ABSTRACT

Title of Document: CHILDREN'S INTERFACE DESIGN FOR SEARCHING AND BROWSING

Hilary Browne Hutchinson,
Doctor of Philosophy, 2005

Directed By: Professor Benjamin B. Bederson,
Department of Computer Science

Elementary-age children are among the largest user groups of computers and the Internet, so it is important to design searching and browsing interfaces to support them. However, many interfaces for children do not consider their skills and preferences. Children can perform simple, single item searches, and are also capable of conducting Boolean searches involving multiple search criteria. However, they have difficulty creating Boolean searches using hierarchical structures found in many interfaces. These interfaces often employ a sequential presentation of the category structure, where only one branch or facet at a time can be explored. This combination of structure and presentation keeps the screen from becoming cluttered, but requires a lot of navigation to explore categories in different areas and an understanding of potentially abstract high-level categories.

Based on previous research with adults, I believed that a simultaneous presentation of a flat category structure, where users could explore multiple, single-layer categories simultaneously, would better facilitate searching and browsing for children. This

method reduces the amount of navigation and removes abstract categories. However, it introduces more visual clutter and sometimes the need for paging or scrolling. My research investigated these tradeoffs in two studies comparing searching and browsing in two interfaces with children in first, third, and fifth grade. Children did free browsing tasks, searched for a single item, and searched for two items to create conjunctive Boolean queries. The results indicate that a flat, simultaneous interface was significantly faster, easier, likeable, and preferred to a hierarchical, sequential interface for the Boolean search tasks. The simultaneous interface also allowed children to create significantly more conjunctive Boolean searches of multiple items while browsing than the sequential interface. These results suggest design guidelines for others who create children’s interfaces, and inform design changes in the interfaces used in the International Children’s Digital Library.

CHILDREN'S INTERFACE DESIGN FOR SEARCHING AND BROWSING

By

Hilary Browne Hutchinson

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2005

Advisory Committee:
Professor Benjamin B. Bederson, Chair
Professor Allison Druin
Professor Emeritus Jack Minker
Professor Kent Norman
Professor Ben Shneiderman

Acknowledgements

Completing a PhD takes a lot of individual work and sacrifice, but it also requires the support, understanding, and advice of family, friends, advisors, and colleagues. I'd like to thank my husband for tolerating numerous years of stress and angst, often about subject matter he didn't have any desire to understand. I'd like to thank my parents for arranging their lives around getting their children into the best public schools and private colleges money could buy, only to have both children flee to warmer climates for graduate school.

My PhD committee members have been a great source of support. I'd like to thank my advisor, Ben Bederson, for supporting me through not one but two PhD projects, a great job, assorted crazy trips to foreign countries, and large amounts of chocolate. I'd like to thank Allison Druin for essentially being my co-advisor, schooling me on how to work with children, and setting a wonderful example for women in technology everywhere. I'd like to thank Jack Minker for providing the fellowship that allowed me to complete my first two years of study, and for always checking up on my progress. I'd like to thank Kent Norman for his advice on experiment design and analysis, and on how to discipline unruly technology. I'd like to thank Doug Oard for his advice on digital library use and design. I'd like to thank Ben Shneiderman for always expressing an interest in anything I was doing, and for all his contributions to the field, which I've cited in every paper I've written. Although not an official member of my committee, I'd also like to thank Chip Denman, OIT Statistician Extraordinaire, for all his advice on how to analyze the results of my study.

The research study I conducted would not have been possible without the cooperation of the four elementary schools I worked with. I'd like to thank Bunny Egerton of the Columbia Association for putting me in touch with the students and staff in the after care programs at Northfield and Clemens Crossing elementary schools in Howard County. I'd like to thank Eric Sloan from Hillcrest elementary school for letting me work with the students in his technology classes. I'd like to thank Beth Hadley from Holy Trinity elementary school for allowing me to work with the after care program in her school. I'd also like to thank all the children who participated and their parents for allowing them to participate.

Finally, I'd like to acknowledge all of the wonderful staff and students who make up the Human-Computer Interaction Lab. The HCIL was the reason I chose to come to the University of Maryland, and the quality of the research and the people continue to make it a wonderful place to work. I'd especially like to thank Catherine Plaisant and Anne Rose, who do great research, keep the lab running, and are always up for a little ice cream to take the edge off.

Table of Contents

Acknowledgements ii
Table of Contents iii
List of Tables vii
List of Figures viii
Chapter 1: Introduction 1
1.1 Motivation 1
1.2 Research Contributions 5
1.2.1 Children's Use and Preference of Search and Browse Interfaces 5
1.2.2 Design Guidelines for Children's Search and Browse Interfaces 6
1.2.3 Working Examples of Interfaces 6
Chapter 2: ICDL Background 8
2.1 Project Description 8
2.2 Research Issues 14
2.3 Software Implementation 16
2.3.1 History 16
2.3.2 ICDL Architecture 17
2.3.3 Software Enhancements 19
Chapter 3: ICDL Interface Design Research 25
3.1 Early Designs 25
3.2 Simple Interface 29
3.3 Advanced Interface 34
Chapter 4: Related Work 38
4.1 Children, Computers, and the Internet 38
4.1.1 Computer and Internet Use by Children: Growth and Concerns 38
4.1.2 Child Development and Computers 42
4.1.3 Children as Computer Users, Testers, Informants, and Partners 47
4.2 Information Visualization for Searching and Browsing Interfaces 51
4.2.1 Psychology of Information Visualization 51
4.2.2 Interface Techniques for Browsing 54
4.2.3 Structure and Presentation in Category Browsers 57
4.2.4 Hierarchies vs. Other Forms of Organization for Children 62
4.2.5 Boolean Search 65
4.2.6 Paging vs. Scrolling on Computer Screens 69
4.2.7 Design of Icons 71
4.3 Digital Libraries for Children 73
4.3.1 Book Selection 73
4.3.2 Category Browsing vs. Keyword Searching 75
4.3.3 Previous Interface Solutions 78
4.3.4 ICDL Interface Solutions 79
4.3.5 Other Current Digital Library Solutions 80
Chapter 5: Controlled Studies 83
5.1 Research Goals and Questions 83

5.2 Hypotheses....84
5.3 Participants....86
5.4 Materials....87
5.4.1 Permission and Assent Forms....87
5.4.2 Interfaces....88
5.4.3 Technology....90
5.4.4 Experimenter Worksheets....91
5.5 Procedure....91
5.6 Hints....94
5.7 Simple Search Tasks....97
5.8 Boolean Search Tasks....98
5.9 Pilot Testing....99
5.10 Analysis Methodology....100
5.10.1 Timed Tasks....101
5.10.2 Difficulty and Like/Dislike Questions....104
5.10.3 Preference Questions....104
5.10.4 Hint Counts....105
5.10.5 Understanding of Boolean Question....105
5.10.6 Browsing Boolean Search Creation....106
5.10.7 Browsing Navigation Button Use....107
5.11 Results: Simple Study....107
5.11.1 Overall....107
5.11.2 Grade 1....116
5.11.3 Grade 3....120
5.11.4 Grade 5....124
5.12 Results: Boolean Study....128
5.12.1 Overall....128
5.12.2 Grade 1....139
5.12.3 Grade 3....143
5.12.4 Grade 5....147
5.13 Results: Usability Issues....151
5.13.1 Understanding Top-Level Categories....152
5.13.2 Confusion of Similar Categories....152
5.13.3 Activation of Category Buttons....153
5.13.4 Icon Design....153
5.13.5 Color Buttons....154
5.13.6 Difficult Words....154
5.14 Results: Most Popular Categories....154
5.14.1 Overall....155
5.14.2 By Gender....156
5.14.3 Grade 1....157
5.14.4 Grade 3....157
5.14.5 Grade 5....158
5.15 Discussion of Results....159
5.15.1 Simple Searches....159
5.15.2 Boolean Searches....160

5.15.3 Browsing Boolean Search Creation 162
5.15.4 Browsing Navigation Tools 162
5.15.5 Usability Issues 163
5.15.6 Most Popular Categories 164
5.16 Limitations of Results 165
5.16.1 Interface Design 165
5.16.2 Structure and Presentation 166
5.16.3 Boolean Logic 166
5.16.4 Browsing 166
5.16.5 Cultural Differences 167
5.16.6 Statistical Power 167
Chapter 6: Contributions 168
6.1 Children's Use and Preference of Search and Browse Interfaces 168
6.1.1 Contributions by age group 168
6.1.2 Contributions for simple searches 169
6.1.3 Contributions for Boolean searches 170
6.1.4 Contributions for comprehension of Boolean search 171
6.1.5 Contributions for browsing 171
6.1.6 Contributions for book selection 172
6.1.7 Contributions for other searching and browsing tools 172
6.2 Design Guidelines 173
6.2.1 Using category browsers 173
6.2.2 Designing classification schemes 174
6.2.3 Designing navigation schemes 174
6.2.4 Supporting simple searches 175
6.2.5 Supporting Boolean searches 175
6.2.6 Supporting browsing 176
6.3 Working Examples 176
6.3.1 ICDL Servlet architecture 176
6.3.2 Simultaneous interface design 177
6.3.3 Sequential interface design 179
Chapter 7: Future Work 181
7.1 Research Questions 181
7.1.1 Interface Design 181
7.1.2 Structure and Presentation 181
7.1.3 Boolean Logic 182
7.1.4 Browsing 183
7.1.5 Cultural Differences 184
7.1.6 Statistical Power 184
7.2 ICDL Design Changes 185
Appendix 1: Letter to Parents 187
Appendix 2: Parental Permission Form 188
Appendix 3: Child Assent Form 189
Appendix 4: General Experiment Instructions 190
Appendix 5: Simultaneous Interface Instructions 191
Appendix 6: Sequential Interface Instructions 192

Appendix 7: Simple Experiment Tasks 193
Appendix 8: Boolean Experiment Tasks 194
Appendix 9: Simple Experiment Worksheet 195
Appendix 10: Boolean Experiment Worksheet 197
Bibliography 199

List of Tables

Table 1. Combinations of structure and presentation 59
Table 2. Hint protocol for not remembering where the book count was found 96
Table 3. Hint protocol for selecting one or more wrong categories 96
Table 4. Hint protocol for forgetting the More Choices button in the Simultaneous interface 97
Table 5. Hint protocol for not knowing which top-level category to look under........97
Table 6. Simple tasks, phrased “How many … are there?” 115
Table 7. Boolean tasks, phrased “How many … are there?” 138
Table 8. Top categories selected overall 155
Table 9. Top categories selected by boys 156
Table 10. Top categories selected by girls 156
Table 11. Top categories selected by 1st graders 157
Table 12. Top categories selected by 3rd graders 158
Table 13. Top categories selected by 5th graders 159

List of Figures

Figure 1. The world search of the ICDL Basic interface..10
Figure 2. The old category browser of the ICDL Basic Interface 11
Figure 3. Search results page in the ICDL Basic interface...11
Figure 4. Book preview page in the ICDL Basic interface 12
Figure 5. The Standard book reader in the ICDL Basic interface 12
Figure 6. The Comic book reader in the ICDL Basic interface..................................13
Figure 7. The Spiral book reader in the ICDL Basic Interface..................................13
Figure 8. *Where's the Bear?* with metadata information in English 21
Figure 9. *Where's the Bear?* with metadata information in Japanese 21
Figure 10. *Where's the Bear?* shown with an Arabic interface.................................22
Figure 11. Book Preview page with button to add book to bookshelf 24
Figure 12. Bookshelf page showing favorite books saved by a user...........................24
Figure 13. Early simultaneous interface idea. 27
Figure 14. The Simple interface 28
Figure 15. The Advanced interface 28
Figure 16. The Simple interface after Rainbow and Fairy Tales have been selected..32
Figure 17. The Advanced interface with Color selected 35
Figure 18. The Advanced Interface with Red selected...36
Figure 19. The first page of the simultaneous test interface.......................................88
Figure 20. The second page of the simultaneous test interface 89
Figure 21. The top level of the sequential test interface...89
Figure 22. The leaf level of the color category in the sequential test interface...........90
Figure 23. Average time per task by grade for simple tasks (n=36) 109
Figure 24. Average time per task by interface and task type for simple tasks (n=36) 109
Figure 25. Difficulty by interface for simple tasks (n=36).......................................110
Figure 26. Like/dislike by interface for simple tasks (n=36) 111
Figure 27. Preference for simple tasks (n=36) 112
Figure 28. Preference by grade for simple tasks (n=36) 112
Figure 29. Average time to complete individual simple tasks (n=36) 116
Figure 30. Total hints needed to complete individual simple tasks (n=36)...............116
Figure 31. Average time per task by interface and task type for simple tasks for 1st grade (n=12) 117
Figure 32. Difficulty by interface for simple tasks for 1st grade (n=12) 118
Figure 33. Like/dislike by interface for simple tasks for 1st grade (n=12) 119
Figure 34. Average time per task by interface and task type for simple tasks for 3rd grade (n=12) 121
Figure 35. Difficulty by interface for simple tasks for 3rd grade (n=12) 122
Figure 36. Like/dislike by interface for simple tasks for 3rd grade (n=12)................123
Figure 37. Average time per task by interface and task type for simple tasks for 5th grade (n=12) 125
Figure 38. Difficulty by interface for simple tasks for 5th grade (n=12) 126
Figure 39. Like/dislike by interface for simple tasks for 5th grade (n=12)................127

Figure 40. Average time per task by interface for Boolean tasks (n=36)..................129
Figure 41. Average time per task by grade for Boolean tasks (n=36).......................130
Figure 42. Average time per task by interface and task type for Boolean tasks (n=36) 130
Figure 43. Difficulty by interface for Boolean tasks (n=36).....................................131
Figure 44. Interface like/dislike for Boolean tasks (n=36).......................................132
Figure 45. Preference for Boolean tasks (n=36)..133
Figure 46. Preference by grade for Boolean tasks (n=36).......................................133
Figure 47. Understood Boolean task by interface and grade for Boolean tasks (n=36) 136
Figure 48. Average time by interface and task number for Boolean tasks (n=36)....139
Figure 49. Total hints by interface and task number for Boolean tasks (n=36)........139
Figure 50. Average time per task by interface and task types for Boolean tasks for 1st grade (n=12) 140
Figure 51. Difficulty by interface for Boolean tasks for 1st grade (n=12).................141
Figure 52. Like/dislike by interface for Boolean tasks for 1st grade (n=12) 142
Figure 53. Average time per task by interface and task type for Boolean tasks for 3rd grade (n=12) 144
Figure 54. Difficulty by interface for Boolean tasks for 3rd grade (n=12) 145
Figure 55. Like/dislike by interface for Boolean tasks for 3rd grade (n=12).............146
Figure 56. Average time per task by interface and task type for Boolean tasks for 5th grade (n=12) 148
Figure 57. Difficulty by interface for Boolean tasks for 5th grade (n=12) 149
Figure 58. Like/dislike by interface for Boolean tasks for 5th grade (n=12).............150

Chapter 1: Introduction

1.1 Motivation

As computers and the Internet continue to make their way more and more into everyday life, one of the largest groups of users is elementary-age children. In 2002, 83% of U.S. homes with children owned a computer and 78% accessed the Internet (CPB, 2002). Nearly 20% of children as young as 3 and 4, 42% of children age 5-9, and 67% of children age 10-13 now use the Internet, (NTIA, 2004). One of the most common ways children use the Internet is for schoolwork, and search engines and digital libraries are popular ways that children can search and browse for information for their assignments. Children spend time playing games and communicating with each other using computers as well, and these activities often require searching and browsing (CPB, 2002). Children also use the Internet for shopping, requiring searching and browsing for merchandise, and they have an enormous impact on the buying decisions of their parents (NIMF, 2000, 2002).

Web sites such as Yahooligans! (yahooligans.yahoo.com) and Ask Jeeves Kids (www.ajkids.com/) are examples of portals that children can use to find age-appropriate content for school projects or consumer purchases. Project Gutenberg (www.promo.net/pg/) and the Rosetta Project (www.childrensbooksonline.org/) are examples of digital libraries that provide access to scans of out-of-copyright children's books from around the world. However, these and many other websites have interfaces with one or more of three crucial problems. First, they do not take into account the information processing and motor skills of children, specifically their difficulties selecting small icons and text links with a mouse. Second, they do not consider children's searching and browsing skills, specifically their difficulties with spelling, typing, navigating, and composing queries. Third, they do not consider how children prefer to search for things, presenting searching and browsing criteria appropriate for adults but not for children. The ability to select content such as reading material on their own is a powerful motivator for children (Kragler and Nolley, 1996), and many of these websites prevent children from doing so.

Recent work in the Human-Computer Interaction Lab at the University of Maryland has focused on designing digital library interfaces that support and scaffold young children's abilities to search and browse for information. The QueryKids interface allowed children to find multimedia information about animals in a zooming user interface (Druin et al., 2001). This interface was scaffolded with large, easily clickable icons rather than a keyword search box that required typing; incremental and clearly visible results to show progress as searches were constructed; and a built-in Boolean protocol to prevent children from having to mentally construct Boolean queries manually. Search categories were based on how children liked to look for animals, such as what they ate or where they lived. Revelle et al. (2002) found that 2nd and 3rd grade children were successfully able to use this interface when prompted to conduct both simple (non-Boolean) and Boolean queries 85% of the time.

Based on the success of the QueryKids interface, the International Children's Digital Library (ICDL, www.icdlbooks.org) software was built using a similar interface with the addition of a hierarchical category browser to allow children to find and read books online (Druin et al., 2003). This interface also consisted of large, easily clickable buttons, automatically constructed Boolean searches, and search categories based on how children like to look for books. Reuter and Druin (2004) found that children in grades 1-5 were able to navigate the category structure to find books in open-ended browsing, but they did not generally use the Boolean capability.

Based on the results of these studies and after observing the use of the ICDL over several years, it is my hypothesis that the structure and presentation of the ICDL category browser discouraged children from creating Boolean searches. While Boolean search is known to be difficult for both children and adults, children are capable of using it both digitally and otherwise (Neimark and Slotnick, 1970; Tversky and Kahneman, 1975). The ICDL category browser was structured using *faceted metadata* (English et al., 2002), a collection of independent classifiers such as shape, color, and genre, each of which was hierarchical in structure. The categories were presented using *sequential or hierarchical menus* (hereafter referred to as sequential) (Norman, 1991; Hochheiser et al., 2000). These presentations only allow users to explore one facet at a time. Creating a Boolean query required navigating to the leaves of one facet and selecting one, backtracking to the top of the hierarchy, and then navigating to the leaves of another facet to add another leaf. In addition, some of the top-level categories in the facets were rather abstract (e.g. Format, Genre), and young children may not have understood them (Rosch et al., 1976). I believed these problems might be alleviated by making two key changes to the category browser, one in the structure and one in the presentation.

For structure, I suggested that collapsing the depth of the hierarchical categories might be easier for children to navigate, as has been found for adults (Miller, 1981). Children also may not naturally use hierarchical categorization (Piaget and Inhelder, 1969) and may have trouble understanding abstract, top-level categories in a hierarchy (Rosch et al., 1976). For presentation, I suggested using *simultaneous menus* (Norman, 1991; Hochheiser et al., 2000), where each facet or branch in a category structure can be explored in parallel. For adults, this design was found to be faster when creating complex queries that required backtracking in the sequential menu design (Hochheiser et al., 2000). However, these two changes yield a design with more categories on the same page, which may be visually overwhelming (Hochheiser et al., 2000). Additionally, it is possible that not all the categories will fit on the screen because of the need to use large, easily clickable category icons, necessitating paging or scrolling to view additional categories.

My research sought to investigate the tradeoffs for children completing searching and browsing tasks between the backtracking and top-level category comprehension required for hierarchical faceted structures presented sequentially and the visual scanning and paging or scrolling required for flattened faceted structures presented simultaneously. Until now, no studies have looked systematically at how children of different ages are able to use hierarchical and faceted structures, simultaneous and sequential menus, and Boolean logic in interfaces designed to support their abilities. This dissertation describes the results of two studies designed to help fill this void.

1.2 Research Contributions

1.2.1 Children's Use and Preference of Search and Browse Interfaces

The major contribution of this research is an analysis of elementary-age children's use and preference of two different combinations of structure and presentation in category searching and browsing interfaces (see Section 6.1). Previous research with adults has explored different combinations of both structure (e.g. Miller, 1981) and presentation (e.g. Hochheiser et al., 2000) with both simple and complex searching tasks. Previous research with children has compared one combination of structure and presentation (sequential hierarchy) to keyword interfaces (e.g. Borgman et al., 1995), and explored simple and complex searching tasks (e.g. Revelle et al., 2002), and open-ended browsing (e.g. Reuter and Druin, 2004), also with sequential hierarchies. However, previous research has not compared different combinations of structure and presentation for children.

I evaluated two combinations of structure and presentation (sequential hierarchy and simultaneous flat) for two types of searching tasks (simple, one-item searches and Boolean, two-item conjunctive searches) as well as open-ended browsing across three different age groups. I report on statistically significant differences in both searching and browsing behavior, as well as qualitative observations and usability issues. I present specific contributions relating to how children of different ages prefer and are able to conduct and understand searching tasks with these interfaces, and how different task types and searching vs. browsing activities influence performance and preference.

These results may be generalizable to other searching and browsing interfaces for children, such as digital libraries, search engines, and e-commerce applications that allow children to browse using categories. These results are not scaleable to large numbers of categories, which would require either placing many categories on the screen, or large amounts of paging or hierarchical navigation to reach many of the categories. However, young children’s shorter attention spans, slower visual information processing speeds, and smaller memory capacities suggest that large numbers of category choices would not be appropriate for children anyway (Baumgarten, 2003; Kail, 1991; Chi, 1976).

1.2.2 Design Guidelines for Children’s Search and Browse Interfaces

As a second contribution, I present design guidelines for designers of children’s searching and browsing software (see Section 6.2). I suggest interface design choices for classification and navigation schemes based on previous research as well as statistical results and qualitative observations from my studies. I also suggest choices for category browser structure and presentation based on my study results depending on whether the target tasks are simple searches, Boolean searches or casual browsing.

1.2.3 Working Examples of Interfaces

The final contributions of this research are the ICDL Servlet technology that I developed and the interface design ideas that I created, implemented, and tested for the ICDL and adapted for the studies (see Section 6.3). The Servlet technology represents two and a half years of development activity and nearly 100 Java class files of approximately 16,000 lines of code running in a live application that supports roughly 25,000 visitors a month from 155 countries. This code connects to a mySQL database maintained by another ICDL project team member containing metadata information for over 800 books of approximately 50,000 total pages, as well as information about search categories and user profiles. The interfaces for the studies were derived from the current ICDL searching and browsing architecture that I helped design, build, and test, and the results of the study will feed back into revisions that will be deployed in the live ICDL software. In addition, I expect that designers of other interfaces will be able to use the ideas from my architecture and interface designs to create and improve their own tools.

Chapter 2: ICDL Background

2.1 Project Description

My research is part of the International Children's Digital Library (ICDL), a 5 year research project initiated in 2002 and funded by the National Science Foundation and the Institute for Museum and Library Services. The ICDL is led by Professor Allison Druin at the University of Maryland's Human-Computer Interaction Lab (HCIL). I am currently a full-time developer and part-time graduate student working on this project. The ICDL has five primary goals, stated on the project website as:

- to create a collection of more than 10,000 books in at least 100 languages that is freely available to children, teachers, librarians, parents, and scholars throughout the world via the Internet;
- to collaborate with children as design partners in the development of computer interface technologies that support children in searching, browsing, reading, and sharing books in electronic form;
- to better understand the concepts of rights management and "fair use" in a digital age;
- to evaluate the impact that access to digital materials may have on collection development and programming practices in school and public libraries;
- to develop a greater understanding of the relationship between children's access to a digital collection of multicultural materials and children's attitudes toward books, libraries, reading, technology, and other countries and cultures.

The project has two main audiences: children age 3-13 and adults such as teachers and librarians who work with them, as well as international scholars who study children's literature. The project draws together an interdisciplinary team of researchers from computer science, information studies, education, and art backgrounds. The research team is also intergenerational – team members also include 6 children age 7-11 who work with the adult members of the team twice a week during the school year and for 2 weeks in the summer to help design the software for the ICDL and other research projects in the lab.

The ICDL initially consisted of two interfaces for accessing the current collection of roughly 800 books in 32 languages. The "Enhanced" interface was a Java application that could be run over the Internet using the Java WebStart plug-in and a broadband connection. The "Basic" interface, which I have been the primary front-end developer of, is implemented with Java Servlets on the server side and HTML and JavaScript on the client side and runs well even on a 56K modem. The Enhanced interface, launched in November 2002 when the project first went live, was phased out over the last year due to its advanced technology requirements and difficulty supporting multi-lingual interfaces. The Basic interface, launched in May, 2003, is based on the same design principles but is more accessible.

In the initial implementation of the Basic interface, users could search for books in three different ways. They could spin a globe using a large, easily clickable arrow and then select a continent to see books from, about, or set in that continent (Figure 1).

They could use a category browser of 14 hierarchical facets and navigate down 2 to 4 levels to select a single leaf-level category (e.g. the color red) (Figure 2). Finally, they could use keywords to find books with matching metadata in title, author, summary, and publication information. All of these methods searched for books with matching metadata and returned a list of books, presented with thumbnail images of their covers (Figure 3). Users could then select a book and get more information about it on a preview page, such as a summary and authors (Figure 4). Finally, users could choose to read the book using one of three book readers – the Standard reader that presents pages one at a time in HTML (Figure 5), the Comic reader that presents an overview of all the pages using Java WebStart (Figure 6), or the Spiral reader that presents the pages in a spiral using Java WebStart (Figure 7).

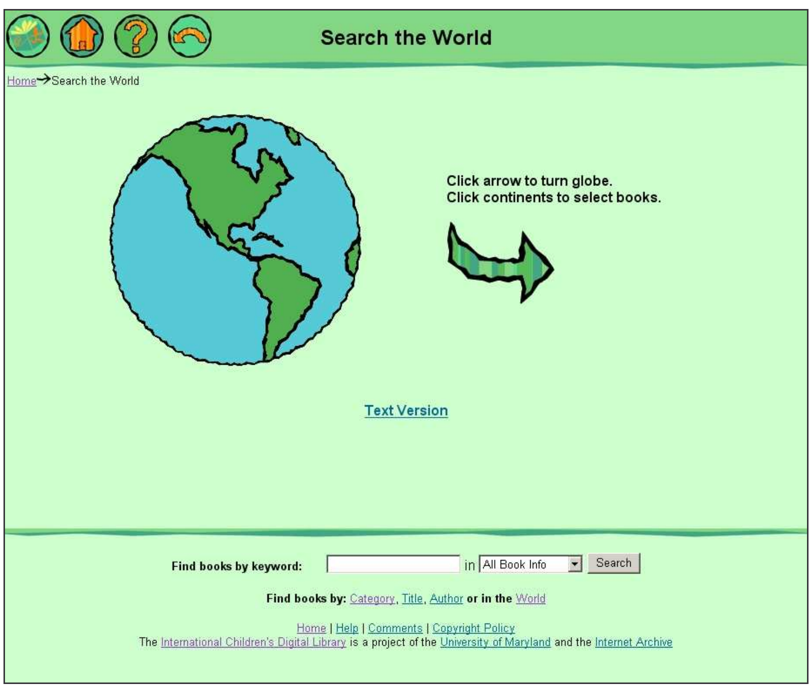

Figure 1. The world search of the ICDL Basic interface

Figure 2. The old category browser of the ICDL Basic Interface

Figure 3. Search results page in the ICDL Basic interface

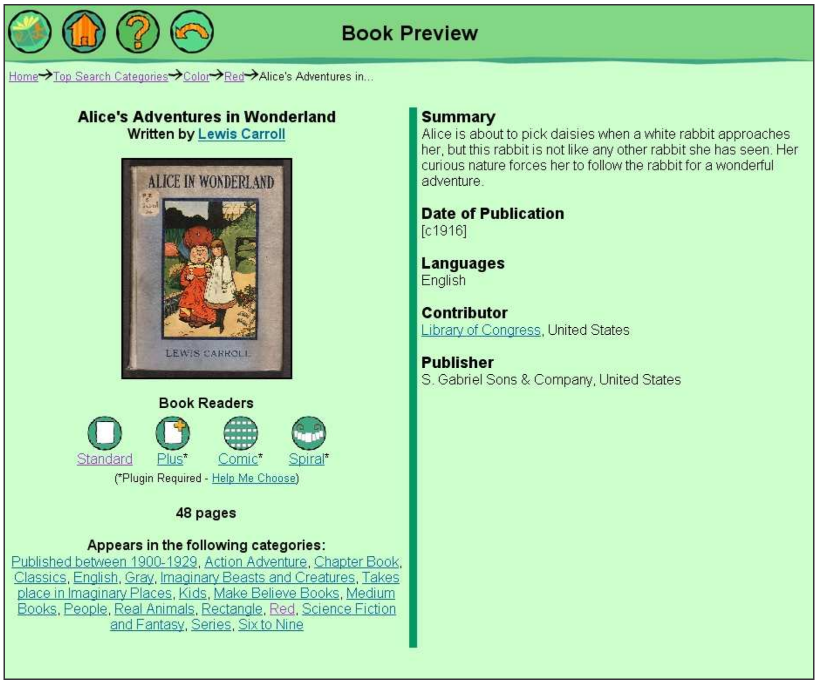

Figure 4. Book preview page in the ICDL Basic interface

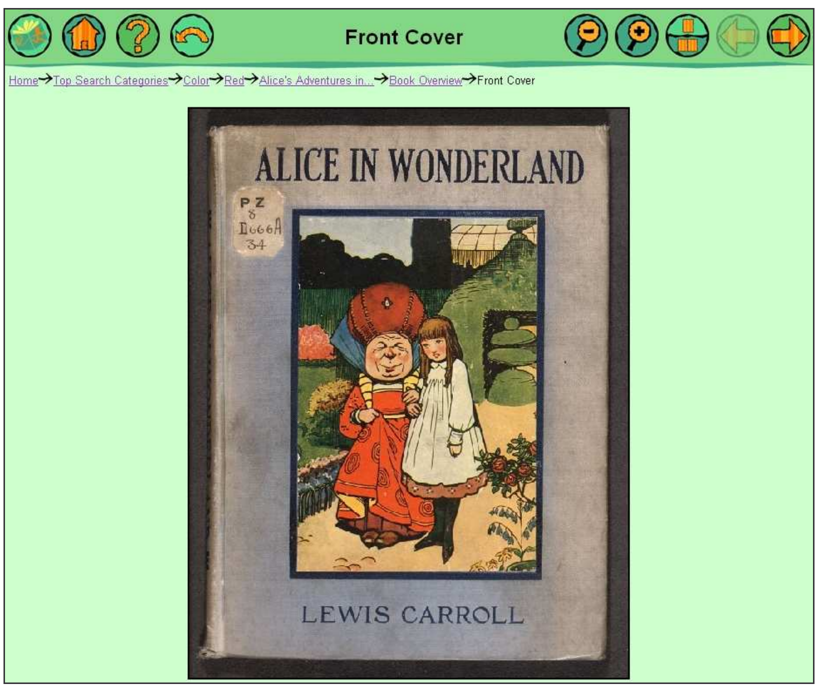

Figure 5. The Standard book reader in the ICDL Basic interface

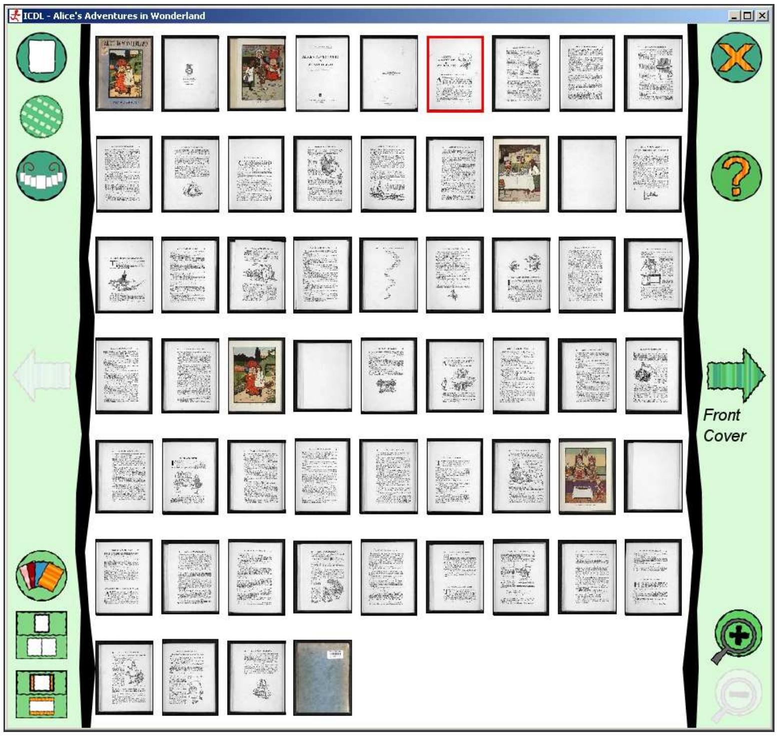

Figure 6. The Comic book reader in the ICDL Basic interface

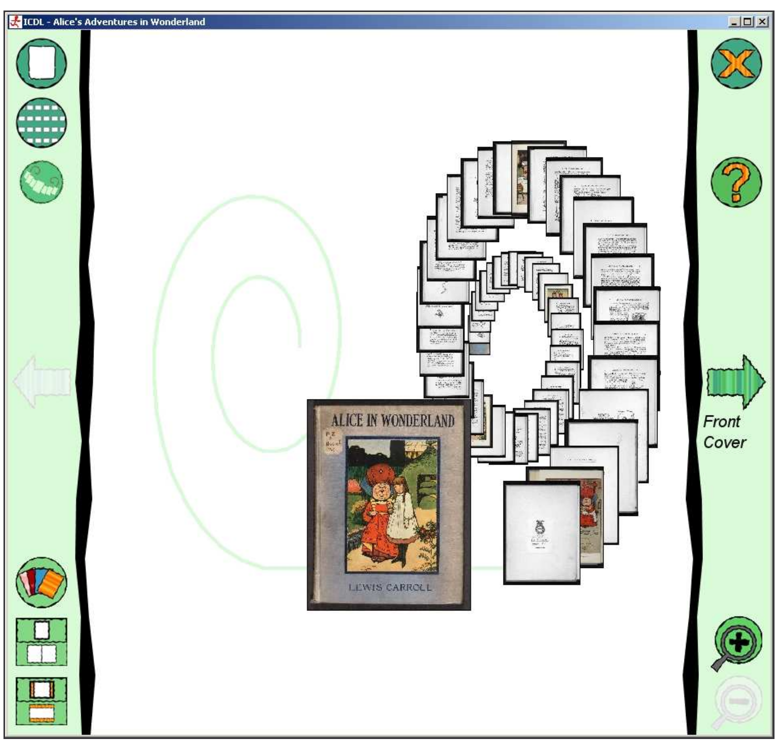

Figure 7. The Spiral book reader in the ICDL Basic Interface

2.2 Research Issues

The ICDL is an extremely fruitful research project. Internationally, our audience is the entire world, which means our users speak different languages and have different customs. This has implications across all aspects of the project. In selecting books to include, our librarians have to deal with different copyright rules for different countries and publishers. Once books are selected, our metadata team has to assist the contributors in providing metadata about the book in its native language and also in English if possible, as well as coordinating a team of volunteer translators to fill in the gaps left by our contributors. Our technology team has to store, process, and deliver book metadata and interface tools in multiple languages, and assist our users to display these languages in their web browsers. Finally, our advisory board keeps tabs on issues of interpretation. They make sure icons, terminology, and book content are understandable and not offensive culturally, religiously, socially, or politically (Hutchinson et al., 2005b).

In addition to being international, our target audience also includes children age 3-13, so we have to design our interface to accommodate their skills and preferences. The interface is icon-based rather than keyword-based, unlike many other digital search environments, and the icons are designed to be large enough so young children can easily click on them. Our interface also provides multiple ways of searching or browsing for books, geared toward different age groups. Previous research indicates that young children prefer the simplicity and concreteness of spinning the globe in the world interface, while older children prefer the category interface, with category choices geared toward their searching preferences (Reuter and Druin, 2004).

While the ICDL is a research project, it is also a service project, and in that regard, we are sensitive to the differing degrees to which our users are digitally enabled. While some of our users are technologically savvy and connect to the Internet with broadband access, many more are computer or Internet novices connecting with 56k modems, often in public locations such as schools or libraries. As such, they may not have the skills or the permission to install browser plug-ins or download large web pages. As a result, much of our design work is focused on making the ICDL broadly accessible by users with different operating systems, browsers, connection speeds, computer skills, and accessibility.

Finally, the ICDL obviously raises many interesting research questions in the library realm. Children search for books in physical libraries differently than adults, and their behavior is similar in digital libraries (Reuter & Druin, 2004). While adults may be interested in bibliographic information such as title or author, children are more likely to focus on physical features of books such as colors and illustrations or genres such as fairy tales or adventures (e.g. Pejtersen, 1986; Cooper, 2002b; Busey and Doerr, 1993; Kragler and Nolley, 1996; Fleener et al., 1997; Reuter and Druin, 2004). The ICDL category browser was designed and continues to be revised to reflect the way children look for books and the terminology they use to do so.

2.3 Software Implementation

2.3.1 History

The ICDL was originally implemented as a Java application that could be downloaded and run over the Internet using the freely available Java WebStart plug-in and a broadband Internet connection. This Enhanced interface used zooming and animation to allow users to spin a globe or search in a category hierarchy for books. The Enhanced category browser allowed users to create Boolean searches by selecting more than one category. Categories with the same parent (e.g. Red and Blue, both colors) were combined disjunctively (or) while categories with different parents (e.g. Red and Happy) were combined conjunctively (and).

After launching the Enhanced version, the team quickly realized that many of our users were unable to install plug-ins and/or didn't have broadband access, so the decision was made to create a static, HTML-only version of the software, known as the Basic version. One member of the team started developing a Java program to generate these pages, but left the project before it was complete. I joined the project in February, 2003 and became the primary developer on the Basic software, released in May, 2003. This interface consisted of simple HTML and JavaScript running on a standard Apache HTTP web server. It used the same searching tools and designs of the Enhanced software but presented them in a format accessible to users with slower Internet connections or who couldn't install plug-ins. The category browser in the Basic interface did not support Boolean search because the team wanted to research how to improve this function before including it.

As the library grew, we knew that generating a static HTML page for every page of every book in the library would not be realistic. As a result, we decided to implement a dynamic version of the Basic software instead. Since our environment was already built using Java, we decided to use Java Servlets. I was the principal architect and programmer for this project, and in July 2003, we republished the Basic software using this technology. Java Servlet technology provides a way to build dynamic web applications using a request-response protocol that extends the standard HTTP request-response protocol. It is available for free on the Java website (www.java.com). Java Servlets are more scaleable and efficient than popular alternatives such as CGI scripting, and unlike both CGI and Microsoft Active Server Pages, Java Servlets are platform independent. Servlet code can either be embedded in an HTML page as script, which is then dynamically assembled into a Servlet class, or else written by extending the Java HttpServlet class to generate HTML. We chose to do the latter because it makes for more modular and reusable code. The downside is that we have to generate our HTML with "print" statements.

2.3.2 ICDL Architecture

The current ICDL Servlet application consists of a package of nearly 100 classes of approximately 16,000 lines of code that I wrote, plus several open source classes that I adapted (e.g. database connection pool). The architecture also includes a separate package of code for the Java book readers that I maintain, and several JavaScript and CSS files that I wrote to control client-side interaction and presentation in a consistent way across multiple browsers and platforms.

Java Servlets require a web server that supports them. There are a number of choices, but the most powerful, freely available one is Apache Tomcat (tomcat.apache.org). Tomcat can be run independently as a complete web server, or be integrated into the standard, more powerful Apache HTTP web server, which is what the ICDL chose to do because the team maintains an HTML-only website on the Apache side with information about the project. Both servers run on a dedicated Linux machine. The Servlets make use of standard JDBC drivers to connect to a MySQL database maintained by another team member to run queries against many of the 41 tables that contain information about books, categories, and users. The application currently supports roughly 25,000 visitors a month from 155 countries.

On startup, Tomcat can be configured to run a context listener class to initialize application-level variables and read application-level data structures into memory or external files. To reduce run-time calls to the database, I created a context listener that accesses the database to create hash tables for a number of commonly accessed structures. These include mappings from search category and book id numbers generated by web page requests to objects containing more information about these objects (e.g. icons and book titles). The context listener also builds a searchable index file of book metadata for each language that we have book metadata for in the library.

Both Apache and Tomcat can be configured to generate log files for web page accesses, but Tomcat also allows you to add special filter classes that intercept every Servlet page request, which you can then use to generate your own log files. I added a filter class to do this so that I could reject page requests from malicious web bots, and create separate log files with additional, application-specific information. In particular, ICDL users are able to register with the site and create personal accounts, where they are asked for demographic information including their age and gender. Using the log filter, I tag entries from users that are logged in with this information so we can analyze usage patterns for different demographic groups.

The remaining classes in the package are either Servlets or classes that support storage or manipulation of data objects in the Servlets. These include object classes for books and categories, comparator classes for sorting books (e.g. by title or author), database connection tools, and a library class of methods I wrote for generating HTML and application constructs (e.g. tables, images, and page headers and footers). The Servlet classes all inherit from a generic Servlet class that contains references to all the constructs built by the context listener on startup. The general design pattern for these classes is to accept an HTTP GET request for the page, read in the url parameters, and reject the request if the parameters are malformed or accept the request and generate HTML based on the information in the parameters.

2.3.3 Software Enhancements

Over the past two years, I have worked on many improvements to the ICDL software, including optimizing database calls with a connection pool, fixing memory leaks, adding the Comic and Spiral book readers as optional interface widgets, and adding an indexed text search. The text search required the integration of Apache Lucene, a freely available Java search engine library that allows you to index and search the content of your web site with advanced feature such as ranked, Boolean, and field-dependent searches. Using Lucene, the Servlet context listener builds indices of the book metadata in multiple languages, allowing users to conduct searches in any one of these languages.

In 2004, I worked with our database developer to redesign our database schema to store book metadata in multiple languages, and then redesigned the ICDL software to allow users to view and search for information about books in multiple languages. For instance, a user can look at metadata information about the book *Where's the Bear* (Harris, 1997) in both English (Figure 8) and Japanese (Figure 9). This was a challenging project because of the need to handle data from multiple character sets. Our solution was to use the Unicode character set, which contains a unique encoding for nearly every character in every language. On the server side, we had to make sure our database software and drivers were updated to be Unicode-compliant and our Servlet code specified that data be handled in the Unicode format. On the client side, we had to generate HTML pages with headers indicating that the content being delivered was encoded using Unicode, provide help pages to assist users in installing fonts for character sets not available on their computers, and design interface tools for searching, sorting, and changing the display language.

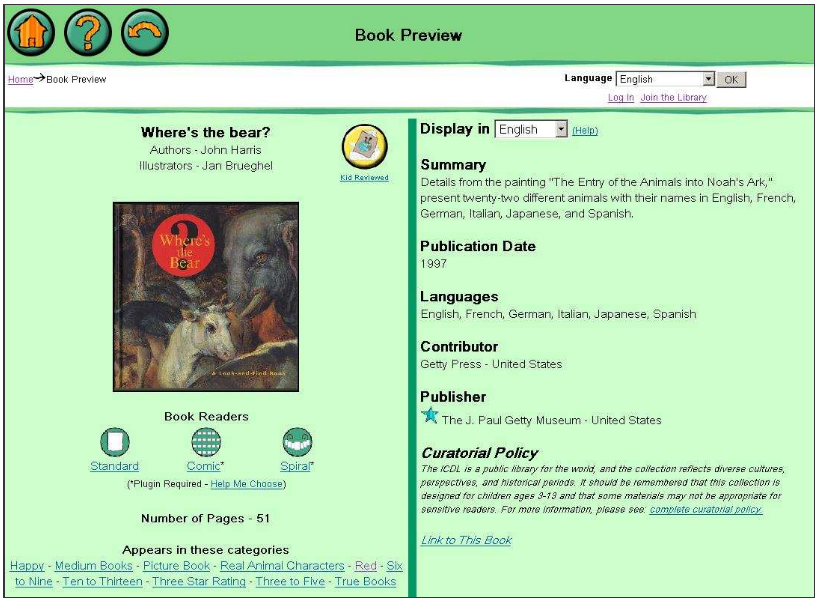

Figure 8. *Where's the Bear?* with metadata information in English

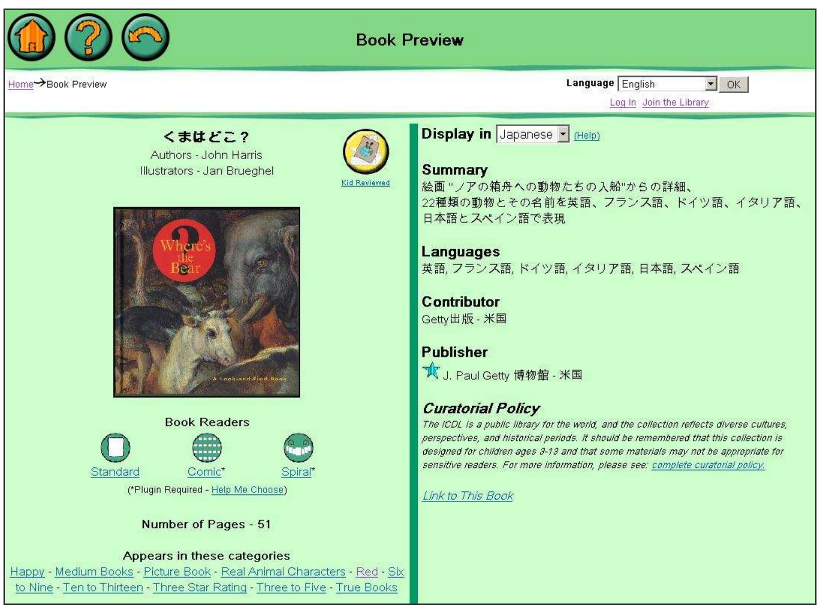

Figure 9. *Where's the Bear?* with metadata information in Japanese

In early 2005, I worked with our technology team to convert the ICDL interface into eight additional languages besides English: Arabic, Chinese, French, Filipino/Tagalog, German, Hebrew, Persian/Farsi, and Spanish. This was a multi-step process, which I began by extracting every interface word and phrase shown on the site and placing them in separate property files. We then had volunteers and a paid translation service translate the words and phrases into each of the eight languages. I then updated the Servlet code to use the translated words and phrases from the properties files, depending on the language a user selected to view the site. Finally, I updated the Servlet code to create HTML that would display pages properly in right to left languages (e.g. Arabic, Hebrew, and Persian/Farsi). To accomplish this in a general way without having to have two cases for every page on the site, I relied on a number of built-in features of HTML, most notably the RTL tag, which automatically does things like display the columns of a table right to left instead of left to right. For instance, Figure 10 shows the Book Preview page for *Where's the Bear* displayed in Arabic. Not only is the interface translated to Arabic, but also the entire layout of the page is mirrored from the English version to be read from right to left.

Figure 10. *Where's the Bear?* shown with an Arabic interface

At the same time we were translating the interface, I also implemented registration and log-in functionality on the site using persistent session information available in Java Servlets. This feature allows users to register with the site to select their preferred interface language and searching interface. This information is stored in our database and loaded into the user's Servlet session when they log in. Users can access their profile from any computer by logging in, and the demographic information they provide when they register (e.g. age, gender) is recorded in our log files as they access different pages, helping the ICDL team learn more about their users and tailor the interface according to their needs. The interface translation and registration features were both launched in May, 2005.

Most recently, I worked on allowing registered users to save their favorite books to a personal bookshelf. This change was bundled with a complete cosmetic redesign of the ICDL home page and informational portion of the website created by another team member and launched in October, 2005. I incorporated the new color scheme and icon designs into the library portion of site controlled by the Servlet code. For the bookshelf feature, I added database fields in the user table to store a list of books and a list of the last accessed page in each stored book. Users who are logged in can choose to add a book to their bookshelf with a button on the Book Preview page (Figure 11). After adding books to their shelf, users can select the bookshelf icon in the header of the page to access all of the books on their shelf (Figure 12). Users can also select background themes to customize their bookshelf with different monsters, who protect their books. Each time a user views a page in a book on their shelf, the

Book Page Servlet automatically records it as the most recently accessed page so that the user can return to where they left off at a later time.

Figure 11. Book Preview page with button to add book to bookshelf

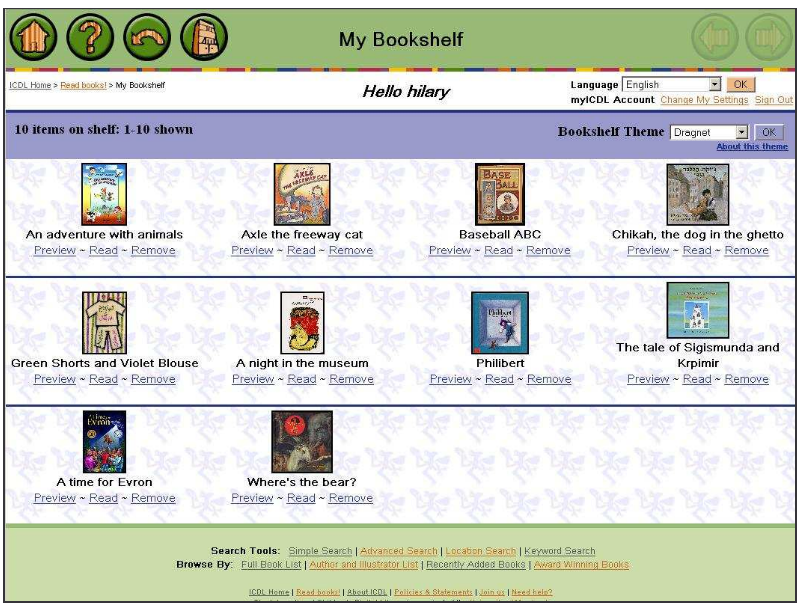

Figure 12. Bookshelf page showing favorite books saved by a user

Chapter 3: ICDL Interface Design Research

3.1 Early Designs

Based on our belief that the original category browser in the ICDL Enhanced interface was difficult for children to use to create Boolean searches, which was later confirmed by a research study conducted by our team (Reuter and Druin, 2004), we held off including Boolean search in the Basic category browser to research the issue further. This research issue became the focus of my dissertation. I analyzed and identified two major problems related to these concerns, plus two more design issues that I thought could be improved in the category browser design.

Structurally, our youngest users may be more inclined to think perceptually than hierarchically (Piaget and Inhelder, 1969; Nazzi and Gopnik, 2000; Deák et al., 2002). In order to find a leaf-level category in a hierarchy, users had to rely on their hierarchical knowledge. In addition, our youngest users may have had problems understanding the more abstract, top-level categories in the hierarchy. Many researchers have demonstrated that preschoolers and early elementary age children have difficulty categorizing and drawing inferences about high-level subjects (Rosch et al., 1976; Tversky, 1985; Gelman and O'Reilly, 1988). Presentation-wise, users could only explore one facet in the browser at a time, so navigation and backtracking were required to select leaves in different branches.

In addition to these two problems, previous research and our own observations indicated that children often did not differentiate between the leaves and the interior nodes in the facet hierarchies because they were visually identical (Reuter and Druin, 2004). This made it difficult to know whether clicking on a category would descend into the hierarchy or add the category to the current search. Finally, the results of a search in the Enhanced category browser were isolated in a small box at the top of the screen, where users might not know to inspect the results more closely. In the Basic version of the category browser, the results were on a different page altogether.

In the first prototype I developed to address these problems, rather than navigating each facet sequentially, users could open different facets simultaneously by clicking on them and having their leaves radiate from behind them (Hutchinson, 2004) (Figure 13). This interface provided a partial solution to the presentation problem by allowing users to view multiple facets simultaneously. However, it did not address the other problems. The interior and leaf nodes were still difficult to tell apart, and there was not enough room on the screen for results. In pilot testing with our kids team and with some pre-school age children at the university's Center for Young Children, I also found that the animation was distracting. Finally, it was unwieldy to maintain the JavaScript required to implement it and obviously would not work for all browsers.

Figure 13. Early simultaneous interface idea.

Based on the problems with this interface, the adult team members held a number of meetings and design sessions in the winter of 2004 to come up with other ideas for new searching interfaces. These included a treasure hunt, a book building tool, and a design I created. This last interface, now called "Simple," consisted of a ring of category icons arranged around a collection of books (Figure 14). These categories were presented as simultaneous menus, and the collection of books shown matched the categories selected. Over the next few months, together with our kids team partners, we critiqued initial sketches of these designs, brainstormed about improvements, and sketched our own versions of new features. While the initial treasure hunt and book building ideas continue to evolve, the Simple interface was an immediate hit, so I chose to pursue this one for immediate implementation. In addition, I also decided to create a more adult version of the Simple interface using text instead of icons and a larger collection of categories. This interface is now called "Advanced" (Figure 15).

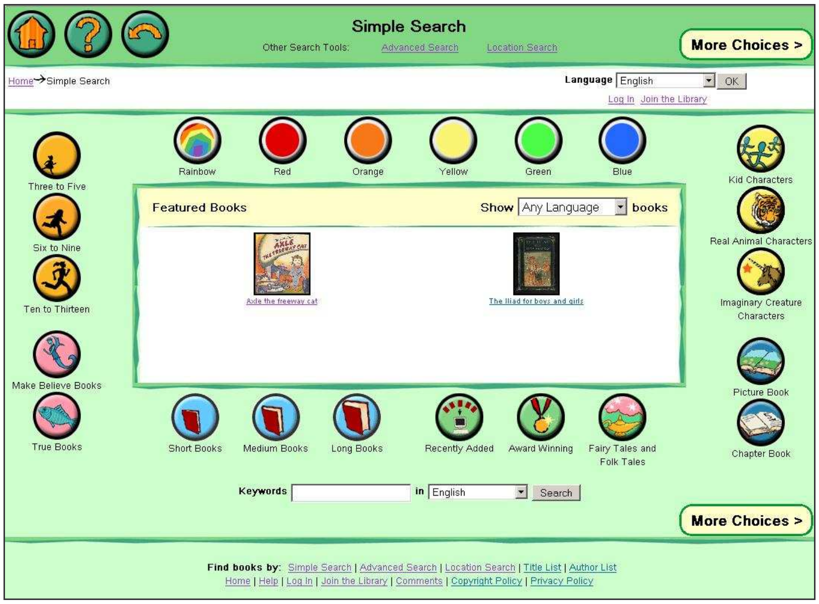

Figure 14. The Simple interface

Figure 15. The Advanced interface

3.2 Simple Interface

The design goal for the Simple interface was to create a tool that elementary-age children could use on their own or with some assistance from an adult. The original ICDL category browser had already addressed issues relating to large icon sizes to support developing motor skills and age-appropriate category choices to support children's searching preferences. What was missing was more attention to children's searching and browsing skills. Structurally, I flattened the hierarchies in each of the category facets to a single layer. I then presented a subset of the most popular leaf-level category icons as buttons that functioned as simultaneous menus, arranged around the perimeter of a box showing matching books. The selected categories, which change to a depressed version when clicked, are joined conjunctively, so the Boolean capability is limited to conjunctive 'and' searches. Clicking a selected category button unselects it and removes it from the search.

I chose to support only conjunctive Boolean searches for three reasons. First, a number of studies indicate that children have an easier time with conjunction than with disjunction (Neimark et al., 1970; Bloom et al., 1980). Second, the goal of the interface is to narrow down the number of books from a large collection so that children can easily select from a few books. Disjunction will increase or keep constant the number of results while conjunction will decrease or keep constant the number of results. Finally, while the Enhanced ICDL used conjunction between categories and disjunction within category groups, I felt that this would be confusing in an interface where all the categories appear on the same level. When categories are selected, they are combined in an “equation” across the top of the results section to indicate that their combination adds up to the count of the results. This visual tool makes the effect of selecting multiple categories concrete, which is important for children learning to reason logically. Trying to indicate both conjunction and disjunction in this equation would be difficult.

For the design of the category icons, I used round icons rather than the existing rectangular ones because children sometimes got confused about whether they were looking at categories or books since both were rectangles in the old category browser (Reuter and Druin, 2004). Frequent observation of children using software also informed my choice to implement a JavaScript progress bar in the results section as searches are built. I observed that children are impatient if an interface does not respond immediately, and may click a button multiple times if they don’t get immediate feedback, generating undesired or unpredictable results. For users with slow Internet connections or days when the software is receiving a lot of traffic, searches may not be instantaneous. The progress bar lets children know that their action has worked and that the results will appear momentarily.

Placing many icons on the same page meant that they needed to be as small as possible so I could fit a lot on the page, but not so small that they were difficult to click. Hourcade et al. (2003) found that 64 pixel icons are sufficient for children as young as 4 to be able to click, so I chose this size. I followed the advice of Plaisant et al. (1997) to bring the “treasures” of the library to the surface by having books appear on the same page with the search tools. I chose to place the books in the middle of the page, rather than having categories on one side and result on the other, as is common in other interfaces, for two reasons. First, I felt that the books were the most important part of the interface and should be the main focus of the page. Even if a user doesn't understand how to use the categories, it is clear that the books are the important part of the page. Second, the inspiration for this design, which I originally called "Fisher Price", came from my observation of toys for young children, which often have a central feature with large buttons around the outside. I felt that using this familiar design might make children more comfortable with the interface. This design turned out to be a nice choice when we translated the interface to languages that are read right to left, because the metaphor remains the same. The downside is that the categories span the far edges of the entire screen, requiring a lot of visual scanning to view all of them.

When no categories are selected, a group of 2 or 3 featured books appears (Figure 14). The results are incrementally updated whenever new categories are added or removed from the search. For instance, if a user selects Rainbow and Fairy Tales, the results show books that match both of these categories (Figure 16). Categories that these results do not appear in are grayed out and unclickable, while categories that these results do appear in remain selectable. This design prevents the creation of no-hit searches.

Figure 16. The Simple interface after Rainbow and Fairy Tales have been selected

Even when the results contain only one book, the remaining categories that this book appears in can still be added to the search. We did this because pilot testing with our kids team indicated that a favorite activity was seeing how many categories could be added to the search. The children frequently would not look at the book(s) selected until they had systematically gone through the interface and added all the possible categories. Keeping these categories selectable also indicates which other categories a book appears in. In addition to selecting categories, users can also refine their search by including keywords and limiting the results to a particular language. The keyword appears as part of the search equation in the results section. The language selection menu is always present in the equation and contains only the languages that appear in the current result set, preventing the creation of no-hit searches. However, it is possible to create no-hit searches when keywords are included in a search.

I analyzed a years' worth of web log data and research on how young children search for books in both physical and digital libraries (Pejtersen, 1989; Busey and Doerr, 1993; Kragler and Nolley, 1996; Fleener et al., 1997; Cooper, 2002b) to determine what subset of the over 100 existing leaf level categories to present in the interface. Even with this smaller size, I found that I could not fit all the categories I wanted to use on a single screen, so I had to introduce paging or scrolling to accommodate them. I chose paging (over 2 pages) because it is believed to be superior to scrolling in many situations (Mills and Weldon, 1987), and because I wanted the interface to fit on a single screen and download quickly. I designed the interface so all the controls fit on the same page at 1024x768 pixel resolution. An alternative solution might show only a single page of categories, and as selections are made, new category choices could replace old ones that were no longer selectable. However, this design would make the location of categories inconsistent and unpredictable.

This interface addresses all of the concerns I had with the Enhanced category browser interface. Children can rely on perception rather than hierarchical knowledge to find categories because the hierarchy is flattened. Children need only select from concrete, leaf-level categories because the more abstract, top-level categories are removed. Children do not have to navigate and backtrack constantly because the categories are viewed simultaneously. Children do not have to distinguish leaves from interior nodes because there are only leaf nodes. Finally, children can more easily find books because the results are prominently displayed in the center of the page. After usability testing, this design replaced the old ICDL category browser in October, 2004.

3.3 Advanced Interface

The Advanced interface is based on the design of U.C. Berkeley's Flamenco interface (English et al., 2002; Yee et al., 2003), as well as many consumer websites such as Sears (www.sears.com) and Epicurious (www.epicurious.com). These sites use various orthogonal category facets (dubbed "faceted metadata" by English et al.) to describe their data. The facets might all appear on one level, or they might be hierarchical, requiring simultaneous or sequential menus within facets to access leaf-level categories. In consumer web sites, this design is an effective way to allow users to specify different features in product, such as cost, manufacturer, and size.

In my implementation, I took the original ICDL category facets and reorganized them into six top level facets: Audience, Appearance, Content, Type, and Subject. Each of these has 3 to 8 of the original ICDL categories underneath it. These all appear on the left side of the search interface, along with a keyword search that can be used separately or in conjunction with a category search. On the right side of the page, I present books that match currently selected categories, or featured books if no categories are selected, as in the Simple interface. Unlike the Flamenco interface, where the home page only shows categories, the results section is always present, keeping the layout of the page consistent and always allowing users to access books.

Selecting any of the links in the 6 sections on the left replaces the links in that section with the subcategories of the selected link. For instance, in Figure 17, I have selected the Color link under Appearance. The other Appearance categories (Format, Length, and Shape) remain accessible as smaller links under the Appearance heading. The Appearance section links are replaced with the various colors available, each indicating how many books match that color. The results side of the page remains unchanged until one of the leaf subcategories (e.g. Red) is actually selected (Figure 18). Selecting Red will change the results section to show all of the Red books in the library. Selectable leaf categories are distinguished from their parent categories by a smaller, italicized font and a count of matching books. I flattened the original category hierarchy in some places so that the subcategories under the top-level facets are always leaves. This provides a consistent, 2-level hierarchy for all facets.

Figure 17. The Advanced interface with Color selected

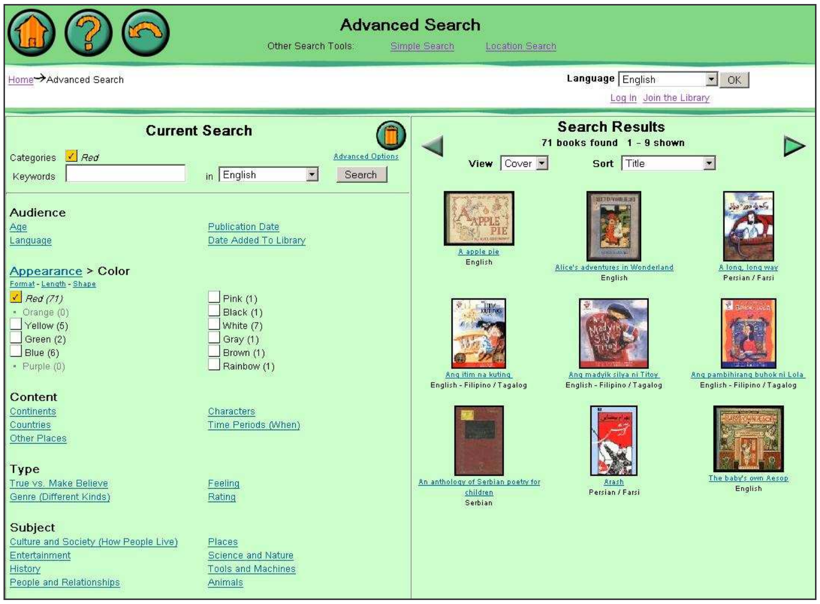

Figure 18. The Advanced Interface with Red selected

By default, combining multiple categories in a search produces a conjunction. For example, selecting Red under Color and English under Language would return only books that are both Red and English. Keywords can also be added to searches, so one can search for Red, English books about "cats." A simple form of query preview (Doan et al., 1996) is used as a search is constructed, where categories and subcategories that have no matching books in the current search criteria are grayed out. This feature prevents the accidental creation of no-hit searches and gives users a sense of the size and scope of the library. Using the "Advanced Options" link in the Keyword section, users can narrow their search results to specific aspects of book metadata (e.g. title, author, and publisher). Users can also change the default conjunctive nature of the search to be disjunctive instead by changing the Match menu from "all" to "any". For example, an "any" search using Red, English, and “cats” would match books that matched any one of these features. After usability testing, this design was added to the ICDL as a new search tool in October, 2004.

Chapter 4: Related Work

4.1 Children, Computers, and the Internet

4.1.1 Computer and Internet Use by Children: Growth and Concerns

Children, regardless of their age, income, or ethnicity, are using computers and the Internet more and more every year. In 2002, 83% of homes with children owned a computer and 78% of homes with children accessed the Internet (CPB, 2002). In 2003, 20% of children age 3-4 used the Internet, 42% of children age 5-9 used the Internet, and 67% of children age 10-13 used the Internet (NTIA, 2004).

So, what are kids doing with computers and the Internet? Not surprisingly, children are most likely to use the Internet for schoolwork and for games. Top activities include exploration (e.g. surfing or searching), communication (e.g. email or IM), and entertainment (e.g. games or downloading music) (CPB, 2002). Children are also spending their own money and their parents' online. In 2003, U.S. children spent 15% of their $172 billion annual income online (Harris Interactive, 2003). In 2002, 56% of parents with Internet access were asked by their children for Christmas presents they saw online (Nua Internet Surveys, November 2002). Clearly, the Internet is having a growing impact on children. 33% of children in the U.S. would choose the Internet over all other media, including television and telephone (Nua Internet Surveys, April, 2002), and 60% of U.K children know what a homepage is but not what a preface to a book is (Nua Internet Surveys, October 2002).

Amidst all of this activity however, some flags have been raised about the appropriateness of computers for children. In a much quoted and rebutted report, the Alliance for Childhood laid out the concerns of many over the focus of parents and educators on getting children to use computers (Alliance for Childhood, 2000). They cited health risks such as repetitive stress injuries, eye strain, and obesity for children who use computers too much. They discussed child development literature that indicates that young children need physical, emotional, and social experiences to develop properly, not advanced, socially isolated intellectual experiences often required by computer software. They criticized school systems for pouring money into buying computers at the expense of teacher training when there is little research to indicate it is helping children learn better. Finally, they noted that much of what children learn is how to operate the computer and how to do "drill and practice" exercises, not how the computer works or how to use it as a tool to think with.

In fact, this report has some valid points (Hourcade, 2004). Many studies have linked childhood obesity to the similar media of television (The Henry J. Kaiser Family Foundation, 2004). Violent television programming, similar to the violence found in computer games, is linked to aggression and desensitization (The Henry J. Kaiser Family Foundation, 2003). However, these fears are grounded in a worst case scenario. Parents are encouraged to set limits on the amount of time their children spend with media and which programming they allow (American Academy of Pediatrics, n.d.) Not all computer software is violent or "drill and practice". For instance, as far back as the 1980's, the LOGO programming language developed at

MIT encouraged children to think on their own about geometry (Papert, 1980) and provided valuable learning experiences (e.g. Clements, 1999).

Haugland (1992) compared preschool children using developmental software to those using non-developmental (i.e. drill and practice) software over a 7 month period. She found that those exposed to developmental software used the computer for a third of the time as those exposed to drill and practices software, but showed significant gains in intelligence, verbal skills, non-verbal skills, problem solving, abstraction, conceptualization, structural knowledge, long-term memory, complex manual dexterity, and self-esteem. The children exposed to the non-developmental software showed gains in concentration, short-term memory, and self-esteem, but showed significant losses in creativity. Children are also not necessarily isolated in their use of computers: 76% of children 6-12 report that there is an adult in the room with them when they go online at home (CPB, 2002). In addition, despite the concerns raised, parents still believe computers are valuable tools for their children: 81% of parents believe the Internet is valuable for their children's learning (CPB 2002).

Fortunately, researchers have also responded to these criticisms and are looking harder at ways to make computers and technology developmentally appropriate and demonstrate its usefulness. Druin and Inkpen (2001) nicely summarize how designers of children's technology can do so by asking three questions:

- Why can technology be appropriate for children?
- What activities for children can technology support?
- What changes in technology should be considered for the future?

To answer the first question, they note that technology can provide children with social experiences, control of their world, and ways to be creative. For instance, children can email with other children around the world, construct virtual worlds with certain game software, and tell stories or draw pictures with other kinds of software. All of these activities support the popular educational and curricular theories of constructivist and constructionist learning, discussed in the next section (Piaget and Inhelder, 1969; Papert 1980).

To answer the second question, they note that children frequently cluster around the same computer, even if there are enough for each child to have their own. Other researchers report similar findings, where children prefer to work with a friend, make new friends, and teach each other when working on computers (Druin et al., 1997; Stewart et al., 1999; Clements, 1999). Certain curricular standards encourage group work and social learning (Vygotsky, 1978), so computers that are properly arranged with multiple chairs and integrated into the classroom can support collaborative activity. In addition, computers with multiple mice and Single Display Groupware (Stewart et al., 1999) can support even more collaboration.

Finally, they note that technology of the future should move away from the desktop and into the everyday world. While Clements (1999) notes that children are indeed capable of using desktop computers and that they are able to understand the symbolic objects on the screen as long as they are concrete, embedding technology in the larger world is even more powerful for young children who are so focused on learning through the physical world. Examples have already been developed, including story-telling robots (Druin et al., 1999), stuffed animal playmates (Strommen, 1998), and programmable physical environments called StoryRooms (Alborzi et al., 2000).

4.1.2 Child Development and Computers

To understand how computers can be used to help children learn and grow, it is necessary to understand a bit about child development. Probably the most influential name in this field is Jean Piaget, a Swiss psychologist whose studies of children in the early 20th century are widely cited and have been used to create developmentally appropriate educational curricula. Although some of his findings and methods have since been challenged, his influence remains strong. Piaget’s main contributions were suggesting that children progress through 4 major developmental stages, and that rather than simply acquiring new skills and information by being taught in these stages, children construct their own models of reality through their experiences with the world around them (Piaget and Inhelder, 1969).

In the first stage, the sensory-motor stage, children from birth to 2 years old use their senses and motor skills to progress from an undifferentiated view of themselves and the world to an understanding of themselves as separate from other objects and as an agent that can act on these objects. By the end of this stage, children achieve object permanence, recognizing that objects continue to exist even if they are out of sight. In the second stage, the preoperational stage, children from 2-7, learn to use play, images, and finally language and text as symbols to represent objects. This stage is characterized by a certain degree of ego-centrism, where children have difficulty taking the viewpoints of others. For instance, if you ask a child to select a picture of how an object will look if you look at it from a certain position, he will pick the picture of how it looks to him presently. Baumgarten (2003) suggests that for children in the preoperational stage, Internet activities that encourage learning and silly fun are good, and that the activities should be brief and simple, as their attention spans are short and their motor skills are not fine-tuned.

In the third stage, the concrete operations stage, children from 7-11 learn to use the symbols they acquire in the previous stage logically. They discover how to reason about conservation of number, mass, and volume, and about how to classify and order things. For instance, a younger child will believe that if you pour a liquid from short, wide glass to a tall, skinny glass, there is more liquid in the taller glass, whereas a child in the concrete operations stage will recognize that the amount of liquid is the same. However, children in this stage can only do this sort of reasoning with concrete objects, and have difficulty with operations that require multiple, systematic steps to complete. In this stage, Baumgarten (2003) notes that children in this age group have improved their motor skills and reading ability, and so enjoy more complex activities that take advantage of these skills, as well as desiring challenge and competition.

In the final stage, the formal operations stage, children 12 and older learn to think logically about abstract things, which often involves testing hypotheses using systematic steps. For children of this age, Baumgarten (2003) notes that in addition to activities that take advantage of these new skills, children of this age are more attuned to the opinions of their peers, and thus enjoy Internet activities that foster social learning and communication, such as chat rooms.

As children progress through these stages, they are not simply acquiring skills by osmosis, but by constructing their own reality, or mental model, of how and why things work (Piaget, 1955). This theory, known as constructivism, is based on two processes that work together but in opposition to help children build their models: accommodation and assimilation. As children experience the world around them and encounter new experiences, they assimilate this knowledge into their existing models. On the other hand, certain experiences contradict their existing models, so they must accommodate these experiences by changing their models to make sense of them.

More recently, a number of researchers have taken issue with some aspects of Piaget's work (Burman, 1994). Some researchers criticized his informal observation techniques, preferring more rigorous scientific studies. Others considered some of the tasks he had children do to be difficult or confusing, and demonstrated that children were able to accomplish certain tasks at younger ages than he predicted. With respect to children's abilities to categorize objects, which is of particular interest in designing category browsers, recent research indicates that Piaget's findings that children progress developmentally from grouping objects according to perceptual features to more abstract concepts like hierarchies may not paint a complete picture of the process. Researchers have found that in addition to developmental skills, both specific domain expertise and cultural norms may influence children's abilities to categorize. Young children are able to develop expertise in areas of personal interest (e.g. dinosaurs) that lead to more sophisticated categorization skills than developmental theory would predict (Chi et al, 1989; Johnson and Eilers, 1998), and children from different cultures sometimes choose to categorize things differently (Cole et al., 1971; Lucy and Gaskins, 2001). Nonetheless, Piaget's research and findings continue to influence the fields of child psychology and education.

A second important figure in child development is Lev Vygotsky, a Russian psychologist whose major contribution to the field was the idea that social interaction heavily influences cognitive development (Vygotsky, 1978). He defined the "zone of proximal development", the time period in which a child could not solve a problem by himself, but could do it if he received help from an adult or peer. Vygotsky argued that psychologists should study children by observing them in this stage, because this is when developmental processes were taking place. In the education world, this theory lead to the idea of scaffolding, whereby adults provide children with more or less assistance depending on their needs, and gradually reduce their assistance as the child becomes more capable (Wood et al., 1976). Wood et al. noted that in a block construction task with 3, 4, and 5 year olds, the youngest children needed to be enticed, assisted, and reassured about their progress, while older children needed only assistance and reassurance, and finally only reassurance.

In the world of computer software for children, the ideas of Piaget and Vygotsky have been appropriated to design developmentally appropriate software and to assist children in learning new skills. In his seminal 1980 book *Mindstorms*, Seymour Papert takes Piaget's concept of children building knowledge and extends it by proposing that this knowledge building can be accomplished best by interacting with the environment to actually build things (Papert, 1980). He argues that the "math-phobic" culture that exists in schools is a result of there not being enough materials in our culture for people to work with to help build their mathematical mental models. Math is instead taught in the abstract, with no reference to anything people can relate to. As an alternative, he offered the LOGO programming language, which allows children to construct geometric shapes with a computer. Papert's theory, known as Constructionism, thus contrasts with Piaget's Constructivism in that it places more emphasis on learning in a concrete situation, rather than on the eventual movement from concrete operations to more formal, abstract thinking (Ackerman, 2001).

While Papert's theories have had a profound influence on educational technology, the idea of scaffolding in computer environments has also taken off. In 1994, Soloway et al. proposed that the field of HCI move away from the idea of "user-centered design" and instead focus on "learner-centered design", using scaffolding in software in the form of coaching, adaptable tools, and different interfaces for different levels. They built a number of applications as exemplars of these techniques, including an editor for learning programming and a physics simulator called Emile for high school students. Guzdial (1995) conducted a study of students' use of Emile and found that students used and tailored the scaffolding to their needs, learned to program physics models, and learned new ways of looking at concepts like velocity and acceleration. Strommen (1998) used scaffolding to guide the design of interactions with ActiMates Barney, an animated stuffed animal for children 2 to 5. Barney encourages learning by facilitating social play with children alone, with a PC, or with television. Revelle (2003) discusses the use of scaffolding including levels of difficulty and hints in interactive products produced by the Sesame Workshop. She also notes the use of scaffolding in the ICDL in the form of a direct manipulation interface rather than a keyword-based query protocol, incremental and clearly visible results, and a built-in Boolean protocol that prevents children from having to choose between conjunctive and disjunctive queries (Revelle et al., 2002).

4.1.3 Children as Computer Users, Testers, Informants, and Partners

Given the emphasis in the human-computer interaction world on understanding and working with users, one would expect that designers of technologies for children would work with children. However, it is only recently that this idea has really taken off, because the obstacles to working with children are many. Cognitively, we know children have short attention spans and limited capability to verbalize thoughts and think abstractly. Practically, children go to school during the day and can't transport themselves to a lab for usability testing or focus groups. Finally, it doesn't quite fit in with the traditional adult-child power structure (Druin, 1999b).

Druin (1999b) describes the various roles that children can take in informing designs: user, tester, informant, and design partner. The oldest and most common role for children is as a user, with adults observing and recording activity. The strength of this approach is that it is relatively easy to incorporate into the design process, but it is limited by the fact that it usually takes place too late in the process for the findings to change the technology, which gives little input to children and means that it is used more by researchers than industry practitioners. The role of child as a tester was popularized by Seymour Papert at the MIT Media Lab in the development of LOGO (Papert, 1980). In this role, children are observed using the software in the same way as a user, but their feedback is requested earlier in the design process. For instance, Hanna et al. (1998) describe usability testing methods at Microsoft, which include site visits, surveys, card sorting, and paper and live prototype tests. As a result of such testing, children's ideas may be integrated into the final design, giving them a sense of empowerment. However, the children don't really have any input into overall design of the technology, which has already been decided by adults.

The idea of children as informants in design emerged in the 1990s. Technologies including a drawing program for kids called KidPad (Druin et al., 1997), a personal communication device for girls (Oosterholt et al., 1996), and an interactive learning environment for teaching ecology (Scaife et al., 1997) were all designed with children as informants. In this process, children are brought in to give input about a technology at different stages of the design process. In addition to testing at the end of the design process, they might brainstorm about new ideas at the beginning of the process by sketching ideas or trying out existing software. The benefit of this role is that children are involved from the beginning, so their ideas are likely to influence the final design to a greater extent, and they will feel more empowered in this role. The downsides are that the adult-child power structure is maintained with adults in charge, and it also takes more time to work with children in this way.

Finally, the role of child as design partner was developed by Druin to address some of the shortcomings of the other roles. The idea of partnering with users grew out of research methods known as cooperative design in Scandinavia (Greenbaum and Kyng, 1991) and participatory and contextual design in the U.S. (Schuler and Namioka, 1993; Beyer and Holtzblatt, 1998). The former built on the socially democratic ideals of the time to allow collaboration between trade workers and researchers to create new technologies for the workplace, while the latter adapted these ideas to use in the integration of technology into the corporate workplace. Druin (1999a) adapted these methods to use in working in partnership with children aged 7-11 throughout the entire design process in a process called Cooperative Inquiry. This age group, in Piaget's concrete operational stage, is old enough to be able to verbalize their thoughts, but is young enough to not be too set in their thinking about the way technologies should look or function.

Cooperative inquiry adapts the idea of contextual inquiry from adults observing adults in the workplace to kids observing each other use technology. Children take notes or draw pictures with Post It notes rather than writing extensively. Frequently they simply write about likes and dislikes, and then work together with adults to organize them into affinity diagrams to extract the main issues with the software. From participatory design, cooperative inquiry adapts the idea of low-tech prototyping to brainstorm about new technologies by building them first with art supplies like pipe cleaners, toilet paper tubes, and socks before working on actual technology prototypes that may look too “finished” to change or critique. Finally, cooperative inquiry makes use of technology immersion by observing what children do with technologies of the future, before adults even have much idea about what these technologies might be good for.

The advantages of working with children as design partners are that the children are equals in the process from the beginning, giving them a huge sense of empowerment and a big influence on the final design of a technology. The downsides are that children and adults must learn to work together as a team, which can take many months, and researchers must work around the limits of children’s schedules, attention spans, and appetites for junk food. Despite these issues, a number of successful technologies have been created with the help of children as design partners: PETS, and QueryKids (Druin et al., 1997; Druin et al., 1999; Druin et al., 2001), as well as a web authoring tool (Gibson et al., 2003), a digital library for 11-14 year-old children to author stories (Theng et al., 2001), and of course the ICDL.

4.2 Information Visualization for Searching and Browsing Interfaces

4.2.1 Psychology of Information Visualization

To understand why we design most user interfaces the way we do, with an emphasis on taking advantage of the human perceptual system, it is useful to understand a bit about how this system works. In 1981, Thomas Moran anchored the idea of understanding the psychology of the user firmly into the field of computer science with an introduction to a special issue on the topic in *Computing Surveys* (Moran, 1981). He noted that users of computers are engaged in goal-oriented activities, but are limited by their short term memory capacity and their tendency to make errors. He divided users into novices and experts, and noted the tradeoffs in defining the success of the system according to various measures such as learning, time, errors, and functionality. He noted that the user's conceptual (or mental) model of how a system works will influence his success in operating it, and that calculational models of a user's mental operations were necessary for helping designers of computer systems try to design with the users' abilities in mind.

Two years later, together with Stuart Card and Allen Newell, he published *The Psychology of Human-Computer Interaction* (Card et al, 1983), where such a model was presented. The Model Human Processor was presented as three interacting subsystems: the perceptual system, the motor system, and the cognitive system. The perceptual system takes input from the senses, such as the eyes, and transfers it into short-term memory. For the eyes, this process takes on the order of 100 milliseconds (ms). The motor system activates muscles, such as those in the fingers for moving a mouse or typing. These movements occur as a series of small micro movements, each taking about 70 ms. Finally, the cognitive system connects input from the perceptual system to output for the motor system.

The cognitive system consists of two memory areas: short-term memory and long-term memory. Short-term memory holds input from the perceptual system for short periods of time. People can generally hold 7 plus or minus 2 chunks of information in short-term memory at any one time (Miller, 1956). Long term memory holds all of a person's available knowledge, and while very large, the ability to retrieve information from it depends on whether associations can be made between the information desired as represented in short-term memory and the information as it is stored in long term memory, a process that usually takes about 100 ms.

It has been shown that young children process information more slowly than adults, and that this in turn affects their motor skills, which rely on rapid processing of perceptual input to make adjustments in motor responses. Kail (1991) studied results from over 70 experiments and found that information processing speed increases exponentially with age from young children to young adults. Thomas (1980) noted that this has a direct effect on motor skills, because the slower speed with which children can process information affects how quickly they can adjust their movements. Chi (1976) attributes the deficit in speed to undeveloped processing strategies such as rehearsal and grouping for moving information between short and long term memory and to children's' smaller long-term memories.

For motor skills that involve moving a mouse, the total time is governed by Fitts Law, which says that the time T to move the mouse is directly proportional to the distance D to the target and inversely proportional to the size S of the target: T = c * log (2D/S). The constant c includes the times for the perceptual, cognitive, and motor systems to each complete one cycle. For children, this constant will be larger than for adults (Hourcade et al, 2003a). Strommen (1994) also found that children have difficulty holding down a mouse button for long and coordinating dragging and clicking. Inkpen (2001) confirmed this result by showing that children perform better and prefer interfaces with point-and-click interaction to those with drag-and-drop style interaction. Children also struggle with double clicking and multi-button mice (Bederson et al., 1996), and with the idea of multiple buttons, because they aren't always able to tell left from right (Strommen, 1998; Hourcade et al., 2003a).

In addition to processing speeds and motor skills, the knowledge that a user has about a system will influence their performance. This knowledge, or mental model, will influence whether the user is able to make connections between his perceptions of the environment and the knowledge he has stored in memory. If people don't yet have a mental model of a system, they will rely on previous knowledge of similar systems, which may or may not match up (Van der Veer and Melguizo, 2002). Two of the earliest discussions of mental models for computer systems were presented by Young (1983) and Norman (1983). Norman noted that mental models are generally incomplete and vague, constantly evolving through the acquisition of new information, unscientific, and sometimes superstitious. To operate a computer system, users run their mental model and adjust it depending on the result. Young noted that people form different kinds of mental models depending on the situation, including analogies, surrogates, and mappings. For young children, Piagetian theory suggests that children of different ages are likely to form different mental models of computer systems, depending on their previous experiences, ability to think abstractly, and the degree to which their understanding of the world is still self-centered.

4.2.2 Interface Techniques for Browsing

Based on the knowledge of how the human perceptual, cognitive, and motor systems operate, a number of user interface techniques have been designed to take advantage of their strengths and weaknesses. Ben Shneiderman was a pioneer in this area, recognizing that as long as it was not overloaded, relying on the recognition ability of the visual perceptual system was much faster than waiting for one or more cognitive cycles to recall information or process text (Shneiderman, 1983; Shneiderman, 1998; Card et al., 1999). For visual user interfaces, Shneiderman helped develop a number of techniques to support browsing activities, which he distinguished from searching or information retrieval, because of it’s emphasis on rapid, progressive filtering of results on the fly based on visual scanning of the current result set, rather than more goal-oriented, methodical searching (Ahlberg and Shneiderman, 1994). Chang and Rice (1993) provide a more thorough definition of browsing, which takes into account the context, influences, process, and consequences that affect the user. For the purposes of my research, I am considering browsing an open-ended exploration of an information space and searching a more goal-oriented, task-driven activity.

In 1983, Shneiderman presented the idea of direct manipulation, now a staple of icon-based user interfaces (Shneiderman, 1983). Based on users' satisfaction with various computer systems, he noted that people were much happier and more productive with systems where they didn't have to remember programming syntax, but could just remember the semantics of the operations required to accomplish their tasks. These interfaces were characterized by continuous representations of objects of interest and physical actions with mice or joysticks to manipulate them, rather than keyword commands to invoke or act on them. In addition, these actions were rapid, incremental, and reversible, so the user immediately saw the result of his action, understood what happened, and could undo it if he made a mistake. For children who have reached the concrete operations stages and understand symbols, direct manipulation of objects on the screen has all the same advantages as for adults. Schneider (1996) notes that there are additional benefits, particularly for those in the preoperational and concrete operations stages, given children's smaller long term memories and shorter attention spans, provided the interface is not overwhelmed with too many objects, colors, or motions.

In 1992, Ahlberg et al. took the idea of direct manipulation and applied it to database querying to create dynamic queries (Ahlberg et al., 1992). Rather than requiring users to remember database language syntax or fill in forms to construct queries, they used graphical widgets such as sliders or buttons to allow users to directly control the values they desired for items in the database. In an experiment comparing a query previews interface to two form fill-in interfaces about the periodic table, 18 undergraduate students were faster with and preferred the query preview interface to the other two. Since then, dynamic queries have been used successfully in many interfaces (Shneiderman, 1994; Fishkin and Stone, 1995; Plaisant et al., 1997).

In addition to problems with textual search, two other problems users of database query systems face is getting too many hits, often overwhelming the database, or getting none at all, leading to frustration. In 1996, Doan et al. presented the idea of query previews as a way to avoid these problems. When a user presents a query to the database, rather than returning complete information about the results, the system instead returns summary information about the results, such as the number of hits and certain important features of each result. These previews provide a number of advantages. First, users are less likely to generate no hit queries because these intermediate results indicate what information will be useful to refine their query. Second, it reduces the load on the network and the database by reducing the amount of information returned. Finally, it provides information about the database contents to aid the user in their searching and browsing.

A final problem in database query systems is the separation of query interface and results interface in many systems. Users often must navigate between the two when refining their query, which takes time and requires them to remember what was going on in the interface that is not currently active. In 1994, Ahlberg and Shneiderman introduced the idea of tight coupling, where dynamic query controls and results are presented together on the same screen, and both are rapidly updated to reflect the current state of the query. As users adjust query controls, they are updated to reflect valid choices that remain given the current search. At the same time, results or query previews are updated to indicate how many and what type of results remain. As a result, users avoid no hit queries because they are able to progressively refine their search based on the feedback they get at each step.

The original ICDL Basic and Enhanced category browsers used direct manipulation of icons representing categories to retrieve books. They used a simplified version of dynamic queries that relied on these icons rather than more complex interface widgets like sliders. They used query previews by only enabling category buttons for which there were matching books in the library. Categories for which there were no books were grayed out and unclickable to avoid generating zero-hit queries. In the Enhanced category browser, the results were tightly coupled to the search by presenting on the same screen. The new Simple and Advanced ICDL category browsing interfaces also use all of these methods to make the searching and browsing experience easy.

4.2.3 Structure and Presentation in Category Browsers

Three of the most common structures for classifying information are hierarchies, trees, and facets (Kwasnik, 1999). Hierarchies and trees have long been used to organize items in meaningful ways, from the biological taxonomies to corporate organization charts. Both hierarchies and trees subdivide a set of data using specific rules for distinction between and across levels, but hierarchies also enforce inheritance relationships between parents and children.

Facets, on the other hand, do not require any type of relationship across levels, but are used to classify a set of data in different, equally meaningful ways. For instance, a user searching in a census database might want to search according to age, location, or income, all unrelated but equally useful ways of thinking about data depending on the task at hand. Each individual facet itself might have a single layer of information (e.g. age) or multiple layers arranged in a hierarchy or tree (e.g. country->state->city). The idea of faceted classification in libraries was developed by the Indian scholar Ranganathan in the 1960s (Kwasnik, 1999). English et al. (2002) extended this idea to digital libraries, coining the term "faceted metadata" to describe the orthogonal metadata descriptors by which a collection of items might be cataloged.

In searching and browsing interfaces, hierarchies, trees, and faceted structures are often presented using sequential or simultaneous menus (Norman, 1991; Hochheiser et al. (2000). There are three different possible combinations of these structure and presentation methods (Table 1). In a sequential presentation, users can only navigate down a single branch or facet at a time. If the interface supports backtracking, they must then backtrack to explore other branches or facets. The original Enhanced and Basic ICDL category browsers are examples of this combination. In a simultaneous presentation of a hierarchy, tree, or hierarchical facets, multiple branches or facets can be explored in parallel. Users can navigate within each branch or facet independently without having to backtrack to explore other areas. Microsoft Windows Explorer is an example of a simultaneous presentation of a file hierarchy. Finally, in a simultaneous presentation of flat facets, all the facets are on the same level and can be explored in parallel. The ICDL Simple search category browser is an example of a simultaneous presentation of single flattened layer of facets.

	Simultaneous Presentation	**Sequential Presentation**
Flat Structure	ICDL Simple Search	Not Applicable
Hierarchical Structure	Microsoft Windows Explorer	Old ICDL Category Browser

Table 1. Combinations of structure and presentation

The sequential presentation has the advantage of allowing users to contend with only a small amount of information at a time, at the expense of backtracking to explore other areas. By contrast, the simultaneous presentations have the advantage of avoiding backtracking between branches or facets, at the expense of a more complex visual presentation consisting of many branches and facets.

In hierarchical structures, tradeoffs must be made between the depth of the individual branches or facets – how many levels – and the breadth – how many items per level when the structure has more than one level. Miller (1981) noted that large breadth will increase search time because the number of items is large, while large depth will also increase search time not only because of the increased number of selections that will need to be made, but also because of limitations of short term memory in keeping track of location in the structure. Many studies have been conducted with adults to try to understand the optimal depth/breadth ratio, and all are in agreement that broad, shallow presentations seem to be better than deep, narrow ones.

Miller (1981) was the first to establish this claim, comparing 4 hierarchies of 64 English words that varied in depth from 1 to 6 levels and in breadth from 2 to 64 items. He obtained U-shaped results curves for both speed and errors, with the best performance on a hierarchy of 2 levels with 8 choices per level. Snowberry et al. (1983) replicated Miller's experiment with similar results, but also found that if the condition with all 64 items on one screen was organized categorically rather than randomly, it had the best performance times. Kiger (1984) also conducted a variation of Miller's experiment that confirmed his results, and also measured user preference, which was consistent with performance data.

Lee and MacGregor (1985) used minimization formulae on the human and computer factors of searching, including visual scanning, key pressing, and computer response time to conclude that the optimal number of items per level was 4-8, which is also consistent with Miller (1956) and his theory of 7 plus or minus 2 items in short-term memory at once. Jacko and Salvendy (1996) conducted similar experiments but also asked users to judge the relative complexity of hierarchies of different depths. They found that users perceived that complexity increased as depth increased. More recently, Zaphiris and Mtei (1997) and Larson and Czerwinski (1998) performed similar experiments on the web, and also found that performance decreased as depth increased.

These previous experiments all dealt with sequential menu presentations, but I believed the results would also be true for individual hierarchical branches or facets in a simultaneous menu presentation. The remaining question in my mind was under what circumstances, if any, are sequential menus superior to simultaneous menus, and vice versa. Hochheiser et al. (2000) compared sequential and simultaneous menus in a web application for browsing census data using hierarchical facets with adults. They found that for simple tasks that did not require backtracking, adults were significantly faster with sequential menus, but for more complex tasks requiring backtracking and selection from multiple facets, they were significantly faster with simultaneous menus. There was no significant difference in preference between the two menu presentations.

While sequential menu designs are ubiquitous, it is only recently that simultaneous menus have been used more extensively in computer interfaces. Ahlberg (1992) was among the first to use the idea of simultaneous menus, using them as dynamic queries in a faceted database about movies. Plaisant et al. (1997) used them in the National Digital Library web interface. Shneiderman et al. (2000) used hierarchical simultaneous menus arranged on two axes, called hieraxes, to visualize data points in a digital library. Marchionini and Geisler (2000) employed simultaneous menus in the Open Video Digital Library to allow users to search for video clips by various facets. Yee et al. (2003) developed the Flamenco browser, using simultaneous menus to allow users to browse the facets of a fine arts database and found that users were more successful and preferred this interface to keyword searching. Naaman et al. (2004) used the Flamenco toolkit to present the results of automatically generated metadata for digital photographs. Reti et al. (2004) used a similar technique to present multimedia metadata in a P2P search system. Gibson (2004) used simultaneous menus to create an overview browsing system for the world wide web.

Today, consumer web sites such as www.sears.com, www.bizrate.com, and www.epicurious.com have found that simultaneous menus are an effective way to allow adult users to specify different features they would like in product. Information technology companies such as Endeca (www.endeca.com), Inxight (www.inxight.com), and i411 (www.i411.com) all offer customized software for business to create simultaneous menu-based search interfaces for their web sites.

4.2.4 Hierarchies vs. Other Forms of Organization for Children

While there is no previous research to indicate if or when sequential or simultaneous menus are a superior presentation tool for children, some research indicates that flatter structures based on simple features may have some advantages over hierarchical structures in searching and browsing interfaces for children. A number of studies in cognitive psychology indicate that hierarchical organization is not the initial way young children group objects. Piaget was among the first to note young children's reliance on concrete, perceptual features when understanding the world around them. He suggested that it is only in later stages of development that they begin to think more abstractly and learn about things such as relational or functional hierarchies (Piaget and Inhelder, 1969).

A number of other studies support these findings. Tversky (1985) studied 3, 4, 6, and 9 year olds and found that when grouping objects, perceptual groupings by facets such as shape and color decreased with age and taxonomic groupings by shared category increased with age. Gentner and Namy (1999) found that 4 year olds were equally likely to select a perceptual match or a categorical match for a particular object. Nazzi and Gopnik (2000) found that when categorizing objects, 3 ½ year olds preferred to group by similar perceptual facets like shape and color than by similar causal features like function. Deák et al (2002) found that both 3 and 4 year olds preferred to sort objects by shape when given no instruction.

However, researchers have also found that young children do accept other ways of categorization besides perception, including simple hierarchies based on more abstract concepts like cause and function. Gentner and Namy (1999) found that when presented with multiple instances of a given category object instead of just one, 4 year olds were more likely to select a categorical match than a perceptual match. Nazzi and Gopnik (2000) found that by age 4 ½, children preferred grouping by causal features to perceptual features. Deák et al (2002) found that 4 year olds could group objects by function if instructed to do so. Nguyen and Murphy (2003) found that children as young as 4 could group objects related by taxonomy, theme (e.g. dog and leash), script (breakfast foods), and evaluation (e.g. junk foods). Hayes and Younger (2004) found that both 6 and 10 year old children were more likely to recall category information that was both functional and perceptual about imaginary aliens than information that was merely perceptual, and were also better able to categorize aliens according to the functional and perceptual information.

While Piaget asserted that children's categorization skills develop through well-defined stages and that these skills apply globally to all domains, more recent research indicates that localized domain expertise and cultural norms can also influence children's categorization skills. Chi et al. (1989) found in two studies that 6 and 7 year-olds with dinosaur knowledge categorized dinosaurs using hierarchical, domain-related information that were not necessarily visual (e.g. where it lives, how it defends itself), whereas children with little dinosaur knowledge sorted more often based on non-hierarchical, visual attributes. Johnson and Eilers (1998) found that adults and 5-9 year olds with similar dinosaur expertise were able to categorize unfamiliar dinosaurs equally well. However, adults performed better when categorizing an unfamiliar domain (birds), indicating that developmental differences still play an important role in categorization skills. Cole et al (1971) found that Kpelle children from Liberia were more likely to sort everyday objects by function, while American children sorted the same objects taxonomically. Lucy and Gaskins (2001) found that while both English speakers from the U.S. and Yucatec speakers from Mexico categorized everyday objects by shape at age 7, by age 9, linguistic and cultural differences between the two groups manifested themselves and 9 year old Yucatec speakers categorized the objects according to material, while English speakers continued to categorize the objects according to shape.

While it is thus possible for young children to categorize using abstract principles, when young children are asked to group hierarchically, some difficulties can arise. Rosch et al. (1976) were among the first to demonstrate young children's difficulties with higher levels of categories. Preschool, kindergarten, and 1st graders could all sort basic level categories (e.g. cats and cars) more than 90% correctly but could only sort super ordinate level categories (e.g. animals and vehicles) correctly less than 60% of the time. Gelman and O'Reilly (1988) found that both pre-schoolers and 2nd graders could draw inferences about objects in basic level categories, but that the older children drew more inferences about super ordinate level categories.

4.2.5 Boolean Search

It has long been known that people have difficulty with Boolean logic, the use of the connectives AND, OR, and NOT to determine whether statements are true or false (e.g. Tversky and Kahneman, 1975). With the advent of computer databases, this problem showed up in query languages for databases for finding matching records (Zloof, 1975), and later in digital library catalogs (Hildreth, 1983), where people frequently misused this feature or didn't bother to use it at all to retrieve bibliographic records (Borgman, 1986). The crux of the issue is that in conversational language, AND is an inclusive term, while in logic, it is exclusive, and vice versa for OR (Johansson and Sjolin, 1975).

Children also have difficulty with Boolean logic, particularly disjunction, though they are still capable of using it. Children as young as 2 years old use and understand conjunction in conversational language (Bloom et al., 1980), and by age 4 use and understand disjunction in conversational language (Johansson and Sjolin, 1975). However, Neimark et al. (1970) found that children didn't understand the use of Boolean conjunction until the 4th grade and Boolean disjunction until high school. For children age 5-13, Snow and Rabinovitch (1969) found that children's ability to complete Boolean tasks, both conjunctive (AND) and disjunctive (OR), using card matching increases with age, but that across all age groups, they made more errors on disjunctive tasks. Rawson et al. (1973) demonstrated that performance improves with age with disjunctive card matching tasks for preschool children, and that even the youngest children performed better than chance. However, even by high school, children struggle with using keyword-based interfaces to create Boolean searches in digital libraries. Nahl and Harada (1996) found that nearly half of 191 students confused AND and OR when creating Boolean keyword searches in a digital library.

Many attempts have been made to simplify the specification of Boolean queries in adult computer interfaces to databases and digital libraries. Zloof (1975) is credited with the first example, known as Query-By-Example. In this interface, rather than typing a query using a database language such as SQL, users specify the query by placing examples of results they would like in skeleton tables that match the database structure. Pane and Meyers (2000) created a similar system and compared it to generating queries with text with adults and children age 10-14. They found that subjects performed better with the tabular system, with no significant differences between children and adults. However, it is clear that such systems are beyond the skills of young children, as they require reading, typing, and abstract knowledge of a database schema. A number of other researchers have created interfaces using the concept of Venn diagrams, where overlapping circles are used to show unions and intersections (e.g. Spoerri, 1993; Michard, 1982; Jones, 1998). A number of these interfaces yielded better performance in both speed and errors than more traditional textual Boolean query languages. However, these interfaces still required typing, which is problematic for children.

Fishkin and Stone (1988) built on the idea of dynamic queries by adding the idea of magic lens filters. Rather than having multiple filters for the various categories on the screen, users create a unique filter for each category and decide whether it should be combined conjunctively or disjunctively with other filters when they are overlapped. While this method provides the ability to create more general queries than the conjunction of disjunctions allowed by dynamic queries, it is even more complicated and not appropriate for young children. Young and Shneiderman (1993) created a "filter/flow" interface based on the idea of water flowing through filters. While they found that adult users performed better in both comprehension and composition tasks using this interface compared to a text-only SQL interface, this interface would likely be too complex for young children to use. Furnas and Rauch (1998) created an infinite, zoomable space where users can place datasets and queries that can be dragged to run a query, and whose results can be reused to run more queries. Druin (2001b) speculates that such an interface is too abstract for children to understand.

For children, the first example of a graphical interface that allowed Boolean searching was the QueryKids interface, developed at the University of Maryland as the predecessor of the ICDL (Druin et al., 2001). In this interface, children could search for animals in a graphical, zooming environment by clicking on sequentially presented categories such as where they lived, what they ate, and how they moved. Boolean searches were accomplished automatically with conjunctions between these categories and disjunctions within them. Revelle et al. (2002) conducted a study with QueryKids where 2nd and 3rd grade children were asked to construct queries including single factor, conjunction, and disjunction. Overall, the children were able to construct the queries 85% of the time, though disjunctive queries were more successful than conjunctive queries. This result is contrary to the previous studies mentioned where children did better with conjunction. I believe this is likely because conjunctive queries required much more navigation to create than disjunctive queries. Nonetheless, Revelle et al. speculate that the scaffolding in the interface was an important factor in the high success rate, and this seems reasonable given previous research indicating that children have difficulty with Boolean logic until they reach later elementary and even high school grades (Neimark and Slotnick, 1970).

The Enhanced version of the ICDL was built based on the QueryKids interface and allowed children to combine categories about book metadata in a similar way (Druin et al., 2003). However, Reuter and Druin (2004) found that when children in grades 1-5 were asked to look for a book, without any direction to perform a Boolean query, most children did not use the Boolean feature, and the few who did only did so by accident. I believe that this is because doing so required backtracking in the sequentially presented category hierarchy to select another category.

4.2.6 Paging vs. Scrolling on Computer Screens

One downside to simultaneous menus compared to sequential menus for category browsers is that if you have a lot of categories, you can only fit so many on the screen. In children's interface design, the burden is even greater because of the need to rely on large, easily clickable targets rather than compact text to display these choices. A decision has to be made about how to present all the categories if there are more than fit on a single screen. The choice comes down to putting all or most of the categories on a single page, requiring users to scroll to view them all, or else dividing the categories into smaller groups and presenting them on different pages, requiring users to navigate between pages to view them all.

A number of studies have suggested that paging may be a superior strategy to scrolling, or at least not any worse, when performing a variety of tasks involving text on a computer screen (Mills and Weldon, 1987). Kolers et al. (1981) performed an eye tracking study that compared subjects reading scrolled or paged text. They found that users were able to read faster, with fewer eye fixations, and with more words acquired per fixation using paging compared to scrolling at their preferred rate. Schwarz et al. (1983) found that inexperienced computer users preferred paging to scrolling when reading on screen, though there was no significant time difference between the two. They also found that users committed more errors on a sorting task when using scrolling compared to paging, and attributed this to the fact that paging allows absolute spatial orientation of items on the screen, while scrolling only allows relative orientation. Using paging, users can recall the position of an item on the screen, while using scrolling they cannot. Piolat et al. (1997) found that when reading and revising text on screen, users were better able to build mental representations of the text and to locate information using paging compared to scrolling.

In a web environment, Byrne et al. (1999) found that users spent 13% of their time scrolling during a 5 hour session, a large portion of time that could have been devoted to other activities. The U.S. Department of Health and Human Services usability.gov website suggests using paging instead of scrolling based on this and other studies. Bernard et al. (2002) studied adults looking for specific links on web pages with 100 search results, presented 10, 50, or 100 results per page. The 10 link condition required paging through 10 pages, the 50 link condition required paging through 2 pages with some scrolling, and the 100 link condition required no paging but much scrolling. They found that users were fastest and preferred the pages with 50 results, and indicated that the 100 link condition presented too many choices, was harder to use, and looked less professional than the 10 link condition. Thus, on web pages with search results, some paging and some scrolling both appear to be acceptable, but too much scrolling is not.

However, another set of studies gives the edge to scrolling over paging in terms of time to complete tasks. Baker (2003) studied adults reading passages in a web environment in paging and scrolling interfaces and found that users were significantly slower with text presented in a paged environment than in a scrolling environment. Manfreda et al. (2002) studied adults completing web surveys and found that users were 30% slower completing multiple paged surveys than single paged, scrolling surveys. In these cases however, the task to be completed required movement only in one direction – from start to finish – to complete the task. The advantages noted by Schwarz and Piolat of paging in creating mental models and locating information spatially were not relevant. To date, I know of no studies that have looked at children using paging and scrolling environments to select search criteria, so the choice made for the ICDL weighed these findings as well as children's cognitive and visual abilities and the need to provide a fast downloading interface.

For the ICDL category interface, I felt that paging would be a better choice for three reasons. First, young children are likely to be overwhelmed with many categories on the same page (Schneider, 1996), which would require scrolling. Second, browsing and searching tasks are likely to be helped by an absolute presentation of searching tools so that users can recall where to find things as they navigate the site and return at later times. Third, one of the goals of the ICDL is accessibility by all users, including those using slow dial-up modems. Splitting the category icons over multiple pages means that each separate page in a paging interface will download faster than one giant page with many icons in a scrolling interface.

4.2.7 Design of Icons

Given the proven benefits of direct manipulation, the fact that young children cannot read yet, and the international audience, using icons rather than text links or form fill-in to access the books in the ICDL database was an obvious choice. However, a number of decisions needed to be made about the appearance of these icons.

Numerous studies indicate that icons on their own are less useful than icons with accompanying text labels. In 1987, Brems and Whitten conducted an experiment that showed users preferred labeled icons to both unlabeled icons and labels only. In 1988, Egido and Patterson had users search through a catalog hierarchy using labels, icons, and labeled icons, and found that users were fastest with labeled icons. In 1991, Kacmar and Carey got similar results in an experiment where users had to match textual descriptions of a function with a label, an icon, or a labeled icon. Finally, in 1993, Byrne found that users performed better in recalling simple icons than complex icons, and that complex icons were no better than empty squares.

In addition to the format of the icon, the size is also important. Given Fitt's Law and the fact that children's motor and cognitive speeds are slower than adults, it is necessary to consider how much distance a user must move the mouse and how large the target is. Hourcade et al. (2003a) studied 4 and 5 year old children and their ability to click on a target using different distances and icon sizes. He found that 64 pixel targets offered significant advantages over 32 and 16 pixels targets for both accuracy and avoiding target reentry once the target was already acquired. For the ICDL, we chose to use labeled icons of at least 64 pixels on either dimension. While the labels won't help pre-literate children, they benefit older children and adults and can be translated into other languages.

Finally, given that the ICDL has an international audience, it was important to consider different cultural interpretations of icons. Murrell (1998) notes that certain images may offend some users for cultural or religious reasons. In the ICDL, we have had to change images that we discovered were culturally insensitive: an icon representing funny or silly with a person sticking their tongue out was offensive in Chinese cultures. Murrell also notes that not all images and metaphors are universally understood (e.g. arrows). This makes the use of labels with icons even more important in an international environment.

4.3 Digital Libraries for Children

4.3.1 Book Selection

Given the physical and cognitive differences between children of different ages and adults, it should come as no surprise that young children look for books differently than older children and adults. With physical books, pre-school and early elementary children choose books based on the appearance of the cover and by flipping through to look at internal illustrations (Pejtersen, 1986; Moore and St. George, 1991; Kragler and Nolley, 1996; Fleener et al., 1997; Robinson et al., 1997). Older elementary and middle school children focus more on textual summary information in jackets, covers, and indices (Wendelin and Zinck, 1983).

Younger children tend not to make a distinction between fiction and non-fiction books, and prefer books about certain genres like fantasy to fiction or learning books (Kuhlthau, 1988; Fleener et al, 1997; Robinson et al., 1997; Cooper, 2002b; Cooper, 2004). Older children are more focused on particular genres that interest them, such as sports and animals (Wendelin and Zinck, 1983; Kuhlthau, 1988; Cooper, 2004). Young children tend to search and browse for books in the physical library by returning to shelves where they have been before, rather than venturing to new areas (Borgman et al., 1995). Children of all ages also enjoy rereading books they have read before, and older children like reading books by the same authors (Wendelin and Zinck, 1983; Fleener et al. 1997; Robinson et al., 1997). Finally, recommendations by peers and teachers also have an important influence on children's book selections (Kragler and Nolley, 1996; Fleener et al., 1997).

One of the earliest researchers to consider book searching in the digital world rather than the physical world was Pejtersen, who noted that the book retrieval process was simply a mapping between book content and users' needs (Pejtersen, 1986, 1989). She studied how librarians mediated this mapping to inform the design of computer interfaces to do the same. She found that there were 4 main facets that users employ to describe their needs, regardless of age: author's intention, frame/setting, subject matter, and accessibility. She found that children were highly focused on accessibility (i.e. how much text vs. pictures), emotional content, and physical appearance. She also found that children relied on three different search strategies: analytical search, search by analogy, and browsing. She postulated that these general characteristics would apply to digital searching as well.

Later studies tend to support her hypotheses about the similarities between searching for physical and digital books. Younger children tended to open books more frequently in the ICDL than older children so they could see illustrations before deciding to read the book, whereas older children relied on textual summary information in the book summary to make this decision (Reuter and Druin, 2004). In addition, younger children like to search by physical attributes such as color, while older children search using genres such as animals (Busey and Dorr, 1993; Reuter and Druin, 2004). In the areas of repetition, the digital also reflects the physical. In both search engines and digital libraries, children exhibit more backtracking and looping behavior than adults, returning to searches they have already run rather than running new ones (Reuter and Druin, 2004; Bilal, 2002, Bilal and Kirby, 2002). The top 5 books in the ICDL accounted for 20% of all book selections (Reuter and Druin, 2004), and the 100 most frequently used search terms accounted for 51% of all search terms used in Solomon's study of an online catalog (Solomon, 1990).

4.3.2 Category Browsing vs. Keyword Searching

In both digital libraries and search interfaces, two interfaces are generally supported: textual keyword entry to support directed searches and selection of pre-defined, categories to support browsing. Many studies have shown that children are capable of using both techniques, but generally prefer and are more successful with category browsing. Shneider (1996) notes that children become experts at browsing during the concrete operations stage. Borgman et al. (1995) explain this result as a combination of children's' "natural tendency to explore" and the relative ease of recognition of categories rather than recall or formulation of keywords. On the other hand, children are not always efficient in navigating hierarchical contexts. Marchionini and Teague (1987) observed children in grades 2-6 using an online encyclopedia and noted that the children often returned to the top of a hierarchy and drilled back down to get to intermediate levels, rather than moving up one level at a time.

Researchers have also noted that young children tend not to plan out their searches, and simply react to the results they receive, a strategy dubbed “interactive browsing” (Marchionini, 1989). This result is consistent with young children’s’ behavior in other activities. Vygotsky (1978) noted that young children solving problems talk themselves through the task as it takes place. Only older children are able to formulate a plan ahead of time, initially with external speech and later with internal speech, or thought. As a result, children tend to perform better using category browsers with open-ended, ill-defined tasks, and equally or better using keywords with directed tasks where the plan is provided for them (Borgman et al., 1995; Hirsh, 1999; Schacter et al., 1998). Solomon (1993) found that children were most successful in keyword searches when using simple, concrete terms, which required little planning or else were provided to them.

In Borgman et al’s extensive studies comparing the hierarchical category browser of the Science Library Catalog to more traditional keyword-based interfaces with 9-12 year olds, the authors found that children performed equally well on directed tasks where keyword searches were easily formulated, but that children did better with the browsing interface on open-ended tasks and those that required difficult spelling. Cooper (2002a) studied 21 second graders using a CD-ROM encyclopedia and found that they browsed more than searched. In their studies of the Yahooligans search engine interface for children, Bilal et al. found that middle school students searched with keywords and browsed the category hierarchy, but were more successful with and more frequent users of the latter (Bilal, 2002; Bilal and Kirby, 2002).

The reasons for children’s’ preference for and better performance with browsing interfaces compared to keyword searching are related to both their physical and cognitive development. Some of these problems abate with age, while others are also observed in novice adult users. At a functional level, simply spelling and typing keywords are difficult for young children (Edmonds et al., 1990; Solomon, 1993; Borgman et al., 1995). Before a child can even get to entering his keyword though, he faces two other obstacles. First, he must have sufficient knowledge of whatever topic he is searching about to come up with a useful query about it (Moore and St. George, 1991). At this point, many children, lacking an appropriate mental model for how a keyword system works, will simply enter a full natural language query into a keyword field (Bilal, 2002; Marchionini, 1989; Schacter et al., 1998; Solomon, 1993).

For children who know they need to use keywords, the second step is to extract keywords from their query. Cognitively, this can be a difficult task for young children in the concrete operations stage who don’t yet think abstractly (Abbas et al., 2002; Large and Beheshti, 2000; Spavold, 1990). Even for those children who do extract what they consider to be appropriate keywords, the search engine or library may use different terminology, resulting in no hits (Abbas et al., 2002). Finally, in systems where Boolean searches are allowed, children who might use these options are often confused between the meaning of AND and OR (Marchionini, 1989). Nahl and Harada (1996) confirm that these steps are problematic even at the high school level. Every student in their study had lexical errors in the searches they created (e.g.

misspelling or failure to pluralize terms), and nearly half had difficulty selecting appropriate search terms and using Boolean logic properly.

4.3.3 Previous Interface Solutions

Based on these findings, a number of researchers have built digital library systems for children that attempt to address some of the shortcomings of early, keyword-based systems by using category browsers instead. Pejtersen (1989) created the BookHouse interface with a metaphor of rooms in a house to support the different types of searching she found children used. Children could browse and search using category icons for different facets of the book classification scheme or find a book they had read before and have the system find books that were classified similarly. Borgman et al. (1995) used a book shelf metaphor with iconic representations of categories in the Dewey decimal hierarchy for the Science Library Catalog rather than keyword search and found children used it equally well or better than keyword searching. Busey and Doerr (1993) worked with children in grades 1-5 to create the Kids Catalog, which provided multiple modes of access. Both Borgman et al. and Busey and Doerr found that the Dewey system didn’t capture the search needs of children well. Both renamed the categories with more child-appropriate terminology, and Busey and Doerr designed their category hierarchy to include categories like animals and fairy tales that were missing from Dewey but desired by children.

Kulper et al. (1997) designed the Bucherschatz interface for 8-10 year olds using a treasure hunt metaphor and a category hierarchy, which they designed so it was impossible to get no hits. While the treasure hunt metaphor didn’t make sense to many of the children, they understood the hierarchy. Finally, Druin et al. (2001) designed the QueryKids interface for finding information about animals. Using cooperative inquiry methods, the team of children and adults on this project designed a category hierarchy for finding the animals. The children also indicated that there should be a purpose to the search, so the metaphor of going on a journey was introduced to give the searching a destination.

4.3.4 ICDL Interface Solutions

While all of the systems described in the previous section were improvements over keyword search, none was actually a publicly accessible digital library in the truest sense of the word. Some provided access to bibliographic records of physical books, and others were small, local collections of specialized media about particular topics. In 2002, this changed with the launch of the ICDL, which provided access to scans of actual physical children's books from all over the world (Druin et al., 2003, 2004; Hourcade et al., 2003b). The ICDL was built based on the same Java-based software as QueryKids, using similar cooperative inquiry methods to inform design. The researchers confirmed previous research by showing that younger children (4-7) liked to look for books based on physical appearance, while older children (8-11) chose books by how they made them feel. Using this information, as well as other information gleaned from observation and other studies, the team created a faceted category browsing interface, as well as a globe interface for searching the world and a variety of interfaces to use for reading books.

4.3.5 Other Current Digital Library Solutions

To date, the ICDL provides the largest, most extensive collection of children's books with an age appropriate interface. A number of online library projects provide access to many books, but inappropriate interfaces. Project Gutenberg (http://www.promo.net/pg/) and the Rosetta Project (http://www.childrensbooksonline.org/) both provide access to scans of out-of-copyright books from around the world, but both have mainly text interfaces geared toward adult searching and don't provide access to more current books. Likewise, the University of Florida's Literature for Children collection (http://palmm.fcla.edu/juv) contains roughly 600 mostly out of copyright books from the U.S. and Britain, accessible with an adult-oriented textual interface.

The Fairrosa Cyber Library of Children's Literature (http://www.fairrosa.info) provides links to digitized books around the web and to lists of various recommended books for children, but also has an interface geared toward adults. Children's Books Online (http://www.magickeys.com/books/) provides access to about 30 English books created specifically for the web environment and published only online by their authors. This interface requires a fair bit of textual navigation to reach these books. The University of Virginia's E-Book Library (http://etext.lib.virginia.edu/ebooks/subjects/subjects-young.html) provides access to over 100 out-of-copyright English children books via Microsoft's Reader and the Palm OS, both interfaces geared toward adult users.

Other web sites have made their content more accessible to children, but don't have very large or diverse collections. Stories from the Web (http://www.storiesfromtheweb.org) contains access to about 30 published English books and book excerpts for children age 8-11 and about 600 for children age 11-14. However, these are extracted text and pictures, not the scans. The web-site for 8-11 year olds organizes the books into kid-friendly categories like funny, scary, and magical, but requires clicking of small text links. StoryPlace (www.storyplace.org) presents interfaces designed for use by preschool and elementary-age children, but only has lists of books and book-related activities. Busey and Doerr's Kids Catalog was updated and turned into a website (http://kcweb.tlcdelivers.com/kcweb/kcHome) with a hierarchical category browser of kid-friendly terms and keyword search tools for finding book call numbers but not actual scans.

A number of popular online and media outlets have children's reading sections that are small but relatively accessible by older elementary-age children. The Internet Public Library's children's area, KidSpace, has a Story Hour section (http://www.ipl.org/div/kidspace/storyhour/) with about 10 online-only stories in English. The BBC's children's area, CBeebies, has a Story Circle section (http://www.bbc.co.uk/cbeebies/storycircle/index.shtml) with about 30 flash-animated original and adapted stories in English. The Sesame Workshop website has about 20 English stories created for the web (http://www.sesameworkshop.org/sesamestreet/sitemap/?sectionId=stories).

Theng et al. (2000, 2001) have used the Greenstone Digital Library software (Witten et al., 2000), an open-source software project that allows people to create and customize digital libraries, to create a digital library to allow children age 11-14 to author and critique each other’s stories. Like the ICDL, the authors used participatory design techniques to work with children to design an environment in which they and their peers could post and review their work. However, this library is restricted to the members of a classroom and contains only materials authored by the students, not a broad collection of published literature.

Chapter 5: Controlled Studies

5.1 Research Goals and Questions

The broad goal of my research is to understand children's abilities and preferences for different searching and browsing interfaces at different ages. As a first step toward this broader understanding, I used the ICDL as a test bed to conduct two controlled research studies. Both compared a flat category structure presented simultaneously (hereafter called *simultaneous*) to a hierarchical category structure presented sequentially (hereafter called *sequential*) with children in 1^{st}, 3^{rd}, and 5^{th} grade.

Since Hochheiser et al. (2000) found a difference in speed between simultaneous and sequential menu interfaces for adults depending on whether participants did simple searches that did not require backtracking or more complex searches that did require backtracking, one of my goals was to explore whether this would be true for children as well. Since the amount of time I could work with any one child was limited, I decided to conduct two different studies. In the first study, children did simple, one-item searches (e.g. How many red books are there?), hereafter referred to as *simple searches*. In the second study, children did more complex, Boolean, two-item conjunctive searches that would require backtracking in the sequential menu design (e.g. How many red happy books are there?), hereafter referred to as *Boolean searches*. Since Reuter and Druin (2004) found that children did not often create Boolean searches on their own when browsing in a sequential hierarchy, my other goal was to see if children would create more conjunctive Boolean searches by selecting two or more categories while browsing using the simultaneous interface. In addition to the directed searching tasks, I also had the children do open-ended browsing tasks with each interface in each study to see if either one generated more conjunctive Boolean searches than the other.

I had five specific research questions to consider, both overall, and within each of the three grades. First, I wanted to know whether completing the same search by navigating in the sequential hierarchy or using paging in the simultaneous design would take longer. Second, I wanted to know whether children would have more difficulty finding categories by guessing where to look in the hierarchy or visually scanning and paging in the simultaneous design. Third, I wanted to know if the answers to these questions varied depending on whether children were selecting a single category or multiple categories. Fourth, when selecting multiple categories, I wanted to know if children understood that they were creating a conjunctive Boolean search, and if this understanding was related to the interface they used or not. Finally, when browsing without a direct goal, I wanted to know if the simultaneous interface facilitated the creation of more Boolean searches than the sequential interface.

5.2 Hypotheses

I had the following hypotheses about my experimental results:

H1. *The simultaneous interface would be faster for simple tasks using one page.* This hypothesis was based on the fact that simple tasks that could be completed on the first page of the simultaneous interface required only one mouse click. The same task required two clicks in the sequential interface. In this case, I believed visually scanning the first page and clicking once in the simultaneous interface would be faster than identifying the appropriate top-level category and clicking twice in the sequential interface.

H2. *The simultaneous interface would be faster, easier to use, and preferred for all Boolean tasks.*

This hypothesis was based on the fact that Boolean tasks would require either two or three clicks in the simultaneous interface, depending on whether paging was required. The same task required five clicks every time in the sequential interface. In this case, I believed in addition to requiring fewer clicks, visually scanning one or two pages in the simultaneous interface would be faster than having to identify two top-level categories and navigate between them in the sequential interface. I also believed the extra navigation and clicking in the sequential interface would make it less preferable than the simultaneous interface.

H3. *Boolean tasks would be harder for younger children to complete.*

This hypothesis was highly likely to be true, but I was interested to see how much better the older children understood than the younger children, and at what age children could reasonably be expected to understand the interfaces well.

H4. *The simultaneous interface would better support creation of Boolean tasks.*

This hypothesis was based on the fact that creating a Boolean search in the simultaneous interface could be completed with a minimum of two clicks, while it required a minimum of three clicks in the sequential interface for a within-facet search and five clicks for a between facet search.

H5. *The simultaneous interface would better support understanding of Boolean tasks.* This hypothesis was based on the fact that creating a Boolean search in the simultaneous interface was more likely to occur on a single screen, where the selected categories would all remain visible. In the sequential interface, navigation between facets meant that previously selected categories were no longer visible on the screen. I thought that this loss of context in the sequential interface might confuse children, or cause them to forget what they had already selected.

5.3 Participants

Seventy two children equally split between 1st grade (age 6-7), 3rd grade (age 8-9) and 5th grade (age 10-11) participated in the two studies. Thirty six children participated in each study, 12 in each grade, equally split between boys and girls in each grade. The participants came from four suburban Maryland elementary schools: Holy Trinity Episcopal Day School in Bowie, Northfield Elementary School and Clemens Crossing Elementary School in Columbia, and Hillcrest Elementary School in Catonsville. Official demographics were not available, but all of the schools were racially and ethnically diverse. Income information was not available, but Holy Trinity accepts both paying and scholarship students. According to the U.S. Census Bureau, the median household income in 1999 for Howard County, where Northfield and Clemens Crossing are located, was 40% above the state average, but for

Baltimore County, where Hillcrest is located, it was 4% below the state average (quickfacts.census.gov/qfd/states/24/24005.html).

Participants had all used computers before but had never used the ICDL. Participants were mostly right handed, but some were left handed. In either case, participants were always prompted to use the mouse with whatever hand they felt comfortable. Participants were not screened for vision ability or color blindness, and participants who wore glasses were allowed to wear them during the study. No vision-related problems were observed during the studies. Participants were not screened for reading ability. In some cases, first graders needed assistance reading questions and icon labels (see Section 5.6). In three of the schools, the study was conducted during after-care programs. In the fourth school, the study was conducted during regularly scheduled technology class time.

5.4 Materials

5.4.1 Permission and Assent Forms

The University of Maryland Institutional Review Board (IRB) approved the use of child participants for these studies. At the two schools in Columbia, the studies were approved by the after care program run by the Columbia Association. At the other two schools, the principals approved the studies. The parents of participating children received a letter explaining the studies (Appendix 1) and signed permission slips (Appendix 2), and each participating child signed an assent form (Appendix 3).

5.4.2 Interfaces

The two interfaces were based on the ICDL Simple Search, with the keyword and language search options removed so that children could only create searches using categories. The simultaneous interface presented 44 leaf-level categories on a single level, but spread over two pages (Figure 19, Figure 20). The sequential interface divided the same 44 leaves into 9 top-level categories, each with a second level of 2-12 leaves (Figure 21, Figure 22). The simultaneous interface required less navigation to select multiple categories, but presented more categories on a single page and required paging to reach half of the categories. The sequential interface presented fewer categories per page, but required more navigation and backtracking to select multiple categories. In both interfaces, 3 featured books were presented in the results area if no categories were selected. When categories were selected, the results were displayed 8 books at a time with paging arrows to view more.

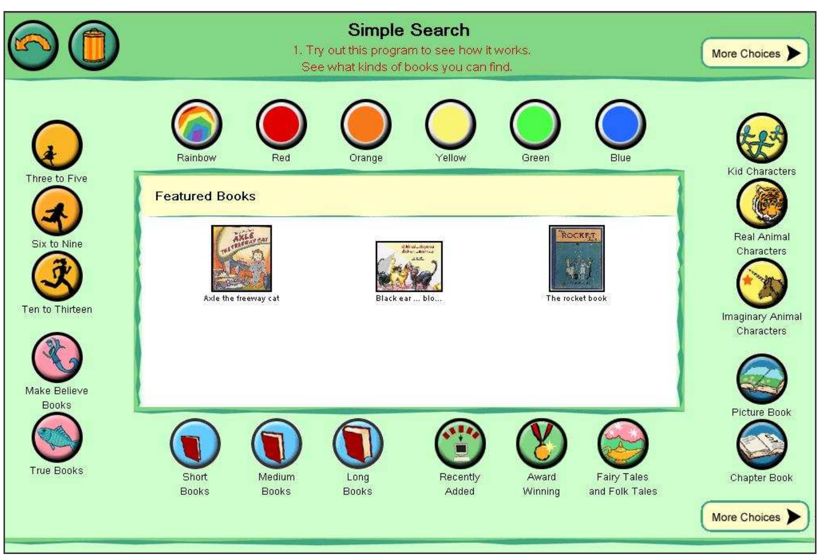

Figure 19. The first page of the simultaneous test interface

Figure 20. The second page of the simultaneous test interface

Figure 21. The top level of the sequential test interface

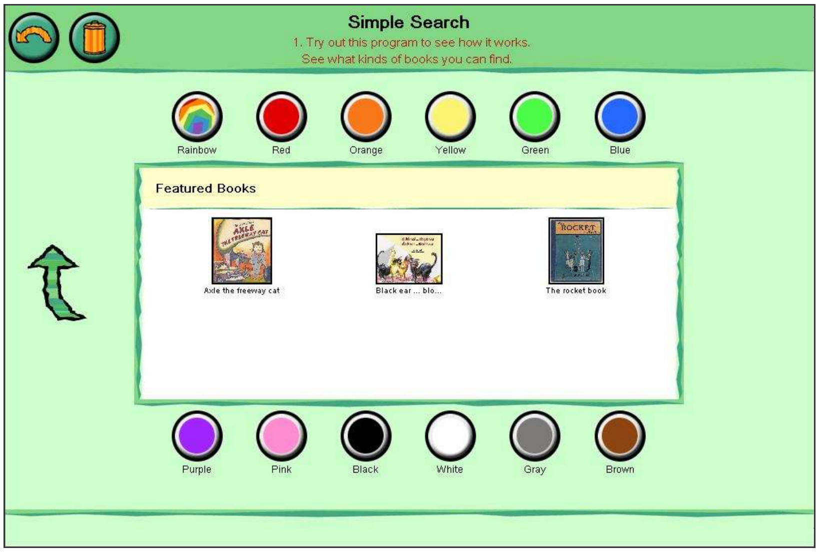

Figure 22. The leaf level of the color category in the sequential test interface

5.4.3 Technology

Participants used a Dell laptop with a 12 inch display and an Intel Pentium 3 Mobile CPU running Windows XP. The processor ran at 800 MHz with 512 M of RAM and a resolution of 1024x768 pixels. Participants used a Kensington USB single-button mouse. The software used for both studies was adapted from the existing ICDL implementation, and consisted of Java class files, JavaScript, CSS and image files, a Tomcat server and a MySQL database of 573 children's books. The software enabled the children to use the simultaneous and sequential search interfaces to look for books. However, I disabled the ability to click on the books in the results sections of the search interfaces to prevent children from getting off task by going to look at individual books. All software was loaded on to the laptop so that no Internet connection was required. The software was instrumented to record the time of each mouse click during timed tasks. Participants used Microsoft Internet Explorer 6 in full screen mode with the address and task bars hidden.

5.4.4 Experimenter Worksheets

During each of the studies, I recorded information about each participant, made notes about what they did on each task, and kept track of how many hints, if any, they needed to complete a task. I recorded this information on specially designed log forms, one for the simple study (Appendix 9) and one for the Boolean study (Appendix 10). Participants were informed that I would be making notes so I would remember what they did. This was done to avoid participants thinking that I was judging them while they used the software.

5.5 Procedure

In each study, the participants worked with me one at a time in a quiet room or hallway in their school. The room or hallway was set up with 2 chairs and a table of appropriate height for the children. During the study, participants sat on a chair in front of the laptop and I sat to their left. I briefly explained to the children what we were going to do: try out two new computer programs for finding books to see how well they worked so that they could help me make them better. I did not tell the children they were being timed to avoid making them nervous. Next I explained that their parents said it was ok for them to help me, but that they had to agree to help me too. Third and fifth graders then read and signed the assent form. I read the assent form to the first graders and then had them sign it.

After signing the assent form, third and fifth graders read a screen of general instructions (Appendix 4) that explained the purpose of the study. I read these instructions out loud to the first graders. Participants then began a series of tasks with the first interface. The order of interface presentation was counterbalanced; half the participants used the simultaneous interface first and half used the sequential interface first. In both studies, the first task was a free browsing task: “Try out this program to see how it works. See what kinds of books you can find.” The goal of this task was to see how participants used the interface to browse without any instruction.

After 2 or 3 minutes of exploration, I stopped the exploration and went on to the next task if the participant hadn’t already indicated they were done. The next task involved my taking control of the mouse and demonstrating the interface. In both studies, I memorized a demonstration script, one for the simultaneous interface (Appendix 5) and one for the sequential interface (Appendix 6). I used the same script for both studies. This demonstration included instruction on how to select a single category to find matching books, how to unselect categories and start over, how to use the results paging arrows, and how to use the More Choices and Up Arrow buttons. During the demonstration, and during the entire session, I did not show how the Boolean functionality worked unless a child specifically asked me about it. I did this so that I could observe whether or not children discovered how this worked on their own.

Participants then completed a series of timed tasks. Since each participant used both interfaces, I created two different but structurally and cognitively similar sets of timed tasks for each study (Appendix 7 and Appendix 8). For example, children were asked to find Pink books in one task set and Purple books in the other. These category buttons were located near each other in both interfaces and involved the same category concept (Color). The interface that each child used a task set with was counterbalanced and the order of the tasks in each task set was randomized for each child. Each timed task was a question of the form “How many X books are there?”, where X was a single category in the simple study and two categories in the Boolean study. The question was presented in the center of the screen with a “Begin” button below it. I read the task aloud to the participant and they pressed the button when they were ready to start.

The “Begin” button served two purposes. It started a timer to record how long it took the participant to complete the task after reading the task. It also centered the mouse on the screen so that each task would have the cursor starting in the roughly the same place. The task was repeated in the top navigation bar of the interface as well in case the participant forgot what they were looking for after beginning the task. When a participant successfully completed a timed task by telling me the correct number of books found, I pressed a button on the keyboard to stop the timer, which also advanced the screen to the presentation of the next task question. Participants could not move on to the next task until they had successfully completed the previous task, which sometimes required hints (see Section 5.6). Overall, this protocol made for longer than might be expected task times in a simple click-counting model. Participants frequently re-read the question to remind themselves what they were looking for, sometimes needed hints or had to correct mistakes, and had to tell me how many books they had found before moving on to the next task.

After completing the timed tasks in an interface, participants were presented with a screen with two preference questions: “Did you like using this program?” and “How hard was this program to use?” Participants selected one of three answers for each of these questions (“Not much, A little, A lot” and “Hard, Medium, Easy”). After answering these questions, the participant repeated the same protocol with the second interface, beginning with the free browsing task. After completing all of the tasks in the second interface, participants were presented with a screen with one final question: “Which program did you like better?” Participants selected one of three answers for this question (“First, Second, Both”). After completing the tasks, I thanked the participant for helping me and offered them an ICDL sticker or globe in appreciation. The entire process took 15-30 minutes depending on the age of the child and whether they were doing the simple study or the Boolean study.

5.6 Hints

University Institutional Review Board (IRB) rules required that the studies take no more than 30 minutes per child, which pilot testing indicated would require either giving some children hints on searching tasks if they had trouble with the tasks, or enforcing strict time limits on the tasks. I chose to give hints so that I would have complete time data, making the data analysis easier, and because I thought I would learn more by talking with children who were having problems rather than watching them struggle. We have observed that working with an adult is a common use scenario for elementary-age children using the ICDL, and 76% of children age 6 to 12 report that there is an adult in the room while they are online (CPB, 2002). Children also often work together on both computer and non-computer projects in school environments, and often receive help from their peers in this way (Druin et al., 1997; Stewart et al., 1999; Clements et al., 1999). These collaborative learning styles led to the idea of scaffolding, where adults or technologies such as help systems and adaptive environments provide children with more or less assistance to complete a task until they can do it on their own (Wood et al., 1976; Soloway et al., 1994).

While having an adult work with a child and provide hints is thus a reasonable use scenario to provide scaffolding, it also introduces the potential for both inconsistency and bias on the part of the hint provider. I was careful to minimize the possibility for both by developing a protocol for giving hints in a consistent way across both interfaces based on the problems children encountered during pilot testing. I recorded the number of hints given for each search task and report on them in the analysis. The five types of hints that children required were: (1) younger children needing help reading a category label; (2) not remembering how to find the count of books found; (3) selecting the wrong category on a search task; (4) inability to find a category because they forgot about the “More Choices” paging button in the simultaneous interface; and (5) not knowing which top-level category to look under in the sequential interface.

For the first issue, reading the category label, I read the label for the child. For the second issue, not remembering where the count of books was on the screen, I went through two gradually more helpful hints until they reported the correct count of books (Table 2). Some children would start manually counting the books on the screen. In this case, I asked the children if they remembered where the count was, and if not, pointed it out on the screen.

Hint #1	Do you remember if there is an easier way to get the count?
Hint #2	The program shows you the count right here above the books.

Table 2. Hint protocol for not remembering where the book count was found

For the third issue, selecting the wrong category, I went through four gradually more helpful hints until they were able to identify and correct the problem (Table 3). For the first hint, I indicated that they didn't have the right answer, and asked if they saw why. If not, I pointed out which category they had selected and reminded them of which category we were actually looking for. Finally, if they did not recall how to unselect the wrong category, I asked them if they remembered how and then reminded them how to do this if not.

Hint #1	That's not quite right. Do you see why?
Hint #2	We're looking for red books. You picked blue books.
Hint #3	Do you remember how to unselect the wrong category?
Hint #4	You can click the blue button again, or use the trash can.

Table 3. Hint protocol for selecting one or more wrong categories

For the fourth issue, forgetting about the More Choices button, I went through three gradually more helpful hints until they were able to find the More Choices button so they could look on the second page of categories (Table 4). I first asked if there was someplace else they could look for more categories. If this didn't help, I then asked if they remembered the More Choices button and then pointed it out if not.

Hint #1	Is there someplace else you could look for that category?
Hint #2	Do you remember the More Choices button I showed you?
Hint #3	Why don't you try clicking the More Choices button.

Table 4. Hint protocol for forgetting the More Choices button in the Simultaneous interface

For the fifth issue, not knowing which top-level category to look under, I went through three gradually more helpful hints until they were able to find the correct top level category to look under (Table 5). I first encouraged the children to guess where to look. Some were willing to employ trial and error until they found the right category, sometimes with a hint from me about looking someplace else. Others became frustrated and gave up at some point and I then told them which category to try looking under.

Hint #1	Where do you think you might find picture books?
Hint #2	Is there someplace else you could look?
Hint #3	I think you should try looking under the Format button.

Table 5. Hint protocol for not knowing which top-level category to look under

5.7 Simple Search Tasks

In the simple study, children did 6 timed tasks with each interface (Appendix 7, tasks 3-8). These tasks were all questions of the form "How many X books are there?" where X was a single category (e.g. red). The questions were of two types: one-page and two-page. In one-page tasks, the category to be found was on the first page when presented with the simultaneous interface. The task could be completed optimally with a single mouse click. All mouse click counts given are optimal. Participants sometimes used more clicks (e.g. if they didn't see a category or selected the wrong one). In two-page tasks, the category was on the second page, requiring two clicks to complete the task optimally. This distinction only mattered for the simultaneous interface. In the sequential interface, all tasks required selecting a top-level category, then the requested leaf-level category, for a total of two clicks optimally.

5.8 Boolean Search Tasks

In the Boolean study, children did 8 tasks with each interface. Six of these tasks were questions of the form "How many X books are there?" where X was two categories (e.g. red happy) (Appendix 8, tasks 4-9). The questions were of two different types: one-page and two-page. In one-page tasks, the categories to be found were both on the first page when presented with the simultaneous interface, requiring two clicks optimally to complete the task. In two-page tasks, the categories to be found required navigating to the second page for one or both categories in the simultaneous interface, requiring three clicks optimally to complete the task. This distinction only mattered for the simultaneous interface. In the sequential interface, all tasks required selecting a top-level category, then the requested leaf-level category for the first category, backtracking, and repeating the process for the second category, which was always in a different facet. This always required a total of five clicks optimally. We chose not to include tasks where both categories were in the same facet since these sometimes were not possible (e.g. books cannot be both short and long), and to avoid adding another variable to the study.

The remaining two tasks were designed to elicit whether or not the children understood what they were doing when they selected two categories (Appendix 8, tasks 3 and 10). One was done just before the 6 timed tasks, and the other was done just after the 6 timed tasks. In these tasks, I asked the children to find and click the buttons for two specific categories. Once they had done this, I asked them questions to see if they understood what they had done (Appendix 10, tasks 3 and 10). I first asked them what kinds of books they had found. If they didn't understand the question, or seemed unsure of their answer, I followed up with a second question, asking them if the books had anything in common. Sometimes children gave an answer that involved both of the selected categories, but it was unclear whether they thought they were combined conjunctively or disjunctively. In these cases, I asked the children if all the books matched the first category and if all the books matched the second category to clarify.

5.9 Pilot Testing

Before conducting the study, I pilot-tested the experiment protocol with 3 children on the Maryland kids team, a group of children we work with as design partners at the HCIL, and 10 additional children at one of the elementary schools used for the study. The pilot testing with the kids team helped establish the timing of the experiment, how many questions I could reasonably ask within a 30 minute time period, and which questions children would need hints with. The pilot testing with the children at the elementary school was helpful in a number of ways. First, I had originally planned to have another graduate student run some of the trials, but we discovered that the laptops we planned to use, though physically identical on the outside and purchased at the same time, had different processor speeds. As a result, I ran all the trials myself on a single laptop to avoid a confounding variable.

Second, I decided to ask the children how they liked an interface and how hard it was to use immediately after they used it, rather than at the end of the study, so that they didn't have to remember the difference. Third, I discovered that when I asked children the comprehension question during the Boolean study, they sometimes gave an answer that could be interpreted either as a conjunctive or disjunctive combination. As a result, I developed the follow up question protocol to elicit which they thought it was. Finally, I discovered that a question that involved finding the "short" books was confusing because the interface had both "Short Length" buttons and "Short Stories" buttons. I replaced this question with one that required finding a clearer category.

5.10 Analysis Methodology

I consulted a professional statistician to ensure that I analyzed each set of data properly. A brief description of the tests used and the rationale for choosing them is given below. Some of the tests are based on data collected during timed searching tasks, and some are based on data collected during browsing tasks. For timed tasks, I used repeated measures analysis of variance (ANOVA) under the assumption that time was a normally distributed, interval dependent variable. For tasks measuring counts of things, such as the number of subjects who preferred a certain interface or the number of hints given to complete a task, I used non-parametric statistics, where the requirement of a normally distributed dependent variable may not be met but it is still ordinal, interval, or categorical in nature. Throughout this thesis, I use the word "significant" to mean statistically significant.

For all statistics, I test a null hypothesis that there is no difference between the groups being compared and I use a probability value (p value) cut-off of 0.05 to reject this hypothesis in favor of the alternative hypothesis that the two groups differ. A p value of 0.05 or less indicates statistical significance, meaning that there is only a 5% probability that the result is a type 1 error (false positive). All p values reported are two-sided unless otherwise noted. In two-sided tests, no direction for the difference between the groups is assumed for the alternative hypothesis. In one-sided tests, the alternative hypothesis is that groups differ in one direction. For all statistics, I looked at differences overall with 36 subjects and within each grade with 12 subjects. The non-parametric statistics within each grade are robust enough that any significant effects found are likely valid, but non-significant effects may indicate either non-significance or a type 2 error (false negative) due to the small number of participants.

5.10.1 Timed Tasks

Analysis of variance calculations were performed on the dependent variable, time, for the timed search tasks in both experiments. For each of the 2 interfaces, children performed 2 different task types: 3 tasks that required viewing both pages in the simultaneous interface (two-page tasks) and 3 that did not (one-page tasks). The order of these 6 tasks was randomized for each child, and a different, but structurally and cognitively similar, task set was used in a counterbalanced way for each interface. Preliminary analysis indicated no effect by task set, so this factor was excluded from analysis. For analysis, each set of 3 times in each interface was averaged together for an overall analysis. Thus, each child had 4 times that were compared in the ANOVAs:

- average time for 3 one-page tasks in simultaneous
- average time for 3 one-page tasks in sequential
- average time for 3 two-page tasks in simultaneous
- average time for 3 two-page tasks in sequential

I used the PROC MIXED function in the SAS statistical software package to compute repeated measures ANOVAs with residual maximum likelihood (REML) estimation, unstructured covariance, and Satterthwaite denominator degrees of freedom. This combination of settings was chosen because it makes no assumptions about the correlation between variables (unstructured covariance) or the variances of the populations studied (Satterthwaite degrees of freedom). Because there was no missing data and all effects were fixed, this method provided results that were similar or identical to ordinary least squares ANOVA but allowed more flexibility to specify covariance and variance and conduct post-hoc testing.

ANOVA results are reported with an F-test, which tests if the means of the groups formed by values of the independent variable(s) are different enough not to have occurred by chance. The F test is a ratio of the between group to the within group variance, so the larger the result, the more likely the difference is significant. The numerator and denominator of this ratio each have a certain number of degrees of freedom, or numbers of pieces of information, that are used to estimate the means, based on the number of subjects and the number of groups. F statistics are reported with the degrees of freedom for the numerator and denominator, respectively, in parentheses, followed by the ratio, and then the p-value for that particular combination of result and degrees of freedom. For instance, $F(1,11)=11.85$, $p<0.01$, indicates an F test using 1 and 11 degrees of freedom and a result of 11.85, with a p-value of less than 0.01, indicating a significant result.

For post-hoc testing, I used t-tests with the Tukey adjustment, a common adjustment used to consider pairwise differences in variables or interactions with more than two levels or groups. The t-test measures if the means of two groups or levels being compared are different enough not to have occurred by chance, and the Tukey adjustment accounts for the fact that there are more than two levels or groups in the sample. Customarily, only those pairs for which the Tukey tests are significant are reported. Graphs indicating mean times are shown with 95% confidence intervals. These intervals indicate the range around this particular sample mean that would contain the means of 95% of other samples were the experiment repeated many times. The narrower the confidence interval, the more confident we can be that the sample mean is close to the true population mean.

In addition to grouping the timed tasks by type, I also looked at the times and number of hints required for the 6 tasks individually in each interface to understand which categories used in the tasks caused more difficulty than others. I did not perform ANOVAs per task since the tasks were highly specific to our application and the results would therefore not have general interface design implications. I was also skeptical of the validity of 12-way post-hoc pairwise comparisons given the sample size. However, this information was useful for feeding back into the design of the ICDL category structure.

5.10.2 Difficulty and Like/Dislike Questions

The difficulty and like/dislike questions involved asking children for their subjective opinions about each interface after they used it in both experiments. Children selected from among 3 choices for each of these questions. These choices were ordinal in nature (e.g. hard, medium, easy), so I assigned them numerical values (e.g. hard=0, medium=1, easy=2) and compared the responses for each interface to each other using Wilcoxon signed rank tests for all grades combined and for each grade individually. The Wilcoxon test is the non-parametric equivalent of a paired t-test for ordinal dependent variables. In this case, the null hypothesis was that there would be no difference in the choices between the two interfaces.

5.10.3 Preference Questions

At the end of the experiment, children were asked which interface they liked using better. Children selected from among three choices, which were then ordered and numbered for analysis (sequential=0, both=1, simultaneous=2). I performed a one-sample median test for all grades combined and each grade individually. The one-sample median test is the non-parametric equivalent of a one sample t-test for ordinal dependent variables. The null hypothesis in this case was that there was no preference – children should prefer both equally. To see if there was a difference in preference by grade, I performed a Kruskal-Wallis test, the non-parametric equivalent of a one-way ANOVA for ordinal dependent variables. In this case, the null hypothesis was that there would be no difference by grade.

5.10.4 Hint Counts

Throughout the experiments, I recorded the number of hints a child needed, if any, to complete a task. To analyze whether there was a difference in the number of hints by interface, I conducted a Wilcoxon signed rank test for all grades combined and for each grade individually, assuming number of hints was an interval dependent variable. In this case, the null hypothesis was that there was no difference in the number of hints by interface. To analyze whether there was a difference in the number of hints by grade, I combined the hints for both interfaces and conducted a Kruskal Wallis test. In this case, the null hypothesis was that there was no difference in the number of hints by grade.

5.10.5 Understanding of Boolean Question

For the Boolean experiment, each child did two additional tasks per interface where they selected 2 categories and then were asked what kinds of books they found. One of these tasks was done at the beginning of the child's use of an interface and one was done at the end. I decided to analyze whether a child seemed to understand that they had created a conjunctive Boolean search on the second of these tasks or not. To see if there was a difference in understanding by interface, I wanted to conduct a McNemar chi-square test. However, our sample size was not large enough. Instead, rather than comparing understanding in each interface to each other, I looked at each interface individually both overall and within each grade. I did Binomial tests to see if the number of children who understood the task while using an interface differed from the null hypothesis that half the children would understand and half would not. The Binomial test is the non-parametric equivalent of a one sample t-test for categorical dependent variables with 2 levels. To see if there was a difference by grade, I conducted a Fisher exact test, using the null hypothesis that there was no relationship between grade and understanding. A Fisher exact test is a form of Chi-square test used for small sample sizes. It is the non-parametric equivalent of a one-way ANOVA for categorical dependent variables.

5.10.6 Browsing Boolean Search Creation

During the free browsing task at the beginning of their session with each interface, I recorded all of the buttons a child clicked on. The browsing sessions were not timed exactly (roughly 2-3 minutes per interface), and children were given no instruction or assistance unless they asked, so even if a child created a Boolean search by selecting more than one leaf-level category, they may not have realized or understood that this is what they were doing. However, by counting the number of times a child selected or unselected a leaf-level category button that resulted in there being at least 2 leaf-level categories selected in each interface, I was able to get a feel for whether either interface supported the spontaneous creation of Boolean searches better.

To analyze whether there was a difference in the number of Boolean searches created by interface, I conducted a Wilcoxon signed rank test for all grades combined and for each grade individually, assuming number of searches created was an interval dependent variable. The null hypothesis was that there was no difference in the number of searches created by interface. To analyze whether there was a difference in the number of Boolean searches created by grade, I combined the searches for both interfaces and conducted a Kruskal Wallis test. In this case, the null hypothesis was that there was no difference in the number of Boolean searches by grade.

5.10.7 Browsing Navigation Button Use

Each of the interfaces required navigational clicking to find different categories, using the “More Choices” button in the simultaneous interface and the “Up Arrow” in the sequential. During the free browsing tasks, I recorded whether or not a child used these buttons. To see if there was a difference in use of any navigation button by interface, I wanted to conduct a McNemar chi-square test, the non-parametric equivalent of a paired t-test. However, the sample size was not large enough for this test. Instead, I looked at each interface individually both overall and within each grade. I did Binomial tests to see if the number of children who used a button in an interface differed from the null hypothesis that half the children would use the button and half would not. To see if there was a difference by grade, I combined all the uses of any navigation button in either interface by grade and conducted a Fisher exact test. In this case, the null hypothesis was that there would be no difference by grade.

5.11 Results: Simple Study

5.11.1 Overall

Search times for all 36 children were submitted to a 3 (grade) x 2 (interface) x 2 (task type) ANOVA. Results of this analysis indicated a significant difference by grade $F(2,33)=23.99$, $p<0.01$ and a significant interaction effect between interface and task type, $F(1,33)=26.64$, $p<0.01$. Tukey post-hoc tests on grade indicated that there were significant differences between two out of the three grade pairs. The 5th graders were nearly 3 times faster than 1st graders, and the 3rd graders were nearly 2 ½ times faster than 1st graders (Figure 23).

Tukey post-hoc tests on the interaction between interface and task type indicate that 3 of the 6 possible comparisons of pairings between levels of these factors were significant (Figure 24). In particular, one-page tasks in the simultaneous interface were nearly twice as fast as one-page tasks in the sequential interface and nearly 79 percent faster than two-page tasks in the simultaneous interface. Two-page tasks in the sequential interface were 70 percent faster than one-page tasks in the sequential interface. We expected to find no difference between one-page and two-page tasks in the sequential interface since the number of steps required to complete both tasks is the same using this interface. However, the one-page task sets contained a task that involved the Format category, a relatively abstract term. Children took more than twice as long and required more than twice as many hints to complete the tasks involving this category compared to the averages of the other tasks done with the sequential interface.

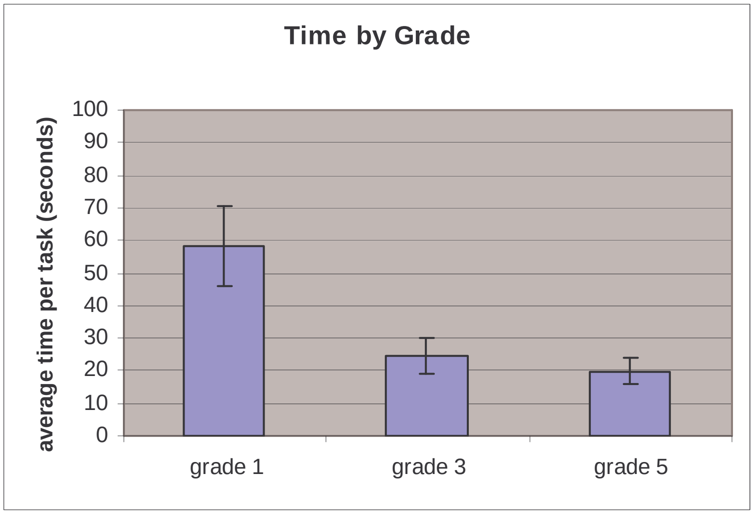

Figure 23. Average time per task by grade for simple tasks (n=36)
(Note: error bars on all time graphs are 95% confidence intervals)

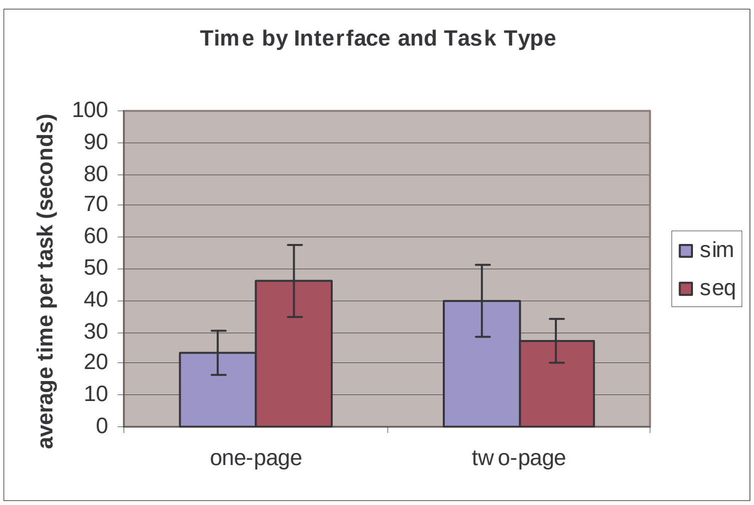

Figure 24. Average time per task by interface and task type for simple tasks (n=36)

For interface difficulty, 25 children thought the simultaneous interface was easy to use, compared with 21 children who thought the sequential interface was easy to use (Figure 25). 10 children thought the simultaneous interface was medium to use, compared with 13 for the sequential interface. 1 child thought the simultaneous interface was hard to use, compared with 2 for the sequential interface. I assigned the choices integer values (hard=0, medium=1, easy=2) and compared the responses for each interface to each other using a Wilcoxon signed rank test. This results of this test were not significant, indicating that there was no difference in perceived difficulty between the two interfaces overall.

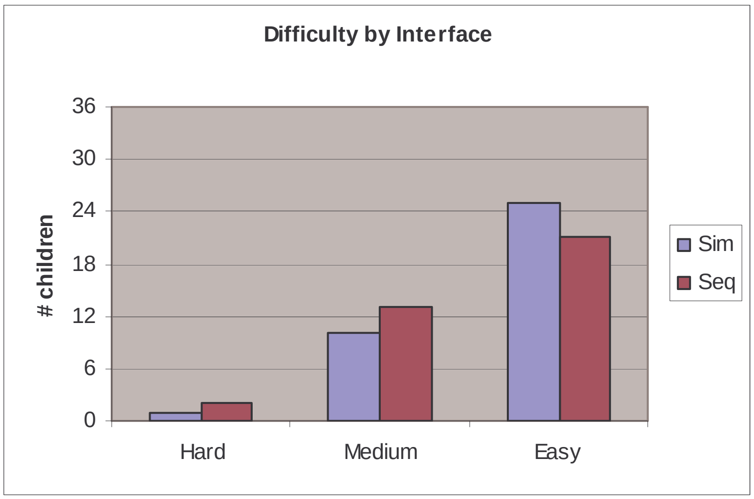

Figure 25. Difficulty by interface for simple tasks (n=36)

For interface like/dislike, 31 children said they liked using the simultaneous interface a lot, compared with 27 who liked using the sequential a lot (Figure 26). 5 children said they liked using the simultaneous interface a little, compared with 8 who liked using the sequential a little. No children said they did not like using the simultaneous interface much, compared with 1 child who did not like using the sequential much. I assigned the choices integer values (not much=0, a little=1, a lot=2) and compared the responses for each interface to each other using a Wilcoxon signed rank test. The results of this test were not significant, indicating that there was no difference in likeability between the two interfaces overall.

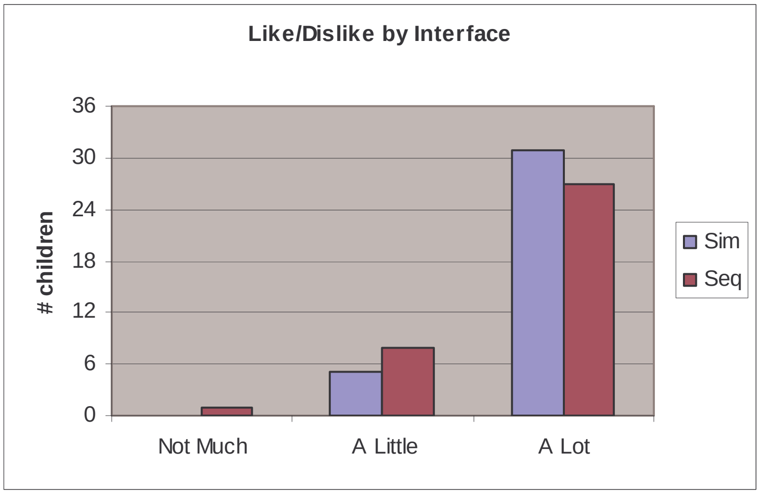

Figure 26. Like/dislike by interface for simple tasks (n=36)

Overall, 11 children preferred the simultaneous interface, 5 preferred the sequential interface, and 20 like both equally (Figure 27). I ordered these choices (sequential=0, both=1, simultaneous=2) and performed a one-sample median test to see if the actual median value differed from the null hypothesis that children would prefer both equally. The results of this test were not significant, indicating that neither interface was preferable to the other overall. To see if there was an affect by grade on preference (Figure 28), I performed a Kruskal-Wallis test. The results of this test were also not significant, indicating that preference did not change by grade.

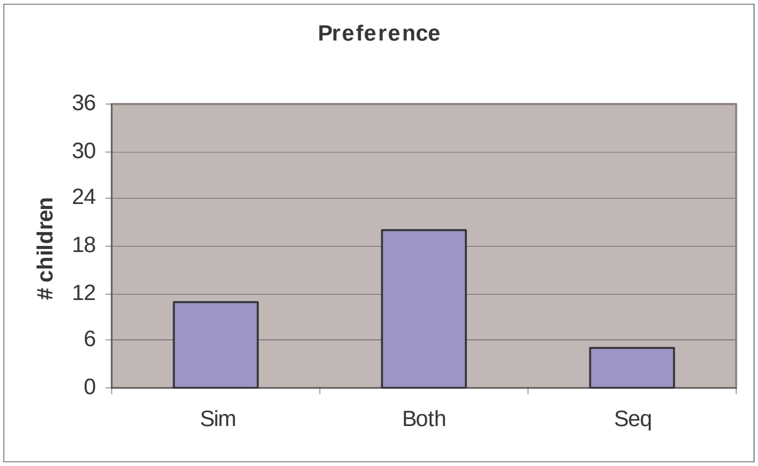

Figure 27. Preference for simple tasks (n=36)

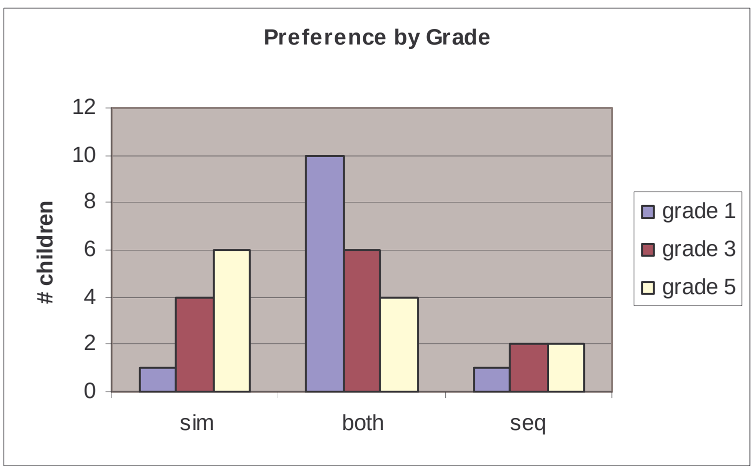

Figure 28. Preference by grade for simple tasks (n=36)

In the simultaneous interface, children needed 79 total hints to complete the timed tasks, compared to 84 total hints in the sequential interface. To analyze whether there was a difference in the number of hints overall by interface, I conducted a Wilcoxon signed rank test, using a null hypothesis that there would be no difference in the number of hints between the 2 interfaces. This results of this test were not significant, indicating that there was no difference between the two interfaces overall. Combining both interfaces and looking at the data by grade, 1st graders needed an average of 10.1 hints, 3rd graders needed an average of 2.25 hints, 5th graders need an average of 1.25 hints. To analyze whether there was a difference in the number of hints by grade, I conducted a Kruskal Wallis test. This results of this test were significant ($p<0.01$), indicating that younger children needed significantly more hints than older children.

Looking at just the hints in the simultaneous interface by task type (one-page vs two-page), the results of a Wilcoxon signed rank test were not significant, indicating that two-page tasks did not require more hints than one-page tasks. The same was also true for hints in the simultaneous interface for each grade.

During the free browsing tasks, I counted the number of Boolean searches a child created in each interface, where a Boolean search was defined as having at least 2 leaf-level categories selected. Children created 183 Boolean searches in the simultaneous interface and 79 in the sequential interface. To analyze whether there was a difference in the number of searches created by interface, I conducted a Wilcoxon signed rank test, using a null hypothesis that there would be no difference between the 2 interfaces. The results of this test were significant ($p<0.01$), indicating that significantly more Boolean searches were created in the simultaneous interface than in the sequential interface. Combining both interfaces and looking at the data by grade, 1st graders did an average of 5.8 searches, 3rd graders did an average of 6.7 searches, and 5th graders did an average of 9.3 searches. To analyze whether there was a difference in the number searches by grade, I conducted a Kruskal Wallis test.

The results of this test were not significant, indicating that neither older nor younger children created more Boolean searches.

During the free browsing tasks, I counted the number of children who used the More Choices and Up Arrow buttons on their own. 9 children used the More Choices button in the simultaneous interface, while 27 did not. To see if these values differed from the null hypothesis that half the children would use this button, I conducted a Binomial test. The results of this test were significant (one-sided $p<0.01$), indicating that significantly more children did not find this button than did find it. In the sequential interface, 21 children used the Up Arrow and 15 did not. To see if these values differed from the null hypothesis, I conducted a Binomial test. The results of this test were not significant. Combining both interfaces and looking at the data by grade, there were 4 uses of any navigation tool in first grade, and 13 each in 3rd and 5th grade. To analyze this data, I conducted a Fisher exact test to see if there was a difference between use of any navigation tool by grade. This difference was significant ($p<0.01$), indicating that significantly more children in 3rd and 5th grade used the navigation tools on their own than 1st grade children.

Finally, I looked at both time and hints for individual tasks to understand how hard or easy our own category structure was to use (Table 6, Figure 29 and Figure 30). In the sequential interface, task 1 appears to take relatively longer and require more hints. This task involved finding picture books (task set 1) or chapter books (task set 2). These buttons were located on the first page in the simultaneous interface, and were thus relatively easy to find with visual scanning. However, they required looking under the "Format" category in the sequential interface, a rather abstract term that many children were not familiar with.

In the simultaneous interface, task 5 appears to take relatively longer and require more hints. This task required finding 4-star rated books (task set 1) or 5-star rated books (task set 2). I observed that a number of children did not understand what was meant by rating, and asked for clarification. In the sequential interface, the top level category they had to look under to find these buttons was labeled "Rating", while in the simultaneous interface, these buttons required navigating to the second page using the "More Choices" button and then scanning for the relevant button. These trends held true at each grade level as well.

Task Number	**Task Set 1**	**Task Set 2**
1	Picture books	Chapter books
2	Books for six to nine year olds	Books for three to five year olds
3	Medium books	Long books
4	Purple books	Pink books
5	Four star rated books	Five star rated books
6	Sad books	Happy books

Table 6. Simple tasks, phrased "How many … are there?"

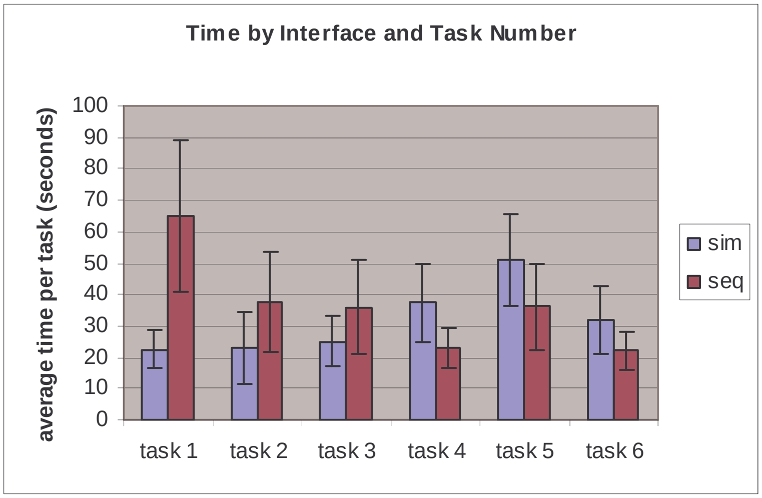

Figure 29. Average time to complete individual simple tasks (n=36)

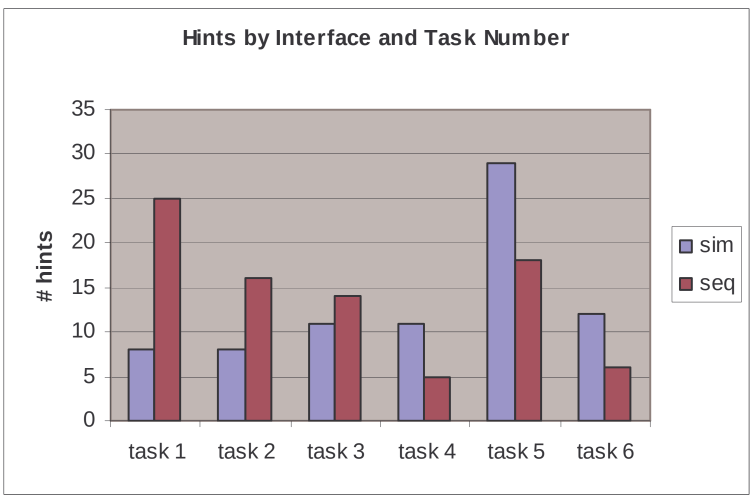

Figure 30. Total hints needed to complete individual simple tasks (n=36)

5.11.2 Grade 1

Search times for all 12 children were submitted to a 2 (interface) x 2 (task type) ANOVA. Results of this analysis indicated a significant interaction effect between interface and task type, $F(1,11)=11.85$, $p<0.01$. Tukey post-hoc tests on the interaction effect indicate a significant difference between one-page tasks in the simultaneous interface and one-page tasks in the sequential interface, with the simultaneous being nearly twice as fast (Figure 31).

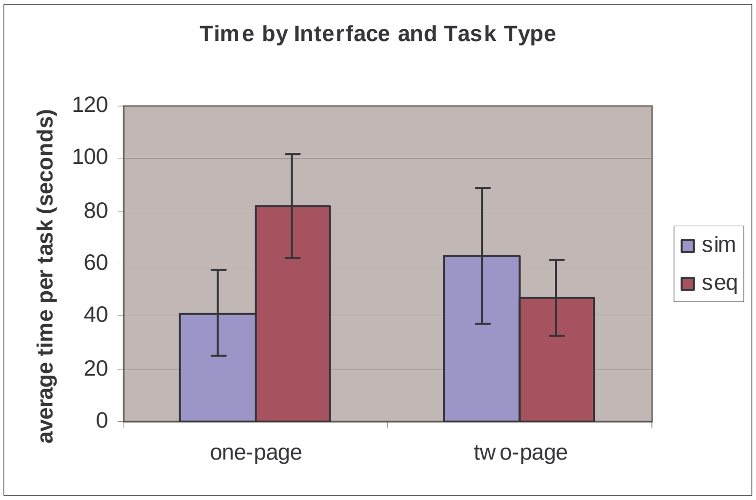

Figure 31. Average time per task by interface and task type for simple tasks for 1st grade (n=12)

For interface difficulty, 9 children thought the simultaneous interface was easy to use, compared with 8 children who thought the sequential was easy to use (Figure 32). 2 children thought the simultaneous interface was medium to use, compared with 3 for the sequential. 1 child thought the simultaneous interface was hard to use, compared with 1 for the sequential. I assigned the choices integer values and compared the responses using a Wilcoxon signed rank test. The results of this test were not significant, indicating that there was no difference in perceived difficulty between the two interfaces.

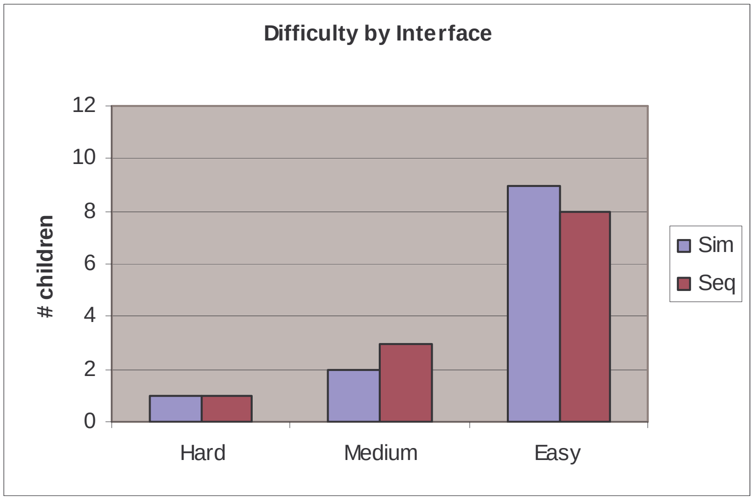

Figure 32. Difficulty by interface for simple tasks for 1st grade (n=12)

For interface like/dislike, 11 children said they liked using the simultaneous interface a lot, compared with 9 who liked using the sequential a lot (Figure 33). 1 child said they liked using the simultaneous interface a little, compared with 2 who liked using the sequential interface a little. No child said they did not like using the simultaneous interface much, compared with 1 child who did not like using the sequential interface much. I assigned the choices integer values and compared the responses for each interface to each other using Wilcoxon signed rank tests. The results of this test were not significant, indicating that there was no difference in likeability between the two interfaces.

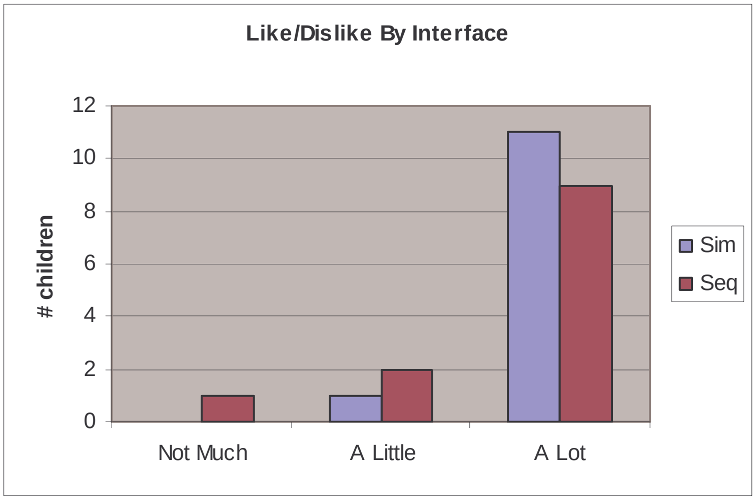

Figure 33. Like/dislike by interface for simple tasks for 1st grade (n=12)

In this grade, 1 child preferred the simultaneous interface, 1 preferred the sequential, and 10 liked both equally (Figure 28). I ordered these choices and performed a one-sample median test to see if the median value differed from the null hypothesis that children would prefer both equally. The results of this test were not statistically significant, indicating that neither interface was preferable.

In the simultaneous interface, children needed 57 total hints to complete the timed tasks, compared to 64 total hints in the sequential. To analyze whether there was a difference in the number of hints by interface, I conducted a Wilcoxon signed rank test. The results of this test were not significant, indicating that there was no difference between the two interfaces for the number of hints required.

For Boolean search creation during the free browsing session, 1st graders created 47 Boolean searches in the simultaneous interface and 23 in the sequential. To analyze whether there was a difference in the number of searches created by interface, I conducted a Wilcoxon signed rank test, using a null hypothesis that there would be no difference between the 2 interfaces. The results of this test were significant ($p=0.02$), indicating that significantly more Boolean searches were created in the simultaneous interface than in the sequential.

For navigation tool use during the free browsing session, no child used the More Choices button in the simultaneous interface. To see if this value differed from the null hypothesis that half the children would use this button, I conducted a Binomial test. The results of this test were significant (one-sided $p<0.01$), indicating that significantly more children did not use this button than did use it. For the sequential navigation tool, 4 children used the Up Arrow and 8 did not. The results of a Binomial test were not significant, indicating that children were equally likely to use this button as not.

5.11.3 Grade 3

Search times for all 12 children were submitted to a 2 (interface) x 2 (task type) ANOVA. Results indicated a significant interaction effect between interface and task type, $F(1,11)=6.37$, $p=0.03$. Tukey post-hoc tests on the interaction effect indicate that none of the comparisons between pairings of interface and task type were significant, but one-page tasks in the simultaneous interface were marginally faster than two-page tasks in the simultaneous interface, $p=0.08$ (Figure 34).

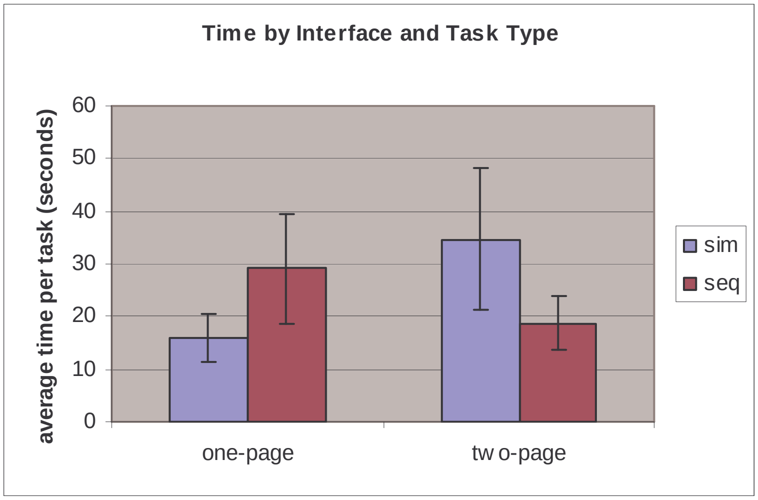

Figure 34. Average time per task by interface and task type for simple tasks for 3rd grade (n=12)

For interface difficulty, 7 children thought the simultaneous interface was easy to use, compared with 6 children who thought the sequential was easy to use (Figure 35). 5 children thought the simultaneous interface was medium to use, compared with 6 for the sequential. No child thought either the simultaneous or sequential interface was hard to use. I assigned the choices integer values and compared the responses using a Wilcoxon signed rank test. The results of this test were not significant, indicating that there was no difference in perceived difficulty between the two interfaces.

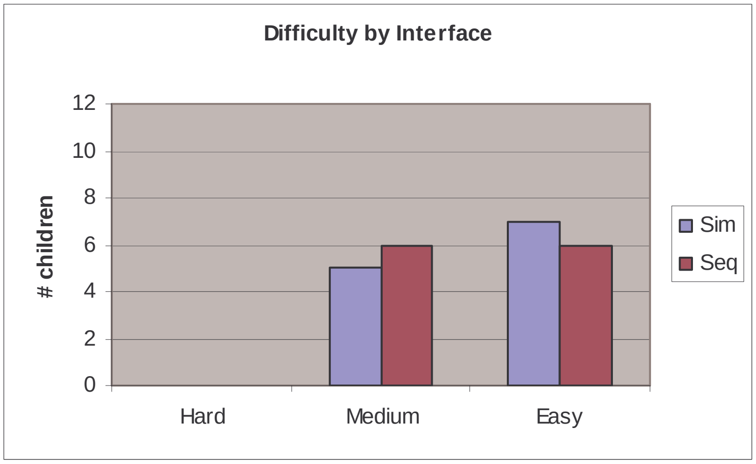

Figure 35. Difficulty by interface for simple tasks for 3rd grade (n=12)

For interface like/dislike, 10 children said they liked using the simultaneous interface a lot, compared with 10 who liked using the sequential a lot (Figure 36). 2 children said they liked using the simultaneous interface a little, compared with 2 who liked using the sequential a little. No children said they did not like using either the interface much. I assigned the choices integer values and compared the responses using a Wilcoxon signed rank test. The results of this test were not significant, indicating that there was no difference in likeability between the two interfaces.

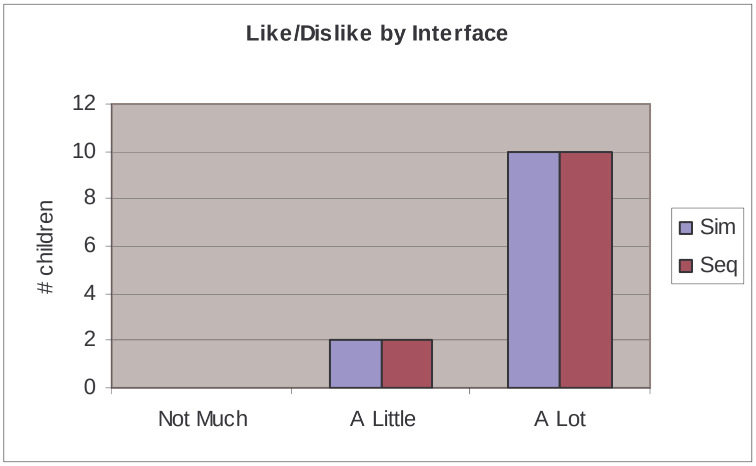

Figure 36. Like/dislike by interface for simple tasks for 3rd grade (n=12)

In this grade, 4 children preferred the simultaneous interface, 2 preferred the sequential, and 6 like both equally (Figure 28). I ordered these choices and performed a one-sample median test to see if the median value differed from the null hypothesis that children would prefer both equally. The results of this test were not significant, indicating that neither interface was preferable.

In the simultaneous interface, children needed 15 total hints to complete the timed tasks, compared to 12 total hints in the sequential. To analyze whether there was a difference in the number of hints by interface, I conducted a Wilcoxon signed rank test. The results of this test were not significant, indicating that there was no difference between the two interfaces for the number of hints required.

For Boolean search creation during the free browsing session, 3rd graders created 59 Boolean searches in the simultaneous interface and 21 in the sequential. To analyze whether there was a difference in the number of searches created by interface, I conducted a Wilcoxon signed rank test, using a null hypothesis that there would be no difference between the 2 interfaces. The results of this test were significant ($p=0.02$), indicating that significantly more Boolean searches were created in the simultaneous interface than in the sequential.

For navigation tool use during the free browsing session, 5 children used the More Choices button and 7 did not. To see if this value differed from the null hypothesis that half the children would use this button, I conducted a Binomial test. The results of this test were not significant. For the sequential navigation tool, 8 children used the Up Arrow and 4 did not. The results of a Binomial test were not significant.

5.11.4 Grade 5

Search times for all 12 children were submitted to a 2 (interface) x 2 (task type) ANOVA. Results of this analysis indicated a significant interaction effect between interface and task type, $F(1,11)=17.83$, $p<0.01$. Tukey post-hoc tests on the interaction effect indicate that two of the pairwise comparisons were significant. One-page tasks in the simultaneous interface were twice as fast as one-page tasks in the sequential interface and 66% faster than two-page tasks in the simultaneous interface (Figure 37).

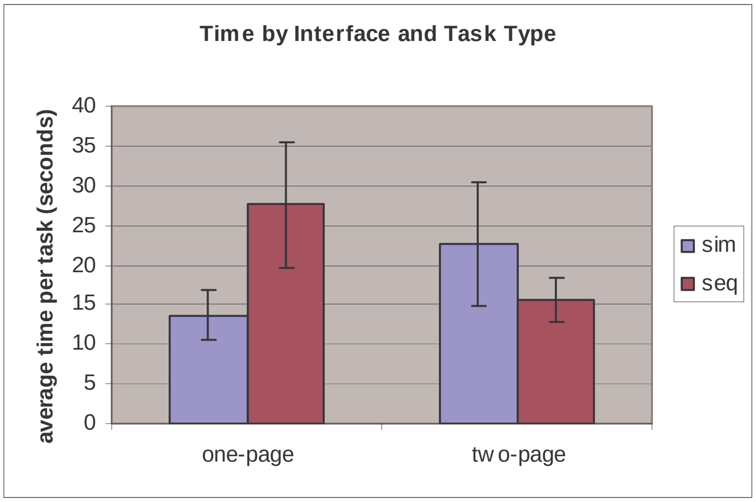

Figure 37. Average time per task by interface and task type for simple tasks for 5th grade (n=12)

For interface difficulty, 9 children thought the simultaneous interface was easy to use, compared with 7 children who thought the sequential was easy to use (Figure 38). 3 children thought the simultaneous interface was medium to use, compared with 4 for the sequential. No child thought the simultaneous interface was hard to use, compared with 1 for the sequential. I assigned the choices integer values and compared the responses using a Wilcoxon signed rank test. The results of this test were not significant, indicating that there was no difference in perceived difficulty between the two interfaces.

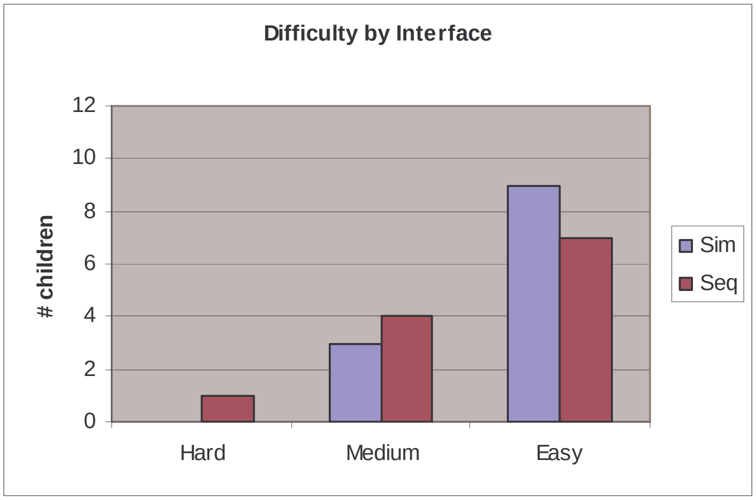

Figure 38. Difficulty by interface for simple tasks for 5th grade (n=12)

For interface like/dislike, 10 children said they liked using the simultaneous interface a lot, compared with 8 who liked using the sequential a lot (Figure 39). 2 children said they liked using the simultaneous interface a little, compared with 4 who liked using the sequential a little. No child said they did not like using either the simultaneous or sequential interface much. I assigned the choices integer values and compared the responses using a Wilcoxon signed rank test. The results of this test were not significant, indicating no difference in likeability between the two interfaces.

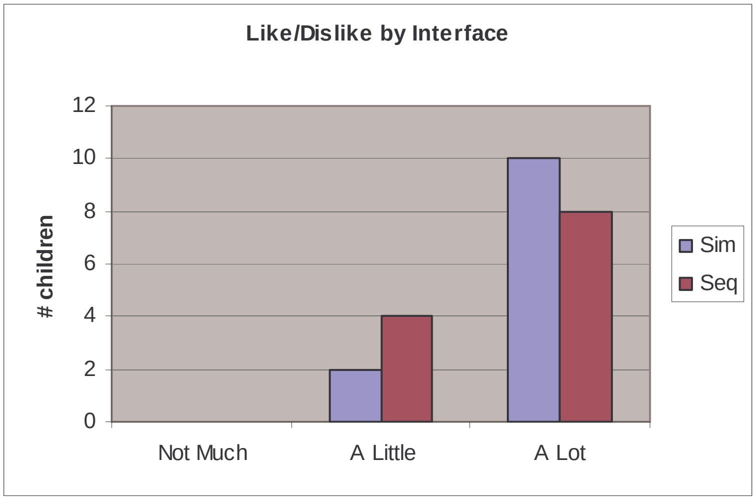

Figure 39. Like/dislike by interface for simple tasks for 5th grade (n=12)

In this grade, 6 children preferred the simultaneous interface, 2 preferred the sequential, and 4 liked both equally (Figure 28). I ordered these choices and performed a one-sample median test to see if the median value differed from the null hypothesis that children would prefer both equally. The results of this test were not significant, indicating that neither interface was preferable.

In the simultaneous interface, children needed 7 total hints to complete the timed tasks, compared to 8 total hints in the sequential interface. To analyze whether there was a difference in the number of hints by interface, I conducted a Wilcoxon signed rank test. The results of this test were not significant, indicating that there was no difference between the two interfaces for the number of hints required.

For Boolean search creation during the free browsing session, 5th graders created 77 Boolean searches in the simultaneous interface and 35 in the sequential. To analyze whether there was a difference in the number of searches created by interface, I conducted a Wilcoxon signed rank test, using a null hypothesis that there would be no difference between the 2 interfaces. The results of this test were significant ($p=0.02$), indicating that significantly more Boolean searches were created in the simultaneous interface than in the sequential.

For navigation tool use during the free browsing session, 4 children used the More Choices button and 8 did not. To see if this value differed from the null hypothesis that half the children would use this button, I conducted a Binomial test. The results of this test were not significant. For the sequential navigation tool, 9 children used the Up Arrow and 3 did not. The results of a Binomial test were significant (one-sided $p=0.04$), indicating that more children used the Up Arrow button than not.

5.12 Results: Boolean Study

5.12.1 Overall

Search times for all 36 children were submitted to a 3 (grade) x 2 (interface) x 2 (task type) ANOVA. Results of this analysis indicated significant differences by grade $F(2,33)=19.96$, $p<0.01$, and interface, $F(1,33)=53.25$, $p<0.01$, and a significant interaction effect between interface and task type, $F(1,33)=18.71$, $p<0.01$. For the interface effect, the simultaneous interface was 68% faster than the sequential interface (Figure 40).

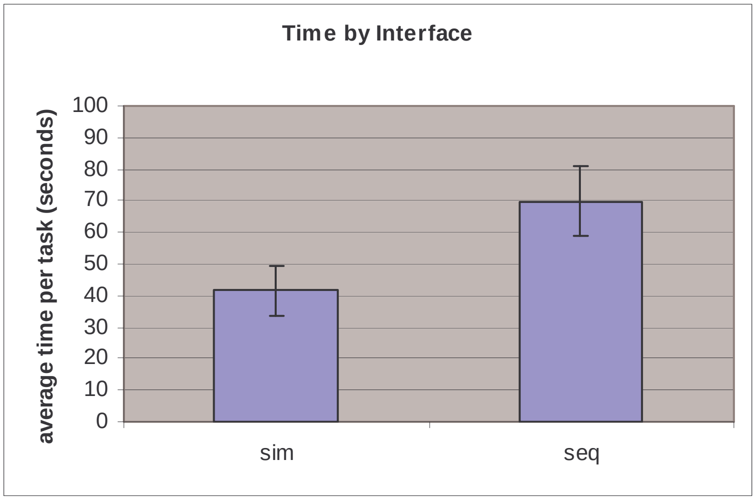

Figure 40. Average time per task by interface for Boolean tasks (n=36)

Tukey post-hoc tests on grade indicated that there were significant differences between all 3 grades (Figure 41). The 5th graders were 72% faster than the 3rd graders and more than twice as fast as the 1st graders. The 3rd graders were 31% faster than the 1st graders. Tukey post-hoc tests on the interaction between interface and task type indicate that 5 out of the 6 possible comparisons of pairings between levels of these 2 factors were significant (Figure 42). In particular, one-page tasks in the simultaneous interface were more than twice as fast as one-page tasks in the sequential interface, 32% faster than two-page tasks in the simultaneous interface, and 77% faster than two page tasks in the sequential interface. Two-page tasks in the simultaneous interface were 62% faster than one-page tasks in the sequential interface and 35% faster than two-page tasks in the sequential interface.

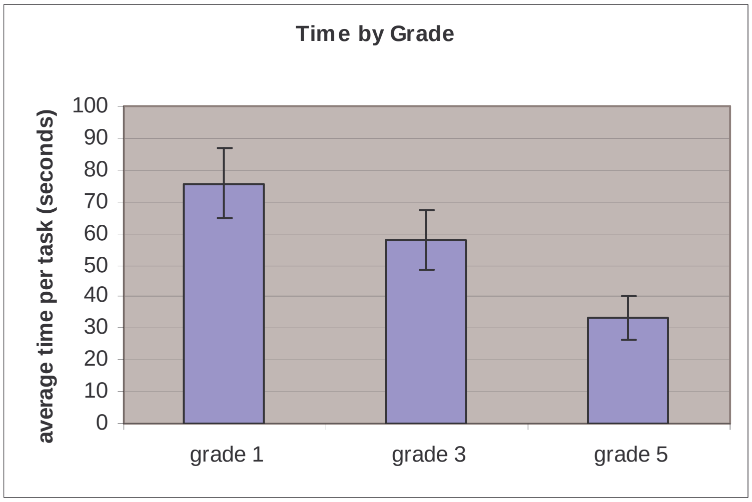

Figure 41. Average time per task by grade for Boolean tasks (n=36)

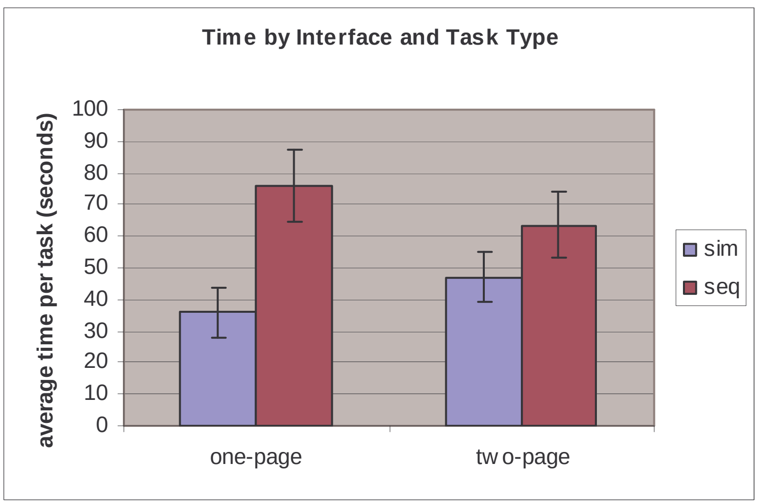

Figure 42. Average time per task by interface and task type for Boolean tasks (n=36)

For interface difficulty, 28 children thought the simultaneous interface was easy to use, compared with 12 children who thought the sequential interface was easy to use (Figure 43). 7 children thought the simultaneous interface was medium to use, compared with 19 for the sequential interface. 1 child thought the simultaneous interface was hard to use, compared with 5 for the sequential interface. I assigned the choices integer values and compared the responses for each interface to each other using a Wilcoxon signed rank test. The results of this test were significant ($p<0.01$), indicating that the simultaneous interface was considered significantly easier than the sequential interface.

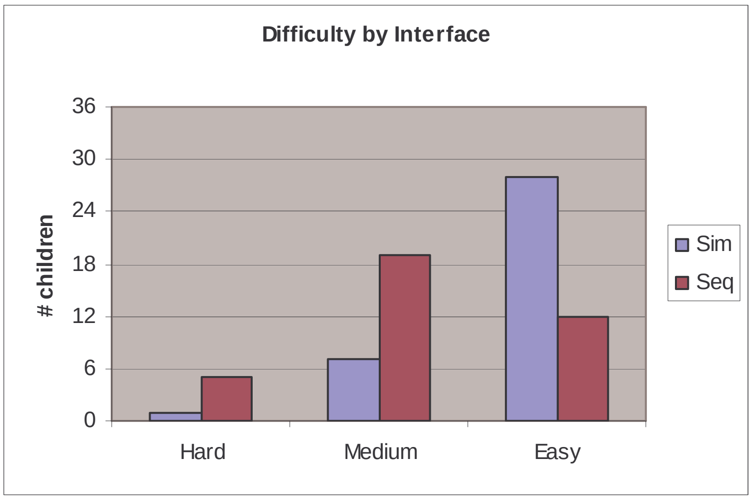

Figure 43. Difficulty by interface for Boolean tasks (n=36)

For interface like/dislike, 33 children said they liked using the simultaneous interface a lot, compared with 24 who liked using the sequential interface a lot (Figure 44). 3 children said they liked using the simultaneous interface a little, compared with 10 who liked using the sequential interface a little. No children said they did not like using the simultaneous interface much, compared with 2 children who did not like using the sequential interface much. I assigned the choices integer values and compared the responses for each interface to each other using a Wilcoxon signed rank test. The results of this test were significant ($p<0.01$), indicating that the simultaneous interface was better liked than the sequential interface.

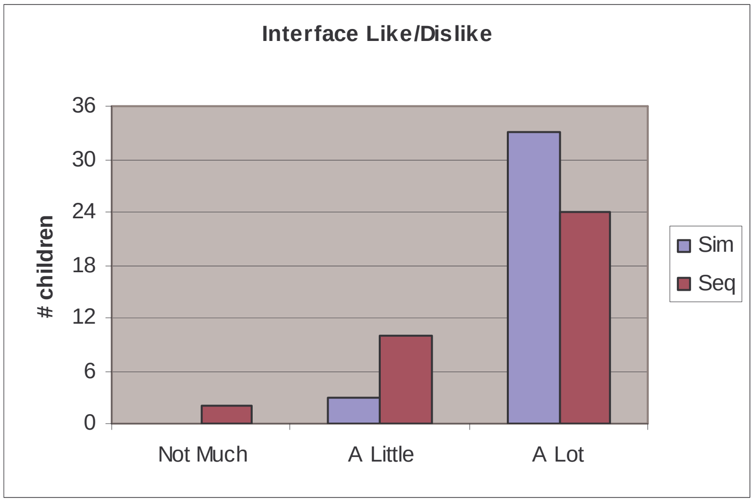

Figure 44. Interface like/dislike for Boolean tasks (n=36)

Overall, 19 children preferred the simultaneous interface, 4 preferred the sequential, and 13 like both equally (Figure 45). I ordered these choices and performed a one-sample median test to see if the median value differed from the null hypothesis that children would prefer both equally. The results of this test were significant ($p<0.01$), indicating that significantly more children preferred the simultaneous interface. To see if there was an affect by grade on preference (Figure 46), I performed a Kruskal-Wallis test. The results of this test were not significant, indicating that there was no difference in preference by grade.

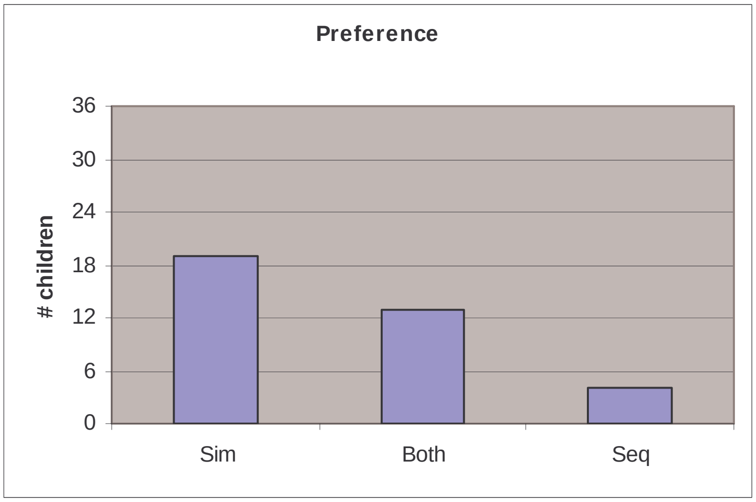

Figure 45. Preference for Boolean tasks (n=36)

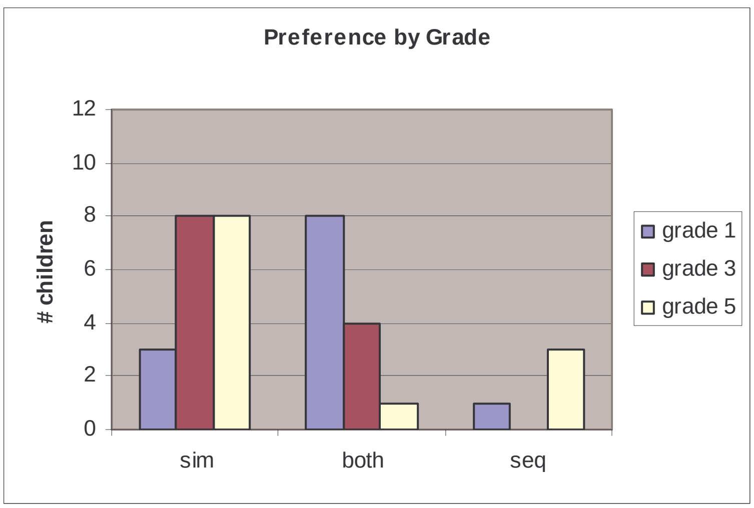

Figure 46. Preference by grade for Boolean tasks (n=36)

In the simultaneous interface, children needed 116 total hints to complete the timed tasks, compared to 221 total hints in the sequential interface. To analyze whether there was a difference in the number of hints overall by interface, I conducted a Wilcoxon signed rank test, using a null hypothesis that there would be no difference in the number of hints between the 2 interfaces. The results of this test were significant ($p<0.01$), indicating that significantly more hints were required in the sequential interface. Combining both interfaces and looking at the data by grade, 1st graders needed an average of 18.2 hints 3rd graders needed an average of 7.4 hints, 5th graders needed an average of 2.5 hints. To analyze whether there was a difference in the number of hints by grade, I conducted a Kruskal Wallis test. The results of this test were significant ($p<0.01$), indicating that younger children needed significantly more hints than older children.

Looking at just the hints in the simultaneous interface by task type (one-page vs two-page), the results of a Wilcoxon signed rank test were not significant, indicating that two-page tasks did not require more hints than one-page tasks. The same was also true for hints in the simultaneous interface for each grade.

During the free browsing tasks, I counted the number of Boolean searches a child created in each interface, where a Boolean search was defined as having at least 2 leaf-level categories selected. Children created 224 Boolean searches in the simultaneous interface and 104 in the sequential interface. To analyze whether there was a difference in the number of searches created by interface, I conducted a Wilcoxon signed rank test, using a null hypothesis that there would be no difference between the 2 interfaces. The results of this test were significant ($p<0.01$), indicating that significantly more Boolean searches were created in the simultaneous interface than in the sequential interface. Combining both interfaces and looking at the data by grade, 1st graders did an average of 7.9 searches, 3rd graders did an average of 7.3

searches, and 5th graders did an average of 12.2 searches. To analyze whether there was a difference in the number searches by grade, I conducted a Kruskal Wallis test. The results of this test were not significant, indicating that neither older nor younger children created more Boolean searches.

During the free browsing tasks, I counted the number of children who used the More Choices and Up Arrow buttons on their own. 12 children used the More Choices button in the simultaneous interface, while 24 did not. To see if these values differed from the null hypothesis that half the children would use this button, I conducted a Binomial test. The results of this test were significant (one-sided $p<0.01$), indicating that significantly more children did not find this button than did find it. In the sequential interface, 23 children used the Up Arrow and 13 did not. To see if these values differed from the null hypothesis, I conducted a Binomial test. The results of this test were significant (one-sided $p=0.05$), indicating that significantly more children found this button than did not find it. Combining both interfaces and looking at the data by grade, there were 10 uses of a navigation tool in first grade, 8 in 3rd grade, and 17 in 5th grade. To analyze this data, I conducted a Fisher exact test to see if there was a difference in use of any navigation tool by grade. This difference was significant ($p=0.03$), indicating that significantly more children in 5th grade used the navigation tools on their own than 3rd and 1st grade children.

For comprehension, 22 children understood what type of task they were doing (i.e. conjunctive Boolean) in the simultaneous interface, compared with 14 who did not

(Figure 47). To see if these values differed from the null hypothesis that half the children would understand and half would not, I conducted a Binomial test. The results of this test were not significant. In the sequential interface 18 children understood what type of task they were doing and 18 did not, indicating that children were half as likely to understand as not. To see if there was a difference in understanding by grade, I combined both interfaces together and conducted a Fisher exact test. There were 10 instances of understanding in either interface in 1st grade, 7 in 3rd grade, and 23 in 5th grade. The results of this test were significant ($p<0.01$), indicating that 5th graders understood more than 1st and 3rd graders (Figure 47).

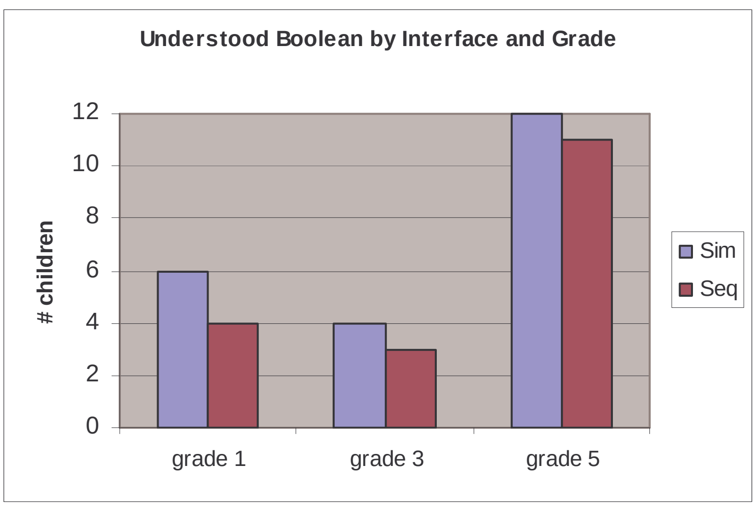

Figure 47. Understood Boolean task by interface and grade for Boolean tasks (n=36)

Finally, I looked at both time and hints for individual tasks to understand how hard or easy our own category structure was to use (Table 7, Figure 48 and Figure 49). In the sequential interface, task 1 appears to take relatively longer and require more hints. This task involved finding award-winning chapter books (task set 1) or fairy tale picture books (task set 2). All of these buttons were located on the first page in the simultaneous interface, and were thus relatively easy to find with visual scanning. However, they required looking under the "Format" and "Genre (Different Kinds)" categories in the sequential interface, both rather abstract terms that many children were not familiar with.

In the sequential interface, task 3 appears to take relatively longer and require more hints. This task involved finding long books about imaginary animal characters (task set 1) or recently added books about real animal characters (task set 2). All of these buttons were located on the first page in the simultaneous interface, and were thus relatively easy to find with visual scanning. However, in the sequential interface, some children were confused between the real vs. imaginary character categories and the true vs. make believe categories. Long books gave some children problems, as they thought perhaps long books would be under the age category, associated with older children, or under the format category, associated with chapter books. Recently added books also caused some problems in the sequential interface, as it wasn't obviously a member of the Genre category.

In the sequential interface, task 4 appears to take relatively longer and require more hints. This task involved finding comic books for six to nine year olds (task set 1) or plays for ten to thirteen year olds (task set 2). In this case, I observed that a number of children had difficulty deciding where to look for comic books and plays in the sequential interface (under Format), and some children didn't understand what the plays category meant, requiring an explanation that plays can be written down in books as well as performed on stage.

Finally, task 2 appears to require more hints in the simultaneous interface than in the sequential interface. This task involved finding blue make believe books (task set 1) or rainbow true books (task set 2). This task was challenging in the simultaneous interface because the first screen of this interface contained buttons for real animal characters, which some children saw first and selected instead of true books, and buttons for fairy tales and imaginary animal characters, which some children saw first and selected instead of make believe books. In the sequential interface, this wasn't an issue because all of these buttons were under different top level categories. The true and make believe categories were located under a button called "True vs. Make Believe", while the real and imaginary animal characters were under a button called "Characters" and the fairy tales were under "Genre (Different Kinds)".

Task Number	**Task Set 1**	**Task Set 2**
1	Award-winning chapter books	Fairy tale picture books
2	Blue make believe books	Rainbow true books
3	Long books about imaginary animal characters	Recently added books about real animal characters
4	Comic books for six to nine year olds	Plays for ten to thirteen year olds
5	White five star rated books	Black four star rated books
6	Short sad books	Medium happy books

Table 7. Boolean tasks, phrased "How many … are there?"

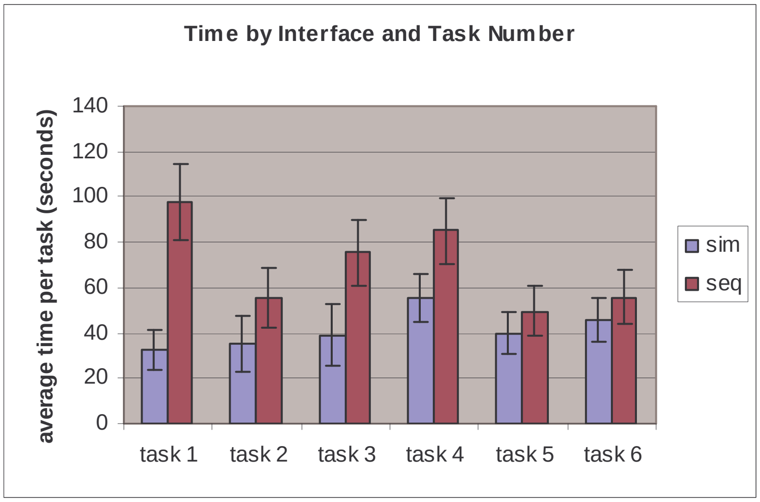

Figure 48. Average time by interface and task number for Boolean tasks (n=36)

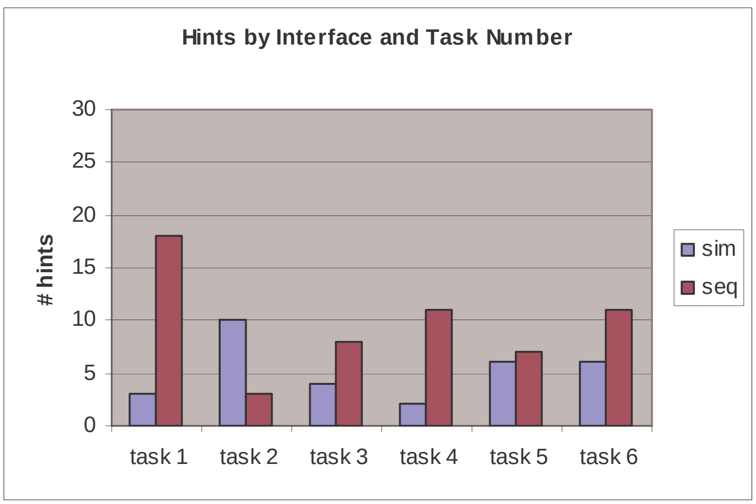

Figure 49. Total hints by interface and task number for Boolean tasks (n=36)

5.12.2 Grade 1

Search times for all 12 children were submitted to a 2 (interface) x 2 (task type) ANOVA. Results of this analysis indicated a significant difference by interface, $F(1,11)=31.08$, $p<0.01$, and a significant interaction effect between interface and task type, $F(1,11)=6.48$, $p=0.03$. For the interface effect, the simultaneous interface was

68% faster than the sequential. Tukey post-hoc tests on the interaction effect indicate that one-page tasks in the simultaneous interface were more than twice as fast as one-page tasks in the sequential and 73% faster than two-page tasks in the sequential. Two-page tasks in the simultaneous interface were 64% faster than one-page tasks in the sequential (Figure 50).

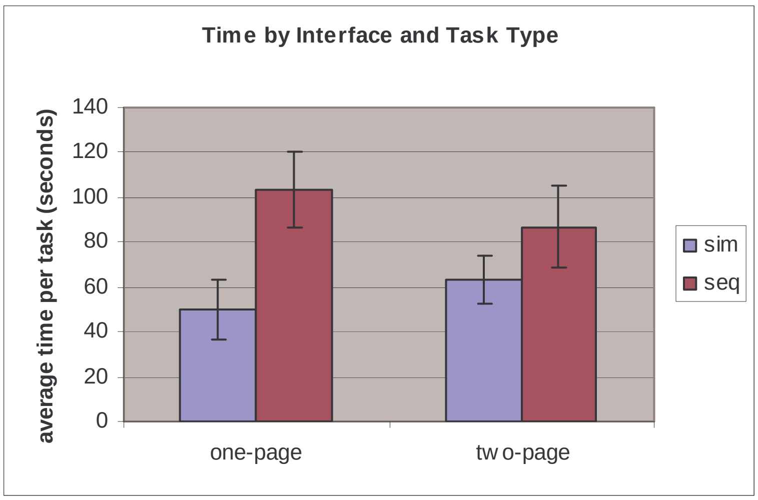

Figure 50. Average time per task by interface and task types for Boolean tasks for 1st grade (n=12)

For interface difficulty, 9 children thought the simultaneous interface was easy to use, compared with 4 children who thought the sequential was easy to use (Figure 51). 2 children thought the simultaneous interface was medium to use, compared with 5 for the sequential. 1 child thought the simultaneous interface was hard to use, compared with 3 for the sequential. I assigned the choices integer values and compared the responses using a Wilcoxon signed rank test. The results of this test were not significant, indicating that there was no difference in perceived difficulty between the two interfaces.

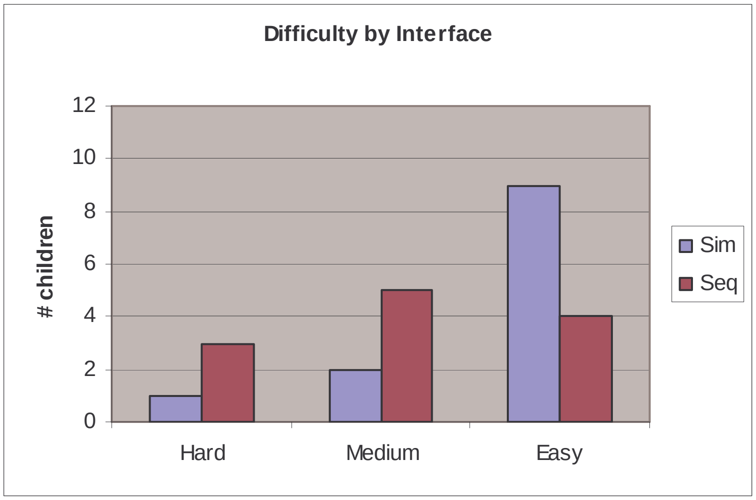

Figure 51. Difficulty by interface for Boolean tasks for 1st grade (n=12)

For interface like/dislike, 10 children said they liked using the simultaneous interface a lot, compared with 8 who liked using the sequential a lot (Figure 52). 2 children said they liked using the simultaneous interface a little, compared with 3 who liked using the sequential a little. No children said they did not like using the simultaneous interface much, compared with 1 child who did not like using the sequential much. I assigned the choices integer values and compared the responses using a Wilcoxon signed rank test. The results of this test were not significant, indicating that there was no difference in likeability between the two interfaces.

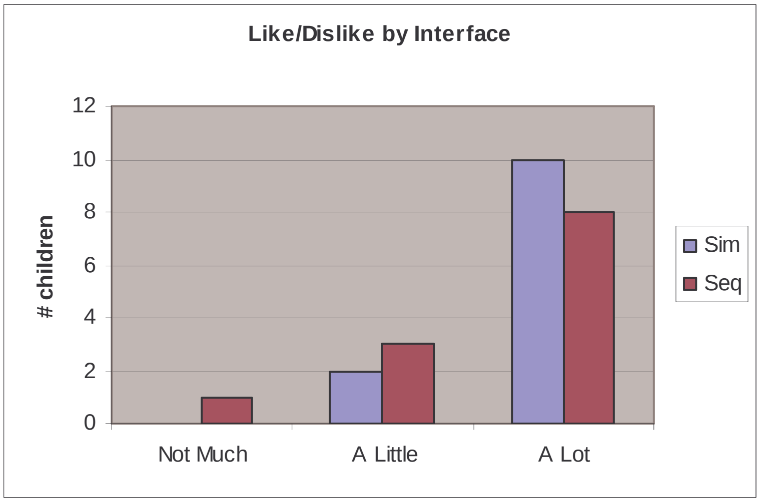

Figure 52. Like/dislike by interface for Boolean tasks for 1st grade (n=12)

In this grade, 3 children preferred the simultaneous interface, 1 preferred the sequential, and 8 like both equally (Figure 46). I ordered these choices and performed a one-sample median test to see if the median value differed from the null hypothesis that children would prefer both equally. The results of this test were not significant, indicating that neither interface was preferable.

In the simultaneous interface, children needed 76 total hints to complete the timed tasks, compared to 142 total hints in the sequential. To analyze whether there was a difference in the number of hints by interface, I conducted a Wilcoxon signed rank test. The results of this test were significant (p=0.01), indicating that significantly more hints were required in the sequential interface than in the simultaneous.

For the simultaneous interface, 6 children understood that they were creating a conjunctive Boolean query and 6 did not. For the sequential interface, 4 children understood and 8 did not. The results of Binomial tests on both of these interfaces were not significant.

For Boolean search creation during the free browsing session, 1st graders created 60 Boolean searches in the simultaneous interface and 34 in the sequential. To analyze whether there was a difference in the number of searches created by interface, I conducted a Wilcoxon signed rank test, using a null hypothesis that there would be no difference between the 2 interfaces. The results of this test were not significant.

For navigation tool use during the free browsing session, 1 child used the More Choices button and 11 did not. To see if this value differed from the null hypothesis that half the children would use this button, I conducted a Binomial test. The results of this test were significant (one-sided $p<0.01$), indicating that significantly more children did not use this button than did use it. For the sequential navigation tool, 9 children used the Up Arrow and 3 did not. The results of a Binomial test were significant (one-sided $p=0.04$), indicating that significantly more children used this button than not.

5.12.3 Grade 3

Search times for all 12 children were submitted to a 2 (interface) x 2 (task type) ANOVA. Results of this analysis indicated a significant difference by interface, $F(1,11)=8.27$, $p=0.02$, and a significant interaction effect between interface and task type, $F(1,11)=6.75$, $p=0.02$. For the interface effect, the simultaneous interface was 55% faster than the sequential. Tukey post-hoc tests on the interaction effect indicate that one-page tasks in the simultaneous interface were significantly faster than one-page tasks in the sequential interface by almost twice (Figure 53).

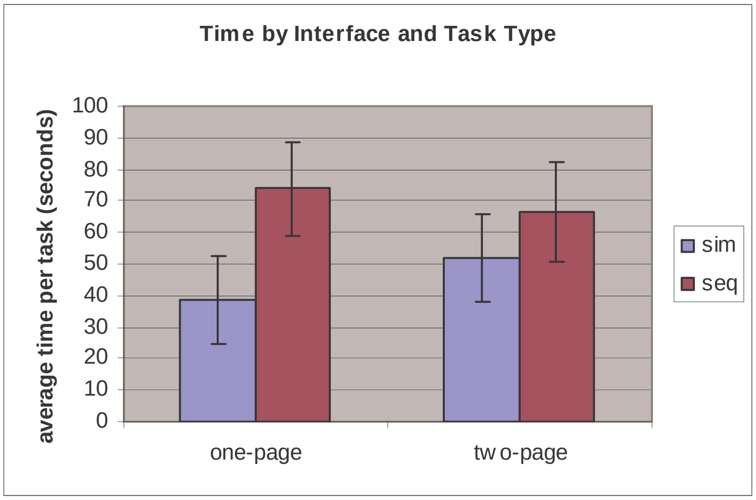

Figure 53. Average time per task by interface and task type for Boolean tasks for 3rd grade (n=12)

For interface difficulty, 8 children thought the simultaneous interface was easy to use, compared with 2 children who thought the sequential was easy to use (Figure 54). 4 children thought the simultaneous interface was medium to use, compared with 9 for the sequential. No child thought the simultaneous interface was hard to use, compared with 1 for the sequential. I assigned the choices integer values and compared the responses using a Wilcoxon signed rank test. The results of this test were borderline significant (p=0.06), indicating that the simultaneous interface may have been easier to use.

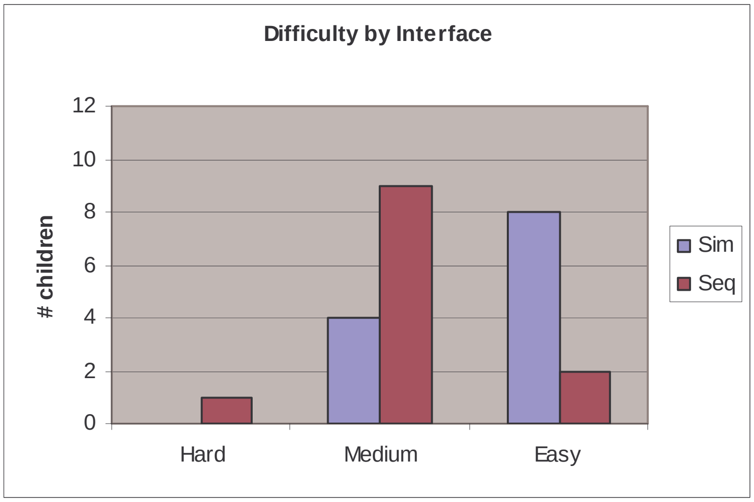

Figure 54. Difficulty by interface for Boolean tasks for 3rd grade (n=12)

For interface like/dislike, 11 children said they liked using the simultaneous interface a lot, compared with 9 who liked using the sequential a lot (Figure 55). 1 child said they liked using the simultaneous interface a little, compared with 3 who liked using the sequential a little. No child said they did not like using either the simultaneous or sequential interface much. I assigned the choices integer values and compared the responses using a Wilcoxon signed rank test. The results of this test were not significant, indicating that there was no difference in likeability between the two interfaces.

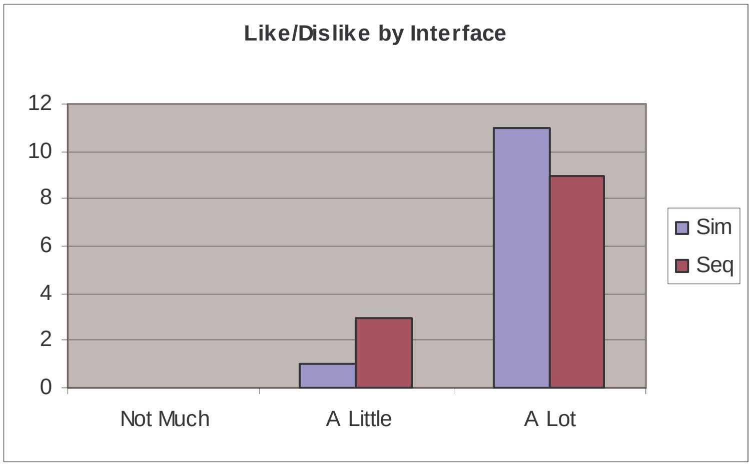

Figure 55. Like/dislike by interface for Boolean tasks for 3rd grade (n=12)

In this grade, 8 children preferred the simultaneous interface, none preferred the sequential interface, and 4 like both equally (Figure 46). I ordered these choices and performed a one-sample median test to see if the median value differed from the null hypothesis that children would prefer both equally. The results of this test were significant ($p<0.01$), indicating that more children preferred the simultaneous interface.

In the simultaneous interface, children needed 31 total hints to complete the timed tasks, compared to 58 total hints in the sequential. To analyze whether there was a difference in the number of hints by interface, I conducted a Wilcoxon signed rank test. The results of this test were not significant, indicating no difference between the two interfaces.

For the simultaneous interface, 4 children understood that they were creating a conjunctive Boolean query and 8 did not. To see if this value differed from the null hypothesis that half the children would understand, I conducted a Binomial test. The results of this test were not significant. For the sequential interface, 3 children understood and 9 did not. The results of a Binomial test were significant (one-sided $p=0.04$), indicating that more children did not understand than did in this interface.

For Boolean search creation during the free browsing session, 3rd graders created 62 Boolean searches in the simultaneous interface and 26 in the sequential. To analyze whether there was a difference in the number of searches created by interface, I conducted a Wilcoxon signed rank test, using a null hypothesis that there would be no difference between the 2 interfaces. The results of this test were borderline significant ($p=0.06$), indicating that Boolean searches may have been more likely to be created in the simultaneous interface than in the sequential.

For navigation tool use during the free browsing session in the simultaneous interface, 4 children used the More Choices button and 8 did not. To see if this value differed from the null hypothesis that half the children would use this button, I conducted a Binomial test. The results of this test were not significant. For the sequential navigation tool, 4 children used the Up Arrow and 8 did not. The results of a Binomial test were not significant.

5.12.4 Grade 5

Search times for all 12 children were submitted to a 2 (interface) x 2 (task type) ANOVA. Results of this analysis indicated a significant difference by interface, $F(1,11)=34.17$, $p<0.01$, and a significant interaction effect between interface and task type, F(1,11)=5.98, p=0.03. For the interface effect, the simultaneous interface was almost twice as fast as the sequential interface. Tukey post-hoc tests on the interaction effect indicate that four out of the six possible pairwise comparisons were significant (Figure 56). One-page tasks in the simultaneous interface were nearly three times faster than one-page tasks in the sequential, nearly twice as fast as two-page tasks in the sequential, and 41% faster than two-page tasks in the simultaneous interface. Two-page tasks in the simultaneous interface were nearly twice as fast as one-page tasks in the sequential interface.

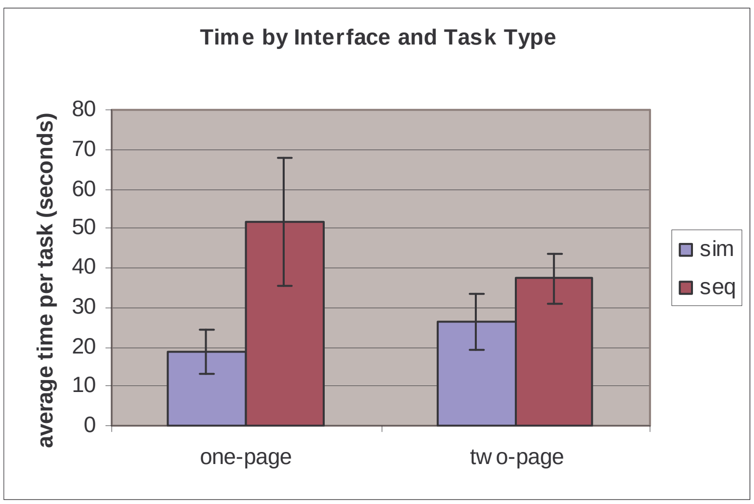

Figure 56. Average time per task by interface and task type for Boolean tasks for 5th grade (n=12)

For interface difficulty, 11 children thought the simultaneous interface was easy to use, compared with 6 children who thought the sequential was easy to use (Figure 57). 1 child thought the simultaneous interface was medium to use, compared with 5 for the sequential. No child thought the simultaneous interface was hard to use, compared with 1 for the sequential. I assigned the choices integer values and compared the responses using a Wilcoxon signed rank test. The results of this test were not significant, indicating that there was no difference in perceived difficulty between the two interfaces.

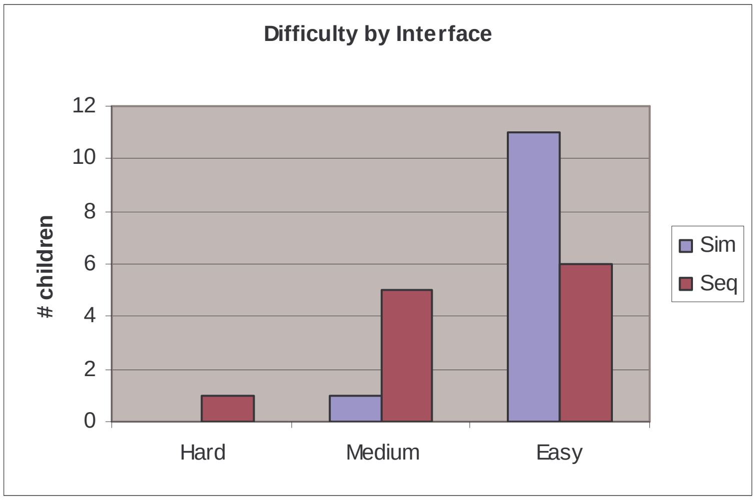

Figure 57. Difficulty by interface for Boolean tasks for 5th grade (n=12)

For interface like/dislike, 12 children said they liked using the simultaneous interface a lot, compared with 7 who liked using the sequential a lot (Figure 58). No child said they liked using the simultaneous interface a little, compared with 4 who liked using the sequential a little. No child said they did not like using the simultaneous interface much, compared with 1 child who did not like using the sequential much. I assigned the choices integer values and compared the responses using a Wilcoxon signed rank test. The results of this test were borderline significant (p=0.06), indicating that children in this grade may have liked the simultaneous interface better.

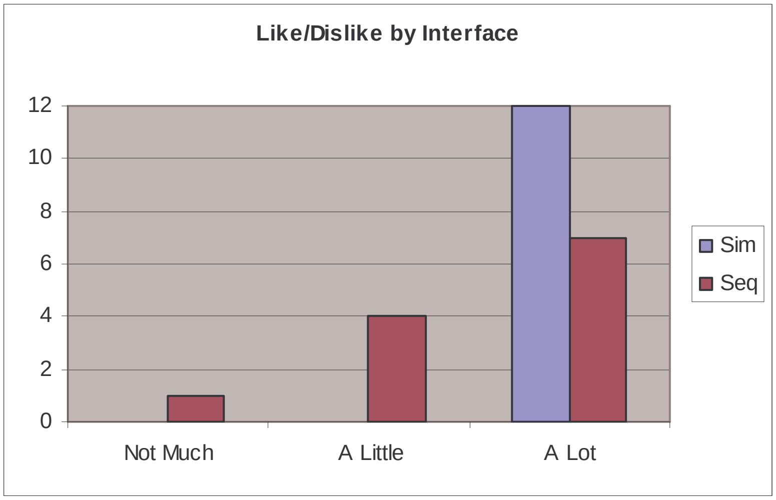

Figure 58. Like/dislike by interface for Boolean tasks for 5th grade (n=12)

In this grade, 8 children preferred the simultaneous interface, 3 preferred the sequential, and 1 liked both equally (Figure 46). I ordered these choices and performed a one-sample median test to see if the median value differed from the null hypothesis that children would prefer both equally. The results of this test were not significant.

In the simultaneous interface, children needed 9 total hints to complete the timed tasks, compared to 21 total hints in the sequential. To analyze whether there was a difference in the number of hints by interface, I conducted a Wilcoxon signed rank test. The results of this test were not significant, indicating no difference between the two interfaces.

For the simultaneous interface, 12 children understood that they were creating a conjunctive Boolean query and 0 did not. To see if this value differed from the null hypothesis that half the children would understand, I conducted a Binomial test. The results of this test were significant (one-sided $p<0.01$), indicating that significantly more children understood than did not in this interface. For the sequential interface, 11 children understood and 1 did not. The results of a Binomial test were significant (one-sided $p<0.01$), indicating that significantly more children understood than did not in this interface.

For Boolean search creation during the free browsing session, 5th graders created 102 Boolean searches in the simultaneous interface and 44 in the sequential. To analyze whether there was a difference in the number of searches created by interface, I conducted a Wilcoxon signed rank test, using a null hypothesis that there would be no difference. The results of this test were significant ($p=0.02$), indicating that Boolean searches were more likely to be created in the simultaneous interface.

For navigation tool use during the free browsing session, 7 children used the More Choices button and 5 did not. To see if this value differed from the null hypothesis that half the children would use this button, I conducted a Binomial test. The results of this test were not significant. For the sequential navigation tool, 10 children used the Up Arrow and 2 did not. The results of a Binomial test were significant (one-sided $p=0.01$), indicating that more children used this button than not.

5.13 Results: Usability Issues

In addition to formally recording and analyzing the above information, I observed a number of usability issues with the interfaces during the course of the experiments.

Exact counts are not available for these issues because I didn't always have time to record them, and for certain issues, stopped recording them because they occurred so often. However, I report the issues here because many of them will feed back into the design of the live ICDL search interface. The first three are relatively important problems, while the others are minor and easily corrected.

5.13.1 Understanding Top-Level Categories

I observed that a number of children did not understand what the Format, Genre, and Rating categories meant in the sequential interface and asked for clarification. For the Format and Genre categories, I encouraged the children to click on the category to see if they could discover the meaning on their own. Some children required additional explanation even after that. I observed that a number of children seemed to have a vague idea about what Format and Genre meant, but not what the difference between them was. These children would look for categories like Picture Books under Genre first, then go to Format, or look for categories like Fairy Tales under Format, then go to Genre. For the Rating category, clicking on the button would not have revealed more information, so I explained that other children gave books ratings about how good the books were. As far as design feedback, this confusion validated my belief that a simultaneous design would avoid problems with abstract top-level categories.

5.13.2 Confusion of Similar Categories

When asked to find categories like Fairy Tales, Make Believe Books, or Imaginary Animal characters, I observed that a number of children selected one of these categories (e.g. Fairy Tales) when the question specifically requested another one (e.g. Make Believe Books). To a lesser extent, this was also sometimes a problem for the Award Winning Books and 5-Star Rated Books categories. These categories were similar enough that until children became familiar with the different choices available, they picked the first one they saw that sounded similar to the question being asked. This seemed to be more of a problem in the simultaneous interface, where the 3 "pretend" categories were all on the first page. For design feedback, I suggest that these categories be combined in some way to avoid confusion.

5.13.3 Activation of Category Buttons

As we observed in our initial usability testing of the simultaneous interface, children in this study also enjoyed seeing how many category buttons they could add to the search. On the other hand, many also became frustrated when they would add another category and the results of the search would not change. This is because we keep a category button active for as long as there are books that match it in the results. We decided to do this because this provides some additional information to the user about what categories their results match. However, this also means that selecting categories may not actually change the search results. For design feedback, I suggest a compromise, where if there is only one book left in the search results, all the remaining unselected category buttons become grayed out.

5.13.4 Icon Design

I observed that two icons gave several children problems. The Purple icon looked more blue than purple to many children, and they missed it when they were asked to find the Purple books. This has already been fixed in the ICDL interface. The Rating icons were confusing to some children because the 3 and 4 star icons had 5 stars on them, some of which were grayed-out to indicate that the highest number of stars a book could get was 5. Some children seemed to miss that these stars were grayed out and thought that all three rating buttons had the same number of stars on them. I suggest that the buttons continue to have the grayed-out stars, but the contrast with the "real" stars be increased. Alternatively, these are all relatively high ratings, so a single "Kid Rated" button including all three rating levels may be sufficient.

5.13.5 Color Buttons

I observed that a number of children didn't initially understand what the color buttons were for. I recommended that the button labels be changed to include the word "Books" or "Covers" since "Red Books" or "Red Covers" is clearer than just "Red". This change has already been made in the live ICDL interface.

5.13.6 Difficult Words

I observed that a number of the younger children were not proficient readers yet and needed help with a few words in the interface. Most notably, the word "characters" was challenging. I suggest using a simpler synonym or simply sticking with "Kids" and "Animals" rather than "Kid Characters" or "Animal Characters".

5.14 Results: Most Popular Categories

During both experiments, the first task in each interface was an open-ended browsing task where children had the opportunity to explore the interfaces for a few minutes without any instruction. I combined all of these browsing tasks together (both experiments, both interfaces), and counted the most popular leaf-level categories chosen during these tasks. The goal of this analysis was to add to our existing knowledge from log file analysis of the most popular categories, and to see whether there were differences by gender and age. I didn't count duplicates within an interface, so if a child used "Red" twice in an interface, I just counted that as one use of Red. I did count duplicates between interfaces – so if a child used "Red" in the simultaneous interface and again in the sequential interface, I counted both uses. I analyzed the counts overall and by gender, and overall for each grade. I present the top 10 categories for each analysis.

5.14.1 Overall

Table 8 shows the top ten categories selected by all 72 children overall. Books falling under the "pretend" categories (make believe books and fairy tales) were quite popular, as were the target age groups of the children in the study (six to nine and ten to thirteen). Children were also most interested in books about particular types of characters, books that won awards, true books, long books, and chapter books.

Category	Count
Make Believe Books	42
Award Winning Books	40
Fairy Tales and Folk Tales	37
Kid Characters	35
Ten to Thirteen	34
Real Animal Characters	30
Chapter Book	29
Six to Nine	29
Long Books	29
True Books	26

Table 8. Top categories selected overall

5.14.2 By Gender

Table 9 and Table 10 show the top ten categories selected by the 36 boys and 36 girls in both studies. Girls chose award winning books the most, while boys chose make believe books most. The age category was more important to girls than boys. Girls selected the true, recently added, and happy categories often, which did not appear in the top ten for boys. Boys selected the long books, imaginary creature characters, and blue categories often, which did not appear in the top ten for girls.

Category	Count
Make Believe Books	24
Real Animal Characters	21
Kid Characters	20
Long Books	18
Fairy Tales and Folk Tales	17
Chapter Book	15
Award Winning	15
Ten to Thirteen	14
Blue	14
Imaginary Creature Characters	13
Rainbow	13

Table 9. Top categories selected by boys

Category	Count
Award Winning	25
Fairy Tales and Folk Tales	20
Ten to Thirteen	20
Six to Nine	19
Make Believe Books	18
True Books	15
Kid Characters	15
Chapter Book	14
Recently Added	12
Happy	12
Rainbow	12

Table 10. Top categories selected by girls

5.14.3 Grade 1

Table 11 shows the top ten categories selected by the 24 1st graders in both studies. Children in this age group differed from the overall counts in several ways. Not surprisingly, children in this age group selected the age category that matched their own (six to nine) but not the older age group. These children, being younger, also selected medium books rather than long books. They were also more interested in perceptual features like color, choosing colors more often than older children. Absent from the top ten list are chapter books and long books, both categories more likely to appeal to older children.

Category	Count
Fairy Tales and Folk Tales	14
Make Believe Books	12
Rainbow	11
Six to Nine	10
Kid Characters	9
True Books	8
Real Animal Characters	8
Green	7
Award Winning	7
Medium Books	6
Blue	6

Table 11. Top categories selected by 1st graders

5.14.4 Grade 3

Table 12 shows the top ten categories selected by the 24 3rd graders. Children in this age group differed from the overall counts in several ways. Like the 1st graders, these children selected their own age group frequently (six to nine). Interestingly, they selected both short and long books, and chapter and picture books, as well as books from a series and happy books.

Category	Count
Six to Nine	14
Make Believe Books	14
Fairy Tales and Folk Tales	13
Award Winning	13
Chapter Book	10
Short Books	8
Happy	8
Series	8
Real Animal Characters	8
Kid Characters	7
Imaginary Creature Characters	7
Picture Book	7
Long Books	7

Table 12. Top categories selected by 3rd graders

5.14.5 Grade 5

Table 13 shows the top ten categories selected by the 24 5th graders. Children in this age group differed from the overall counts in several ways. Again, these children selected the age group category that matched their own (ten to thirteen). The fairy tales category drops out of the top ten, indicating that this category appeals more to younger children. Instead, children in this age group were interested in the recently added category and the color blue.

Category	Count
Ten to Thirteen	25
Award Winning	20
Long Books	20
Kid Characters	19
Chapter Book	16
Make Believe Books	16
Real Animal Characters	14
Recently Added	13
True Books	12
Blue	11

Table 13. Top categories selected by 5th graders

5.15 Discussion of Results

5.15.1 Simple Searches

The simple search study results indicate that *younger children require more time and more hints than older children to find a single category, regardless of the interface presentation*. Fifth graders and third graders were significantly faster than first graders (Figure 23), and first graders needed significantly more hints than third or fifth graders. The results also indicate that *neither the simultaneous nor sequential interface is faster or requires more hints for single category searches*, as no statistically significant differences were found between the two interfaces for either time (Figure 24) or hints overall or within any of the grades.

In addition, *no statistically significant differences were found to indicate that either interface was easier to use, liked more, or preferred* (Figure 25, Figure 26, Figure 27), either overall or within any of the three grades. However, my hypothesis (H1) was supported: one-page tasks are faster in the simultaneous interface than in the sequential interface overall and within grades 1 and 5 (Figure 24, Figure 31, Figure 37). Not surprisingly, one-page tasks were faster than two-page tasks in the simultaneous interface overall (Figure 24), and within grade 5 (Figure 37), though there was no statistically significant difference in the number of hints required to complete one-page vs. two-page tasks in this interface. Interestingly, while one-page and two-page tasks in the sequential interface were structurally equivalent, the two-page tasks were significantly faster overall (Figure 24). This is likely because the one-page tasks involved selecting from the Format category, which many children did not understand and required help with (Figure 29 and Figure 30).

5.15.2 Boolean Searches

The Boolean search study results indicate that *younger children require more time and more hints than older children to find two categories, regardless of the interface presentation*. Fifth graders and third graders were significantly faster than first graders and third graders were significantly faster than first graders (Figure 41). First graders needed significantly more hints than third or fifth graders, and third graders needed significantly more hints than fifth graders.

The results also indicate that *the simultaneous interface is significantly faster and requires fewer hints than the sequential interface overall* (Figure 40). In addition, *the simultaneous interface is significantly easier to use, liked more, and preferred overall* (Figure 43, Figure 44, Figure 45). These differences were also statistically significant in favor of the simultaneous interface or neutral in individual grades as well. Of all the interactions between interface and number of pages required to select a category, only the difference between one-page and two-page tasks in the sequential interface was not significant overall (Figure 42). These results supported my hypotheses (H2) that the simultaneous interface would be significantly faster, easier, and preferred for both one-page and two-page tasks compared to the sequential interface.

In addition to being faster and requiring fewer hints, *fifth grade children understood that they were creating a conjunctive Boolean query significantly more than first and third grade children* (Figure 47), supporting my hypothesis (H3). However, *neither interface was found to be significantly more likely to support this understanding* (Figure 47). My hypothesis (H5) that the simultaneous interface would better support this understanding was not supported. Unfortunately, I did not have enough participants to compare comprehension in the two interfaces directly to each other, which is something I would like to see done in future work.

For all grades, the simultaneous interface was significantly faster than the sequential interface (Figure 50, Figure 53, Figure 56), but only first graders needed significantly more hints in the sequential interface. First and fifth graders did not like either interface significantly more than the other, consider either one significantly easier to use, or prefer one significantly to the other. Third graders did not like either interface significantly more or consider either one significantly easier to use, but did significantly prefer the simultaneous interface. However, with only 12 subjects per grade, I may not have had enough statistical power to detect differences within each grade.

5.15.3 Browsing Boolean Search Creation

In both the simple and Boolean studies, *the browsing tasks seemed to favor the simultaneous interface with regard to children's ability to create conjunctive Boolean searches*, supporting my hypothesis (H4). During the browsing task, children created significantly more Boolean searches using the simultaneous interface than the sequential, both overall and within some of the grades. I did not record whether or not children understood that they were creating a Boolean search while browsing, but these results suggest that creating one is easier in the simultaneous interface.

5.15.4 Browsing Navigation Tools

In both studies, *children were more significantly more likely to find the Up Arrow button than not to navigate in the sequential presentation, and significantly less likely to find the More Choices buttons than to find them to navigate in the simultaneous presentation*. I observed that children did not have a problem understanding or using the More Choices buttons once they found them or I pointed them out during the instructions. I believe that some combination of more icons on the screen, the smaller size of the More Choices buttons, and the placement of the More Choices buttons in the corners made them harder to find than the Up Arrow. This was a serious usability problem with the simultaneous interface for children who did not receive prior instruction on how it worked. Future research is necessary to determine what about these buttons made them hard to find and what can be done to fix the problem.

5.15.5 Usability Issues

The usability issues I observed during both studies yielded three important observations. *In the sequential interface, the hierarchical structure of the category facets was difficult for many of the children to understand when the top-level categories were abstract (e.g. Format and Genre).* These observations were supported when I looked at the time and number of hints for individual tasks in both studies. Figure 29 and Figure 30 illustrate that children may have had more difficulty finding a single category in the Format category using the sequential interface. Figure 48 and Figure 49 illustrate that children may have had more difficulty finding categories in the Format and Genre categories using the sequential interface. These observations confirmed previous research that young children have difficulty categorizing with high-level topics (Rosch et al., 1976; Gelman & O'Reilly, 1998), and provide evidence that a flattened structure of basic-level topics may be preferable in interfaces for young children.

In both interfaces, I found that *we had probably over-classified some of our facets, as children had difficulty making distinctions between categories like Make Believe, Fairy Tales, and Imaginary Animal Characters*. Creating meaningful and differentiable facets is a problem for any category interface, whether for adults or children. However, more care is required for young children, who not only classify things differently than adults, but also have less knowledge of the world, preventing them from differentiating at as fine a level as adults.

Finally, I discovered that *our previous design decision to keep a category button active for as long as it matched something in the results was sometimes confusing and frustrating*. This is a tricky issue, since many children also seem to enjoy seeing how many categories they can include in their search. A compromise may involve deactivating all categories once there is only one book left in the results section.

5.15.6 Most Popular Categories

The categories presented in both interfaces were included because they were likely to appeal to elementary age children, based on ICDL project research, physical library book selection research, and log file analysis. Nonetheless, I was interested to see which of these categories were the most popular in my studies overall, by gender, and by grade. To some extent, this popularity is weighted toward categories that appear on the first page of the simultaneous interface, since they are the most readily accessible. However, these categories were placed on the first page because our previous research indicated that they were probably more popular than the ones on the second page.

My results from this analysis contribute by confirming previous research about the categories preferred by children of different ages (Table 11, Table 12, Table 13). First, *children in all three grades were particularly interested in selecting books that matched their own age group*. Kulper et al. (1997) found a similar trend in their analysis of a category browsing interface for books, and actually removed the category because they thought the children took the category too seriously. We didn't consider this a problem in the ICDL. Books are catalogued in the ICDL according to reading level by age, so selecting an age is likely to be helpful to children. Second, *the first graders were more interested in perceptual features like color than the older children*, confirming previous research to this effect (Pejtersen, 1986; Moore and St. George, 1991; Kragler and Nolley, 1996; Fleener et al., 1997; Robinson et al., 1997). Third, *the younger children were more interested in pretend categories like fairy tales than the older children*, confirming previous research to this effect (Kuhlthau, 1988; Fleener et al, 1997; Robinson et al., 1997; Cooper, 2002b; Cooper, 2004).

5.16 Limitations of Results

5.16.1 Interface Design

This research was limited by the fact that I used a particular interface design idea for the category browser. I chose to place the books in the middle of the page and the categories around the perimeter of the page. I felt this design would place the focus on the books, which are the most important part of the interface, and that it mirrored the design of certain children's toys, which might make children feel more comfortable. However, another design choice might have placed the books on one side of the page and the categories on the other, as is often done in interfaces for adults. This other design choice has the advantage that the categories are in one part of the page, reducing the area that has to be visually scanned to find a category. Using this design choice might have caused the results of the studies to be different, though I believe that a simultaneous presentation of flat facets would still perform better than a sequential presentation of hierarchical facets because the extra navigation steps and abstract, high-level categories would still be problematic.

5.16.2 Structure and Presentation

This research was limited by the fact that I only compared two of the three possible combinations of structure and presentation. Although I found some significant advantages for a flat structure presented simultaneously over a hierarchical structure presented sequentially, I don't know whether the structure or the presentation of the former interface had more to do with its advantages. I did not evaluate the third possible combination, a hierarchical structure presented simultaneously, so I also don't know how this particular combination would perform. However, based on early design work (Figure 13), I believe it would likely be too complicated for children.

5.16.3 Boolean Logic

This research was limited by the fact that I only evaluated conjunctive Boolean searches of two categories between facets. I did not look at other Boolean logic combinations (e.g. disjunction, negation), more than two categories, or within facet searches. I focused only on conjunction because this helps users narrow down the result set faster. I believe that I would get similar results with other Boolean logic combinations and with more than two categories because the navigation and top-level category comprehension issues remain. However, the advantages for the simultaneous interface would likely be reduced for within-facet searches since navigational backtracking is not necessary in the sequential interface for this type of search.

5.16.4 Browsing

This research was limited by the fact that I only had the children do a single browsing task. This task was not strictly timed, and I did not question the children about what they were doing or whether or not they understood what they were doing during this task. As a result, while I was able to find significant differences between the two interfaces as far as the number of Boolean searches created, I can't say whether these differences also included better comprehension. A more structured browsing study that included stricter timing and questioning of the children by the experimenter would lead to a deeper understanding of children's browsing skills.

5.16.5 Cultural Differences

This research was limited by the fact that my study participants came from a similar geographic location in a western culture. Although the participants were somewhat racially, ethnically, and socio-economically diverse, the results of the studies might have been different in a population with a different cultural background. Children from different cultures have been shown to categorize objects differently (Cole et al., 1971; Lucy and Gaskins, 2001), so it would not be surprising if they had different preferences for category browsing interfaces as well.

5.16.6 Statistical Power

This research was limited by the fact that I only had 12 children in each grade in each study. Some of the differences I looked at within each grade suggested an advantage for the simultaneous interface, but the results of non-parametric statistical tests were not significant. I believe that some of these non-significant results may have type two (false negative) errors due to a lack of statistical power.

Chapter 6: Contributions

6.1 Children's Use and Preference of Search and Browse Interfaces

The major contribution of this research is an analysis of elementary-age children's use and preference of two different combinations of structure and presentation in category searching and browsing interfaces. For adults, previous research indicated that a flatter structure was generally faster and preferred to a deep hierarchical structure (Miller, 1981; Kiger, 1984). Previous research also indicated that for adults, a sequential presentation of categories was faster for simple searches that did not require backtracking, while a simultaneous presentation was faster for more complex searches that did require backtracking (Hochheiser et al., 2000). For children, previous research indicated that children were capable of creating both simple, one-item searches that did not require backtracking and complex, multi-item Boolean searches that did in a sequential hierarchy (Revelle et al., 2002), but did not generally do the latter when browsing on their own (Reuter & Druin, 2004). This research adds to the understanding of children's searching and browsing skills by exploring how children are able and prefer to use both sequential hierarchical and simultaneous flat designs for both simple and Boolean searching and open-ended browsing.

6.1.1 Contributions by age group

The results of my studies indicate that children in first, third, and fifth grade become significantly faster and require significantly fewer hints to complete both simple, one-item and Boolean, two-item searches as they get older, regardless of whether they use an interface with a simultaneous presentation of flat facets or a sequential presentation of two-level hierarchical facets (Sections 5.15.1 and 5.15.2). This conclusion is not surprising, but adds to previous research indicating that interfaces designed for elementary-age children need to consider that children of different ages have a wide range of skills (Druin, 1999b). However, since none of the three age groups differed significantly from each other in their preference or perception of the difficulty and likeability of these two interfaces, designers of these types of interfaces may be able to support multiple age groups with similar designs, with the understanding that adult or system scaffolding will be necessary for younger children.

6.1.2 Contributions for simple searches

Children in these three grades are not significantly faster and do not require significantly more hints to complete simple, one-item searches in either of these two interfaces, either overall or within any of the grades (Section 5.15.1). Children also do not significantly like one interface more than the other, consider either one significantly more difficult to use, or express a significant preference for either interface when doing simple, one-item searches, either overall or within any of the grades. These results were interesting because they differed somewhat from the findings of Hochheiser et al. (2000), who found that adults were faster with a sequential presentation when doing similar tasks that did not require backtracking in the sequential presentation. The tasks and interfaces differed enough between these two studies that it is difficult to know whether age or interface design played a greater role in the results.

6.1.3 Contributions for Boolean searches

Children in these three grades are significantly faster and require significantly fewer hints with the simultaneous interface when doing two-item, Boolean tasks, both overall and within each grade (Section 5.15.2). Overall, children also like it significantly more, consider it significantly easier, and significantly prefer it. The Boolean search study did not detect many significant differences for hints, likeability, ease of use, or preference within each grade, but this may be due to a lack of statistical power. These findings are an important contribution because they suggest a clear advantage for one interface over the other when children are asked to select more than one category from different facets.

It is not clear whether the difference in interface structure or presentation had more influence on this advantage. The flat structure may have been better than the hierarchical structure because children struggled with some of the top-level categories in the hierarchy (e.g. Format, Genre) (Section 5.15.5). I found some evidence to support this when I looked at the times and number of hints for individual tasks (Figure 48 and Figure 49). The simultaneous presentation may have been better than the sequential presentation because the sequential presentation required more clicks and also required navigational backtracking, known to be challenging for children (Marchionini and Teague, 1987). More research is needed to explore which of these differences was more important.

6.1.4 Contributions for comprehension of Boolean search

Older children are more likely to understand that they are creating a conjunctive Boolean search when they select more than one category than younger children (Section 5.15.2). This point was underscored by my observation of a number of children who became frustrated when they selected multiple search categories and were surprised that the results of their search did not always change. This finding was not surprising, but provides an important contribution for other interface designers who want to support Boolean search for elementary-age children and need to know which age groups will understand it. Regardless of interface, less than half of first and third graders understood that they were creating a conjunctive Boolean search, while nearly all fifth graders did understand (Figure 47). However, neither the simultaneous nor the sequential interface seemed to support this understanding any better.

6.1.5 Contributions for browsing

When children are browsing without any prior instruction, the simultaneous interface seems to facilitate creation of Boolean searches more easily than the sequential interface. In both of my studies, children created more Boolean searches in the simultaneous interface than in the sequential interface overall and within some grades (Section 5.15.3). This finding is an important contribution because it suggests that children have an easier time narrowing down the result set with the simultaneous interface. Even if the children did not understand that they were creating a conjunctive Boolean search by selecting multiple buttons, they were still making more progress toward finding a manageable result set.

The large, conspicuously placed navigation up arrow in the sequential interface with two levels was found by significantly more children than not when browsing (Section 5.15.4). This finding is an important contribution because previous research suggests that children are inefficient when navigating hierarchies (Marchionini and Teague, 1987). My research suggests that with a large button and only two levels in the hierarchy, they are comfortable. However, the two smaller, less conspicuous paging arrows in the simultaneous interface were not found by most children (5.15.4). This finding is an important contribution because placing paging arrows at the top and bottom of the screen is a common design paradigm in interfaces that require paging. My research indicates that it may not be sufficient for children.

6.1.6 Contributions for book selection

When browsing on their own for books, children in all three age groups often selected the category for their own age group, and younger children selected books based on perceptual features like color and using pretend categories like "Make Believe" and "Fairy Tales" more often than older children (Table 11, Table 12, Table 13). These findings are not surprising given similar results reported by other researchers for physical libraries (e.g. Pejtersen, 1986; Moore and St. George, 1991; Kragler and Nolley, 1996), but add to the smaller body of literature confirming that these trends are also true for digital libraries (e.g. Busey and Dorr, 1993; Reuter and Druin, 2004).

6.1.7 Contributions for other searching and browsing tools

Finally, these results provide a broader contribution because they are likely generalizable to other searching and browsing tools for children, such as search engines and e-commerce applications. Sites in these domains might support both simple and Boolean searches using faceted or hierarchical category browsers, and the results of these studies suggest design guidelines (see below) as well as reinforce some specific points unique to digital libraries. These results are not scaleable to large numbers of categories, which would require placing many categories on the screen, or large amounts of paging or navigation. However, young children's shorter attention spans, slower visual information processing speeds, and smaller memory capacities suggest that large numbers of category choices would not be appropriate for children anyway (Baumgarten, 2003; Kail, 1991; Chi, 1976).

6.2 Design Guidelines

6.2.1 Using category browsers

The results of these studies suggest several guidelines for designers of searching and browsing interfaces for elementary-age children. In general, my studies confirm that elementary-age children are willing and able to use both a simultaneous presentation of flat facets and a sequential presentation of hierarchical facets to search for information. Children in both studies were able to complete all the searching tasks within a reasonable amount of time, though sometimes required hints (Figure 29, Figure 30, Figure 48, Figure 49). More research is necessary to understand if the third possible combination of structure and presentation in a category browser – simultaneous presentation of a hierarchical structure – would be useful. However, given the number and size of the icons that would be on the screen at once in this design, my hypothesis is that it would be visually overwhelming and require too much scrolling, paging, and/or navigation.

6.2.2 Designing classification schemes

Regardless of which category browser is used, designers need to be careful when designing their category classification scheme. Children had difficulty with some of the more abstract, high-level categories in the sequential interface in both of my studies (e.g. Format and Genre) (Figure 29, Figure 30, Figure 48, Figure 49). At the other end of the spectrum, they also had difficulty differentiating between some of the more similar leaf level categories in the Boolean study (e.g. fairy tales and make believe books) (Figure 48, Figure 49). My research supports previous findings that children require a classification scheme that is not too abstract (Rosch et al., 1976; Gelman and O'Reilly, 1988), but also underscores the importance of not making it too finely differentiated either.

6.2.3 Designing navigation schemes

My research suggests that young children are comfortable navigating in a hierarchy if it is two levels deep, provided they have a large, conspicuously placed navigation arrow. Previous research suggested that children do not navigate efficiently in category hierarchies of multiple levels (Marchionini and Teague, 1987), but significantly more children than not found and used the navigation arrow in a two-level hierarchy when browsing without instruction in my studies (Section 5.15.4). On the other hand, smaller, less obviously placed buttons for paging in a simultaneous presentation seem to get lost among the other buttons. Significantly more children did not find the paging arrows than did find them while browsing without instruction in my studies (Section 5.15.4). As a result, I recommend larger, more prominently placed paging buttons, or avoiding paging altogether by decreasing the number of categories in the classification structure.

6.2.4 Supporting simple searches

For searching and browsing interfaces that will only support selecting one category at a time, both a two-level sequential presentation of a hierarchical structure and a two-page simultaneous presentation of a flat structure are appropriate design choices. There were no significant differences in speed, number of hints, perceived difficulty, likeability, or preference between these two interfaces (Section 5.15.1). However, if all of the categories can fit on a single page in the simultaneous presentation, it is likely to be faster for children to navigate. When children completed one-item searches that did not require paging, they were significantly faster with the simultaneous interface (Figure 24).

6.2.5 Supporting Boolean searches

If Boolean search is to be supported, a simultaneous presentation of a flat structure is preferable to a sequential presentation of a two layer hierarchical structure, at least for conjunctive searches of two items. Overall, the simultaneous design was significantly faster, required significantly fewer hints, was considered significantly easier and likeable, and was significantly preferred (Section 5.15.2). More research is needed to understand if this is the case for disjunctive and other logical combination searches.

However, supporting Boolean search does come with a few caveats. Not all elementary-age children, particularly younger ones, will understand this functionality, so the interface needs to work regardless of whether children realize what is going on. Less than half of the first and third graders in the Boolean study understood that they were creating conjunctive Booleans searches (Figure 47). Features like dynamic queries, tightly coupled results, and preventing no-hit searches (see Section 4.2.2) can all make the interaction easy even if a child doesn't understand the underlying logic. In addition, I observed that children were sometimes confused and frustrated that selecting a category didn't always change their search results, so deactivating categories that won't change the search results is also likely to be helpful.

6.2.6 Supporting browsing

In the browsing tasks, children created significantly more Boolean searches in the simultaneous interface than the sequential interface in both of my studies (Section 5.15.3). While it is not clear whether or not they understood that they were creating Boolean conjunctions, they were more likely to narrow down their search results to a manageable size. As a result, for interfaces that are designed to support browsing activities, with the goal of helping children narrow down their search results, the simultaneous interface is likely to be better than the sequential interface.

6.3 Working Examples

6.3.1 ICDL Servlet architecture

The final contributions of this research are the working examples of technology and interfaces that I developed for the ICDL and adapted for use in this study (Chapter 2). The current ICDL Servlet architecture represents two and half years of my development work, leading to the creation of nearly 100 Java class files of approximately 16,000 lines of code in an application that supports roughly 25,000 visitors a month from 155 countries. This code connects to many of the 41 tables in a mySQL database maintained by another ICDL project team member containing information about approximately 800 books of 50,000 total pages, as well as search categories, and user profile information.

This development work entailed learning how Servlets worked, researching and installing appropriate software to support them, integrating database drivers and connection pools, ensuring Unicode compliance of all components, and designing and coding interfaces to support multiple languages and customs. In addition to supporting the ICDL project, I expect that designers of other interfaces will be able to use the ideas from my architecture and interface designs to create and improve their own tools. The ICDL is built with entirely open-source technology, and the interface development methodology and design is detailed here and in other ICDL-related publications (Hutchinson, 2004; Hutchinson et al., 2005a; Hutchinson et al., 2005b).

6.3.2 Simultaneous interface design

In addition to developing the Servlet software, I developed the idea for the design of the Simple Search category browser (Chapter 3), and based the simultaneous interface on the same ideas. This interface makes use of a number of known interface design and interaction features known to work for children. It supports children's less developed motor skills by using large, easily clickable category and result target buttons (Hourcade et al., 2003a). It supports their searching and browsing skills by avoiding spelling and typing (e.g. Edmonds et al., 1990) and minimizing navigation

(Marchionini and Teague, 1987). It adapts techniques developed for adults and shown to work for children as well by preventing zero-hit searches (Doan et al., 1996), using direct manipulation and dynamic queries (Shneiderman, 1983; Ahlberg et al., 1992), and automatically constructing Boolean queries (Revelle et al., 2002). Finally, it supports their preferences by providing book selection criteria appropriate for children (e.g. Pejtersen, 1986; Fleener et al., 1997; Reuter and Druin, 2004).

All of these features were also available in the original ICDL category browser, but I made a number of changes to improve the interface based on problems identified in previous studies. In the Enhanced (Java) ICDL category browser, the books that matched a search were tightly coupled with the browser on the same page, but in a small, inconspicuous box at the top of the screen. In the Basic (HTML) version of this browser, the results could only be reached by going to a new page. The Simple Search interface improved on both of these designs by tightly coupling the results on the same screen, and displaying them prominently in the center of the page. In my studies, I did not observe any children having difficulty finding the results.

Reuter and Druin (2004) found that children sometimes confused the interior and leaf-level category buttons as well as the books in both the Enhanced and Basic category browsers because they were all rectangles of about the same size. In the Simple Search interface, I made the category buttons round to make the distinction between books and categories clearer, and only presented leaf-level categories. In my studies, I did not observe any children who were confused between categories and books in the simultaneous interface.

Reuter and Druin (2004) found that children did not make use of the Boolean functionality in the Enhanced ICDL category browser when browsing on their own. My hypothesis was that the sequential presentation required navigation and backtracking to create a Boolean search, making this activity unwieldy. In addition, the hierarchical structure of the categories meant that children may have been hesitant to explore if they didn't understand some of the more abstract, high level categories (e.g. Format, Genre). In the Simple Search interface, children do not have to navigate and backtrack because the categories are presented simultaneously. I also flattened the hierarchy structure to a single level, so children can rely on perception rather than abstract, hierarchical knowledge to find categories. The results of the Boolean study indicate that children were able to conduct Boolean searches when asked to do so more easily and preferred this interface to a sequentially presented hierarchy (Section 5.15.2). In both the simple and Boolean studies, children also created more Boolean searches when browsing freely using the simultaneous interface than the sequential interface (Section 5.15.3).

6.3.3 Sequential interface design

The sequential interface I designed for the studies also included some improvements over the original ICDL category browser. It uses the same tightly-coupled presentation of results and categories as the Simple Search, making the results obvious and reducing the amount of navigation. I used the same round buttons from the Simple Search for the leaf-level categories in the sequential interface to help distinguish the leaves from the top-level categories. The original category browser consisted of hierarchical facets of depths varying from 2 to 4 levels. I flattened the category hierarchy in a number of places so that it was always 2 levels. Combined with the large, prominently placed navigation arrow, more children than not were able to find and use the arrow to navigate in the hierarchy without any instruction (Section 5.15.4). Though my studies do not demonstrate whether the sequential interface was any better or worse than the original ICDL category browser, children did use it to create Boolean searches while browsing without any instruction, something that Reuter and Druin (2004) did not observe happening in the old version.

Chapter 7: Future Work

7.1 Research Questions

7.1.1 Interface Design

While my studies had the broad goal of comparing two different combinations of structure and presentation, they were limited by the fact that the interfaces I used for this comparison could have been designed in many ways. I chose one particular design metaphor – a ring of categories surrounding a collection of books – because I wanted to place the focus on the books and I felt that this design would be comfortable for children who might be used to using physical toys with similar designs. Other design choices might have led to different results. It would be interesting to repeat the studies with a different design metaphor to evaluate how large a role, if any, the design metaphor plays in the results.

7.1.2 Structure and Presentation

One of my goals in designing the ICDL Simple Search interface was to allow children to more easily create multiple item Boolean searches. I believed that the original ICDL category browser suffered from two crucial problems. Structurally, some of the hierarchical facets had top-level categories that were likely too abstract for young children to understand. In presentation, the sequential menu design required a lot of navigation and backtracking to move between different facets. The Simple Search interface was designed to address both of these problems by flattening the structure and changing the presentation to a simultaneous design. The results of my studies indicate that one or both of these changes was likely successful in making Boolean searches faster and easier to create. However, because I changed both the structure and the presentation, I don't know which of these was more important.

I did not evaluate the third possible combination of structure and presentation, a simultaneous hierarchy. As a result, an interesting future study might compare all three combinations with the goal of trying to tease out whether structure or presentation seems to matter more. My hypothesis is that simultaneous presentation of a hierarchy, as is done in Microsoft Windows Explorer, would be more difficult for children to use than simultaneous presentation of the same categories in a flat structure. I'm not sure whether it would be harder or easier to use than a sequential presentation of a hierarchy. The first prototype I developed (Figure 13) involved a simultaneous presentation of a hierarchy, but the animation and maintenance both proved troubling. Other possible solutions might involve a design like Windows Explorer with icons instead of text, or a fisheye design where the branch of focus gets more screen space than other branches.

7.1.3 Boolean Logic

The original ICDL category browser supported disjunctive Boolean search within a particular top-level category and conjunctive Boolean search between top-level categories. In the Simple Search browser, I chose to only support conjunction. I did this because a number of studies indicate that children have an easier time with conjunction than with disjunction, and conjunction decreases the number of results found, which helps narrow down the list of books for children to choose from. I also wanted to provide a visual indication of the search being created, which the team decided to do with an “equation” to indicate that their combination adds up to the count of the results. This visual tool makes the effect of selecting multiple categories concrete, which is important for children learning to reason logically. Future studies might explore whether other logical combinations were useful and understandable, either independently or together in the same interface. Future studies might also explore conjunctive searches within the same category facet and searches involving more than two categories. My Boolean study only included searching for two items between two different facets.

7.1.4 Browsing

One of the motivations for my research were the findings by Revelle et al. (2002) and Reuter and Druin (2004) that children could do Boolean searches in a sequential category browser when asked to do a directed search, but didn’t generally create Boolean searches on their own when browsing. During the free browsing tasks in my studies, I recorded the number of Boolean searches that children created with each interface. However, I did not ask the children whether or not they understood what they were doing at this time because I wanted this task to be open-ended so I could observe other things as well. The browsing tasks were also not strictly timed, so the results are not entirely controlled. Later on in the Boolean study, I had the children do a directed Boolean search for 2 categories and then questioned them to elicit whether or not they understood what was going on. In the browsing task, more children created Boolean searches with the simultaneous interface than the sequential interface. However, in the searching task, neither interface seemed to promote understanding better. In the future, it would be interesting to do a browsing study where children did multiple browsing tasks and researchers questioned them whenever they created Boolean searches to see if they understood.

7.1.5 Cultural Differences

In my studies, I worked with children in three different counties in suburban Maryland. The children I worked with were racially and ethnically diverse, and census data indicates that they were somewhat diverse socio-economically. However, they were all still being educated in schools from the same western culture even if their family background was from some other culture. The findings from my study are thus only generalizable to children with similar backgrounds. It would be interesting to repeat the study with children from other cultures in other parts of the world to see if there are differences in the results. Previous research with children categorizing objects indicates that there might be interesting and significant differences (Cole et al., 1971; Lucy and Gaskins, 2001).

7.1.6 Statistical Power

One of the tradeoffs I made in doing two different studies was reducing the number of children in each study. I had originally wanted to do a single study, but as the number of variables grew, I decided it would be better to do two studies given the limited time I had with each child. Since Hochheiser et al. (2000) found differences between simultaneous and sequential interfaces depending on whether the tasks were simple or complex, I wanted to study this effect in children as well, and this particular variable made for a clean division between the studies. This proved to be a useful decision because I got different results depending on whether children were doing simple, one-item tasks or complex, two-item Boolean tasks. The downside of this decision is that when looking at some of the non-parametric statistics by grade, with only 12 subjects I may not have had enough power to detect differences that may have been present.

As a result, it would be interesting in the future to repeat some parts of the studies with more children. In particular, a number of overall counts (hints, likeability, difficulty, uses of navigation tools) were significantly different between the two interfaces overall in the Boolean study, but not within a number of the grades. I think that repeating this particular study with more children might yield more significant differences within each grade. Particularly with the younger children, the Boolean study ran up against the 30 minute time limit more often than the simple study, and the children may have felt more hurried in this study. This may have affected the results in some way. A future study that left out the browsing and comprehension tasks would avoid this problem.

7.2 ICDL Design Changes

The final piece of future work that comes out of these studies relates to the results that will feed back into the design of the live ICDL category browsing interface. When we released the Simple Search interface in the fall of 2004, we had done usability testing to ensure that children were able to use it and were somewhat confident that the hypotheses that I had about the structural and presentation changes would be an improvement over the old category browser. The results of my study largely indicate that my hypotheses were on the mark. However, we did identify one major issue during the course of the study that the ICDL team is presently working to fix.

A number of factors suggest that we should reduce the number of categories in the Simple Search and possibly place them all on a single page. The combination of children having difficulty making a distinction between a number of similar categories, their difficulty finding the More Choices buttons, their propensity to select categories mostly on the first page when browsing on their own, and the fact that the simultaneous design was faster for one-page tasks for both simple and Boolean tasks all indicate that a single page of fewer categories might be a better design. The ICDL team is presently evaluating these suggestions.

My studies also exposed a number of smaller usability issues that will be included in this redesign. We plan to deactivate all category buttons when there is only a single book left in the results to avoid confusion over adding categories to the search and not having the results change. We plan to update the stars on the rating buttons to be clearer, and perhaps combine all three rating buttons into a single “Kid Rated” button for simplicity. We have already changed the labels on the color categories to include the word “covers” to make it clearer what the colors are referring to. Finally, we plan to change icons with difficult words, such as “Characters” to be more easily readable.

Appendix 1: Letter to Parents

Dear Parents and Guardians,

Attached is a permission slip that will enable your child to participate in a research project for the International Children's Digital Library (ICDL). The ICDL is a non-profit, National Science Foundation funded research project being run by the University of Maryland. The ICDL collects award-winning children's books from around the world, scans them, and makes them available online to be read for free at **www.icdlbooks.org**.

In addition to developing a world-class book collection, the ICDL also works with children to develop new software that they can use to search and read in the library. It is important that this software be developmentally appropriate and easily usable so that children can use the library with confidence. We have recently designed a new searching tool that we are now evaluating in your child's school.

Children who would like to participate and whose parents sign the permission slip will be asked to use the software while a researcher takes notes and records how long it takes to search for books in the library. The project takes 15-30 minutes and will be done during the after-care program. Each child's identity will remain anonymous for the purposes of reporting the results of the study – we will record only a child's age and gender.

There are a variety of benefits to having your child participate. First, they will have access to a collection of children's literature from around the world. Second, they will learn how to search efficiently for books in this collection using our software. Third, they will discover a new environment in which to spend time reading. Staff and teachers at the school will also learn to use the ICDL so that they can incorporate it into lessons. The broader benefit of the study is that the results will help improve the design of the ICDL for users everywhere. We will also publish our results, so our findings will help other designers of children's software.

We would be extremely grateful if you would allow your child to participate in the study. We look forward to working with your child and introducing them to this project!

Thanks,

Hilary Hutchinson
ICDL Faculty Research Assistant and Graduate Student
Human-Computer Interaction Lab
University of Maryland, College Park
hilary@cs.umd.edu
(301) 405-7445

Appendix 2: Parental Permission Form

Graphical User Interface Performance
<u>Parental Permission Form</u>

DESCRIPTION: Your child is invited to participate in a research study on the performance of different designs within graphical user interfaces.

PROCEDURE: Your child will be asked to perform common computer tasks (such as performing simple commands, navigating within a document, or looking for information). Your child's interactions with the computer will be logged for analysis purposes. This information will not be disclosed to others and the data will be discarded after the study is over.

RISKS AND BENEFITS: There are minimal risks associated with this study as the software is typical of current freely available systems. The benefits that may reasonably be expected to result from this study are better designs of computer programs that your child can use for educational or entertainment purposes.

TIME INVOLVEMENT: Your child's participation in this experiment should take no longer than 30 minutes.

PAYMENTS: Your child will receive no financial compensation for participation in this study.

SUBJECT'S RIGHTS: If you have read this form and have decided that your child can participate in this project, please understand that their participation is voluntary and your child has the right to withdraw consent or discontinue participation at any time without penalty. Your child has the right to refuse to answer particular questions, and may ask any question they wish. Their individual privacy will be maintained in all published and written data resulting from the study. If you have questions about your child's rights as a study subject, are dissatisfied at any time with any aspect of this study, or wish to report a research-related injury, please contact - anonymously, if you wish – the Institutional Review Board Office, 2100 Lee Building, University of Maryland, College Park, MD 20742 USA (e-mail: irb@deans.umd.edu, telephone: 301-405-4212).

I state that I am over 18 years of age, and wish my child to participate in a program of research being conducted by Benjamin Bederson at the University of Maryland, College Park (Tel: (301) 405-2764, email: bederson@cs.umd.edu).

CHILD's Name ____________________

PARENT SIGNATURE _____________________________ DATE ____________

PARENT NAME (PRINT) ____________________________

Appendix 3: Child Assent Form

Computer Program Performance
<u>Assent Form</u>

DESCRIPTION*:* You can help us understand how well these programs work.

PROCEDURE*:* You will be asked to use a computer program to do common things (such as doing simple commands, moving within a document, or looking for things). We will keep track of your interactions with the computer.

RISKS AND BENEFITS*:* This study will be very similar to using other common computer software. As such, there are very few risks associated with this study. By joining us, you can help us design better computer programs.

TIME INVOLVEMENT*:* This will take less than half an hour.

PAYMENTS*:* You will not receive anything for helping us.

SUBJECT'S RIGHTS*:* You may change your mind at any time and stop working with us, or you can ask any questions you like. You also can choose to not do any task if you wish. If you have problems of any kind, please let your parent or teacher know.

☐ Check this box if you agree to participate

Name (PRINT) ______________________________

DATE __________________________

If the subject can not read, then this assent form was read to the child and the child understands its contents:
INVESTIGATOR __________________________________
DATE __________________________________

Appendix 4: General Experiment Instructions

1. Today we are going to look for books using two new computer programs.

2. You can help us make these programs better by using them to answer some questions

3. Some of the questions might be easy and some might be hard.

4. Don't worry - you can ask for help at any time.

5. When you are done with a question, tell me, and we'll move on to the next question.

6. You get to drive the mouse to answer the questions and I get to drive the keyboard.

7. We aren't actually going to read the books, we're just going to find them.

8. But, you can use the programs another time to read books if you want.

Appendix 5: Simultaneous Interface Instructions

In this program, the books in the library that you can read are in the middle of the page. You can see little pictures of their covers.

Around the outside of the page are round buttons with categories that describe the books. If you press one, the program will show you all of the books that match the category.

For instance, if I press the "Red" button, the program will show us all the books with red covers and pictures. You can see at the top of the box that there are XX red books in the library. This page shows you the first 8 books. You can press these arrows to see more.

When buttons are grayed out like the “Orange” one, that means you can’t click the orange one because there aren’t any books that are both orange and red.

If you don't want to see red books, you can press the "Red" button again to make it go away. You can also press the trash can button to make it go away so you can start over.

Not all of the buttons fit on this page, so if you want to see other categories, you can press the "More Choices" button. There is one at the top of the page and one at the bottom.

Now you can see that there are some other categories that you can pick. If you want to go back to the first page of categories, press the "More Choices" button again.

Does that all make sense? Do you have any questions?

OK, now I am going to ask you to find some books again.

Appendix 6: Sequential Interface Instructions

In this program, the books in the library that you can read are in the middle of the page. You can see little pictures of their covers.

Around the outside of the page are square buttons with categories that describe the books. If you press one, the program will show you some round buttons about that category.

For instance, if I press the "Color" button, the program will show us lots of different colors. If I press the "Red" button, the program will show us all the books that have red covers.

You can see at the top of the box that there are XX red books in the library. This page shows you the first 8 books. You can press these arrows to see more.

When buttons are grayed out like the “Orange” one, that means you can’t click the orange one because there aren’t any books that are both orange and red.

If you don't want to see red books, you can press the "Red" button again to make it go away. You can also press the trash can button to make it go away so you can start over.

If you want to go back up to the first screen to look at all the different square categories again, you can press the "Up" arrow. Now we can see all the square categories again.

Does that all make sense? Do you have any questions?

OK, now I am going to ask you to find some books again.

Appendix 7: Simple Experiment Tasks

Task Set 1	**Task Set 2**
1. Try out this program to see how it works. See what kinds of books you can find.	1. Try out this program to see how it works. See what kinds of books you can find.
2. Great job! Now I am going to show you how I use the program.	2. Great job! Now I am going to show you how I use the program.
3. How many picture books are there?	3. How many chapter books are there?
4. How many books for six to nine year olds are there?	4. How many books for three to five year olds are there?
5. How many medium books are there?	5. How many long books are there?
6. How many purple books are there?	6. How many pink books are there?
7. How many four star rated books are there?	7. How many five star rated books are there?
8. How many sad books are there?	8. How many happy books are there?
9. Did you like using this program? How hard was this program to use?	9. Did you like using this program? How hard was this program to use?

Appendix 8: Boolean Experiment Tasks

Task Set 1	Task Set 2
1. Try out this program to see how it works. See what kinds of books you can find.	1. Try out this program to see how it works. See what kinds of books you can find.
2. Great job! Now I am going to show you how I use the program.	2. Great job! Now I am going to show you how I use the program.
3. Can you find and click the buttons for the color yellow and the real animal characters?	3. Can you find and click the buttons for the color orange and the imaginary animal characters?
4. How many award-winning chapter books are there?	4. How many fairy tale picture books are there?
5. How many blue make believe books are there?	5. How rainbow true books are there?
6. How many long books about imaginary animal characters are there?	6. How many recently added books about real animal characters are there?
7. How many comic books for six to nine year olds are there?	7. How many plays for ten to thirteen year olds are there?
8. How many white five star rated books are there?	8. How many black four star rated books are there?
9. How many short sad books are there?	9. How many medium happy books are there?
10. Can you find and click the buttons for medium books and picture books?	10. Can you find and click the buttons for long books and chapter books?
11. Did you like using this program? How hard was this program to use?	11. Did you like using this program? How hard was this program to use?

Appendix 9: Simple Experiment Worksheet

User ID: ________________ Age: ______ Grade_____ Gender______
Date: ___________ School: ______________ Used ICDL Before: o Yes o No
1st Interface: _____________ 1st Task Set: _________ Laptop: ___________

Interface 1

Task 1: Exploration	o Hint o Hint o Hint
Task 2: Instructions	o Hint o Hint o Hint
Task 3: Search	o Hint o Hint o Hint
Task 4: Search	o Hint o Hint o Hint
Task 5: Search	o Hint o Hint o Hint
Task 6: Search	o Hint o Hint o Hint
Task 7: Search	o Hint o Hint o Hint
Task 8: Search	o Hint o Hint o Hint
Task 9: Like/Dislike and Difficulty	

Interface 2

Task 1: Exploration	o Hint o Hint o Hint
Task 2: Instructions	o Hint o Hint o Hint
Task 3: Search	o Hint o Hint o Hint
Task 4: Search	o Hint o Hint o Hint
Task 5: Search	o Hint o Hint o Hint
Task 6: Search	o Hint o Hint o Hint
Task 7: Search	o Hint o Hint o Hint
Task 8: Search	o Hint o Hint o Hint
Task 9: Like/Dislike and Difficulty	

Preference Comments

Appendix 10: Boolean Experiment Worksheet

User ID: ________________ Age: ______ Grade_____ Gender______
Date: ___________ School: ______________ Used ICDL Before: o Yes o No
1st Interface: _____________ 1st Task Set: _________ Laptop: ___________

Interface 1

Task 1: Exploration	o Hint o Hint o Hint
Task 2: Instructions	o Hint o Hint o Hint
Task 3: Identification Q1: What kinds of books did you find? Q2: Do these books have anything in common?	
Task 4: Search	o Hint o Hint o Hint
Task 5: Search	o Hint o Hint o Hint
Task 6: Search	o Hint o Hint o Hint
Task 7: Search	o Hint o Hint o Hint
Task 8: Search	o Hint o Hint o Hint
Task 9: Search	o Hint o Hint o Hint
Task 10: Identification Q1: What kinds of books did you find? Q2: Do these books have anything in common?	
Task 11: Like/Dislike and Difficulty	

Interface 2

Task 1: Exploration	o Hint o Hint o Hint
Task 2: Instructions	o Hint o Hint o Hint
Task 3: Identification Q1: What kinds of books did you find? Q2: Do these books have anything in common?	
Task 4: Search	o Hint o Hint o Hint
Task 5: Search	o Hint o Hint o Hint
Task 6: Search	o Hint o Hint o Hint
Task 7: Search	o Hint o Hint o Hint
Task 8: Search	o Hint o Hint o Hint
Task 9: Search	o Hint o Hint o Hint
Task 10: Identification Q1: What kinds of books did you find? Q2: Do these books have anything in common?	
Task 11: Like/Dislike and Difficulty	

Preference Comments

Bibliography

1. Abbas, J., Norris, C., and Soloway, E. (2002). Middle School Children's Use of the ARTEMIS Digital Library. *Proceedings of the Joint Conference on Digital Libraries*, 98-105.

2. Ackerman, E. (2001) Piaget's Constructivism, Papert's Constructionism: What's the Difference?

3. Ahlberg, C. and Shneiderman, B. (1994). Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays. *Proceedings of Human Factors in Computing*, 313-317.

4. Ahlberg, C., Williamson, C., and Shneiderman, B. (1992). Dynamic Queries for Information Exploration: An Implementation and Evaluation. *Proceedings of Human Factors in Computing*, 619-626.

5. Alborzi, H., Druin, A., Montemayor, J. et al. (2000). Designing StoryRooms: Interactive Storytelling Spaces for Children. *Proceedings of Designing Interactive Systems*, 95-104.

6. Alliance for Childhood (2000). Fools Gold: A Critical Look at Computers in Childhood. Retrieved May 30, 2004, from http://www.allianceforchildhood.net/projects/computers/computers_reports.htm.

7. American Academy of Pediatrics (n.d.). Television and the Family. Retrieved on June 11, 2004 from http://www.aap.org/family/tv1.htm.

8. Baker, J. (2003). The Impact of Paging vs. Scrolling on Reading Online Text Passages. *Usability News 5 (1)*. Retrieved April 11, 2005 from http://psychology.wichita.edu/surl/usabilitynews/51/paging_scrolling.htm.

9. Baumgarten, M. (2003). Kids and the Internet: A Developmental Summary. *ACM Computers in Entertainment*, *1 (1)*, 11.

10. Bederson, B., Hollan, J., Druin, A., et al. (1996). Local Tools: An Alternative to Tool Palettes. *Proceedings of User Interface Software and Technology*, 169-170.

11. Bernard, M., Baker, J., and Fernandez, M. (2002). Paging vs. Scrolling: Looking for the Best Way to Present Search Results. *Usability News, 4 (1)*. Retrieved June 15, 2004 from http://psychology.wichita.edu/surl/usabilitynews/41/paging.htm.

12. Beyer, H. and K. Holtzblatt. (1998). *Contextual Design: Defining Customer-Centered Systems*. San Francisco: Morgan Kaufmann Publishers.

13. Bilal, D. (2002). Children's Use of the Yahooligans! Web Search Engine. III. Cognitive and Physical Behaviors on Fully Self-Generated Search Tasks. *Journal of the American Society for Information Science and Technology, 53 (2)*, pp 1170-1183.

14. Bilal, D. and Kirby, J. (2002). Differences and Similarities in Information Seeking: Children and Adults as Web Users. *Information Processing and Management, 38*, 649-670.

15. Bloom, L., Lahey, M., Hood, L., Lifter, K., and Fiess, K. (1980). "Complex Sentences: Acquisition of Syntactic Connectives and the Semantic Relations They Encode." *Journal of Child Language, 7*, 235-261.

16. Borgman, C. (1986). "The User's Mental Model of an Information Retrieval System: An Experiment on a Prototype Online Catalog." *International Journal of Man-Machine Studies, 24*, 47-64.

17. Borgman, C., Hirsh, S., Walter, A., and Gallagher, A. (1995). Children's Searching Behavior on Browsing and Keyword Online Catalogs: The Science Library Catalog Project. *Journal of the American Society for Information Science and Technology, 46 (9)*, 663-684.

18. Brems, D. and Whitten, W. (1987). Learning and Preference for Icon-Based Interface. *Proceedings of the Human Factors Society*, 125-129.

19. Bruckman, A. and Bandlow, A. (2002). Human-Computer Interaction for Kids. *The Human-Computer Interaction Handbook*. Mahwah, NJ: Lawrence Erlbaum Associates.

20. Burman, E. (1994). *Deconstructing Developmental Psychology*. New York: Routledge.

21. Busey, P. and Doerr, T. (1993). Kid's Catalog: An Information Retrieval System for Children. *Youth Services in Libraries, 7 (1)*, 77-84.

22. Byrne, M. (1993). Using Icons to Find Documents: Simplicity is Critical. *Proceedings of Human Factors in Computing*, 446-453.

23. Byrne, M., John, B., Wehrle, N. and Crow, D. (1999). The Tangled Web We Wove: A Taskonomy of WWW Use. *Proceedings of Human Factors in Computing*, 544-551.

24. Card, S., Mackinlay, J., and Shneiderman, B. (1999). *Readings in Information Visualization: Using Vision to Think*. San Francisco: Morgan Kaufmann.

25. Card, S., Moran, T., and Newell, A. (1983). *The Psychology of Human-Computer Interaction*. Hillsdale, New Jersey: Lawrence Erlbaum.

26. Chang, S. and Rice, R. (1993). Browsing: A Multidimensional Framework. *Annual Review of Information Science and Technology*, 28, 231-276.

27. Chi, M. (1976). Short-Term Memory Limitations in Children: Capacity or Processing Deficits. *Memory and Cognition, 4 (5)*, 559-572.

28. Chi, M., Hutchinson, J., and Robin, A. (1989). How Inferences About Novel Domain-Related Concepts Can Be Constrained By Structural Knowledge. *Merrill-Palmer Quarterly, 35(1)*, pp. 27-62.

29. Clements, D. (1999). The Effective Use of Computers with Young Children. In J.V. Copley, Ed., *Mathematics in the Early Years*. Reston, VA: National Council of Teachers of Mathematics.

30. Cole, M., Gay, J., Glick, J, and Sharp, D. (1971). *The Cultural Context of Learning and Thinking: An Exploration in Experimental Anthropology*. New York: Basic Books.

31. Cooper, L. (2002a). A Case Study of Information-Seeking Behavior in 7-Year-Old Children in a Semistructured Situation. *Journal of the American Society for Information Science and Technology, 53 (11)*, 904-922.

32. Cooper, L. (2002b). Methodology for a Project Examining Cognitive Categories for Library Information in Young Children. *Journal of the American Society for Information Science and Technology, 53 (14)*, 1223-1231.

33. Cooper, L. (2004). The Socialization of Information Behavior: A Case Study of Cognitive Categories for Library Information. *Library Quarterly, 74*(3), 299-336.

34. Corporation for Public Broadcasting (CPB) (2002). Connected to the Future: A Report on Children's Internet Use. Retrieved May 30, 2004 from http://www.cpb.org/ed/resources/connected.

35. Deák, G., Ray, S., & Pick, A. (2002). Matching and Naming Objects by Shape or Function: Age and Context Effects in Preschool Children. *Developmental Psychology, 38*(4), 503-518.

36. Doan, K., Plaisant, C. and Shneiderman, B. (1996). Query Previews in Networked Information Systems. *Proceedings of Research and Technology Advances in Digital Libraries*, 120-129.

37. Druin, A. (1999a). Cooperative Inquiry: Developing New Technologies for Children with Children. *Proceedings of Human Factors in Computing*, 592-599.

38. Druin, A. (1999b). The Role of Children in the Design of New Technology. *Behaviour and Information Technology, 21 (1)*, 1-25.

39. Druin, A. (2005). What Children Can Teach Us: Developing Digital Libraries for Children with Children. *Library Quarterly, 75 (1)*, 20-41.

40. Druin, A., Bederson, B., Hourcade, J., et al. (2001). Designing a Digital Library for Young Children: An Intergenerational Partnership. *Proceedings of the Joint Conference on Digital Libraries*, 398-405.

41. Druin, A., Bederson, B., Weeks, A., et al. (2003). The International Children's Digital Library: Description and Analysis of First Use. *First Monday, 8 (5)*, http://firstmonday.org/issues/issue8_5/.

42. Druin, A., Hendler, J., Montemayor, J. et al. (1999). Designing PETS: A Personal Electronic Teller of Stories. *Proceedings of Human Factors in Computing*, 326-329.

43. Druin, A. and Inkpen, K. (2001). When are Personal Technologies for Children? *Personal Ubiquitous Computing, 5 (3)*, 191-194.

44. Druin, A., Stewart, J, Proft, D., et al. (1997). KidPad: A Design Collaboration Between Children, Technologists, and Educators. *Proceedings of Human Factors in Computing*, 462-470.

45. Edmonds, L., Moore, P, and Balcom, K. (1990). The Effectiveness of an Online Catalog. *School Library Journal, 36 (10)*, 28-33.

46. Egido, C. and Patterson, J. (1988). Pictures and Category Labels as Navigational Aids for Catalog Browsing. *Proceedings of Human Factors in Computing*, 127-132.

47. English, J., Hearst, M., Sinha, R. et al. (2002) Hierarchical Faceted Metadata in Site Search Interfaces. *Extended Abstracts of Human Factors in Computing*, 628-629.

48. Fishkin, K. and Stone, M. (1995). Enhanced Dynamic Queries via Moveable Filters. *Proceedings of Human Factors in Computing*, 415-420.

49. Fleener, C., Morrison, S., Linek, W., and Rasinski, T. (1997). Recreational Reading Choices: How Do Children Select Books? In W. Linek & E. Sturtevant (Eds.), *Exploring Literacy: The 19th Annual Yearbook of the College Reading Association*, 75-84. Platteville: University of Wisconsin.

50. Furnas, G. and Rauch, S. (1998). Considerations for Information Environments and the NaviQue Workspace. *Proceedings of Digital Libraries*, 79-88.

51. Gelman, S. & O'Reilly, A. (1988). Children's Inductive Inferences Within Superordinate Categories: The Role of Language and Category Structure. *Child Development, 59*, 876-887.

52. Gentner, D. & Namy, L. (1999). Comparison in the Development of Categories. *Cognitive Development, 14*, 487-513.

53. Gibson, D. (2004). The Site Browser: Catalyzing Improvements in Hypertext Organization. *Proceedings of Hypertext 2004*, ACM Press, pp. 68-76.

54. Gibson, L., Newall, F., and Gregor, P. (2003). Developing a Web Authoring Tool that Promotes Accessibility in Children's Designs. *Proceedings of Interaction Design and Children*, 23-30.

55. Greenbaum, J. & Kyng, M. (Eds.) (1991). *Design at Work: Cooperative Design of Computer Systems*. New Jersey: Lawrence Erlbaum Associates.

56. Guzdial, M. (1995). Software-Realized Scaffolding to Facilitate Programming for Science Learning. *Interactive Learning Enviroments, 4 (1)*, 1-44.

57. Hanna, L., Risden, K., Czerwinski, M., and Alexander, K. (1998). The Role of Usability Research in Designing Children's Computer Products. In A. Druin (Ed.), *The Design of Children's Technology*. San Francisco, CA: Morgan Kaufman.

58. Harris Interactive (2003). Generation Y Earns $211 Billion and Spends $172 Billion Annually. Retrieved on June 15, 2004 from http://www.harrisinteractive.com.

59. Harris, J. (1997). *Where's the Bear?* Los Angeles: The J. Paul Getty Museum.

60. Haugland, S. (1992). The Effects of Computer Software on Preschool Children's Developmental Gains. *Journal of Computing in Childhood Education, 3*, 15-30.

61. Hayes, B. and Younger, K. (2004). Category-Use Effects in Children. *Child Development, 75 (6)*, pp. 1719-1732.

62. The Henry J. Kaiser Family Foundation (2003). Key Facts: TV Violence. Retrieved on June 11, 2004 from http://www.kff.org.

63. The Henry J. Kaiser Family Foundation (2004). Issue Brief: The Role of Media in Childhood Obesity. Retrieved on June 11, 2004 from http://www.kff.org.

64. Hildreth, C. (1983). To Boolean or Not to Boolean? *Information Technology and Libraries, September 1983*, 235-237.

65. Hirsh, S. (1999). Children's Relevance Criteria and Information Seeking on Electronic Resources. *Journal of the American Society for Information Science and Technology, 50 (14)*, 1265-1283.

66. Hochheiser, H. and Shneiderman, B. (2000). Performance Benefits of Simultaneous Over Sequential Menus as Task Complexity Increases. *International Journal of Human-Computer Interaction, 12(2)*, 173-192.

67. Hourcade, J. (in press). Design for Children. In G. Salvendy Ed., *Handbook of Human Factors and Ergonomics (3rd Ed.)*, New York: Wiley Interscience.

68. Hourcade, J., Bederson, B., Druin, A. and Guimbretiere, F. (2003a). Accuracy, Target Reentry and Fitts' Law Performance of Preschool Children Using Mice. *University of Maryland Technical Report, HCIL-2003-16*.

69. Hourcade, J., Bederson, B, Druin, A., et al. (2003b). The International Children's Digital Library: Viewing Digital Books Online. *Interacting with Computers, 15*, 151-167.

70. Hutchinson, H. (2004). Children's Interface Design for Hierarchical Search and Browse. *ACM SIGCAPH Newsletter*. ACM Press, 11-12.

71. Hutchinson, H., Druin, A., Bederson, B., Reuter, K., Rose, A., & Weeks, A. (2005a). How do I find blue books about dogs? The errors and frustrations of young digital library users. *Proceedings of the 11th International Conference on Human-Computer Interaction (HCII 2005)* (CD-ROM). Mahwah, NJ: Lawrence Erlbaum Associates.

72. Hutchinson, H., Rose, A., Bederson, B., Weeks, A., and Druin, A. (2005b). The International Children's Digital Library: A Case Study in Designing for a Multi-Lingual, Multi-Cultural, Multi-Generational Audience. *Information Technology and Libraries, 24 (1)*, American Library Association, pp. 4-12.

73. Inkpen, K. (2001). Drag-and Drop versus Point-and-Click Mouse Interaction Styles for Children. *ACM Transactions on Computer-Human Interactions, 8 (1)*, 1-33.

74. Jacko, J. and Salvendy, G. (1996). Hierarchical Menu Design: Breadth, Depth, and Task Complexity. *Perceptual and Motor Skills, 82*, 1187-1201.

75. Johansson, B. and Sjolin, B. (1975). Preschool Children's Understanding of the Coordinators 'and' and 'or'. *Journal of Experimental Child Psychology, 19*, 233-240.

76. Johnson, K. and Eilers, A. (1998). Effects of Knowledge and Development on Subordinate Level Categorization. *Cognitive Development, 13 (4)*, pp. 515-545.

77. Jones, S. (1998). Graphical Query Preview Specification and Dynamic Result Previews for a Digital Library. *Proceedings of User Interface Software and Technology*, 143-151.

78. Kacmar, C. and Carey, J. (1991). Assessing the Usability of Icons in User Interfaces. *Behaviour and Information Technology, 10 (6)*, 443-457.

79. Kail, R. (1991). Developmental Change in Speed of Processing During Childhood and Adolescence. *Psychological Bulletin, 109 (3)*, 490-501.

80. Kiger, J. (1984). The Depth/Breadth Trade-Off in the Design of Menu-Driven User Interfaces. *International Journal of Man-Machine Studies, 20*, 201-213.

81. Kragler, S. and Nolley, C. (1996). Student Choices: Book Selection Strategies of Fourth Graders. *Reading Horizons, 36 (4)*, 354-365.

82. Kuhlthau, C. (1988). Meeting the Information Needs of Children and Young Adults: Basing Library Media Programs on Developmental States. *Journal of Youth Services in Libraries*, Fall 1988, 51-57.

83. Külper, U., Schulz, U., and Will, G. (1997). Bücherschatz – A Prototype of a Children's OPAC. *Information Services and Use, (17)*, 201-214.

84. Kwasnik, B. (1999). The Role of Classification in Knowledge Representation and Discovery. *Library Trends, 48 (1)*, pp. 22-47.

85. Large, A. and Beheshti, J. (2000). Primary School Students' Reactions to the Web as a Classroom Resource. *Proceedings of the Canadian Association for Information Science*, May 2000.

86. Larson, K. and Czerwinski, M. (1998). Web Page Design: Implications of Memory, Structure and Scent for Information Retrieval. *Proceedings of Human Factors in Computing*, 25-32.

87. Lee, E. and MacGregor, J. (1985). Minimizing User Search Time in Menu Retrieval Systems. *Human Factors, 27 (2)*, 157-162.

88. Lucy, J. and Gaskins, S. (2001). Grammatical Categories and the Development of Classification Preferences: A Comparative Approach. In M. Bowerman and S. Levinson (Eds.), *Language Acquisition and Conceptual Development* (pp. 257-283). Cambridge: Cambridge University Press.

89. Manfreda, K., Batagelj, Z, & Vehovar, V. (2002). Design of Web Survey Questionnaires: Three Basic Experiments. *Journal of Computer Mediated Communication 7 (3).*

90. Marchionini, G. (1989). Information-Seeking Strategies of Novices Using a Full-Text Electronic Encyclopedia. *Journal of the American Society for Information Science*, *40 (1)*, 54-66.

91. Marchionini, G. and Geisler, G. (2002). The Open Video Digital Library. *D-Lib Magazine*, *December 2002*.

92. Marchionini, G. & Teague, J. (1987). Elementary Students' Use of Electronic Information Services: An Exploratory Study. *Journal of Research on Computing in Education 20*(2), 139-155.

93. Michard, A. (1982). Graphical Presentation of Boolean Expressions in a Database Query Language: Design Notes and an Ergonomic Evaluation. *Behaviour and Information Technology, 1 (3)*, 279-288.

94. Mills, C. and Weldon, L. (1987). Reading Text From Computer Screens. *ACM Computing Surveys*, *19 (4)*, 329-358.

95. Miller, D. (1981). The Depth/Breadth Tradeoff in Hierarchical Computer Menus. *Proceedings of the Human Factors Society*, 296-300.

96. Miller, G. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity fro Processing Information. *The Psychological Review*, *63*, 81-97.

97. Moore, P. and St. George, A. (1991). Children as Information Seekers: The Cognitive Demands of Books and Library Systems. *School Library Media Quarterly, 19*, 161-168.

98. Moran, T. (1981). An Applied Psychology of the User. *Computing Surveys*, *13 (1)*, 1-11.

99. Murrell, K. (1998). Human Computer Interface Design in a Multi-cultural Multi-lingual Environment. Proceeding *of the 13th Annual MSc and PhD Conference in Computer Science*, University of Stellenbosch, pages 50 - 54.

100. Naaman, M., Song, Y., Paepcke, A., and Garcia-Molina, H. (2004). Automatically Generating Metadata for Digital Photographs with Geographic Coordinates. *Proceedings of World Wide Web 2004*, ACM Press, pp. 244-245.

101. Nahl, D. and Harada, V. (1996). Composing Boolean Search Statements: Self Confidence, Concept Analysis, Search Logic, and Errors. *School Library Media Quarterly, 24 (4)*, pp. 199-207.

102. National Institute on Media and the Family (2002). Fact Sheet: Children and Advertising. http://www.mediafamily.org/facts/facts_childadv.shtml.

103. National Institute on Media and the Family (2000). Fact Sheet: Internet Advertising and Children. http://www.mediafamily.org/facts/facts_internetads.shtml

104. National Telecommunications and Information Administration (NTIA) (2002). A Nation Online: How Americans are Expanding Their Use of the Internet. Retrieved May 30, 2004 from http://www.ntia.doc.gov/ntiahome/dn/.

105. National Telecommunications and Information Administration (NTIA) (2004). A Nation Online: Entering the Broadband Age. Retrieved January 5, 2005 from http://www.ntia.doc.gov/reports/anol/index.html.

106. Nazzi, T., & Gopnik, A. (2000). A Shift in Children's Use of Perceptual and Causal Cues to Categorization. *Developmental Science*, *3*(4), 389-396.

107. Neimark, E. and Slotnick, N. (1970). Development of the Understanding of Logical Connectives. *Journal of Educational Psychology*, *61 (6)*, 451-460.

108. Nguyen, S. & Murphy, G. (2003). An Apple is More Than Just a Fruit: Cross-Classification in Children's Concepts. *Child Development*, *74*(6), 1783-1806.

109. Norman, D. (1983). Some Observations on Mental Models. In Gentner, D. and Stevens, A., Eds. *Mental Models*. Hillsdale, NJ: Lawrence Erlbaum.

110. Norman, K. (1991). The Psychology of Menu Selection: Designing Cognitive Control at the Human/Computer Interface. Norwood, NJ: Ablex Publishing Corporation.

111. NUA Internet Surveys (April 2002). US Kids Choose Internet Over Other Media. Retrieved May 25, 2004 from http://www.nua.com/surveys/index.cgi?f=VS&art_id=905357821&rel=true

112. NUA Internet Surveys (October 2002). Kids Know More About Internet than Books. Retrieved May 25, 2004 from http://www.nua.com/surveys/index.cgi?f=VS&art_id=905358419&rel=true

113. NUA Internet Surveys (November 2002). Pester Power Pushes Net Shopping Spree. Retrieved June 15, 2004 from http://www.nua.com/surveys/index.cgi?f=VS&art_id=905358559&rel=true

114. Oosterholt, R., Kusano, M., and deVries, G. (1996). Interaction Design and Human Factors Support in the Development of a Personal Communicator for Children. *Proceedings of Human Factors in Computing*, 450-457.

115. Pane, J. and Myers, B. (2000). Improving User Performance on Boolean Queries. *Proceedings of Human Factors in Computing*, 269-270.

116. Papert, S. (1980). *Mindstorms: Children, Computers, and Powerful Ideas*. New York: Basic Books.

117. Pejtersen, A. (1986). Design and Test of a Database for Fiction Based on an Analysis of Children's Search Behaviour. In P. Ingwersen & A. Pejtersen (Eds.) *Information Technology and Information Use: Towards a Unified View of Information and Information Technology*, 125-147. London: Taylor Graham.

118. Pejtersen, A. (1989). A Library System for Information Retrieval Based on a Cognitive Task Analysis and Supported by an Icon-Based Interface. *ACM Conference on Information Retrieval*, ACM Press, 40-47.

119. Piaget, J. and Inhelder, B. (1969). *The Psychology of the Child*. New York: Basic Books.

120. Piolat, A., Roussey, J., and Thunin, O. (1998). Effects of Screen Presentation on Text Reading and Revising. *International Journal of Human-Computer Studies, 47*, 565-589.

121. Plaisant, C., Marchionini, G., and Bruns, T. et al. (1997). Bringing Treasures to the Surface: Iterative Design for the Library of Congress National Digital Library Program. *Proceedings of Human Factors in Computing*, 518-525.

122. Rawson, L, Tamayo, F., Vehle, M., and Willemsen, E. (1973). Disjunctive Concept Utilization in Preschool Children. *The Journal of Genetic Psychology, 122*, 211-216.

123. Reti, T. and Sarva, R. (2004). DiMaS: Distributing Multimedia on Peer-to-Peer File Sharing Networks. *Proceedings of Multimedia 2004,* ACM Press, pp. 166-167.

124. Reuter, K. (2004). A Portrait of the Child as an Information Seeker. Unpublished manuscript for a qualifying exam at the University of Maryland.

125. Reuter, K. and Druin, A. (2004). Bringing Together Children and Books: An Initial Descriptive Study of Children's Book Searching and Selection Behavior in a Digital Library. *Proceedings of American Society for Information Science and Technology Conference*, Providence, RI.

126. Revelle, G. (2003). Educating Via Entertainment Media: The Sesame Workshop Approach. *ACM Computers in Entertainment, 1 (1),* 16.

127. Revelle, G., Druin, A., Platner, M., et al. (2002). A Visual Search Tool for Early Elementary Science Students. *Journal of Science Education and Technology, 11(1)*, 49-57.

128. Rideout, V., Vandewater, E., and Wartella, E. (2003). Zero to Six: Electronic Media in the Lives of Infants, Toddlers, and Preschoolers. The Henry J. Kaiser Foundation, on June 15, 2004 from http://www.kff.org.

129. Robinson, C., Larsen, J., and Haupt, J. (1997). Picture Book Selection Behaviors of Emergent Readers: Influence of Genre, Familiarity, and Book Attributes. *Reading Research and Instruction, 36 (4)*, 287-304.

130. Rosch, E., Mervis, C., Gray, W., Johnson, D., & Boyes-Braem, P. (1976). Basic Objects in Natural Categories. *Cognitive Psychology, 8*, 382-439.

131. Scaife, M., Rogers, Y., Aldrich, F., and Davies, M. (1997). Designing For of Designing With? Informant Design for Interactive Learning Environments. *Proceedings of Human Factors in Computing*, 343-350.

132. Schacter, J., Chung, G., & Dorr, A. (1998). Children's Internet Searching on Complex Problems: Performance and Process Analysis. *Journal of the American Society for Information Science, 49*, 840-849.

133. Schneider, K. (1996). Children and Information Visualization Technologies. *Interactions, September-October 1996*, 68-74.

134. Schuler, D. & Namioka, A. (Eds.) (1993). *Participatory Design: Principles and Practice*. New Jersey: Lawrence Erlbaum.

135. Schwarz, E., Beldie, I., and Pastoor, S. (1983). A Comparison of Paging and Scrolling for Changing Screen Contents by Inexperienced Users. *Human Factors, 25 (3)*, 279-282.

136. Shneiderman, B. (1983). Direct Manipulation: A Step Beyond Programming Languages. *IEEE Computer, 16 (8)*, 57-68.

137. Shneiderman, B. (1994). Dynamic Queries for Visual Information Seeking. *IEEE Software, 11 (6)*, 70-77.

138. Shneiderman, B. (1998). *Designing the User Interface, 3rd Ed.* Reading, MA: Addison Wesley.

139. Shneiderman, B., Feldman, D., Rose, A., and Grau, X. (2000). Visualizing Digital Library Search Results with Categorical and Hierarchical Axes. *Proceedings of Digital Libraries*, 57-65.

140. Snow, C. and Rabinovitch, M. (1969). Conjunctive and Disjunctive Thinking in Children. *Journal of Experimental Child Psychology, 7*, 1-9.

141. Snowberry, J., Parkinson, S., and Sisson, N. (1983). Computer Display Menus. *Ergonomics, 26 (7),* 699-712.

142. Solomon, P. (1993). Children's Information Retrieval Behavior: A Case Analysis of an OPAC. *Journal of the American Society for Information Science and Technology, 44 (5)*, 245-264.

143. Soloway, E., Guzdial, M., and Hay, K. (1994). Learner-Centered Design: The Challenge for HCI in the 21st Century. *Interactions, April 1994*, 36-48.

144. Spavold, J. (1990). The Child as Naïve User: A Study of Database Use with Young Children. *International Journal of Man-Machine Studies, 32*, 603-625.

145. Spoerri, A. (1993). InfoCrystal: A Visual Tool for Information Retrieval and Management. *Proceedings of Information Knowledge and Management*, 11-20.

146. Stewart, J., Bederson, B., and Druin, A. (1999). Single Display Groupware: A Model for Co-Present Collaboration. *Proceedings of Human Factors in Computing*, 286-293.

147. Strommen, E. (1994). Children's Use of Mouse-Based Interfaces to Control Virtual Travel. *Proceedings of Human Factors in Computing*, 405-410.

148. Strommen, E. (1998). When the Interface is a Talking Dinosaur: Learning Across Media with ActiMates Barney. *Proceedings of Human Factors in Computing*, 288-295.

149. Theng, Y., Nasir, N. Thimbleby, H., et al. (2000). Children as Design Partners and Testers for a Children's Digital Library. *Proceedings of the European Conference on Digital Libraries*, 249-258.

150. Theng, Y., Nasir, N., Buchanan, G., et al. (2001). Dynamic Digital Libraries for Children. *Proceedings of the Joint Conference on Digital Libraries*, 406-415.

151. Thomas, J. (1980). Acquisition of Motor Skills: Information Processing Differences Between Children and Adults. *Research Quarterly for Exercise and Sport, 51 (1)*, 158-173.

152. Tversky, B. (1985). Development of Taxonomic Organization of Named and Pictured Categories. *Developmental Psychology, 21*(6), 1111-1119.

153. Tversky, A. and Kahneman, D. (1974). Judgments Under Uncertainty: Heuristics and Biases. *Science, 185*, 1124-1131.

154. Van der Veer, G. and Melguizo, M. (2002). Mental Models. *The Human-Computer Interaction Handbook*. Mahwah, NJ: Lawrence Erlbaum Associates.

155. Van Kesteren, I., Bekker, M., Vermeeren, A., and Lloyd, P. (2003). Assessing Usability Evaluation Methods On Their Effectiveness to Elicit Verbal Comments From Children Subjects. *Proceedings of Interaction Design and Children*, 41-49.

156. Vygotsky, L.S. (1978). *Mind and Society: The Development of Higher Mental Processes*. Cambridge, MA: Harvard University Press.

157. Wendelin, K. and Zinck, R. (1983). How Students Make Book Choices. *Reading Horizons*, *23*, 84-88.

158. Witten I., McNab R., Boddie S. and Bainbridge D. (2000). Greenstone: A Comprehensive Open-Source Digital Library Software System. *Proceedings of Digital Libraries 2000*, pp 113-121.

159. Wood, D., Bruner, J., and Ross, G. (1976). The Role of Tutoring in Problem Solving. *Journal of Child Psychology, Psychiatry, and Allied Disciplines, 17 (2)*, 89-100.

160. Yee, K., Swearingen, K., Li, K. and Hearst, M. (2003). Faceted Metadata for Image Searching and Browsing. *Proceedings of Human Factors in Computing*, 401-408.

161. Young, R. (1983). Surrogates and Mappings: Two Kinds of Conceptual Models for Interactive Devices. In Gentner, D. and Stevens, A., Eds. *Mental Models*. Hillsdale, NJ: Lawrence Erlbaum.

162. Young, D. and Shneiderman, B. (1993). A Graphical Filter/Flow Representation of Boolean Queries: A Prototype Implementation and Evaluation . *Journal of American Society for Information Science*, 44 (6), 327-339.

163. Zaphiris, P. and Mtei, l. (1997). Depth v. Breadth in the Arrangement of Web Links. Retrieved May 30, 2004 from http://otal.umd.edu/SHORE/bs04/.

164. Zloof, M. (1975). Query-By-Example: A Data Base Language. *Proceedings of AFIPS Conference, 44*, 324-343.