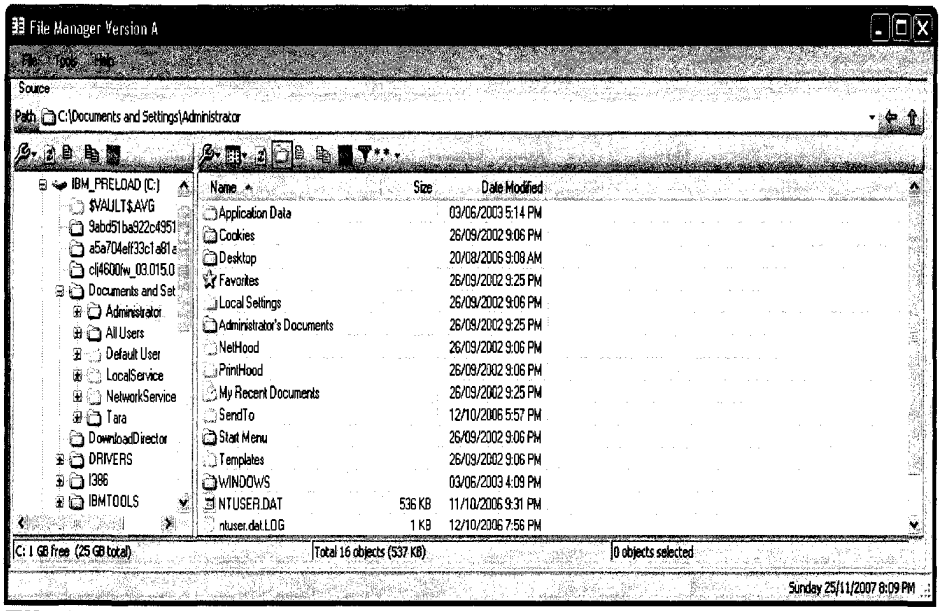
Why?

Questionnaire: Post-Session Comparison (Session Two)

Note that the order in which the two managers were presented varied depending on the order actually used in that session. The example below is for an even-numbered participant, who used the new manager (Helium) first. In addition, the questionnaire was printed in landscape orientation, so the managers were shown side by side; they are re-arranged in the example below due to space constraints.



File Manager "Helium"



File Manager "Carbon"

The file managers are shown above, in the order you used them. For each of the tasks you performed today, which file manager do you think best helped you complete that task? Check the appropriate box for each task below:

Task: Checking file permissions
Helium__ Carbon__ both were the same ___

Task: Determining who a file was sent to, if anyone
Helium__ Carbon__ both were the same ___

Task: Finding information about past accesses on a file
Helium__ Carbon__ both were the same ___

{screenshots reprinted here in original questionnaire, for convenience; omitted for space}

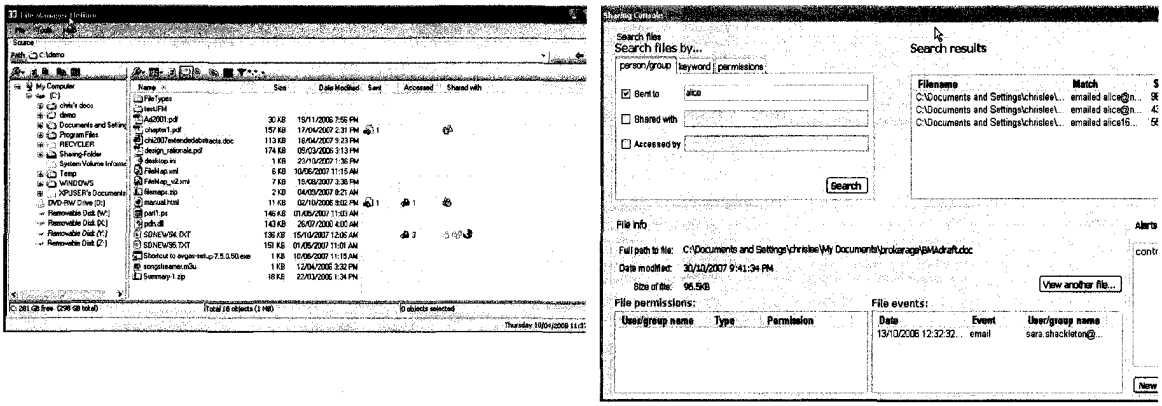
Please answer these questions about the two file managers you used today. They are shown above in the order you used them.

1. Overall, it was simpler to use this file manager
   Helium__ Carbon__
2. I could more effectively complete the tasks using this file manager
   Helium__ Carbon__
3. I was able to complete the tasks quicker using this file manager
   Helium__ Carbon__
4. It was easier to find the information I needed using this file manager
   Helium__ Carbon__
5. I was better able to find out how the workspace had changed during my absence (e.g., the "week" since the last session) using this file manager
   Helium__ Carbon__
6. I was better able to keep track of events on shared files using this file manager
   Helium__ Carbon__
7. I was better able to find out who had made changes to shared files using this file manager
   Helium__ Carbon__
8. I had greater awareness of which files were shared with whom using this file manager
   Helium__ Carbon__
9. I had more complete information about shared files using this file manager
   Helium__ Carbon__
10. I believe I could manage my shared files better using this file manager
    Helium__ Carbon__

11. Overall, I prefer: Helium__ Carbon__

Why?

Design Critique and Suggestions (Helium Only)



The following questions are about the "Helium" file manager (shown above). Indicate your level of agreement with each statement below:

		1	2	3	4	5	6	7	
1. I found it difficult to understand the information about shared files provided by the file manager.	strongly agree	O	O	O	O	O	O	O	strongly disagree
2. I was able to easily find the information I needed, about shared files, using this file manager.	strongly agree	O	O	O	O	O	O	O	strongly disagree
3. I found it difficult to learn to use this file manager.	strongly agree	O	O	O	O	O	O	O	strongly disagree
4. I was able to perceive information about shared files "at a glance" using this file manager	strongly agree	O	O	O	O	O	O	O	strongly disagree
5. The information provided by the file manager was ***not*** useful for completing the tasks.	strongly agree	O	O	O	O	O	O	O	strongly disagree
6. The information displayed by this file manager is presented in a non-intrusive manner.	strongly agree	O	O	O	O	O	O	O	strongly disagree
7. There is too much information displayed in this file manager.	strongly agree	O	O	O	O	O	O	O	strongly disagree
Comments:									

This final set of questions asks you to discuss what changes, if any, you would make to improve the features of the file manager (including the Sharing Console). You may also discuss any problems you had with these features; feel free to draw on this sheet.

Describe improvements for and/or problems with:

1. The file manager columns ("sent"; "accessed"; "shared with"), including pop-ups

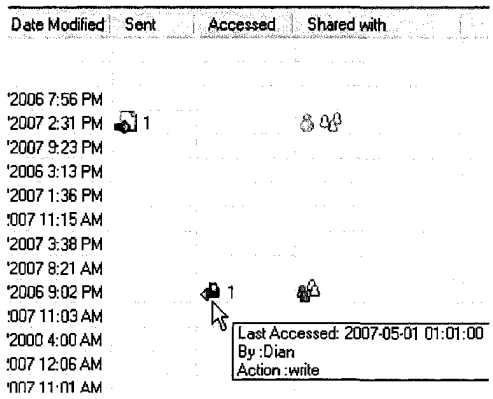
2. The search feature in the Sharing Console:

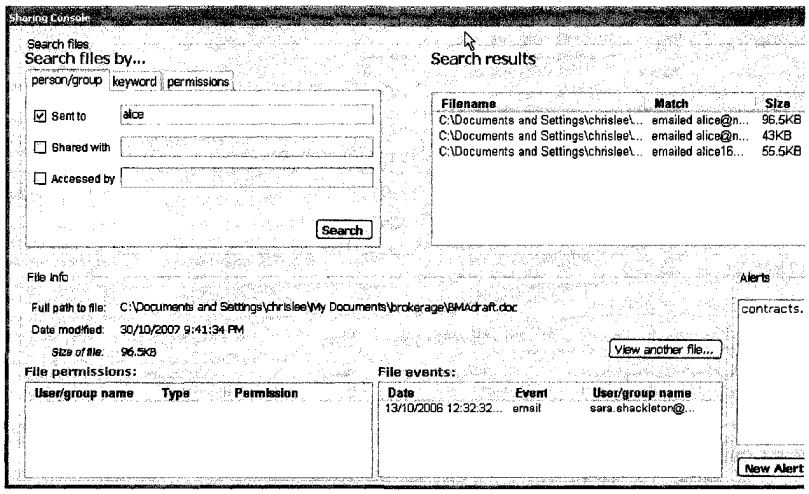
3. The file details in the Sharing Console:

4. The alert in the Sharing Console:

5. Are there any features you think were missing from the file manager and/or the console, that you think would help you to manage your shared files? If so, please describe:

Screenshots: these were embedded within the original questionnaire (landscape format) for reference and for drawing on; moved to bottom of this page due to space constraints.





Technology Acceptance Model (TAM) Questionnaire

	strongly disagree 1	2	3	4	5	6	strongly agree 7
1. Using this file manager in my job would enable me to accomplish tasks more quickly	o	o	o	o	o	o	o
2. Using this file manager would improve my job performance	o	o	o	o	o	o	o
3. Using this file manager in my job would increase my productivity	o	o	o	o	o	o	o
4. Using this file manager would enhance my effectiveness on the job	o	o	o	o	o	o	o
5. Using this file manager would make it easier to do my job	o	o	o	o	o	o	o
6. I would find this file manager useful in my job	o	o	o	o	o	o	o
7. Learning to operate this file manager would be easy for me.	o	o	o	o	o	o	o
8. I would find it easy to get this file manager to do what I want it to do	o	o	o	o	o	o	o
9. My interaction with this file manager would be clear and understandable	o	o	o	o	o	o	o
10. I would find this file manager to be flexible to interact with	o	o	o	o	o	o	o
11. It would be easy for me to become skilful at using this file manager	o	o	o	o	o	o	o
12. I would find this file manager easy to use	o	o	o	o	o	o	o
13. Assuming this file manager would be available on my job, I predict that I will use it on a regular basis	o	o	o	o	o	o	o