May, 1995

The effect of domain knowledge on elementary school children's search behavior on an information retrieval system: The Science Library Catalog

Sandra Hirsh, *University of Arizona*

The Effect of Domain Knowledge on Elementary School Children's Search Behavior on an Information Retrieval System: The Science Library Catalog

Sandra Goldstein Hirsh

Graduate School of Education and Information Studies
University of California, Los Angeles
Los Angeles, CA 90024, USA
ibkgsan@mvs.oac.ucla.edu

ABSTRACT

Few information retrieval systems are designed with children's special needs and capabilities in mind. We need to learn more about children's information-seeking behavior in order to provide them with information-based tools which support exploratory learning. This dissertation examines children's search behavior on a hypertext-based automated library catalog designed for elementary school children. The focus of this research is on the effect of domain knowledge on children's search performance, search behavior, and learning as they look for science books on this system. Research has shown that level of domain knowledge influences the way people search for information. Data was collected through one-on-one interviews, direct observation, and online monitoring of search sessions. This dissertation will contribute to our understanding of children's search behavior and the factors which influence their behavior. This research also has implications for information retrieval system evaluation and interface design.

KEYWORDS

Elementary school children, Information retrieval, Search behavior, Science education, Learning, Online catalogs, Domain knowledge, Browsing, Keyword searching, Online monitoring, User interface design.

SUMMARY

A number of recent information retrieval studies have found that adult search behavior is affected by knowledge of the particular domain being researched. High levels of domain knowledge enable users to formulate appropriate search queries, generate synonyms for their query terms, determine alternative strategies when their first attempt is unsuccessful, and recognize relevant materials. Experts in particular domains utilize different search strategies and have higher success rates than people who know little about the domain being searched. However, no studies found have examined the influence of domain knowledge on children's search behavior on an information retrieval system. My dissertation is taking the first look at the effect of children's prior subject knowledge on their search behavior, performance, and learning on an information retrieval system. The primary research questions are:

- Does level of prior science knowledge affect children's abilities to locate science materials on an automated library catalog?

- Does familiarity with the system affect how effectively and efficiently children perform search tasks?

- Does level of prior science knowledge affect children's selection of search methods?

- Does familiarity with the system affect children's selection of search methods?

- Do children learn about the organization of science knowledge from using the system?

This dissertation builds on the baseline data collected in a joint UCLA-Cal Tech research project studying elementary school children's search behavior on an electronic information retrieval system, the Science Library Catalog [1]. This study employs a more advanced version of the Science Library Catalog interface than previously available. Adding to the existing browsing-oriented search method which allows children to navigate through knowledge hierarchies, the new interface provides children with more sophisticated keyword search options. (See Figure 1)
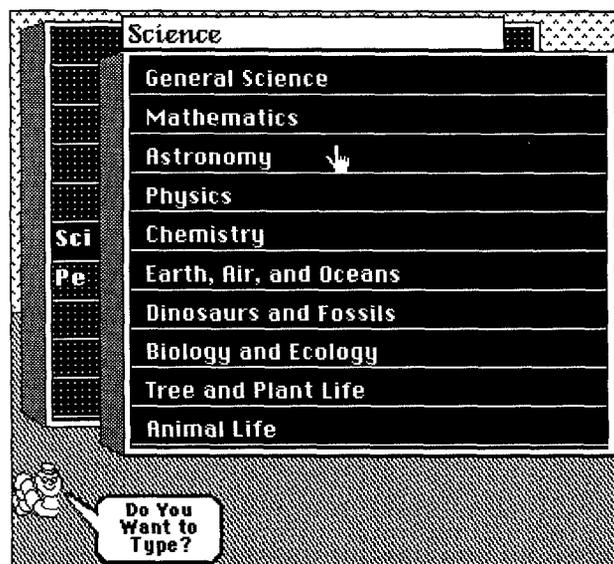
Figure 1. Science Library Catalog Screen

The browsing approach allows children to recognize science topics they want to explore by clicking on the bookshelf topics with the mouse. The keyword approach requires children to recall appropriate search terms and type in their query using the keyboard. The interface employs a bookshelf metaphor to make the search process more tangible.

Sixty-four fifth-grade children participated in this study. The sample was balanced by level of domain knowledge and gender. The domain area under study in this dissertation was science. Science knowledge was measured by science grades which teachers assigned using standardized methods across the school district. Students were drawn individually from their classrooms for two one-on-one interviews, each interview separated by one week. During each interview session, children were asked to find books on four different science topics on the Science Library Catalog. Children were instructed to find books as if they needed them to write a school paper on each of the topics. None of the children had previous exposure to the system. The effects of other factors (e.g. gender and computer experience) on search success, search behavior, and learning were considered in the dissertation, but are not reported in this summary.

Search Success
Domain knowledge is expected to influence children's success in finding books on assigned topics. Children with high domain knowledge are expected to be more successful (e.g., to find more books on the assigned topics in the automated library catalog) than children with low domain knowledge. Success is also expected to vary by the difficulty of the assigned tasks.

Search Patterns
Search behavior data were collected through direct observation and online monitoring of each user's search session. The monitoring logs captured keystrokes unobtrusively during search sessions. Sequences of search moves will be evaluated by using a pattern matching program developed with researchers at the Computer Science Department at the University of New South Wales. This program asks the researcher to specify the particular pattern to look for within the monitoring data; the output will be a set of frequency counts of the number of times that a particular pattern was matched within the file specified, in addition to the amount of time spent on these sequences of moves. The pattern matching program will be used to identify user searching models based on patterns exhibited by different users under varying conditions. The effect of domain knowledge on these patterns will also be examined.

Learning
This dissertation also explores two types of learning which may occur during the search sessions. Because of the unique design of the Science Library Catalog, which allows children to browse knowledge hierarchies based on the Dewey Decimal Classification, children are expected to learn about the organization of science. Science hierarchy learning is evaluated through a card sorting task administered as a pre- and post-test.

Most information retrieval research evaluates users' performance during their first encounter with a new system.

Few studies have examined what happens to people's search strategies and performance as they become more familiar with the system. In this dissertation, children's search performance is compared over two search sessions.

Data for the full study have been collected and are currently being analyzed. Statistical tests will be used to analyze the relationship of children's domain knowledge to: success rates in completing search tasks; the search strategies used to complete search tasks; the sequences of search moves utilized to complete search tasks; learning of science classification through the use of the Science Library Catalog; and, learning of the automated library catalog.

Findings from a Pilot Study, involving eight gifted-and-talented fifth-grade children, revealed interesting differences in children's search behavior and performance based on domain knowledge. For example, one finding indicated a complex relationship between domain knowledge and the difficulty of the assigned search task. As expected, children with high domain knowledge performed better than children with low domain knowledge on simple tasks. However, children with low domain knowledge performed surprisingly better on complex tasks than children with high domain knowledge. This finding may be explained by the search methods children used. Children with high domain knowledge used more browsing techniques than children with low domain knowledge on both simple and complex tasks. However, browsing led to fewer successful matches than keyword search techniques.

Another finding from the pilot study was the amount of learning that takes place during the search process. Children appear to learn about the hierarchical organization of science knowledge from using the Science Library Catalog, as indicated by a decrease in card sorting errors. As expected, the findings indicate that children with lower domain knowledge make consistently more errors on both card sorting tasks than children with high domain knowledge. This finding suggests that children with high domain knowledge have a better understanding of the hierarchical organization of science than children with low domain knowledge. Children also appear to search more successfully as they become more familiar with the Science Library Catalog, abandoning half as many searches in their second search session as compared to their first session.

This dissertation will result in a model of children's search behavior on an information retrieval system under varying circumstances and tasks. Understanding children's search behavior and how different factors affect that behavior is necessary for designing appropriate interfaces for children's information retrieval systems. This dissertation will also contribute to the broader methodological knowledge on online monitoring data for computer-human interaction, applying an established linear technique to a non-linear, graphical environment.

REFERENCES
1. Borgman, C.L., Gallagher, A.L., Hirsh, S.G., and Walter, V.A. (in press). Children's Searching Behavior on Browsing and Keyword Online Catalogs: The Science Library Catalog Project. *J. Am. Soc. Info. Sci.*