SMOOTHING THE INFORMATION SEEKING PATH: REMOVING REPRESENTATIONAL OBSTACLES IN THE MIDDLE SCHOOL DIGITAL LIBRARY ENVIRONMENT

June M. Abbas, B.A., M.L.S.

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

May 2002

APPROVED:

Brian C. O'Connor, Major Professor and Associate Director of the Interdisciplinary Information Science Program
Cathleen Norris, Committee Member
Linda Schamber, Committee Member
Elliot Soloway, Committee Member
Jud Copeland, Committee Member
Philip Turner, Dean of the School of Library and Information Science
C. Neal Tate, Dean of the Robert B. Toulouse School of Graduate Studies

Abbas, June M., Smoothing the information seeking path: Removing representational obstacles in the middle-school digital library. Doctor of Philosophy (Information Science), May 2002, 225 pp., 31 tables, 21 figures, 12 appendices, references, 74 titles.

Middle school student's interaction within a digital library is explored. Issues of interface features used, obstacles encountered, search strategies and search techniques used, and representation obstacles are examined. A mechanism for evaluating user's descriptors is tested and effects of augmenting the system's resource descriptions with these descriptors on retrieval is explored.

Transaction log data analysis (TLA) was used, with external corroborating achievement data provided by teachers. Analysis was conducted using quantitative and qualitative methods. Coding schemes for the failure analysis, search strategies and techniques analysis, as well as extent of match analysis between terms in student's questions and their search terms, and extent of match analysis between search terms and controlled vocabulary were developed.

There are five chapters with twelve supporting appendixes. Chapter One presents an introduction to the problem and reviews the pilot study. Chapter Two presents the literature review and theoretical basis for the study. Chapter Three describes the research questions, hypotheses and methods. Chapter Four presents findings. Chapter Five presents a summary of the findings and their support of the hypotheses. Unanticipated findings, limitations, speculations, and areas of further research are indicated.

Findings indicate that middle school users interact with the system in various sequences of patterns. User groups' interactions and scaffold use are influenced by the teacher's objectives for using the ADL. Users preferred to use single word searches over Boolean, phrase or natural language searches. Users tended to use a strategy of repeating the same exact search, instead of using the advanced scaffolds. A high percent of users attempted at least one search that included spelling or typographical errors, punctuation, or sequentially repeated searches. Search terms matched the DQ's in some instantiation 54% of all searches. Terms used by the system to represent the resources do not adequately represent the user groups' information needs, however, using student generated keywords to augment resource descriptions can have a positive effect on retrieval.

ACKNOWLEDGEMENTS

It is often said that the dissertation process is a solitary journey, one that must be made alone. However, without the proper preparation and continued guidance from those we revere, the traveler may lose their way or turn back out of frustration. To all of my guides, I thank you. Without you, I fear I would not have begun this journey, or I would have turned back long ago.

I wish to give acknowledgement to my committee members for their time, effort, and devotion to me and my continued growth as a researcher. You have kept me focused, inspired, and helped me through many forks in the road. Thank you to Elliot Soloway of the University of Michigan and Cathleen Norris for providing me the exciting opportunity and guided exploration of examining the ARTEMIS Digital Library. I extend continual gratitude to the National Science Foundation for providing the funding for the ADL project and this research.

I wish to give especial acknowledgement to the Chair of my Committee, Brian C. O'Connor who has been my mentor, friend, and trail guide for many years. He has taught me to persevere in the face of many frustrations, to always look beyond the immediate or easily seen picture to what is beneath, and to stretch my thinking to accommodate new ideas, implications, and possibilities. Without his continued support, I would have lost my way long ago.

Most importantly, I thank my family for their love, support, and patience.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Trail Markers in Search Space

Children have unique information needs and information seeking strategies (Walters, 1994). Maximizing the utility of children's information seeking requires appropriate representations of both their questions and the system documents. Children learn about their world through diverse resources and in many contexts. Two of the primary influences in children's lives today are their parents and schools. Within these contexts children learn language, social skills, academic skills, and the wealth of knowledge that is their world. There has been a push in recent years to have access to the Internet and World Wide Web (WWW) in every school. Access to the growing body of information on the Internet and WWW has presented educators with the additional challenges of 1) how to access the best appropriate information available, and 2) how to use the digital resources successfully. Integrating technology and digital resources into the classroom demands that educators understand and teach information seeking skills and information retrieval strategies. Integrating technology and digital resources also requires educators to attempt to understand how this new technology (resource) can provide the best educational benefit for their students. In order to understand how best to integrate the technology, it is important that we examine how children interact with the information retrieval system.

This study brings together theories and practices from the field of information science in the hopes of developing a deeper understanding of representation of questions and documents by middle school children. It is constructed as an examination of how children make use of one specific digital library learning environment in the hopes of learning more about children's question states. This study will also explore how this knowledge can then be applied to representation and retrieval theories and practices, in order to smooth or clear the information seeking path of obstacles encountered by children as they engage with the system. It is an information seeking study in the sense that construction and use of representations significantly impact search space, search time, and search quality.

Context for the Study

Fundamental definitions and an explanation of the ARTEMIS digital library learning environment set the stage of inquiry.

Definitions

Digital library: Many definitions have been proposed for the term digital library. Lynch & Garcia-Molina (1995) define a digital library as "a system that provides a community of users with coherent access to a large, organized depository of information and knowledge." Bishop and Starr (1996) outline three elements necessary to a digital library: "(1) some sense of a collection, with some kind of organization; the content may be partly physical and partly electronic, or entirely electronic; (2) a collection that is not entirely bibliographic or exclusively a set of pointers to other material, it must contain some 'full-form online material' and may be in a variety of formats; and (3) a goal exists to link ‘audience, group, patron, or community with attributes of the collection” (p. 350). Waters (1998) defines a digital library as “organizations that provide the resources, including the specialized staff, to select, structure, offer intellectual access to, interpret, distribute, preserve the integrity of and ensure the persistence over time of collections of digital works so that they are readily and economically available for use by a defined community or set of communities.”

Borgman (1999) reviews the definitions proposed by library and information science researchers and practitioners in the digital environment. She sees two distinct uses of the term by researchers and practicing librarians. Researchers take a narrow approach and define a digital library as “a set of electronic resources . . . for creating, searching, and using information. In this sense they are an extension and enhancement of information storage and retrieval systems that manipulate digital data in any medium (text, images, sounds; static or dynamic images) and exist in distributed networks. The content of digital libraries includes data, metadata that describe various aspects of the data (eg. representation, creator, owner, reproduction rights) and metadata that consist of links or relationships to other data or metadata, whether internal or external to the digital library. Digital libraries are constructed, collected and organized, by (and for) a community of users, and their functional capabilities support the information needs and uses of that community” (p. 234). While each of the above definitions was formulated for a specific purpose or agenda; they all include the important elements of selecting and evaluating resources, and providing access to the resources for a community of users.

Borgman further contends that digital libraries are becoming "enabling technologies for many applications" (p. 239).

For the purpose of the study digital library is defined as a preselected and organized collection of electronic resources designed to serve the information needs of a specific community of users. The resources may include any form of media such as text, images, video, and sound. The collection exists within a distributed networked environment. A further essential element to the ARTEMIS digital library is that the interface includes additional enabling technologies that assist the community of users in locating, organizing, evaluating, and learning science related content. For this reason, the ARTEMIS digital library will be referred to as a digital library learning environment.

Information: In the field of information science the term information has been defined in various contexts and to serve specific agendas. Buckland (1991) defines the term as "information-as-knowledge," or a change in the person's cognitive state; "information-as-thing," or the physical representation of someone's cognitive thoughts or ideas"; and "information-as-process," or the act of informing or facilitating informing" (p.5). Borgman (1999) describes information as "a relation of at least four terms: A PERSON is informed by a SIGN about some THING within a certain CONTEXT" (p. 20). He further explains that even if the signs and the contexts are plain, one more ingredient is needed in order to understand the sign. A person must possess the INTELLIGENCE or the code to understand the sign. Borgman's definition re-emphasizes Eco's meaning of information within the context of representation. He states "there is no sign without a code" (O'Connor, 2001), which means that unless the person encountering the information understands the code in which it exists, they will have little or no use for the "thing" being encountered. The terminology and domain knowledge that students learn within the classroom are important factors that could affect their information seeking. Within the context of the study, information is defined as that "thing" or representation (sign) of the "thing" which serves to inform the student. The informing takes place within a specific context, which can in itself influence how the student encodes or learns the information.

Representation: Representation is a *central* concept in information organization and information seeking. It has been defined by O'Connor (1996) as "the set of means by which one thing stands for another . . . . [It is] a complex web of attributes of disparate objects and concepts, idiosyncratic and socially constructed codes and agreements, and neurological abilities" (p. 11). Marr (1982) defines the concept as "a system for extracting or highlighting some characteristics of concepts or things along with an explanation of the rules and reasons for the extraction" (p. 20). The term can also be used to refer to the process or activity of representing an object (Jacob and Shaw, 1998). The representation process "seeks to establish [a] systematic correspondence between the target domain and the modeling domain" (Jacob and Shaw, p. 146). In essence, representation creates a *surrogate* to stand for or take the place of the object for the same purposes.

Blair (1990) sees the problem of representation and information retrieval as linguistic in nature. How effectively we utilize language to represent an object, determines the success or failure of the information retrieval process. Blair's perspective shifts the focus of user-centered information retrieval away from interface or design issues and towards a sociocognitive emphasis on “how language is used to communicate in various social activities” (Jacob and Shaw, 1998 p. 146). According to Blair (as cited in Jacob and Shaw, p. 205) language [is] not the *product* of thought, but the *vehicle* of thought”. Where information retrieval fails is when the “congruence” between the representation, or the indexer’s thoughts, and the use or the user’s thoughts of the representation is not achieved (Shera, 1965; Goodrum, 1997).

O’Connor (1996) further cautions that representations are not just another instance of the original. They are generally smaller, shorter, or less time consuming, so something is missing. There are also limitations to the use of representations. If someone does not know or understand the rules and procedures for the representation, the representation will be useless to the user.

Representation is also central to the process of indexing or abstracting an object. The term indexing is derived from the Greek word meaning “to point” (O’Connor, 1996 p. 8). Indexing may be defined as “the process of deciding what some item is about and of giving it a label to represent this decision” (Ornager,1997 p. 202). We might say the main process of indexing is to create a representation. The information an index presents is generally of two types: **explicit** or that which is expressed by the author, and **implicit** or that which is not directly expressed by the author, but is understood by the person using the object (Ornager, 1997). An index may be attached to a document, such as the index in the back of a book or embedded into a retrieval system, such as a controlled vocabulary. The function of an index may therefore be described as a system which points to the essence of the object.

Indexing points to the essence of a thing by determining its *attributes* or “any kind of feature, component, or property of a stimulus that can be represented by an information processing system” (Jorgensen, 1997, p. 209). Attributes may be thought of as the “characteristics” of an object. Attributes are not limited to physical or perceptual characteristics of the object. They also include other cognitive, affective, or interpretive responses or feelings about the object. There are also two categories of attributes that are important to the representation of objects: diachronic attributes, or those which remain the same across time, and synchronic attributes, or those which may change with time and context (O’Connor, 1996).

These categories become very important not only in creating representations for the object, but also in retrieving the object. However, traditional means or systems for retrieving objects, do not incorporate these elements within their design. These systems are instead concerned with the physical access to objects, not to the *intellectual* access to objects.

Abstracting, or the process of distilling the document down to its fundamental essence, is another form of representation used within some information retrieval systems. The ARTEMIS digital librarians create abstracts of each document chosen for the collection and provide the abstracts, as well as additional keywords they choose to represent the intellectual content of the document, within a searchable database. Retrieval within the ARTEMIS environment consists of the students querying the database of abstracts and keywords. Successful retrieval is achieved when the student query terms match exactly with those within the abstract or keywords representing the documents. Figure 1 provides an example of an ARTEMIS abstract retrieved by a student search for “space stations”.

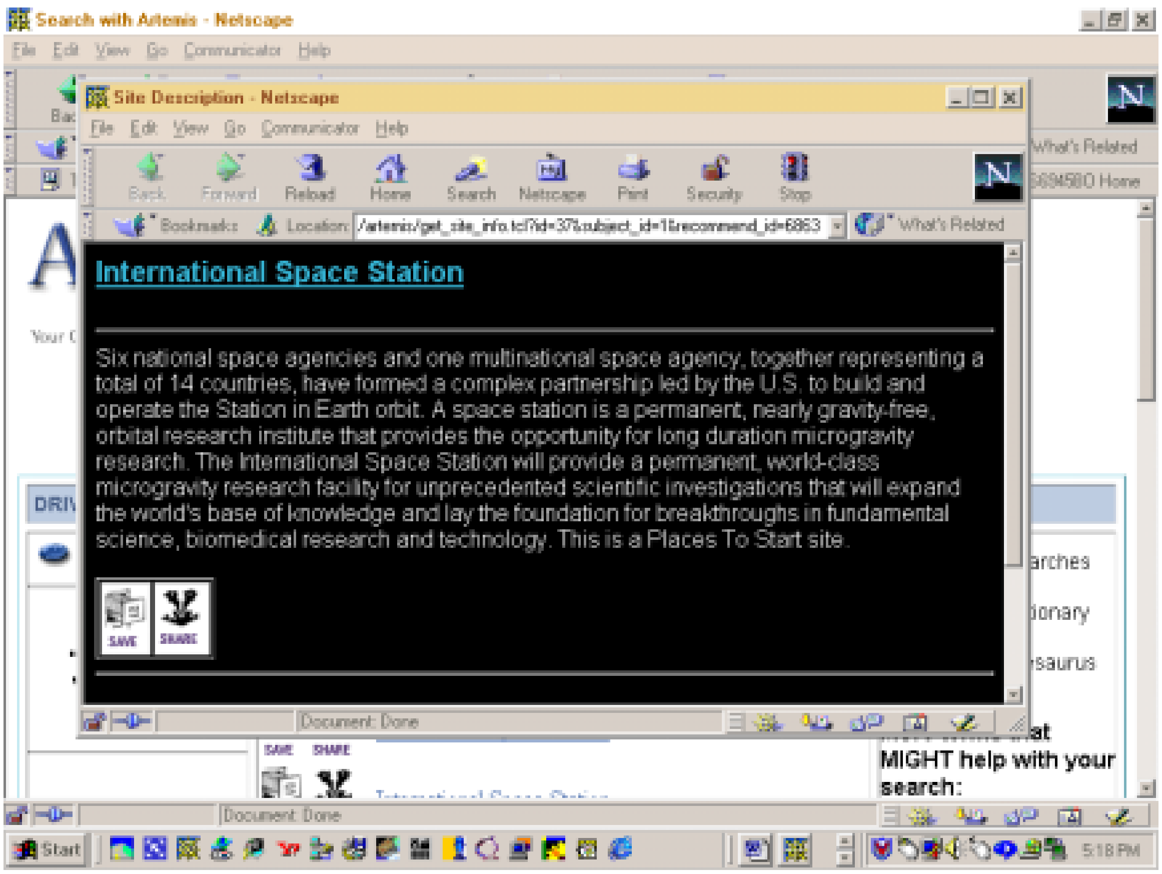


Figure 1: Example of ADL Abstract

Question/question state: Asking a question and looking for the answer to that question is at the heart of the information retrieval transaction. Paul Churchland (1996), philosopher and human-machine cognition researcher believes that we possess no real robust models of this state. He proposes that if we did image the brain during a question state, it might appear much like a landscape, with depressions and differing "wells of attraction". The activity of the question might be represented as gatherings of impulses or "things" lying on the edges of the wells in a dynamic tension state of falling into the wells or remaining on the edges. Presumably, what needs to occur to resolve this tension state is connecting with the existing knowledge to answer the question or realizing there is a knowledge gap that needs resolution.

Questions may be the most powerful technology we have ever created. Questions and questioning allow us to make sense of a confusing world. They are the tools that lead to insight and understanding (Scherer, 1997). Questions allow us a glimpse into the user's mind in the attempt to help them realize what it is that they do not know or what is required to fill in the gap in their knowledge state. Questions thus can be viewed as an external index of internal affective and cognitive processes (Graesser, Person & Huber, 1992). Furthermore, questions can be thought of as representations of our information needs. Olson, Duffy, & Mack (1985) state questioning is a "device for seeking new knowledge that is to be related to an existing knowledge structure" (p. 219). Whether expressed orally or communicated within an information retrieval system, questions are the vehicle for expressing our information needs.

Pairing our own cognitive technology, questions, with the powerful technology of digital learning environments could have profound results. The problem lies in how best to represent questions and to coordinate them with representations of the resources in the digital library learning environment. Questions and/or question state will then be defined for the purpose of the study as cognitive representations of our information gap or information need expressed either orally or communicated within an information retrieval system.

Cognitive information retrieval: The central theory behind cognitive information retrieval is that retrieval of information should actually be thought of as an information seeking process which begins with an information need--either expressed or unexpressed--and the underlying processes or internal representations in the mind of the seeker which occur during the process. Also important are the cognitive states such as memory, recall, recognition, perception, and previous knowledge states which influence information retrieval, as well as problem-solving processes that determine usefulness of retrieved documents. According to Ingwersen (1996) the essential "kernel" of the cognitive information retrieval viewpoint is that "both the reception *and* the generation of information are acts of information processing . . . . [which are] dependent on the world model of the actor – whether human or machine" (p. 5). Furthermore, in "human information processing the world model constitutes the individual cognitive space which, consisting of highly dynamic and interchangeable cognitive structures, controls the perception and further processing of external input" (Ingwersen, 1996 p. 5). Our individual cognitive space and structures are "determined by the experiences gained through time in a social and historical context" (Ingwersen, 1996 p. 6). Ingwersen also explains information retrieval systems contain similar cognitive structures which are embedded within the system's algorithms or textual symbolic strings. Information processing within computerized information retrieval systems may then be thought of as the interaction between the seeker's cognitive structures and the system's embedded cognitive structures. When these two states achieve a useful match successful cognitive information retrieval results. Cognitive information retrieval, however, is not as simple as it might seem. There are many human cognitive components, as well as system components, which are involved. For example, there is a large host of cognitive structures--or transformed representations such as abstracts, indexes, full text, rules of indexing, algorithms, etc.--which are generated by a variety of human cognitive sources such as systems designers and producers, indexers, authors of text and images, indexing rule constructors, intermediary mechanism designers, and users all within a domain-related social context (Ingwersen, 1996 p. 8). The following model by O'Connor and Pai, (as presented in O'Connor, 1996 p. 11) represents the cognitive structures and sources present within cognitive information retrieval.

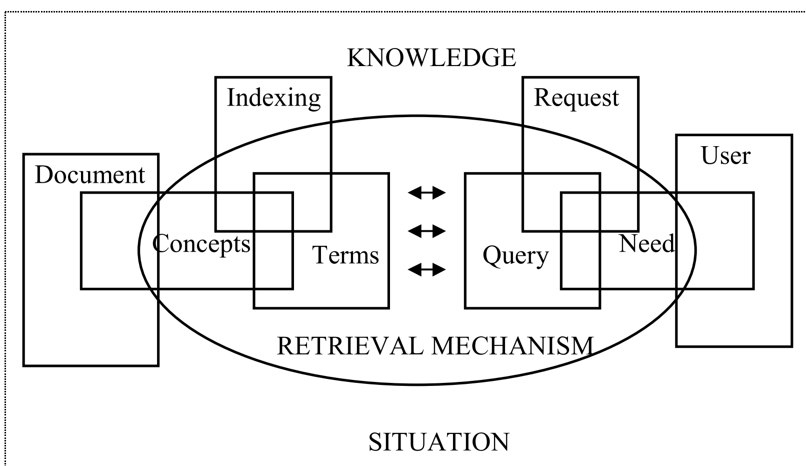


Figure 2: O'Connor and Pai model

The cognitive structures of these cognitive sources are most likely to differ. Well known examples of this are the issues of interindexer inconsistency or inconsistent assessment of relevance. These cognitive structures may also be influenced by the current cognitive state of the seeker, the cognitive task, be it work related or of personal interest (Ingwersen, 1996); by an individual's cognitive abilities such as logical reasoning; and by individual cognitive styles of learning, thinking, and problem-solving (Allen, 1991). Children's cognitive and/or developmental states can affect their use of digital library learning environments. They are subject to all the adult issues and they have different cognitive characteristics than the system designers.

There are also system and design issues that influence the cognitive information retrieval process. In essence, a theory of cognitive information retrieval must somehow integrate the disparate cognitive structures of the various cognitive sources, whether human or machine, to achieve a "cognitive overlap." Where all of these cognitive sources and structures overlap represents effective cognitive information retrieval (see Figure 3 Ingwersen's 1996 overlap model).

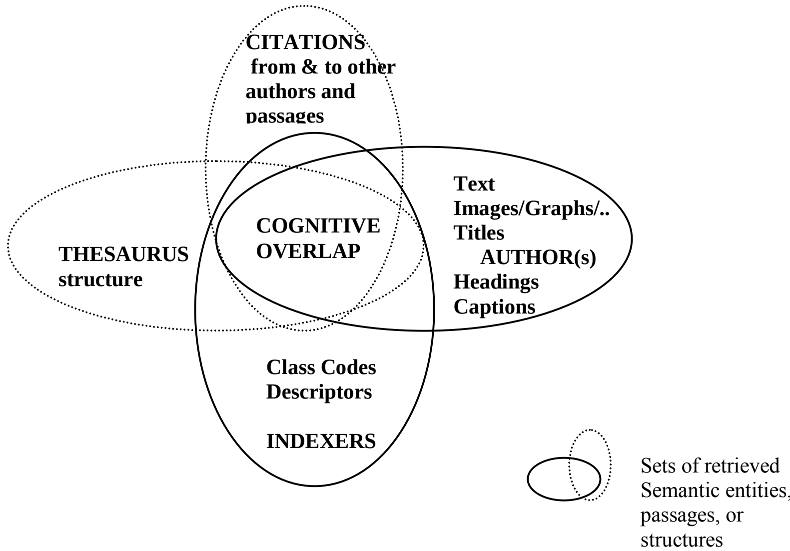


Figure 3. Ingwersen's Overlap Model

Information seeking: Information seeking has been defined as "a problem-solving activity dependent on: the seeker, the problem, the search system and the outcomes. It involves interaction among a number of sub-processes: problem recognition, problem definition, search system selection, query conceptualization, query formulation, query execution, examination of results, and iteration of some or all of these sub-processes if the results suggest this is necessary." (Large et al., 1999). Information seeking is described as a fundamental human activity, in which we engage every day to make sense of the world around us. Marchionini (1995) states that humans purposely engage in information seeking in order to change their state of knowledge. Marchionini prefers the term *seeking* to *retrieval* because it implies a process of acquiring knowledge rather than finding something we already knew or retrieving a known object. Information seeking can be affected by characteristics of searchers such as age, knowledge, and experience (Bellardo, 1985; Fenichel, 1981; and Fidel, 1987). Information seeking in a digital environment adds additional challenges to the seeker, who must also learn the intricacies of the information system being used. For younger users, complicated system language and query input syntax, as well as spelling and vocabulary choice, must be learned. For the purpose of this study, information seeking is defined as the cognitive *process* in which a person engages to search for, evaluate, and make use of information to satisfy some information need or gap in their knowledge state.

ARTEMIS Digital Library Learning Environment (ADL)

The environment in which the study was conducted is the ARTEMIS Digital Library Learning Environment (ADL). It is a web-based interface that provides access to the University of Michigan's Middle Years Digital Library (MYDL) collection. The Artemis Digital Library project is funded by the National Science Foundation, grant number REC9980055. There are three components to the ADL as explained below.

The Library

At the time of the pilot study, MYDL contained approximately 2,000 age-appropriate science-and-space technology-related electronic resources, which had been identified by the four digital librarians of the University of Michigan. The collection was comprised of web pages, online public access catalog (OPAC) records from Ann Arbor Public Library, and science and technology CD Rom encyclopedia and other reference materials. The Artemis Digital Library is a database of abstracts representing the resources within the collection. The database resides on the server at the University of Michigan. The purpose of the collection is to provide quality, age-appropriate resources in order to enable inquiry-based learning among sixth-grade science students. A model of the ADL is presented in Figure 4.

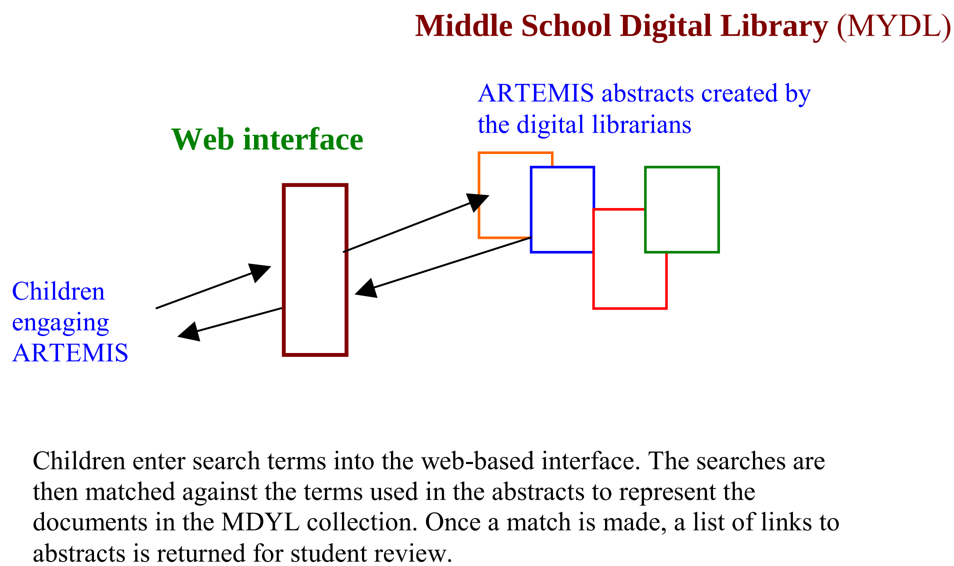


Figure 4: Model of ARTEMIS Engagment

The Interface

The ADL is accessed through a web interface. It is organized by topic and it includes a practice area, called the Scavenger Hunt, in which students learn how to use their mouse, how to navigate a web page, how to click on hyperlinks, and how to use the search features of the interface. The interface includes scaffolding features that enable the students to engage in information seeking activities, by reducing cognitive load. The scaffolds, or components of the system that are designed to provide additional support to the students, are task-related and serve particular functions within the system. The scaffolding provides the students with a persistent workspace in which to save past searches, to save and organize search results and web page links, and to post informational messages and their self-generated driving questions (DQs). Driving Questions are defined by ARTEMIS designers as “a central question that is the focus of a student’s investigation. Driving Questions are large in scope, interesting to students, have real-world meaning or application, and can be broken down into smaller questions. Driving questions require a significant amount of research to investigate and have multiple answers" (Hi-Ce, 1998). See Appendix A. for an example of a worksheet designed to teach the students to formulate DQs. The scaffolds within the system are outlined in the following section. Refer also to Figure 5 for the interface screen and to Figure 6 for the Scavenger Hunt opening screen.

Figure 5: Artemis Interface

Figure 6: Artemis Scavenger Hunt

The Scaffolds

The Artemis interface includes a variety of scaffolding features, which are supported by print-based student and teacher manuals. Designers of the interface identified five dimensions of learner needs: domain knowledge, technology knowledge, strategic knowledge, metacognitive knowledge, and motivation. Artemis was designed with features that would assist each student to develop each of these five areas. The interface also provides support functions, or scaffolds, that enable students to realize they are engaging in a process while seeking information, generate new search terms, stay on task, and evaluate information (Wallace, 1998). Designers of the ADL used the social constructivist Learner Centered Design (LCD) model when creating the system. Within this model, scaffolds serve the function of providing support or structure in order for learners to mindfully engage with the system, to make the task of learning the processes of the new work they are engaged in easier. Science inquiry is new to the users of the

ADL, so the scaffolds provide them with the needed support or structure. Quintana (2001) outlines criteria software features must meet in order to be considered scaffolds. The criteria include: make new tasks doable so learners can *actively engage* in new tasks; *situate* learner in a more authentic representation of the work; make aspects of the work practice and work *community* visible and understandable so the learners can engage in *discourse* with others; and lastly, fade from the software or lessen or discontinue use when the learner no longer needs the support they provide. The above criteria are important when attempting to understand the interaction or activity patterns seen in the data. The LCD model, as explained by Quintana (2001), emphasizes that scaffold use should fade or become less frequently used once the learner no longer needs the scaffold's support. The particular scaffolding features of the system are outlined below and their purposes discussed.

The twelve features can be grouped into four basic functions: Informational, Organization, Maintenance, and Search features.

Informational Features

Post DQ: Constructing questions that are of interest to the student group is an essential part of inquiry-based learning. Students post their group's driving question(s) to the communication space of the interface. Each group, which is comprised of two or three randomly selected students, must formulate a driving question that will guide their research. In this manner they tell the other students what topic they are interested in and can elicit feedback and potential helpful resources from the other students.

Post Cool Site: Students post web sites or other resources that they found that were of interest to them. They may also post or send sites to other students that may be of use to the other students. This feature allows the students to communicate findings, collaborate with each other, and to explain why they thought the site was of interest to them.

Comment: Students can post comments on each other's DQs, post useful sites they found, ask questions of each other or propose suggestions. This feature allows students to collaborate, generate new ideas, share resources, and learn through interaction. Posting DQs, Cool Sites, and Comments helps students stay on task, gives them a chance to evaluate each other's learning and gives students additional feedback from students and teachers, which aids in metacognitive learning.

Organization Features

Create Folder, Rename Folder, and Create Item: Student groups each have their own workspace for use in Artemis. Within their workspace, students create folders to hold their DQs, past searches, and resources. These features help students organize their workspaces by teaching them to categorize and group like items together into folders. This activity enables students to see relationships between topics and subtopics and to place resources for their topics into appropriate folders for use. Generally the first activity a student group will engage in will be to Create Folders for their topic and subtopics. Having a persistent workspace helps keep students on task, and allows them a place to start each time, as opposed to starting over each search session. This reinforces the idea that information seeking is often an ongoing iterative process.

Maintenance Features

Copy and delete: Students can copy and save successful sites to their workspace. They can also delete those no longer needed. If a resource is deemed unneeded or not useful, the students can delete it from their folders, in this way keeping only the resources they will use for their final projects. These features again reduce cognitive load and allow the student groups to evaluate resources and to organize their thoughts and resources.

Search Features

Search and Show Document: These features allow students to search the MYDL collection of age appropriate, topic specific resources. The search engine is available via the web-based HTML interface, which then queries an Oracle database that houses the abstracts and keywords representing the resources. See page 11 for discussion on abstracts and an example of an abstract created by the digital librarians. The students can access and evaluate the resource abstracts and can link directly to the resources via hyperlinks. Searching and accessing only age-appropriate and topic-specific resources enables the students to search with less time and less confusion than searching the entire Internet and World Wide Web. Librarians at the University of Michigan have evaluated and selected only those resources that they feel will be useful and appropriate for the students. Analysis of the search histories indicates that students on the average only view one in five of the actual resources found. Searching a self-contained collection of resources also reduces cognitive load, as students have fewer resources to sift through and evaluate.

Saving and Viewing: These features allow students to save and view past searches, which aids in the construction of Strategy knowledge as students learn that information seeking is a process that may or may not have only one correct answer. Saving searches reduces cognitive load of remembering search terms and strategies they used previously, thereby enabling students to evaluate why past searches did or did not work and how to revise them for more effective retrieval.

The Present System

Since the time the pilot study was conducted, ARTEMIS has been redesigned to make the interface easier to use by the students and also to collect data more efficiently. Several of the previous features are no longer available in the new system. The search capability of the new system has been expanded to include phrase searching, noncase sensitive searching, as well as a link to an online thesaurus students may use to find alternative terms to use in their queries. An ontology called WordNet has also been added as a source of alternative words the system can use when searching. A significant change to the system is that the Artemis collection now contains only web resources. The CD-ROM and OPAC resources are no longer available. The collection now has an estimated 4,800 web pages available for searching and use by the students. See Figure 7 for a look at the new ADL interface.

Figure 7: New Artemis Interface

Pilot Study

Outline of study

A pilot study conducted between September 2000 and December 2000 established the necessity for the research and addressed the methods. The preliminary study assessed the processes in which the students engaged and the features the students used within the ADL environment. Search log data from a total of 73 groups comprised of sixth grade science students was reviewed. The data analyzed consisted of transaction log entries collected by an Oracle database with a web interface. Individual Artemis sessions were collected between 11/98 and 3/2000. A total of 7585 sessions were logged during the time period. For the purposes of the pilot study, sessions conducted by students from two different classes were identified and analyzed. This yielded a subset of data numbering

548 sessions, approximately seven percent of the total sessions. The subset was offered the possibility of detecting obstacles to retrieval, and exploring findings from earlier ARTEMIS research conducted by Hoffman (1999).

Research Questions and Methodology:

Several research questions guided the investigation. The first question was concerned with measuring system effectiveness. The remaining questions deal with student use of Artemis.

Research Questions

1. How long on average does a student engage the system? Average session time? Average search time?
2. How are the students using Artemis? Which activities or features of Artemis are they using? What search techniques are they attempting?
3. Do teacher context variables affect student use of Artemis?
4. Do the results in the Artemis system indicate similar strategies and activities on student web searching?
5. Is it possible for the pilot study to use quantitative data to assess student search satisfaction with Artemis and digital library information retrieval systems?

The first three questions were the focus of the preliminary research, while the last two questions support possible lines of inquiry for future research.

Definitions

Several definitions are offered here to aid explanation of both Artemis and the research.

Session is defined as the time the student used the Artemis digital library beginning with the login and ending with the logout or termination of the session. During a session a student may engage in one or more of the twelve activities that Artemis provides.

Activities include the informational, maintenance, organization or search feature provided within Artemis. Each category of activity is presented below.

1. informational:
    a) Post DQ (driving question)
    b) Post Cool Site
    c) Comment

2. maintenance:
    a) Save search
    b) Copy
    c) Delete item

3. organization:
    a) Create folder
    b) Rename folder
    c) Create item

4. search:
    a) Search
    b) View search
    c) Show document

Search is defined as an activity in which the student's query (search) terms are matched with terms used to describe the document in the document itself or in a database record describing the document.

Student group may include two or three students identified by a user ID number.

Context variables are defined as factors used to explain differences in teaching style and student outcomes.

Digital library is defined as a self-contained collection of resources made available electronically. The resources in the Artemis collection are in a variety of formats such as web pages, CD-ROM encyclopedic programs, and online public access catalog (OPAC) records.

Methodology

The goal of the preliminary research was to examine each student group's Artemis search history from the first time they logged on until the last logout time. Using the web transaction logs, it was possible to get a fairly complete picture of a student group's search history. The data from each student group's search sessions were then plotted into an Excel spreadsheet. Listed below are the quantitative variables measured.

1. Minimum/maximum/mean/median/mode sessions per group/ per class.
2. Minimum/maximum/mean/median/mode activities/session per group/ per class.
3. Minimum/maximum/mean/median/mode searches per group/ per class.
4. Types of searches attempted.
5. Reason searches were revised.
    a. Common mistakes such as misspelling, case sensitive searches, and typographical errors.
    b. Inappropriate search techniques used. Artemis does not allow phrase, Boolean OR and NOT searching, or natural-language searching.

c. Search terms not representative of collection indexing language or the controlled vocabulary used to provide descriptions of the documents within the collection.

Measures from the individual classes were then compared to determine any differences between class search histories within Artemis.

Description of Results

While the data used in this study were quantitative in nature, the results yielded both quantitative and qualitative results. The quantitative results are presented below and the qualitative results will be addressed later in the Further Directions for the Research section.

Question #1: How long on average does a student engage the system? Average session time? Average search time?

Figures for student use of Artemis are presented in Table 1. Figures for average search time were not adequately captured in this dataset and will be a focus for future investigation. Few significant differences are noted with the exception of Class 2 engaging Artemis in longer sessions (80 minutes for Class 2, compared to 47 minutes for Class 1) and viewing more items than Class 1 (73 maximum viewed items for Class 2, compared to 28 maximum viewed items for Class 1).

Table 1. Comparison of Student Groups in Two Classes

Activity	**Class 1**	**Class 2**
Min Time/Session	5 seconds	11 seconds
Max Time/Session	47 minutes	80 minutes
Mean Time/Session	13 minutes	13 minutes
Max SharedSites/Session	2	2
Mean SharedSites/Session	0.236	0.351
Max Comments/Session	2	2
Mean Comments/Session	0.578	0.027
Max SavedItems/Session	11	11
Mean SavedItems/Session	1.47	1.67
Max ViewedItems/Session	28	73
Mean ViewedItems/Session	5.236	15.621

Question #2: How are the students using Artemis? Which activities or features of Artemis are they using? What search techniques are they attempting?

Two sixth grade classes groups' search histories were examined. Class 1 included 38 student groups, while Class 2 included 35 student groups, for a total of 73 student groups. Each student group was comprised of two or three students randomly chosen by the teacher. The entire search sessions for each group were individually plotted and analyzed.

Figures were calculated for individual groups and for class totals. The figures were then compared for the two classes. The results are presented in Table 2. Postings figures include the informational activities, as outlined above. Activities figures include the organization and maintenance activities as outlined above. In every variable group, Class 2 showed a significant increase of Maintenance, Organization, and Search activities used. Only minor differences were noticed in the Informational activities between the two groups. Class 2 made use of more of the features than did Class 1. Class 2 also conducted more searches, and saved and viewed more documents.

Table 2. Activities for Each Class

Activity Totals	Class 1	Class 2
Minimum Sessions	1	1
Maximum Sessions	13	20
Mean Sessions	5.763	9.4
Median Sessions	5	10
Mode Sessions	4	14
Total Sessions	219	329
Minimum Searches	0	0
Maximum Searches	15	13
Mean Searches	2.96	10.224
Median Searches	13	30.5
Mode Searches	10	39
Total Searches	112.5	359.96
Minimum Postings	0	0
Maximum Postings	4	1
Mean Postings	1.668	2.22
Median Postings	1.5	2
Mode Postings	1	1
Total Postings	63.4	77.668
Minimum Activities	0	0
Maximum Activities	21	37

Mean Activities	16.342	38.15
Median Activities	2	10
Mode Activities	1	6
Total Activities	621	1,347.342

Types of Searches Attempted

An interesting result was obtained by tracking the types of search techniques attempted by the students. As mentioned in section I above, Artemis does not allow phrase, Boolean OR and NOT searching, or natural language searching. Overall, the majority of the 1,968 searches examined for the two classes used either Broad Topics alone or Broad Topics in conjunction with one search term. However, there was evidence that phrase searching was attempted, as well as Boolean OR and natural language searching. Exact figures for each technique are not yet calculated.

Reason Searches were Revised

An attempt was made to determine what caused a student's search to be revised. Common mistakes that were evident included: 1) spelling words incorrectly. The student often did not even realize that he/she had misspelled the word, which was evidenced by the student's subsequent use of the same misspelled word. 2) case-sensitive searching. The Artemis system is case sensitive so that inappropriate capitalization of letters will result in failed searches. Again it was noted that often the student did not understand why his/her search failed. Several searches were submitted in all capital letters. 3) typographical errors. Many typographical errors were noted in student search terms. Students again did not seem to notice or to understand why these searches failed.

Student Search Terms not Representative of Collection Indexing Language

The Artemis search engine submits a student query to a database of short abstracts written by the librarians to describe the content and use of each document in the collection. Refer to Figure 1 for an example of an ADL abstract. If the language or terms used by the librarians who create the abstracts does not exactly match the terms entered by the students, then the search will fail. See Table 3 for examples of driving questions and search terms used by the students. Further investigation of the existence of or use of an appropriate age and topic specific controlled vocabulary of terms is warranted. Compiling a list of the actual search terms used by the students may be an appropriate way to construct such a controlled vocabulary. This is an area that requires further research.

Table 3. Driving Questions and Search Terms

Student's Driving Questions	Search terms used
Is it true that Saturn is dense enough to float in a glass of water?	Saturn
What would happen if you put a furby in space?	space travel
How does love and hate come from a muscle?	muscle
Could roaches live on Mars?	Mars and life
what causes an echo?	echo
Do electro magnetic feilds cause tumors in plants? (child's exact wording and spelling)	electro magnetic fields, plants
Why do I have my mom's nose and my dad's hair?	genetics

Question #3. Do teacher context variables affect student use of Artemis?

The comparative data presented in Tables 1 and 2 do indicate that there are distinct differences between the two classes' use of the Artemis system. Teacher context variables or teacher training and knowledge of the features of Artemis, searching, and their individual comfort levels with using digital library technology may account for some of the differences shown. Student training and continued direction by the teachers may also be evident. Because the data analyzed were quantitative in nature, it is necessary to collect qualitative data through direct or video observation in order to explore these differences.

Additional Findings

Patterns and process trends within the data were discernable from the plotted search histories of the two classes. For example, Class 1 followed a very specific sequence of use when engaging Artemis. With very few exceptions, the entire class' groups began their first session by posting a Driving Question to the system. They next engaged in a sequence of maintenance and organizational activities such as creating a folder, and renaming the folder, thereby setting up their workspaces. Class 1 did not perform searches until much later in their session sequence. Class 2 did not follow any particular sequence of use. Approximately one half of the Class 2 groups began with posting a Driving Question, then proceeded to search the collection. Class 2 groups focused more on searching, viewing documents, and revising their searches, while Class 1 focused more on organizing their workspaces, with searching activities being almost of secondary importance. At this time possible reasons for the difference in the two classes groups' activities cannot be attributed to any specific variables. Again, more qualitative data is needed to ascertain if teacher context variables or some other variable account for the differing search sequence characteristics.

Another interesting aspect of this research that is also difficult to attribute to any particular variable, is the stopping behaviors of the students. Based on the search patterns in the data, the students appeared to become frustrated and edited their existing query or ended their information seeking prematurely. For example, as noted above, the groups often revised their queries due to 1) spelling errors, 2) capitalization or case sensitive entry, and 3) typographical errors. In most instances the student did not simply edit the error, but began a new query using a new term, which might imply that they did not realize they had entered the query incorrectly, or they did not know enough about the system and how searching is accomplished to know why their query failed. It might benefit the students to have further training in typical errors to look for when entering queries. These findings also suggest possible system redesign.

The students also had propensity to attempt to use non-Artemis specific search techniques. Boolean OR searching, phrase searching, and natural language search techniques were noticed in the search data. It might, therefore, benefit the students to be trained further in search strategies supported in Artemis and those that are not. It would also be interesting to know why the students chose these other forms of searching. One possible reason could be that many students are accustomed to searching with search engines on the Internet and World Wide Web and they may believe that Artemis can be searched in the same manner. This possibility warrants further investigation.

Limitations of the Study

It is important to note that the data collected for the pilot study presented a somewhat incomplete picture of the student groups' search sessions. While it enabled us to track the students' use of the system, it did not provide data on uniform resource locators (URLs) for the results returned during retrieval. It also did not include data on which documents the children decided to retrieve. Only if a search or a site was saved was it possible to view this information about the search session. Future data collection will need to include a mechanism for capturing this data so that the full search session is available for analysis.

Directions for Further Study

Several of the preliminary findings suggest that obstacles to representation and seeking warrant further examination. Obstacles such as misspelling terms, typographic errors, and unsupported search strategies need to be explored further. Representation of their information needs and matching these representations with the system's representations of the documents is another area that should be explored further. Using student's driving questions and their search terms as a means to glimpse into their minds and as a representation of their knowledge state may lead to findings about representation, questions, and information needs.

Because the students are interacting with a digital library learning system that includes additional scaffolding features to reduce cognitive load, their information seeking process is augmented by the system's features and may, therefore, differ from the traditional information seeking models. For this reason, further investigation of the scaffolding features and how they affect information seeking and retrieval is warranted.

The patterns indicate more than just the search process. They may also provide a metric for student success. For example, is the student who uses more searching features more successful at content understanding and system use? New questions of pattern indicators are summarized in Table 4 below.

Table 4. New Questions to Consider

Pattern	**Outcome**	**Question**
More searches	Successful search	More facility with system?
	Unsuccessful search	Frustration with system?
Fewer searches	Successful search	More facility with system?
	Unsuccessful search	Ended search without finding answer?
More Postings	Successful	Collaborate more?
Fewer Postings	Unsuccessful	Less collaboration?
More saved items and folders created	Successful	Understood features and their functions?
More saved items and folders created	Unsuccessful	Did not understand features and their functions?
Fewer saved items	Successful	Understood features and their

and folders created		functions?
Fewer saved items and folders created	Unsuccessful	Did not understand features and their functions?

Determining successful patterns can help us train teachers and students in better use of the system or processes to follow, as well as identify other problem areas to address. Students using online resources often cannot balance learning to use the system while seeking information to increase their content understanding in a particular subject (Hirsh, 1997). Research based on these preliminary data to determine the nature of obstacles of representations on the part of the system, and on the part of the student, may lead to more effective information seeking behaviors.

Chapter 2 presents a review of the previous studies that have been considered prior to the design of the study and the analysis of the data. These studies have investigated how children interact with systems while seeking information. Previous ADL studies are also presented and linked to the present study.

CHAPTER 2

PRIOR RESEARCH AND THEORETICAL BASIS

Literature Review

We are informed by the work of others who have investigated similar ideas and issues. This research builds on research from information science, in the hopes of developing a clearer understanding of the use of digital resources by children. It is, at once, an information seeking study and a study of representation; for representations form the seeking environment, provide the tools for moving within the environment, and set the paths for seeking. The review outlines the literature on previous ARTEMIS studies, children's information seeking, representation schemes, and cognitive information retrieval and models.

Previous ARTEMIS Studies

Earlier studies conducted at the University of Michigan add to our knowledge of how children engage with information systems such as digital libraries and the WWW. These studies also contributed to the design of the previous version of the ARTEMIS system as well as the current system. They also provided the framework for the first long term study of the ARTEMIS system completed by Hoffman in 1999.

Hoffman (1997); Hoffman, Kupperman, & Wallace (1997) explored how children navigate the Web and what devices or scaffolds could be incorporated into design of a web-based interface that would facilitate content understanding and information seeking. They determined that Web browsers provide a limited environment that is difficult for children to use for the iterative process of information seeking. Their studies determined that providing a collection of online, age-appropriate resources would reduce the cognitive load felt by the children when searching online.

Lyons, Hoffman, Krajcik, & Soloway (1997) explored how sixth and ninth grade science students used the WWW to research student generated questions. They also further evaluated problems encountered by the children when working online. Students in this study used the Web to search for information to complete a classroom assignment. They used resources gathered to produce a small report or booklet explaining the question they investigated. The results of the study added to the design of both the online and the printed resources of the ARTEMIS system. For example, findings indicated that 1) students had difficulty remaining on task, 2) students did not plan out their searching sessions, 3) students had difficulty choosing terms to search with because of a lack of domain knowledge, 4) students preferred keyword searching to browsing, 5) students viewed the WWW as a fact answering system rather than one that helped them to explore multiple facets of a question, and 6) students rarely evaluated what they found on the WWW and possessed little knowledge on how to assess the reliability of the sites they found.

Wallace, Kupperman, Krajcik & Soloway (1997); Wallace and Kupperman (1997) present one of the first information seeking studies of students using the interim design of ARTEMIS. They examined sixth-grade science students as they used the new digital library to find information to complete an inquiry-based science unit. The research focused on how the students used the technologies to navigate, their information seeking strategies, and how they evaluated and used the resources they found. Their findings indicated that while students use the digital library with little difficulty, they used it naively. Information seeking was a complex and difficult process for the students, as they struggled with learning search strategies, navigation, broadening or narrowing searches, and choosing terms. Students also thought of the activities they engaged in as "homework" that had to be completed, rather than the exploration teachers hoped for. Further, understanding the content was made more difficult because the students were also required to learn and use the new technology of the digital library.

Hoffman (1999) built upon each of these previous studies. His study used the then current design of ARTEMIS. His study included a larger sample, took place over an entire school year, and incorporated the collection of documents within the digital library. The earlier studies used the resources of the WWW. The study also included the print and online scaffolds developed as a result of the findings of the earlier studies. The Hoffman study investigated how the student groups engaged with the system and the effect the use of a digital library learning environment had upon the content understandings of the students. He gathered observational (process video) data, as well as survey and interview data from a group including two teachers and eight pairs of students-four pairs from each class. Teaching styles, integration of technology within the curriculum units, and teacher instruction and use of the ARTEMIS system were examined. Student use of ARTEMIS, as well as extent of student content understanding, were evaluated by examining student artifacts. The Hoffman study concluded that learners often do not engage with online environments at adequate levels, and this inadequacy can affect student's content understanding. Furthermore, pedagogical practices of teachers can affect how students engage with the system. Clear differences were evident in the teaching styles and scaffolds provided by the two teachers observed in the study. Hoffman concludes that "it is clear students can benefit from access to on-line resources for inquiry-based activities when objectives are selectively identified, curricula are carefully developed, resources are thoughtfully chosen, support and scaffolding are extensively provided, and sense-making is purposely mediated by the teacher." (p. 207). As teacher pedagogical styles or context variables do appear to have an effect on how children engage with a system, it is important that we understand more about how children learn to formulate the representations of their information needs (their questions and search terms) and also how they learn to use the system. We are provided a glimpse of this process by examining their representations and searching patterns.

Children's Information Seeking Studies

Research on information seeking has been conducted with many different user populations and in many different contexts. The research literature is replete with studies on adult information seeking with a smaller body of research focusing on children and young adults. Walter (1994) points out, children have different information needs than adults do, thus it is appropriate to set aside the large body of research about adult information seeking and focus on children and young adult's information seeking. Research using print resources and online public access catalogs (OPACs), electronic multimedia resources, and the Internet and World Wide Web will each be addressed separately.

Print Resources and OPACs

Prior to Kuhlthau's (1988) study on high school senior's information seeking, there was little research in this area. From this study and further studies by Kuhlthau (1991; Kuhlthau 1997) a model of the information search process of adolescents has been developed. Her research revealed information seeking as a "dynamic, complex process in which basic constructs are construed and reconstrued as progress is made through levels of information need--from ambiguous to specific" (p. 241). Edmonds, Moore, and Balcom (1990) studied fourth, sixth, and eight graders use of both manual card catalogs and an OPAC. The research focused on determining what developmental skills were needed to use the catalogs, as well as obstacles encountered and system preference. It was found that younger students lacked some of the basic skills, such as alphabetizing words, that were necessary to make effective use of the catalogs. It was also determined the sequential nature of the information seeking process was difficult for the children to master and presented too many opportunities for errors to occur. An interesting finding was that 68 percent of the children preferred using the manual card catalog to the 16 percent who preferred the OPAC.

Soloman (1993) presents research conducted in an elementary school library using an OPAC. He focuses on representation issues and the problems or obstacles encountered during information retrieval. He concluded children use simple, concrete terms for retrieval; children encounter obstacles such as choosing terms to search with, understanding of multiple or synonymous forms of words, and misdirected searches due to spelling errors; and children may not possess the developmental knowledge to understand components of information systems and information seeking such as syntax rules, evaluating query results, and revising queries. Soloman also reports that children exhibit three distinct intentions when using information systems: 1) locating materials or goal directed behavior of satisfying a personal need; 2) fact retrieval or the belief that the system could provide them with a factual answer; and 3) need to explore or play with the technology to discover its use and entertainment value. Lastly, Soloman discusses representation issues such as indexing practices used in the system, children's generation of search terms, and the relative problems of matching query terms to document representations.

Large, Beheshti, & Moukdad (1994) assessed sixth grade student's retrieval techniques using both print and CD-ROM versions of Compton's Encyclopedia. The students' searches were timed, the retrieval strategies compared, and the search terms were recorded and analyzed. The research determined that students were able to retrieve information from either form of resource, but they exhibited more ease in using the electronic resources and were more willing to explore the interface than the students using print resources.

Borgman, Hirsch, Walter & Gallagher (1995) and subsequent studies by the authors, report research conducted on a Dewey decimal-based hierarchical browsing system, the Science Library Catalog, implemented in two school libraries and a public library setting. Their research spans a three year period and includes four separate iterations of the catalog's design and use by elementary school children. Earlier tests of the Science Library Catalog revealed that children encounter problems with typing and keyboarding, spelling, choosing search terms, alphabetizing, and Boolean logic. These findings were used in the redesign of the catalog to allow children to browse for subjects through a hierarchical structure instead of typing in search terms. The study also explored whether age, gender, or experience affected system use. The authors tested the children with predefined questions and compared their searching against a traditional keyword OPAC. The authors concluded that there was virtually no difference in search success between the two catalogs, with variations being attributed to children's difficulties in choosing terms to search with or faults within the experimental design such as more difficult questions for the keyword systems as opposed the browsing system. The children were equally satisfied with either OPAC or the Science Library Catalog.

Hirsh (1997) further reports on their study but focuses instead on the task specific differences in information seeking strategies used by children as well as the effect domain knowledge may have on their seeking activities. Results indicated that children were more successful with simple-browsing tasks than complex-browsing tasks. Gender also impacted success with boys performing more successfully on technology related topics than on science topics. Results indicated that domain knowledge influenced search success on all types of tasks. Children with high domain knowledge performed better on all tasks than children with low domain knowledge. Hirsh does, however, caution that variations may also be attributed to student's attitudes and motivation for learning about science. Results also indicated that having more than one search method or strategy available made searching more successful.

Electronic Resources

Marchionini and Teague, 1987; Marchionini (1989) conducted an exploratory study of elementary school students (primarily fourth and sixth graders) searching a full-text electronic encyclopedia on CD-ROM. They determined that older searchers were more successful and took less time than younger searchers. They also determined that searchers used a highly interactive search strategy and experienced little trouble using Boolean search functions.

Large, Behesti, & Breuleux (1998) studied three multimedia CD-ROM programs in two sixth-grade classrooms. They report the students experienced few problems using the programs, but constructing effective search strategies was difficult and the students preferred to browse the programs instead.

Hirsh (1999) reports research on fifth-grade students using a variety of electronic resources including an on-line catalog, an electronic encyclopedia, and electronic magazine index, and the WWW. She examined both children's relevance criteria and search strategies and reported that children spent a large portion of their time to finding pictures. They also did not spend much time evaluating the authority of the resources.

Internet and World Wide Web

Dresang (1999) notes that more research is needed on how youths seek information on the Internet. Earlier research examined information seeking behavior, relevance criteria, resources used, and task related behaviors. Schacter, Chung, & Dorr (1998) report their study on fifth and sixth grade students task related information seeking on the Internet. The students were given two tasks: one well defined, and one illdefined. Results indicated that children preferred to browse and employed few systematic search strategies.

Watson (1998) asked for students' perceptions of using technology and the WWW. Student respondents indicated that high domain and system knowledge was necessary to use these resources effectively. Browsing was often preferred but for certain types of questions, it was ineffective. Her study further notes the importance of understanding student behavior and perspectives of using the technology for bibliographic and searching instruction.

Researchers are responding to Dresang's call for more research in this area. Fidel, et al. (1999) examined high school students' use of the WWW for homework assignments. The study reports that the students enjoyed using this resource, but often had difficulty choosing search terms, using search engines, and evaluating resources.

Bilal (2000a); Bilal (2000b) report her research on students' use of the web search engine Yahooligans! She examined the cognitive, physical, and affective behaviors of the students for fact-based tasks and research oriented tasks. She found that children encountered similar obstacles to those uncovered by researchers of OPAC and other electronic resources. Spelling, typographic errors, and choosing search terms were difficult for the children. She also discovered that children had trouble navigating within web sites and often became disoriented and where unsure of how to return to their starting page.

Obstacles to information seeking by children are numerous and frequently discussed in the reviewed literature. There appears to be significant overlap between the literature and the pilot study for this research. This suggests the validity of exploring ways in which we can augment the system's representation scheme by using terms within the students' driving questions and search statements to describe the content of the documents within the system.

Representation Studies

Representation schemes for children or young adults is an area that has been largely neglected in the research. Development of indexing languages and controlled vocabularies or subject headings lists have focused on the user as either a homogenous group with no age specified, or for a specific discipline or domain. Few efforts to develop controlled vocabularies for children exist. Jansson (as reported by Lundgren, 1998) developed a special thesaurus for children consisting of about 800 simple, concrete words within 21 areas of interest. Librarians using the list to represent documents are encouraged to add to the list as they feel necessary. This list has been distributed to libraries in Sweden where, Lundgren reports, it has met with much approval.

Bucherschatz, a prototype hyperlink catalog for children developed in Germany, uses descriptions written specifically for children. The descriptions are designed to peak the children's interest and to be whimsical, fun, and thrilling. This catalog uses three primary access points into the collection for the children: books for fun and leisure; books on children's life and problems; and other non-fiction books. Each of these three access points is represented by a picture: an octopus, a seagull, and a pirate, respectively. The catalog uses a treasure hunt theme as the metaphor of children searching for information or “treasure”, hence the graphics used for the main access points (Kulper, Schultz, & Will, 1997).

Miralpeix (1994) reports on a small exploratory study of two libraries catalogs (one public and one small private). These systems were used by children between the ages of 10 and 13. Both catalogs used the AACR2 rules to create catalog records, but the catalogers simplified the vocabulary used to describe the children’s books.

Pejtersen developed a Danish interface for children’s materials called the Book House. This interface is icon based and includes very in depth indexing. The bibliographic records include additional information such as level of reading difficulty, time period, geographic location, and the emotional effect the book may produce. It is important to note that these elements are not traditionally found in bibliographic records. Lundgren and Dalgaard augmented the system with an online form that allowed the children to write book descriptions themselves. The book descriptions were primarily written by 11 and 12 year olds and contained very emotive descriptions of the books as well as evaluative comments of the books (Lundgren, 1998).

Systems designed specifically for children’s information needs have been developed. Borgman, et al. (1997) as described above, developed a keyboard independent system that enabled children to browse subject content of a science collection. The focus of their studies examined children’s engagment with the system and the effectiveness of the iterative design of three different interfaces. They did, however, make use of a standard controlled vocabulary to represent the documents within the collection.

Brown (2000) reviews OPACs designed especially for children, as well as reviewing the literature of child-centered design and the challenges children face when using an OPAC. She reviews OPACS such as the Kids Catalog, developed by Sandlian, Busey, and Doerr in 1990; Kids Online, developed by the vendor Innovative Interfaces; DRA Kids, developed by Data Research Associates; Book House, developed by Pejtersen and later tested and augmented by Lundgren and Dalgaard (as mentioned above); as well as other web-based interfaces such as Follett Software Company, Book Systems, Inc., Inspire Kids, and Just for Kids.

Extent of Match Studies

A further body of research that is related to this study is that of the transaction log analysis (TLA) extent of match studies that have been conducted using transaction logs of OPACs. TLA has been used for two primary purposes: as a form of system monitoring or tracking system features and use, and as a means to unobtrusively observe human behavior or use patterns. TLA evolved out of the need to monitor performance of computerized information retrieval systems. One of the earliest, if not the first, research projects conducted using TLA was by Meister and Sullivan (1967). One of the most important advantages to TLA is that it collects data gathered from real users in naturalistic situations, though TLA has also been used in a few experimental contexts. TLA has been used to track commands or features used, response time, session lengths, transitions from one command to the next, failure analysis, use of specific search functions, which access points are used, which advance search features are used, user's printing, downloading, quitting and persistence, and extent of match between search terms and controlled vocabularies or indexing languages used to represent the documents within the collection. See Peters (1993) for a comprehensive review of TLA studies.

Extent of match studies analyze user's search terms and the controlled vocabulary used in the system to represent the documents. Extent of match studies were conducted by Taylor (1984) matching search terms to the authority files at Northwestern University; Markey (1984) matching search terms with Library of Congress Subject Headings (LCSH) she determined that 18 percent of the terms matched LCSH headings and five percent matched cross references; Carlyle (1989) also matching LCSH headings and found 47 percent match between user terms and LCSH headings; Doyen and Wheeler (1989) found that only 21 percent of users terms matched the controlled vocabulary of the system; while Lester (1989) found approximately 40 percent matched LCSH headings; Drabenstott and Vizine-Goetz (1990) revealed a 25 percent match between user terms entered into three OPACs and LCSH headings. Drabenstott and Vizine-Goetz further revealed that the citations retrieved through exact or close matches were not entirely satisfactory or relevant.

A more recent study by Jansen, Spink & Saracevic (2000) used TLA to track users' search sessions on the Excite search engine on the WWW. They used TLA to track sessions, number of terms per search, as well as determined the most frequently used search terms. Out of a total of 113,793 terms from all queries examined (51,473 queries), they discovered a total of 21,862 unique terms. They further constructed a table of the most frequently used search terms, those that had been used at least 100 times and compiled a list of 74 terms that appear a total of 20,698 times as search terms.

Greenberg (2001) used TLA extent of matching to determine effectiveness of automatic query expansion by examining how closely user search terms matched the semantic relationships found within the system's thesaurus. Greenberg compared the search terms of forty-two M.B.A. students using the ABI/Inform database system. Each search term or search string was compared to the broader term, narrower term, and the related term within the thesaurus. Greenberg further used these comparisons to measure the effect on retrieval if queries were expanded using the broader, narrow, or related term for the original search term.

Each of the above studies adds to our knowledge of how inadequate controlled vocabularies and subject headings lists are as devices to represent documents within a collection. This study will look specifically at children's search terms and how they match the controlled vocabulary being used within the digital library. The study will also examine how closely terms in the children's driving questions match the terms present in their search queries. It will then take this research to the next level, using children's search terms to create a list of terms to represent the documents. From this point, searches will be conducted and recall and precision difference calculated to determine the effect user-defined representations have on retrieval.

Theoretical Basis

Theories provide researchers with models to frame our understanding of an idea, problem, or an area of study. This study will bring together theories and models of information science, in an attempt to frame children's representations of their question state as expressions or vehicles of expressing some cognitive gap in their knowledge state, together with the role of external representations in their seeking. This section will discuss 1) cognitive information retrieval and 2) representation theory.

Cognitive Information Retrieval Theory

Two specific theories will serve to illustrate the areas of cognitive information representation and retrieval theory. Each will be explained briefly below.

Cooper and O'Connor Model

The Cooper model of cognitive information retrieval attempts to illustrate how cognitive variables and the very definition of "representation" can affect both representations of a user's information need as well as representations about the documents in an information retrieval system. It takes into account the user's often undefined and unformulated question or question state as well as the variables that influence the expression of the question within the system. O'Connor has adapted this model to include other semantic and cognitive variables that interfere with information representation and retrieval as illustrated in Figure 8.

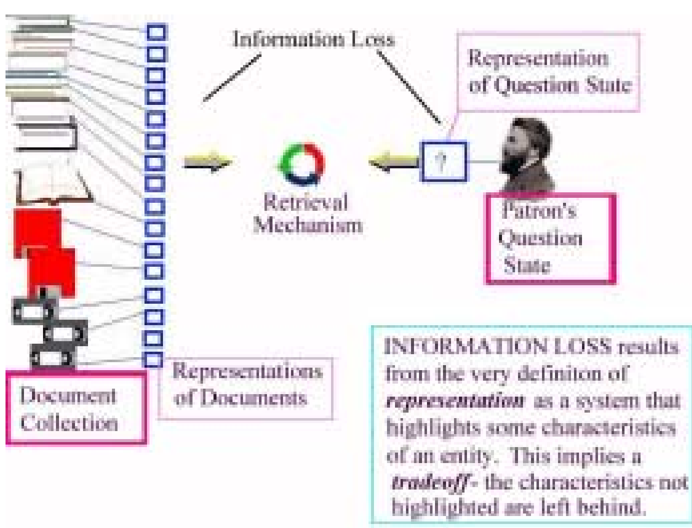


Figure 8: Cooper and O'Connor Model

The Cooper and O’Connor model further emphasizes the inherent problem of representation, that of information loss when a document’s essence or subject is distilled down into two to three subject terms that is often the practice today. Information loss implies a tradeoff, or a loss of potentially important representations of the essence. We can further adapt the Cooper and O’Connor model to illustrate how the user’s developmental and cognitive state, domain and system knowledge, and indexer’s knowledge of the user’s intended purpose(s) for the objects, or the idea of functional representation, can affect representation and retrieval. The indexer’s understanding of the domain knowledge of the users is also important to our understanding of how representations of objects within a topic-specific, age-appropriate collection should be created. Figure 9 presents an adaptation of the model for the purpose of the study.

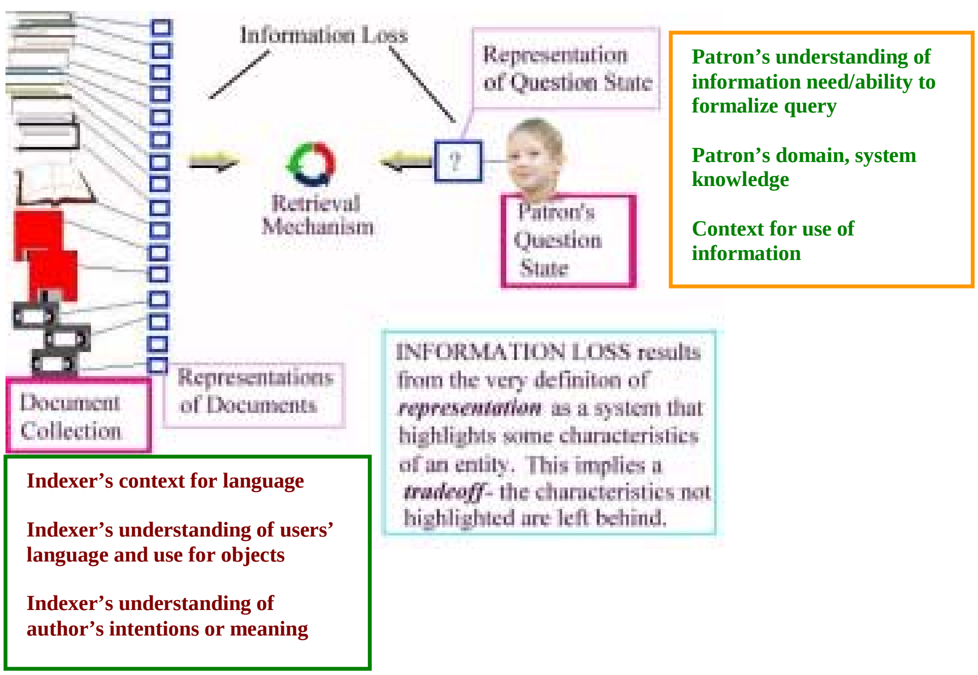


Figure 9: Adaptation of the Cooper and O’Connor Model

Blair's Language Community Model

Blair (1990) provides a further context for examining the problems of information representation and retrieval that are particularly relevant to this study. He posits that the language that we use to represent both our information needs or questions and to index documents is learned in a social context or community. Blair explains the theory of "language games", as first developed by the early twenty century philosopher Ludwig Wittengenstein and the process in which we learn language and meaning. We do not acquire language purely by learning the word and its definition, but instead learn its use and appropriateness within the context of our "forms of life" or everyday experiences. Furthermore, we have to possess some prior understanding of the form of life or the language game context we are engaged in before the words can have meaning. The ARTEMIS system is an information retrieval system that provides access to science related resources for a specific user community, 6^{th} grade students. An important aspect of learning science related concepts is the terminology of the discipline and the uses and functions of the information. Students are engaging with the system within the social context of an elementary science classroom. Each classroom of learners engages in "language games" as they go through their daily "forms of life" or experiences. Direct influences on their learning are their teachers, the documents they engage with, and the information system they interact with. Learning and knowing the appropriate "language" or terminology to use within these contexts is vital to their success both in information retrieval and content understanding.

Chapter 3 presents the design of the study and the methods for exploring the multifaceted dimensions of children's interaction with the ADL. Methods for exploring representation and retrieval issues, as well as the obstacles encountered during information retrieval are also explained.

CHAPTER 3

RESEARCH DESIGN AND METHODS

Representation as a fundamental scaffold within a children's information seeking system presents a multi-faceted challenge and opportunity for discovery. A variety of research methods and instruments is required to explore the children's construction of representations of their own question states and their use of system representations of available resources. The research questions that guided the study are outlined and discussed briefly below. Models of children's use of ARTEMIS during preliminary observations are linked to a foundational model of representation within information seeking systems to guide construction of hypotheses and appropriate methods and instruments. These models informed the evaluation and analysis of the children's use of the system, their questions, the correlation between search terms and resource representations, and how information retrieval would be affected by injecting or incorporating the children's own terms into the system for representing resources.

Research Questions

Listed below are the research questions that helped to frame the study.

1. What strategies are the children engaging in to find information to answer their driving questions? Which scaffolds are being used and do they enable successful searching?

2. How does children's language relate to the language used to represent the documents in the system? Can children's language be used to represent documents within the collection? Will using student language within representations affect retrieval?

Discussion of Research Questions

Question One

Question number one provided the basis for the exploration of the children's use of the system. Log data gathered and analyzed in an earlier pilot study indicated some commonalities of use and representation obstacles while using the ADL that warrant further research. Children have information needs and information seeking strategies that differ from those of adults. Learning more about how they engage with systems can illuminate the cognitive processes in which they engage while seeking information. Modeling how children engage with systems may help in understanding paths, processes, and obstacles they may encounter. Modeling children's engagement may also provide insight into the need for additional scaffolds.

Earlier research conducted on the ADL by Wallace, Kupperman, Krajcik & Soloway (1997); Wallace and Kupperman (1997) and Hoffman (1999), as reported in Ch. 2 pp. 36-38, indicated that while students use the digital library with little difficulty, they used it naively. Information seeking was a complex and difficult process for the students, as they struggled with learning search strategies, navigation, broadening or narrowing searches, and choosing terms. Students also thought of the activities they engaged in as "homework" that had to be completed, rather than the exploration teachers hoped for.

Further, understanding the content was made more difficult because the students were also required to learn and use the new technology of the digital library.

This research explored question number one by first examining the scaffolds used by the children, and secondly, extracting from the data any activity patterns the children engaged in while using the ADL. The scaffolds and activity patterns are explained briefly below.

Scaffolds

The ADL includes specific system features designed to serve as scaffolds, or supports to the children. The scaffolds make Artemis more than a search engine. The three multipurpose scaffolds being examined in this research project include:

1. *Persistent workspace* (space to save driving questions (DQ), comments, bookmarks, past searches, and past results). Having children create DQs before searching helps them to plan their strategies before engaging with the search function. This scaffold also enables students to create a personal workspace that they return to each time they log on, thereby eliminating the need to start over each time, and thus reinforcing the idea that information seeking is an iterative process, not just a search for the one right answer.
2. *Website abstracts* (age and topic specific descriptions of the resources). Age-appropriate representations of the resources enable better information seeking and retrieval of resources because the representations of the documents are created with the children's and the curriculum's goals in mind.

3. *Collaborative space* (area to share Cool Sites, and to view and comment on other's Cool Sites and DQs). The collaborative space allows the children to communicate their findings to each other, make comments on each other's DQs, and to share resources they encountered that they think may be useful to other students.

For a complete list of all of the activities collected by the transaction data logs, please refer to Appendix B.

Activity Patterns

Preliminary examination of the search session histories of the groups reveals that children exhibit some commonalities in patterns of activity as they engage with the system. The following log file models illustrate the complete search session histories of two student groups.

Group One

1. Create DQ folder
2. Open DQ folder
3. Create DQ folder
4. Open DQ folder
5. View Shared DQs
6. View Shared Cool Site
7. Open DQ folder
8. Create DQ folder (7 and 8 repeated 2 times)
9. View Shared Cool Sites
10. End of session
11. View Shared DQs
12. Search
13. View Abstracts (7 total)
14. Search (3 times)
15. View Abstracts (3 total)
16. Save site in DQ
17. View Abstracts (7 total)
18. End of session
19. Open DQ (2 times)
20. End of session
21. Open DQ
22. View saved Abstract
23. Open DQ
24. View Saved Abstract
25. Viewed Shared DQ
26. End of session
27. Open DQ (2 times)
28. View saved Abstract (2 total)
29. End of capture

Group two

1. View Shared DQs
2. Search
3. View Abstract (3 total)
4. Viewed website
5. Shared site
6. Viewed Abstract
7. Viewed website
8. Viewed Abstract
9. Shared site
10. Viewed Shared DQs
11. End of session
12. Search
13. Viewed Abstract
14. Viewed Shared DQs
15. Viewed Abstract
16. Viewed website
17. End of session
18. Search
19. View Abstract
20. View website (9 total)
21. View Abstract
22. End of session
23. This same pattern was repeated for the next 14 search iterations
24. Open Past Search folder
25. View Results from Previous Search
26. View Abstracts
27. View website
28. End of session
29. Search
30. View Abstract
31. View website
32. End of capture

The following models illustrate the two groups' interaction with the system as demonstrated in the log file data shown above. The models are presented to illustrate the complexity of the interactions and also to demonstrate the commonalities seen between the two groups use of the ADL.

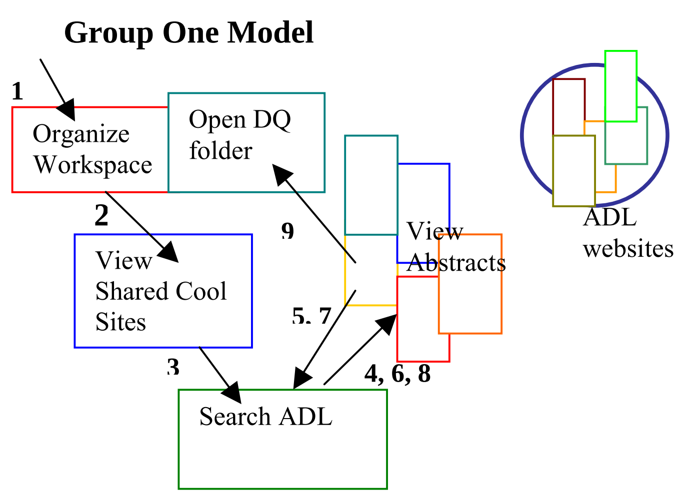


Figure 10: Model of Group One's Interaction with the ADL

The Group One Model illustrates what appears to be the initial stage or interaction the student group had with the system. This may be seen as a familiarization and exploration stage or activity pattern. The group began by setting up their persistent workspace and viewing Cool sites that had been shared by other students. They next began their first searches and viewed abstracts retrieved as a result of the search. This group, however, did not at this point in their system use, retrieve or view any of the websites they were pointed to by the abstracts. Further tracking of this student group may see them begin to retrieve and view the websites indicated in the abstracts.

Group Two Model

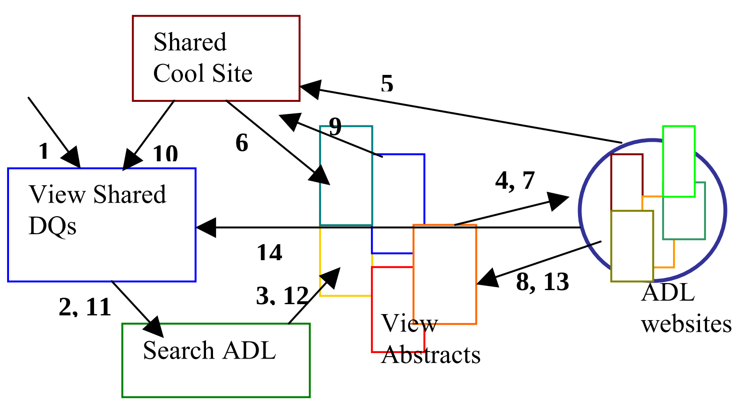


Figure 11: Group Two's Interaction with the ADL

The Group Two Model illustrates a more advanced use of the system. Group two began in a fashion similar to that of group one, but they extended their use to include retrieving and viewing the websites pointed to by the abstracts. They also began to make use of more advanced scaffolds such as Past Search and Past Results folders.

As is evident from the above models, student groups interact with the system in different patterns. These two models may illustrate typical student engagement, or there may be yet other models to uncover. Both models do, however, demonstrate the uses of the scaffolds on the ADL and how they may affect the search process.

From the log files and the above models, it is now possible to extract some common activity patterns exhibited by the children's use of the ADL. These patterns have been grouped into four categories as illustrated in Table 5.

Table 5. Activity Patterns, Scaffolds, and Use

Activity Pattern	Activities in Sequence	Scaffold Use
Exploration (3 activities)	View Shared DQs View Shared Cool Sites Conduct First Initial Search	Each scaffold used between 1-2 times in beginning stages. Not used during intermediate sessions, but used 1-2 times at end of sessions.
Workspace Setup (4 activities)	Create DQ folders Create Past Searches folder Create Past Results folder Post DQ	Each scaffold used and repeated 1 to 3 times at the beginning of sessions. Not used during intermediate sessions, but may be used 1-2 times at end of sessions.
Beginning Search (4 activities)	Conduct Search View Abstract(s) View Website(s) Revise Search	Each repeated 3-4 times in an iterative sequence through majority of sessions.
Extended Search (4 activities)	Open Past Search folder(s) View Results from Past Searches View Abstract(s) View Website(s)	Each of these advanced scaffolds are used 1 to 2 times as the student learns more about the system features and the iterative nature of the search process.

After initial exploration and familiarization with the system, children begin using more of the scaffolds. As they become more proficient in their use, they begin using some of the more advanced scaffolds, such as the Past Search and Past Results folders. Further exploration may reveal new activity patterns. The exploration may also reveal that children move through different stages from beginning system user (novice) to more advanced or proficient user. Table 6 below illustrates this potential process.

Table 6. Stages of Interaction

Stage	Activity Pattern	Activities in Sequence	Scaffold Use
Teacher modeling (rule-based learning)	Exploration (3 activities)	View Shared DQs View Shared Cool Sites Conduct First Initial Search	Each scaffold used between 1-2 times in beginning stages. Not used during intermediate sessions, but used 1-2 times at end of sessions.
Exploration and familiarization (Novice/beginner)	Workspace Setup (4 activities)	Create DQ folders Create Past Searches folder Create Past Results folder Post DQ	Each scaffold used and repeated 1 to 3 times at the beginning of sessions. Not used during intermediate sessions, but may be used 1-2 times at end of sessions.
Advanced Beginner/Competence	Beginning Search (4 activities)	Conduct Search View Abstract(s) View Website(s) Revise Search	Each repeated 3-4 times in an iterative sequence through majority of sessions.
Competence/Proficiency	Extended Search (4 activities)	Open Past Search folder(s) View Results from Past Searches View Abstract(s) View Website(s)	Each of these advanced scaffolds are used 1 to 2 times as the student learns more about the system features and the iterative nature of the search process.

An additional means for interpreting efficiency of the scaffolds was to link teacher evaluation of student outcomes to the scaffolds used and the search history patterns. This correlation between teacher data and interaction patterns provided external corroboration of the analysis. Further design and training issues may also be uncovered. The methods used to collect and analyze the data used to examine question one are discussed later in this chapter.

Question Two

Question number two is the basis for examining one of the most problematic aspects of information seeking and retrieval, choosing the most useful words with which to search for the needed documents. Choosing words with which to search is an obstacle that is reported in much of the literature on information seeking. This problem was also evidenced in the pilot study conducted earlier, as well as the earlier ADL studies conducted by Wallace, Kupperman, Krajcik & Soloway (1997); Wallace and Kupperman (1997) and Hoffman (1999).

One of the most robust models of information seeking, as described by Maron, Cooper and Robertson (1982), illustrates the engagement of the user with a system as an intersection of four events: 1) a set of documents which may be relevant to the user; 2) a user searching for the document(s); 3) index terms used to represent the documents; and 4) query terms used to represent the user's information need. Significant potential problems often arise on both sides of the system interface. On the user side, the user must somehow formulate his/her information need into words and then translate their words into system syntax or query language, while at the same time attempting to predict the terms used by the system to represent the documents. On the system side, it may be possible to predict that document X may be useful to the user, but then the system must somehow predict how the user will ask for the document. If these two predictions do not match, then the document is not retrievable by the user.

Most information retrieval systems have a feedback component that allows for the iterative process of revising and resubmitting search words for another chance at matching the system words. Artemis from the beginning has provided user side scaffolds of representation, the DQ folders and spaces where users can save DQs and past searches. These representations are valuable resources to both the user and the system. With the stored search words and driving questions it is possible to feed back into the system user side descriptors for documents, thereby increasing functionality and successful retrieval and reducing search time and search space. The system's agent, the cataloger, can use these valuable resources as a means to predict possible uses and patron description. If system agents use only system side terms, usually derived from a predetermined list of controlled vocabulary terms, it makes the task harder for user prediction. This study proposes a means in which information scientists can make use of this valuable resource provided within the ADL scaffolds to enhance representations of the documents within the collection.

Preliminary analysis of the data illustrates how the ADL scaffold of Driving Questions and the Persistent Workspace can be used to determine the potential mismatch between user side search terms and system side index terms. It further illustrates the powerful advantage ADL provides researchers to use the user side search terms to augment the descriptions or abstracts that describe the documents. Table 7 illustrates actual student search terms and system terms used in the document abstracts.

Table 7. Student Terms Versus System Terms

Student's Driving Questions	**Search terms used**	**System terms used**
Is it true that Saturn is dense enough to float in a glass of water?	Saturn	saturn, planets "saturn", planets "moons" "saturn", density "volume", astronomy "stars & galaxies" "planets" "universe

What would happen if you put a furby in space?	space travel	astronomy "space exploration" "space shuttle" "orbits", astronomy "astronauts"
Do electro magnetic fields cause tumors in plants	electro magnetic fields, plants	biology, electromagnetism
Why do I have my mom's nose and my dad's hair?	genetics	genes, genetics "deoxyribonucleic acid -- dna"

Hypotheses

Based upon the preliminary analysis of the data as presented above, the following hypotheses were developed.

1. Children will engage with the system in various ways but common activity patterns will also be evident.

    Sub-hypothesis A: Teacher effect will have an impact on a user group's use of scaffolds.

    Sub-hypothesis B: Teacher effect will have an impact on a user group's time spent using the ADL.

    Sub-hypothesis C: Teacher effect will have an impact on a user group's mode of engagement (activity patterns).

2. Children will use a variety of scaffolds while engaging with the system.

    Sub-hypothesis A: Differing levels of achievement will correlate with various patterns of interaction with the ADL.

3. There will be significant differences in strategies used by the user groups.

    Sub-hypothesis A: Overall, children will use more single word searches than phrase, Boolean, or natural-language searches.

Sub-hypothesis B: In a significant number of searches, obstructions due to mechanical errors, such as spelling errors, typographic errors, repeated searches, and repeating searches sequentially, will occur.

2. In a significant number of searches, system representations of documents will not match children's representations of their information needs.

3. Representing documents using the language of student's searches will have a positive effect on retrieval.

Research Methods

Quantitative and qualitative methods were used in order to provide validity and reliability to the study. The sample for each method used for the study varied also. Each method employed a different means of collecting data and used different samples of either children's transaction log data, teacher-assigned evaluation of student outcomes, cataloging instruments used to create the representations, or system librarians as respondents.

Method One: Revisit Pilot Study with New Data Collection

System Use by Children

This method explored research question number one. The sample included the search transaction logs from the ADL system. The sample size was expanded from the sample of the pilot study. Data were collected from January 2001 through May 2001 for *all* schools using ADL. The only limitations placed on the sample were to deselect all user groups that did not engage the system at least a minimum of three times. This rate was determined by analysis of the data gathered from the pilot study to be the mode total that student user groups engaged the system. It was also determined that system use of less than three times was not an adequate indication of student interaction. The deselection was also necessary due to the practice of the teachers to reassign students to other groups if they felt the original group members were not working well together.

1. This method attempted to discover how children used the redesigned system, scaffolds they used, obstacles they encountered, and search terms and search strategies they employed. The entire search history for each student group was plotted into a SPSS spreadsheet and an Access database and search strategies, search terms, and information seeking patterns were examined. Refer to Appendixes B. and C. for the coding schemes and semantics.

2. A failure analysis of the obstacles to information retrieval such as spelling errors, typographical errors, and incorrect search techniques, uncovered in the pilot study was also conducted. Refer to Appendix D. for the coding scheme.

Method Two: Teacher-Assigned Evaluation of Student Outcomes

This method explored research question number one. Teachers whose classes used the ADL within the current data collection period were asked to provide their evaluation of each student group's final project or student-generated artifact. Teacher evaluations were then correlated to use and time figures, as well as the activity patterns evidenced by the individual groups. This correlation provided yet another picture of what constitutes successful use of the ADL, as well as indicated design and training issues.

Method Three: Representation Issues

This method explored research question number two. Quantitative and qualitative analysis of search transaction logs for all of the students using ADL examined two central issues: 1) student issues of representation and 2) system representation scheme issues. Each is outlined below.

Student Issues

Three methods were used to explore student question state and representation issues. This method set used as a sample the search transaction logs from the ADL system. The sample size was expanded from the sample of the pilot study. Data were collected from January 2001 through May 2001 for *all* schools using ADL. Data were again limited to those user groups that engaged the system a total of three or more times.

1. Student search terms and the terms within the student groups' driving questions were correlated. Following the extent of match rules as outlined in App. E., as defined by Greenberg (2001) and later modified by Abbas, all user groups' driving questions were compared to the search terms students use to retrieve the documents within the collection.

2. Student search terms were compared to the keywords and abstracts used in ADL. Following the extent of match rules as outlined in App. E., as defined by Greenberg (2001) and later modified by Abbas, all user groups' search terms were compared to the keyword list used by ADL catalogers to index the documents. The search terms were also compared to the abstracts created by the catalogers to represent the documents within the collection.

3. Term frequency rates were determined and a list of frequently used terms were compiled. The term frequency was to be determined within the context of the individual searches and

subject topics for which the search is being conducted. However, because there was a low degree of variation in the search terms, contextual frequency was not required. See Ch. 4 for further discussion of the change to the original intention of this method.

System Representation Scheme Issues

Two methods were used to determine current representation schemes used within ADL. The sample for this method was the ADL librarians and the instruments they currently use to create representations within ADL.

1. Content analysis of the cataloging instruments and instructions to catalogers was to be conducted on each of the cataloging instruments used by the ADL catalogers to determine the process or rules they follow to create the representations. The focus was to be on: how to choose terms; source of terms; number of terms to assign; and other pertinent rules and instructions to the catalogers. See Ch. 4 for discussion of the change to the original intention of this method.

2. ADL librarians were interviewed to clarify any ambiguous interpretation of the content analysis of the cataloging instruments. Follow up e-mail interviews were conducted to clarify ambiguities that emerged as a result of the above content analysis. These interviews were further analyzed using the focus of the above content analysis of the cataloging instruments. See Ch. 4 for discussion of the change to the original intention of this method.

Method Four: Retrieval Issues

This method explored research question number two. The possibility of using student search terms as representations or index terms for the resources within the collection was addressed. The sample was the list compiled in Method Three: Student Issues, student searches from Method One above, and the database of ADL resource descriptions.

1. The list of compiled student terms was used to index the ADL resources. Terms were added to the existing keywords field in the abstracts for each appropriate resource.

2. A sample of the students' original searches were re-entered by the researcher and the results analyzed to determine effects of student-generated keywords on retrieval. The sample size was limited to one tenth of all searches conducted within the data collection period. In order to determine the effect on retrieval, results from the sample queries were compared to the student's original queries and retrieved documents were examined for their utility in answering the students' original driving questions.

Limitations and Expectations

It is important to note that the study used various methods to measure the different aspects of student interaction with the ADL. The primary data collection method, TLA, helps to protect reliability and validity, but also limits the variables that can be measured. TLA provides an unobtrusive method to collect data from real users within real contexts. TLA also guards against problems of researcher subjectivity of selective data gathering or bias.

Reliability, or consistency in the results, was also be assisted by TLA. This method was used to collect data on all of the users. The data set was then filtered and segmented into each class and then further divided into grade groups. The same method of analyzing the data was used, with consistent results.

Other methods employed the use of coding schemes. Each of these schemes were reviewed for ambiguity and clarity. A small sample of the data was also coded by two other coders to test the reliability of the results, ambiguity of the rules, and guard against researcher's expectations and bias.

The primary limitation of TLA is that it does limit the variables that can be measured. It also doesn't provide a means to clarify ambiguous data. The earlier pilot study tested this method and allow the researcher to tweak the system used to collect the data, therby lessening ambiguous or incomplete data. It also illustrated the variables that could not be measured. Many factors such as affective and cognitive variables cannot be measured with this method. Reasons for children's revision of search terms are not always apparent within TLA data and assumptions can only be formulated based on the patterns displayed in the data. TLA does, however, remain a very powerful tool for collecting real data within a natural context that is not changed as a result of imposing a researcher into the environment.

CHAPTER 4

PRESENTATION OF THE DATA

Analysis of the data presented findings that help us to understand children's interaction with the system and representation issues specific to children's information seeking and retrieval. This chapter presents the analyses and links each of the findings to the pertinent research question and the hypothesis.

Question One Revisited

1. What strategies are the children engaging in to find information to answer their driving questions? Which scaffolds are being used and do they enable successful searching?

Question One provided the basis for the exploration of the children's use of the system. Log data gathered and analyzed in the pilot study indicated some commonalities of use and representation obstacles while using the ADL that warranted further research. Question One also suggests the need to understand more about the system features used by the students in the course of their information seeking and retrieval activities.

Based on observations and analysis of preliminary data the following hypotheses were formulated to explore Research Question One:

1. Children will engage with the system in various ways but common activity patterns will also be evident.

   Sub-hypothesis A: Teacher effect will have an impact on a user group's use of scaffolds.

Sub-hypothesis B: Teacher effect will have an impact on a user group's time spent using the ADL.

Sub-hypothesis C: Teacher effect will have an impact on a user group's mode of engagement (activity patterns).

2. Children will use a variety of scaffolds while engaging with the system.

   Sub-hypothesis A: Differing levels of achievement will correlate with various patterns of interaction with the ADL.

The sample for this portion of the analysis included all student user groups that had logged in to the system a minimum of three times, as recorded in the system transaction logs (Action ID #1). Refer to Appendix B for a complete list of Action ID codes and associated activities. Any student group that did not log in a minimum of three times was deselected from the sample analyzed. Teacher activities, as well as those conducted for the testing and design of the system, were also not included in the sample. Initially, the transaction logs included a total of 64, 941 recorded activities. After the data were filtered to remove the user groups that did not meet the selection criteria and to remove the teacher and system designer activities, the total number of activities analyzed was 52, 781 (81.3% of the initial total). At the beginning of the data analysis, a total of 1,648 user groups had used the Artemis Digital Library (ADL). After the user group data were filtered for those that did not meet the criteria for selection, a total of 754 user groups were used in the analysis (45.8% of the initial total). See Table 8 for a breakdown of user groups by class and region.

Table 8. User Groups Divided by Class and Region

Class Number	Region	User groups (filtered)
C1	Michigan	28
C2	Michigan	1
C3	Michigan	43
C4	South Carolina	3
C5	Michigan	3
C6	Michigan	11
C7	Michigan	47
C8	Michigan	16
C9	Michigan	33
C10	New York	23
C11	Arizona	1
C12	Michigan	37
C13	Michigan	3
C14	Michigan	42
C15	Michigan	1
C16	Michigan	1
C17	Michigan	84
C18	Michigan	18
C19	Michigan	51
C20	Michigan	5
C21	New York	24
C22	California	1
C23	Michigan	47
C24	Michigan	1

C25	Michigan	1
C26	Maryland	62
C27	Michigan	97
C28	Michigan	14
C29	Michigan	1
C30	Michigan	18
C31	Michigan	1
C32	Michigan	36
		754 total user groups

In order to ascertain a robust picture of student interaction, the data were analyzed in several ways, as defined in Chapter 3: Method 1 (pp.62 -63) and as described below. This approach focused on the use of various mechanisms to uncover interactions with the system including: scaffolds used, teacher influence on students' use of the ADL, change over time in usage patterns, and the relationship between grades achieved by the user groups for the ADL assignment and scaffold and system activity use. Each analysis is presented and discussed separately below.

Scaffolds Used

The ADL was designed with specific features to reduce cognitive load while the students complete their information seeking and retrieval tasks. These features are called scaffolds. Refer to Ch. 1 (pp. 17-21) for discussion of the available scaffolds. The data for the 754 user groups presents a picture of how the students engage with the system, in particular which activities and scaffolds were used and which were not used or used only minimally. Table 9 illustrates activity and scaffold use by the users.

Table 9. Activity and Scaffold Use by All User Groups

Activity	Frequency	Percent	Cumulative Percent
Viewed website description (abstract)	12465	23.6	23.6
Performing a search	11487	21.8	45.4
View website	4963	9.4	54.8
Login	4917	9.3	64.1
Open DQ Folder	4344	8.2	72.3
Saved MYDL Site in DQ Folder	2284	4.3	76.7
Create DQ from Search Page	1851	3.5	80.2
Logout	1641	3.1	83.3
Saved DQ Notes in DQ Folder	1305	2.5	85.7
Viewed website description of site saved in DQ Folder	1202	2.3	88.0
Viewed Shared DQ	1081	2.0	90.1
Viewed Shared Cool Sites	856	1.6	91.7
Performed Dictionary Query on Term	757	1.4	93.1
Opened Past Searches Window	610	1.2	94.3
Performed Search With Wordnet	446	.8	95.1
Shared MYDL Site with Class	385	.7	95.9
Deleted Bookmark for DQ Folder	373	.7	96.6
Deleted DQ Bin	340	.6	97.2
Opened Dictionary Search Tool	287	.5	97.8
Deleted Past Search	281	.5	98.3
Edited DQ	258	.5	98.8
Opened Results of Previous Search	191	.4	99.1
Login Teacher	154	.3	99.4
Saved Bookmark in DQ Folder. Site not MYDL	116	.2	99.6
Created DQ from View Page	68	.1	99.8
Opened Thesaurus Page	55	.1	99.9
Performed Thesaurus Query on Term	44	.1	100.0
Shared NonMYDL Site with Class	20	.0	100.0
Total	52781	100.0	100.0

Table 9 illustrates the frequency of use of each activity. Five activities account for 67.3% of all activities (with the exception of Action ID numbers 1,2,3 which record log in and log out activities by both students and teachers): 1) View full website description (abstract); 2) Performing a search; 3) Viewing website; 4) Open DQ folder; and 5) Saved site from MYDL in DQ folder. The scaffolds used with the highest frequency include: 1) View full website description (abstract); 2) Open DQ folder; 3) Saved site from MYDL in DQ folder; 4) Creating DQ from Search Page; and 5) Saving Notes in DQ Folder. The use of these scaffolds comprises 42.1% of scaffold activity. The scaffolds used with lowest frequency include: 1) Shared site not from MYDL with class; 2) Performed a thesaurus query on term; 3) Opened the thesaurus page; 4) Create a DQ from the View page; and 5) Saved NonMYDL bookmark in DQ folder. The lower frequency activities comprise a total of 0.5% of all activities. Scaffolds that were designed for collaborative purposes such as Viewing DQs , and Sharing and Viewing Cool Sites, comprise only 4.3% of the total activities. Table 9 illustrates the searching activities used most frequently by the users in the sample: conducting a search, viewing resource abstracts, and viewing websites. Furthermore, collaborative scaffolds mentioned above, as well as the more advanced scaffolds of opening and retrieving past searches and results, using the thesaurus, Wordnet, and dictionary scaffolds are used less frequently or very minimally. Use of the more advanced scaffolds comprises only 4.5% of the total activities. Organizational activities including deleting bookmarks from DQ folders, deleting DQ bins, deleting past searches, and editing DQs comprise only 2.3% of all activities.

Teacher Effect on Scaffold Use

To examine the variable of teacher effect on system and scaffold use, it was necessary to segment the data into individual classes. Method Two (refer to Ch. 3 p. 63) outlines a means of providing external corroboration for the interaction patterns that are present in the data. The ADL teachers were asked to provide grade information for their user groups for the data collection period. Four teachers volunteered to provide the study with this data. There are a total of 122 user groups in Class 1, 127 user groups in Class 2, 51 user groups in Class 3, and 51 user groups in Class 4 for which data were examined. The size of the sample within each class varies as a result of filtering out all user groups that did not log into the system three or more times. The sample size was also affected by the incomplete records correlating user identification with grade information that were kept by the teachers. The results of activity and scaffold use are presented in Table 10.

Table 10. Total Activities Conducted by Each Class

Activity and Action ID Code	Class One	Class Two	Class Three	Class Four
Login (1)	714	555	277	225
Logout (2)	294	345	159	40
Login teacher (3)	0	0	0	0
Open Results of Previous Search (4)	9	24	6	11
Performing a Search (5)	2259	1143	641	604
Performing a Search with Wordnet (6)	31	18	5	27
Open Dictionary Search Tool (7)	83	23	12	21
Performed Dictionary query on term (8)	179	39	11	123
Opened Thesaurus page (9)	17	5	2	2
Performed thesaurus query on term (10)	17	4	1	2
Open Past Searches Window (11)	94	45	20	23
Deleted a Past Search (12)	45	0	4	15
Created DQ from Search Page (13)	322	285	74	141
Created DQ from View Page (14)	0	1	0	3
Deleted a DQ Bin (15)	56	40	10	39
Edited DQ (16)	18	37	27	16

Opened DQ Folder (17)	374	1448	298	86
Saved Notes in DQ Folder (18)	39	721	157	11
Deleted a Bookmark from a DQ Folder (19)	28	193	7	4
Saved Site from the MYDL in DQ Folder (20)	281	1026	45	73
Saved Bookmark in DQ Folder. Site not MYDL (21)	0	7	1	1
Viewed Full Website Description (22)	2192	1612	630	593
Viewed Website (23)	943	513	0	126
Shared Site from the MYDL with class (24)	9	44	0	24
Shared Site NOT from the MYDL with class (25)	0	0	4	0
Viewed Shared Cool Sites (26)	45	24	48	77
Viewed Shared DQ (27)	19	14	98	66
Viewed Website Description of Site Saved in DQ Folder (28)	77	222	60	23
Total Activities	8145	8388	2568	2376

Table 10 illustrates distinct differences between the classes' use of the ADL. For example, Class One conducted the most searches, viewed more abstracts and websites, created more DQs, and used more of the advanced scaffolds, such as performing Wordnet searches, using the dictionary and thesaurus tools, and opening the past searches window. Class Two opened the DQ folders more frequently, saved sites and notes into the DQ folders more (both MYDL and nonMYDL), viewed website descriptions saved in DQ folders more, reviewed results of previous searches, and edited DQs and deleted bookmarks from DQ folders more than the other classes. Class Three shared more nonMYDL sites with the class, and viewed more shared DQs than the other classes. Class Four created more DQs from the View Page, and viewed more shared Cool Sites than the other three classes.

Similarities between the classes' ADL use can be noted. By examining the scaffolds used with the greatest degree of frequency, we see, with some variations, each class used similar scaffolds, but to differing degrees. Table 11 shows the scaffolds that comprise approximately 80% of all scaffold activity by each class.

Table 11. Top 80 Percent of Scaffolds Used by Each Class

Class One	%	Class Two	%	Class Three	%	Class Four	%
Viewed website descriptions (abstracts)	55.7	Viewed website descriptions (abstracts)	27.6	Viewed website descriptions (abstracts)	42.2	Viewed website descriptions (abstracts)	39.0
Opened DQ Folder	9.5	Opened DQ Folder	24.8	Opened DQ Folder	19.9	Created DQ from Search Page	10.2
Created DQ from Search Page	8.18	Saved MYDL site in DQ Folder	17.6	Viewed shared DQs	10.5	Performed Dictionary query on term	8.9
Saved MYDL site in DQ Folder	7.14	Saved Notes in DQ folder	12.4			Opened DQ Folder	6.2
						Viewed saved Cool Sites	5.57
						Saved site in DQ Folder	5.28
						Viewed Saved DQ	4.7

Table 11 shows the top scaffold used by all classes was to view the website descriptions (abstracts). Each class's second highest scaffold used, with the exception of Class Four, was to open the DQ folders.

A further means to explore the variable of teacher effect is to examine the time used to access the ADL by each class. The sample of the four teachers' class data, as described above, will also remain the sample in this analysis. Results of this analysis are presented in Table 12.

Table 12. Time Spent Using ADL Activities/ Divided by Class (total time/second)

Activity and ActionID by Class	Class One	Class Two	Class Three	Class Four
Login (1)	56935	31473	43003	27549
Logout (2)	292189	112078	99515	53132
Login teacher (3)	82719	100808	0	0
Open Results of Previous Search (4)	260	644	155	248
Performing a Search (5)	85724	54417	18960	62587
Performing a Search with Wordnet (6)	951	318	88	1411
Open Dictionary Search Tool (7)	3265	828	181	1268
Performed Dictionary query on term (8)	10766	6924	27895	49658
Opened Thesaurus page (9)	676	141	19	80
Performed thesaurus query on term (10)	965	89	33	49
Open Past Searches Window (11)	7280	8426	588	13614
Deleted a Past Search (12)	132	0	3	24
Created DQ from Search Page (13)	40542	13448	2443	7286
Created DQ from View Page (14)	0	2	0	3
Deleted a DQ Bin (15)	2785	1928	201	5647
Edited DQ (16)	502	653	744	1414
Opened DQ Folder (17)	19798	110163	44642	5458
Saved Notes in DQ Folder (18)	1691	35885	10878	306
Deleted a Bookmark from a DQ Folder (19)	375	512	6	68
Saved Site from the MYDL in DQ Folder (20)	26779	35147	12184	4128
Saved Bookmark in DQ Folder. Site not MYDL (21)	0	27	0	1
Viewed Full Website Description (22)	366001	106720	125115	67188
Viewed Website (23)	241708	85216	0	57396
Shared Site from the MYDL with class (24)	678	789	0	2582
Shared Site NOT from the MYDL with class (25)	0	0	0	0
Viewed Shared Cool Sites (26)	14098	2031	3957	14286
Viewed Shared DQ (27)	1358	1240	11395	5560
Viewed Website Description of Site Saved in DQ Folder (28)	5300	15316	14764	19840
Time used for all activities	1263477	725223	416769	400783

Table 12 illustrates distinct differences between the amount of time each of the classes engaged in ADL activities. For example, Class One spent more time overall using ADL, conducting searches, viewing abstracts and websites, creating DQs, as well as more time using the advanced scaffolds, such as opening the Dictionary search tool and performing searches using the thesaurus tool. Class Two spent more time opening DQ folders, opening the results of past searches, saving sites, notes, and bookmarks into the DQ folders (both MYDL and nonMYDL), and deleting bookmarks from DQ folders. Class Three spent more time viewing shared DQs than the other classes. Class Four spent more time creating DQs from the View Page, viewing shared Cool Sites and saved website descriptions, deleting DQ Bins, editing DQs, opening the past search window, and using the advanced scaffolds of Wordnet and the Dictionary tool.

Analysis of the scaffolds used approximately 80 percent of the total class ADL time can also illustrate the impact teacher effect has on ADL use. The data presented in Table 13 indicates how long each class spent engaging with the specific scaffolds while using the ADL.

Table 13. Time Spent Using Scaffolds a Total of 80 Percent of the Time

Class One	%	Class Two	%	Class Three	%	Class Four	%
View website descriptions (abstracts)	72.5	Open DQ Folder	32.2	View website descriptions (abstracts)	49.0	View website descriptions (abstracts)	33.5
Create DQ from Search Page	8.0	View website descriptions (abstracts)	31.2	Open DQ Folder	17.4	Performed Dictionary query on term	24.8
		Save Notes in DQ Folder	10.5	Performed Dictionary query on term	10.9	View Saved website descriptions (abstracts)	9.9
		Saved Sites in DQ Folder	10.3	View Saved website descriptions (abstracts)	5.7	View Shared Cool Sites	7.1
						Open Past Search window	6.9

It is also important to note that there are fewer similarities in scaffold use when we look at time figures as opposed to use figures. The top scaffold of Viewing website descriptions (abstracts) is only evident in Classes One, Three, and Four. It is of secondary time use by Class Two. Table 13 shows Class One spent the majority of its time (72.5%) viewing website descriptions (abstracts), while the other classes spent less time using this scaffold and more time using other scaffolds such as creating, opening and saving sites in DQ folders. Class Four spent more time using collaborative scaffolds such as viewing shared Cool Sites and/or shared DQs than the other three classes. Classes Three and Four spent more time using the advanced scaffolds of conducting dictionary queries and opening past searches.

Further categorization of the scaffolds by function allows us to see yet another picture of scaffold use by each class. Table 14 presents scaffold use within the functional categories of: 1) Maintenance: Deleting Past Searches, DQ Bins and Bookmarks, and Editing DQs; 2) Organizational: Creating DQs, Opening DQ Folders, Saving Notes in DQ Folders, Saving Sites in DQ Folders, and Saving Bookmarks in DQ Folders; 3) Searching: Viewing Website Descriptions (abstracts); 4) Collaborative: Sharing MYDL and nonMYDL sites with the Class, Viewing Shared Cool Sites, and Viewing Shared DQs; and 5) Advanced: Performing Search with Wordnet, Opening and Performing a search using the Dictionary and Thesaurus tools, Opening Past Searches window, Opening Results of a Previous Search, and Viewing Website Descriptions Saved in DQ Folders. Use and Time figures have both been included in this analysis. The figures in red bold type indicate the class with the highest level of that functional category's scaffold use.

Table 14. Functional Categorization of Total Scaffold Use

Scaffold Category	Class One Use/Time%	Class Two Use/Time%	Class Three Use/Time%	Class Four Use/Time%
Maintenance	3.7 / 0.75	4.6 / 0.9	3.0 / .37	5.3 / 3.57
Organizational	25.8 / 17.6	**59.8 / 57.0**	38.0 / 27.4	22.8 / 8.58
Searching	**57.6 / 72.5**	27.6 / 31.2	**42.0 / 49.0**	**42.9** / 33.6
Collaborative	3.8 / 3.19	1.4 / 1.2	10.0 / 6.0	12.0 / 11.2
Advanced	10.5 / 5.84	6.5 / 9.5	7.8 / 17.0	16.7 / **43.0**

Table 14 illustrates that Classes One and Three spent more use and time in search related scaffolds, Class Two spent more use and time in organizational related scaffolds, and Class Four spent more use in search related scaffolds, but more time using advanced scaffolds.

Examining each class's use of the scaffolds presents a persuasive view of how teacher effect impacted the students' use and time spent engaging with ADL scaffolds. However, it might also be useful to examine the "average" user group's interaction with the system, in order to compare across classes. Table 15 presents a normalized view of a user group's interaction with the ADL. The numbers in bold red type indicate the highest use or time figure for the activity across the classes.

Table 15. Average Student's Interaction with the ADL Divided by Class

Activity and ActionID	Class1 Mean Use/Time in seconds	Class 2 Mean Use/Time in seconds	Class 3 Mean Use/Time in seconds	Class 4 Mean Use/Time in seconds
Login (1)	**5.8** / 84.9	4.37 / 59.8	5.43 / **338.6**	4.41 / 85.8
Logout (2)	2.4 / 3075.6	2.71 / 1205.1	**3.11 / 6634.3**	0.78 / 4830.1
Login teacher (3)	0	0	0	0
Open Results of Previous Search (4)	0.07 / **28.8**	0.18 / 26.8	0.11 / 25.8	**0.21** / 17.7
Performing a Search (5)	**18.5** / 43.1	9.0 / 54.5	12.56 / 30.7	11.84 / **85.9**
Performing a Search with Wordnet (6)	0.25 / 35.2	0.14 / 18.7	0.09 / **44**	**0.52** / 38.1
Open Dictionary Search Tool (7)	**0.68** / 43.5	0.18 / 39.4	0.23 / 20.1	0.41 / **43.7**
Performed Dictionary query on term (8)	1.46 / 74.7	0.30 / 197.8	0.21 / **1743.4**	**2.41** / 379.0
Opened Thesaurus page (9)	**0.13 / 45.0**	0.03/ 35.2	0.03 / 9.5	0.03 / 26.6
Performed thesaurus query on term (10)	**0.13** / 1	0.03 / 29.6	0.01 / **33**	0.03 / 24.5
Open Past Searches Window (11)	0.**77** / 80.8	0.35 / 168.5	0.39 / 28	0.45 / **340.3**
Deleted a Past Search (12)	**0.36 / 2.93**	0	0.07 / 0.75	0.29 / 1.6
Created DQ from Search Page (13)	2.63 / **151.2**	2.24 / 57.4	1.45 / 38.7	**2.76** / 45.8
Created DQ from View Page (14)	0	0	0	0.05 / 1.4
Deleted a DQ Bin (15)	0.45 / 54.6	0.31 / 49.4	0.19 / 22.3	**0.76 / 117.6**
Edited DQ (16)	0.14 / 29.5	0.29 / 21.7	**0.52** / 31	0.31 / **54.3**
Opened DQ Folder (17)	3.06 / 55.3	**11.40** / 84.4	5.84 / **157.7**	1.68 / 40.1
Saved Notes in DQ Folder (18)	0.31 / 49.7	**5.67** / 51.7	3.07 / **74.5**	0.21 / 18
Deleted a Bookmark from a DQ Folder (19)	0.22 / 11.3	**1.51** / 3.06	0.13 / 0.8	0.07 / **5.6**
Saved Site from the MYDL in DQ Folder (20)	2.30 / 110.2	**8.07** / 41.4	0.88 / **297.1**	1.43 / 59.8
Saved Bookmark in DQ Folder. Site not MYDL (21)	0	**0.05 / 3.8**	0.01 / 0	0.01 / 0.2
Viewed Full Website Description (22)	**17.96** / 194.8	12.69 / 72.6	12.35 / **213.5**	11.62 / 92.4
Viewed Website (23)	**7.72 / 325.7**	4.03 / 204.8	0	2.47 / 257.3
Shared Site from the	0.07 / 61.63	0.34 / 24.6	0	**0.47 / 122.9**

MYDL with class (24)				
Shared Site NOT from the MYDL with class (25)	0	0	0	0
Viewed Shared Cool Sites (26)	0.36 / 391.6	0.18 / 81.2	0.94 / **989.2**	**1.50** / 207.0
Viewed Shared DQ (27)	0.15/ 113.1	0.11 / 59.0	**1.92 / 130.9**	1.29 / 85.5
Viewed Website Description of Site Saved in DQ Folder (28)	0.63 / 71.6	**1.74** / 84.6	1.17 / 278.5	0.45 / **734.8**
Average Total Uses and Time in seconds	**66.76** /11083	66.04 / 8336	50.3 /**13444**	46.58 /11788

Table 15 provides us with a picture of the average user within each class. For example, the Class One user conducted more searches than the other classes, yet spent less time using the search feature than Class Four. Table 15 also illustrates that there is a large degree of variability between use and time figures. Use of a feature is not necessarily dependent on time spent using the feature.

Data in the above tables illustrates that there are distinct differences and similarities between each classes' use of the ADL. Class use of the ADL may be attributed to each teacher's objectives for the use of, and the integration of the ADL into classroom activities, as well as the teacher's individual emphasis of the use of particular scaffolds during the information seeking process.

Time spent using the different activities and scaffolds of the ADL may also be attributed to each teacher's objectives for using the ADL. It might be further attributed to system factors, or cognitive factors, especially as related to the collaborative scaffolds of viewing shared DQs and Cool Sites. These are areas that require further study. See Appendix F and G for mode and mean times for each activity.

While the above analysis does indicate strong evidence of teacher effect upon student use of the ADL, it does not indicate how teacher effect can impact the students' mode of engagement with the system or the activity patterns of their interaction. The above analysis also does not indicate the impact teacher effect has on the successful use of the ADL. Findings concerning the activities patterns and the correlation between ADL use and student achievement will be explored below.

Teacher Effect on Activity Patterns

As is evident from the above analyses, the ADL users interact with the system in many diverse ways. In order to continue to build a robust picture of an ADL user, it is necessary to look for common interaction sequences or activity patterns in the transaction logs. Various activity patterns are evident from the examination of the data. Individual classes' data were examined to determine the potential effect teacher instruction had on these activity patterns. Each class's total ADL sessions were divided into thirds and each third examined for patterns. Table 16 synthesizes activity patterns by each stage of interaction. (It should be noted, however, that this analysis presents a qualitative look at the data, rather than a quantitative view. Future work in this area may make use of more quantitative approaches at analyzing the activity patterns.)

Table 16. Revised Table 5: Activity Patterns, Scaffolds, and Use

Activity Pattern Name	Activity Pattern(s)	Activities in Sequence	Scaffold Use
Exploration/Beginning of ADL Activities (three or four activities)	13 1 13 2 13 5 1 (5) (22) 5	Create DQ from Search Page Log out/Log back in Conduct First Initial Search	Scaffold of DQ used in beginning of first session. May be used in middle sessions as well. Not often used in end sessions.

Workspace Setup and Organization/Middle sessions (two or three activities in iterative sequence)	17 18 Repeat 17 and 18 in sequence of two or more iterations, with the addition of 19 at times.	Open DQ Folder Saved DQ Notes in DQ Folder Deleted Bookmark from a DQ Folder	Scaffold of DQ Folders used throughout sessions. Folders used to organize students information seeking activities.
Beginning Search/Middle sessions (two or three activities)	4 22 23	Open Results of Previous Search View Abstracts View Websites	Advanced scaffold of Previous Search Results used to reduce cognitive load of student while searching.
Beginning Search/Middle sessions (two or three activities)	5 22 23 (20) May or may not include 23 and/or 20.	Conduct Search View Abstracts View Websites Saved MYDL Site in DQ Folder	Advanced searching scaffolds used to reduce cognitive load of student while searching. **May occur at any point in sessions.**
Beginning Search/Middle sessions (two to four activities)	7 8 with the addition of 9 and 10 at times.	Open Dictionary Search Tool Performed Dictionary Query on Term Opened Thesaurus Page Performed a Thesaurus Query on Term	Advanced scaffolds of Dictionary and Thesaurus used to aid student in choosing terms to search with.
Workspace Organization/Middle and End sessions (two to four activities in iterative sequence)	17 15 or 18 Repeat 17 and 18 in sequence of two or more iterations, with the addition of 19 at times.	Open DQ Folder Delete DQ Bin Saved DQ Notes in DQ Folder Deleted Bookmark from a DQ Folder	Scaffold of DQ Folders used throughout sessions. Folders used to organize students information seeking activities.
Extended Search/Middle and End sessions (two activities. May be repeated)	4 6	Open Results of Previous Search Performing a Search with Wordnet	Advanced scaffolds of Previous Search Results and Wordnet to choose terms used to reduce cognitive load of student while searching.
Extended Search/Middle and End sessions (two to four activities in iterative sequence. May not include all five activities)	5 7 8 9 10 in sequence of two or more iterations and alternating 7 and 8.	Performing Search Open Dictionary Search tool Performed Dictionary Query on Term Opened Thesaurus page Performed Thesaurus Query on Term	Advanced scaffolds of Dictionary and Thesaurus used to aid student in choosing terms to search with.
Extended Search/Middle and	6 7	Performing a Search with Wordnet on	Advanced scaffolds of Dictionary and

End sessions (two to three activities in iterative sequence)	10	Open Dictionary Search tool Performed a Thesaurus Query on Term	Thesaurus used to aid student in choosing terms to search with.
Extended Search/Middle and End sessions (two activities in iterative sequence)	11 4	Open Past Searches Window Opened Results of Previous Search	Advanced scaffolds of Past Searches and Previous Search Results used to reduce cognitive load of student while searching.
Extended Search/Middle and End Sessions	20 22 in sequence of two or more iterations	Saved Site from MYDL Viewed Abstract	Organizational and Searching scaffolds used to reduce cognitive load of student while searching. May occur at any point in the sessions, but most frequently at middle and end sessions.
Extended Search/End sessions (two to three activities in iterative sequence)	17 28 (23) May or may not include 23.	Opened DQ Folder Viewed Abstracts Saved in DQ Folder Viewed Website	Scaffolds of DQ Folder and items saved in DQ Folder used to aid student in further information seeking. May occur at any point in the sessions, but most frequently in end of sessions.
Ending Activities/End sessions (two activities)	5 26	Performing Search Viewed Shared Cool Sites	Searching and Collaborative scaffolds used by few students to bring information seeking to conclusion.

Table 16 summarizes the most often used activity patterns by all groups. It further characterizes the patterns within the context of the search session. It is also interesting to note that several activity patterns appear to be class specific, or were used by groups within a particular class. For example: Class 1 user groups began their first session by Creating a DQ (13), Logging Out (2), Logging In (1), and Searching (5). Class 2 A and B users began their first session by Creating a DQ (13), Searching (5), and Viewing Abstracts (22).

It is also important to note that student modeling is evident in the activity patterns in the data. Several instances of almost exact sequencing of Action ID's between student groups were noted. User groups also used the collaborative scaffold of View DQs to look for ideas of what driving question to create. There is evidence that students modeled each other's DQs. Several instances show that student groups within the same class used the exact same DQ as other group(s) in the class. In order for the student modeling to be an effective strategy within the ADL, the use of the DQ scaffold needs to be encouraged and developed further by the teachers. Both of these areas warrant further study. Please refer to Appendix H for a complete breakdown of each class's activity patterns. This appendix is further segmented into the activity patterns evident within each graded user group, as will be discussed below.

While the above analysis does indicate strong evidence of teacher effect upon student use of the ADL, it does not indicate how teacher effect can impact the student's successful use of the ADL. Findings concerning the correlation between ADL use and student achievement will be explored below.

Scaffold Use as a Means of Predicting ADL Success

Method Two (refer to Ch. 3 p. 63) outlines a means of providing external corroboration for the interaction patterns that are present in the data. The ADL teachers were asked to provide grade information for their user groups for the data collection period. Four teachers volunteered to provide the study with this data. The user groups within these four classes were then divided into groups based on the grade they achieved on the ADL assignment. These groups are labeled A-users, B-users, C-users, and either D, F, U-users, depending on how the teacher distinguished the final grade. Their individual user group search histories were then plotted and scaffold use for each graded group was analyzed. Total number of scaffold uses, as well as time spent using the scaffolds was examined. Tables 17 - 20 present a profile of the average graded user group within each class.

Table 17. Average Graded Users in Class One

Activity and ActionID	A-user Mean Use/Time in seconds	B-user Mean Use/Time in seconds	C-User Mean Use/Time in seconds	D-User Mean Use/Time in seconds.
Login (1)	5.7 / 75.3	4.19 / 97.1	3.68 / 76.8	3.00 / 35
Logout (2)	3.6 / 974.6	1.44 / 2020.1	1.57 / 3782.6	1.17 / 121
Login teacher (3)	0	0	0	0
Open Results of Previous Search (4)	0.1 / 21	0.09 / 45.3	0.11 / 20.6	0
Performing a Search (5)	24.1 / 21.7	12.31 / 35.9	11.21 / 55.5	8.17 / 44.4
Performing a Search with Wordnet (6)	0.3 / 12.3	0.06 / 21.5	0.71 / 41.4	0
Open Dictionary Search Tool (7)	1.1/ 27.2	0.58 /47.8	0.43 / 42.8	0.33 / 62
Performed Dictionary query on term (8)	1.4 / 89.0	1.26 / 74.36	0.79 / 74.7	0.50 / 19.3
Opened Thesaurus page (9)	0.1 / 16	0.11 / 48	0.07 / 46	0
Performed thesaurus query on term (10)	0.1 / 495	0.13 / 40.2	0.04 / 18	0
Open Past Searches Window (11)	0.4 / 1199.5	0.26 / 30.6	1.39 / 28.3	0
Deleted a Past Search (12)	0	0.13 / 3	0.46 / 2.9	0
Created DQ from Search Page (13)	5 / 34.4	1.64 / 71.6	1.32 / 301.1	3.17 / 144.8
Created DQ from View Page (14)	0	0	0	0
Deleted a DQ Bin	1.7 / 59.7	0.30 / 54.2	0.29 / 48.5	0.17 / 71

(15)				
Edited DQ (16)	0.3 / 19.3	0.12 /15.8	0.04 / 52.8	0.17 / 0
Opened DQ Folder (17)	7 / 51.1	2.31 / 55.0	1.46 / 59.5	0.67 / 82.3
Saved Notes in DQ Folder (18)	1.3 / 93	0.48 / 33.6	0.14 / 20.3	0
Deleted a Bookmark from a DQ Folder (19)	0.2 / 0.5	0.32 / 16.7	0.04 / 3.4	0
Saved Site from the MYDL in DQ Folder (20)	3.2 / 78.9	1.68 / 155.8	0.89 / 71.3	0.83 / 68
Saved Bookmark in DQ Folder. Site not MYDL (21)	0	0	0	0
Viewed Full Website Description (22)	23.5 / 127.5	12.63 /182.3	10.25 / 225.0	7.83 / 157.3
Viewed Website (23)	13.1 / 308.5	5.56 / 277.0	4.14 / 358.2	1.50 / 366
Shared Site from the MYDL with class (24)	0	0.08 / 115.2	0.04 / 49.3	0.17 / 27
Shared Site NOT from the MYDL with class (25)	0	0	0	0
Viewed Shared Cool Sites (26)	0.2 /774.5	0.38 / 175.4	0.39 / 655.2	0.33 / 0.55
Viewed Shared DQ (27)	0	0.11 / 193	0.11 / 56.1	0.83 / 0
Viewed Website Description of Site Saved in DQ Folder (28)	0.6 / 58.5	0.62 / 51.2	0.07 / 92.2	0.17 / 59
Average Total Uses and Time in seconds	93 / 107.0	45.58 / 118.0	39.64/ 177.6	29.00 / 62.9

Table 19 shows the differences between the graded user groups within the same class. It also shows a comparison of each graded group's use and time figures while using the ADL. For example, the A-users conducted more searches, created more DQs, opened more DQ Folders, and viewed more website descriptions (abstracts) than the other groups. While the A-users may have used more of the features of the ADL, the table also shows that they spent less time engaging in these activities than the other graded groups.

Table 18. Average Graded Users in Class Two

Activity and ActionID	A-User Mean Use/Time in seconds	B-User Mean Use/Time in seconds	C-User Mean Use/Time in seconds	F-User Mean Use/Time in seconds
Login (1)	5.38 / 45.8	2.5 / 91.6	2.44 / 183.4	2.70 / 430.9
Logout (2)	3.27 / 1324.1	2 / 72	1.89 / 74.6	1.60 / 0
Login teacher (3)	0	0	0	0
Open Results of Previous Search (4)	0.18 / 27.4	0	0.33 / 30.5	2.0 / 20.5
Performing a Search (5)	10.08 / 58.5	4 / 131.6	7.56 / 23.2	6.38 / 51.2
Performing a Search with Wordnet (6)	0.14 / 14.6	0	0.17 / 17.3	2.0 / 25.7
Open Dictionary Search Tool (7)	0.24 / 38.4	0	0	1.0 / 103
Performed Dictionary query on term (8)	0.42 / 202.4	0	0	1.0 / 40
Opened Thesaurus page (9)	0.09 / 33	0	0	1.0 / 42
Performed thesaurus query on term (10)	0.07 / 40	0	0	1.0 / 9
Open Past Searches Window (11)	0.36 / 49.8	0	0.56 / 15.8	3.0 / 14.6
Deleted a Past Search (12)	0	0	0	0
Created DQ from Search Page (13)	2.33 / 55.2	5 / 69	2.72 6.22	2.11 / 66.3
Created DQ from View Page (14)	0	0	0	1.0 / 2
Deleted a DQ Bin (15)	0.26 / 40.35	1 / 38	0.78 / 71.8	1.0 / 63
Edited DQ (16)	0.40 / 18.0	0	0.17 / 51	1.50 / 24.5
Opened DQ Folder (17)	14.00 / 87.0	5 / 217.2	9.50 / 56.4	3.50 / 93.1
Saved Notes in DQ Folder (18)	6.87 / 38.5	5.5 /379.0	4.22 / 81.4	4.60 / 79.2
Deleted a Bookmark from a DQ Folder (19)	1.56 / 3.6	0	2.89 / 1.91	2.0 / 0
Saved Site from the MYDL in DQ Folder (20)	9.72 / 40.8	3 / 66.3	7.33 / 40.7	7.67 / 81.2
Saved Bookmark in DQ Folder. Site not MYDL (21)	0.08 / 5.2	0	0.11 / 0.5	0
Viewed Full Website Description (22)	15.85 / 69.1	6 / 49.2	6.44 / 81.1	6.08 / 73.2
Viewed Website (23)	5.30 / 198.6	0.5 / 1983	1.22 / 210.2	2.25 / 0
Shared Site from the MYDL with class	0.35 / 25.5	0.5 / 12	0.61 / 21	1.0 / 46

(24)				
Shared Site NOT from the MYDL with class (25)	0	0	0	0
Viewed Shared Cool Sites (26)	0.27 / 79.3	0	0.17 / 419	0
Viewed Shared DQ (27)	0.10 / 60.4	0	0.11 / 85.5	2.0 / 63
Viewed Website Description of Site Saved in DQ Folder (28)	2.35 / 78.4	0	0.50 / 50.6	2.0 / 85
Average Total Uses and Time in seconds	79.44 / 74.2	35/ 162.9	49.72 / 42.8	19.61 / 46.6

Table 18 shows Class Two's interaction with the ADL. Again differences in the graded user groups can be noted. The A-users conducted more searches, created and opened more DQs and DQ folders, saved more notes and sites in DQ folders, and viewed more abstracts than the other graded user groups. They also spent less time conducting these activities.

Table 19. Average Graded Users in Class Three

Activity and ActionID	A-User Mean Use/Time in seconds	B-User Mean Use/Time in seconds	C-User Mean Use/Time in seconds	U-User Mean Use/Time in seconds
Login (1)	5.68 / 153.6	6.83 / 205.8	6.29 / 163.9	4.42 / 174.5
Logout (2)	2.95 / 24670.5	3.33 / 4098	2.86 / 0	3.32 / 3788
Login teacher (3)	0	0	0	0
Open Results of Previous Search (4)	0.05 / 130	0.50 / 5.6	0.29 / 4	0
Performing a Search (5)	14.63 / 25.2	14.33 / 32.5	15.43 / 909.8	8.89 / 35.5
Performing a Search with Wordnet (6)	0.05 / 38	0	0.14 / 50	0.16 / 0
Open Dictionary Search Tool (7)	0.32 / 17.5	0.50 / 18.6	0.29 / 41	0.05 / 14
Performed Dictionary query on term (8)	0.37 / 65.6	1.33 / 86	0.29 / 13407	0.05 / 65
Opened Thesaurus page (9)	0	0.17 / 11	0.14 / 8	0
Performed thesaurus query on term (10)	0	0.17 / 33	1.43 / 0	0
Open Past Searches	0.11 / 18	1.0 / 18.5	0.57 / 17.5	0.11 / 109

Window (11)				
Deleted a Past Search (12)	0	0	1.43 / 0.75	0
Created DQ from Search Page (13)	1.37 / 25.5	1.17 / 27.6	0	1.63 / 35.5
Created DQ from View Page (14)	0	0	0.14 / 0	0
Deleted a DQ Bin (15)	0.21 / 17.7	0	0.71 / 0	0.26 / 26
Edited DQ (16)	0.37 / 6.3	1.0 / 12.6	6.71 / 62.8	0.47 / 33.5
Opened DQ Folder (17)	5.32 / 142.7	6.83 / 141.6	5.0 / 132.1	5.74 / 187.7
Saved Notes in DQ Folder (18)	2.84 / 150.2	4.0 / 31.9	0.29 / 34.5	2.32 / 37.4
Deleted a Bookmark from a DQ Folder (19)	0.11 / 0	0.17 / 56.6	0.86 / 1	0.11 / 2.5
Saved Site from the MYDL in DQ Folder (20)	0.79 / 44.6	1.17 / 56.6	0 / 27.8	0.89 / 727.3
Saved Bookmark in DQ Folder. Site not MYDL (21)	0	0.17 / 56.6	19.43 / 0	0
Viewed Full Website Description (22)	13.37 / 315.9	12.83 / 165.1	0 / 98.4	8.58 / 180.5
Viewed Website (23)	0	0	0	0
Shared Site from the MYDL with class (24)	0	0	0	0
Shared Site NOT from the MYDL with class (25)	0	0	0.71 / 0	0
Viewed Shared Cool Sites (26)	0.21 / 4.3	0.33 / 1973.5	1.86 / 38	0.26 / 7
Viewed Shared DQ (27)	1.89 / 146.5	1.83 / 182.9	0.86 / 111	2.0 / 112.5
Viewed Website Description of Site Saved in DQ Folder (28)	1.05 / 174.3	1.33 / 244.1	65.71 / 198.1	1.37 / 407.3
Average Total Uses and Time in seconds	51.6 / 367.1	59.0 / 277.5	40.63 / 276.6	40.63 / 367.8

Table 19 shows the differences within Class Three. In this class it is evident that the graded users interactions vary considerably. For example, the C-users conducted more searches than the other groups, the F-users created more DQs and the B-users opened more DQ Folders. The A-users viewed more abstracts, but did not engage in the same level of frequency of searching, creating and opening DQs and DQ folders, as Class One and Two A-users did.

Table 20. Average Graded Users in Class Four

Activity and ActionID	A-User Mean Use/Time in seconds	B-User Mean Use/Time in seconds	C-User Mean Use/Time in seconds
Login (1)	5.79 / 96.0	5 / 104.7	1.58 / 67.3
Logout (2)	0.71 / 13.5	1.08 / 8374.7	0.25 / 0
Login teacher (3)	0	0	0
Open Results of Previous Search (4)	0.57 / 7.14	0.12 / 20.5	0
Performing a Search (5)	16.36 / 104.3	11.96 / 48.2	6.33 / 38.3
Performing a Search with Wordnet (6)	1.29 / 30.5	0.24 / 52.6	0.25 / 33.3
Open Dictionary Search Tool (7)	1.0 / 47.3	0.24 / 47.6	0.08 / 28
Performed Dictionary query on term (8)	6.36 / 588.9	1.24 / 79.6	0.25 7
Opened Thesaurus page (9)	0.07 / 38	0.04 / 32	0
Performed thesaurus query on term (10)	0.07 / 12	0.04 / 37	0
Open Past Searches Window (11)	0.71 / 6.5	0.4 / 19	0.25 / 0
Deleted a Past Search (12)	0.21 / 2	0.12 / 0	0.75 / 0
Created DQ from Search Page (13)	2.64 / 39.0	2.8 / 52.7	2.83 / 42.7
Created DQ from View Page (14)	0	0.08 / 1	0.08 / 1
Deleted a DQ Bin (15)	1.0 / 73.2	0.76 / 254.7	0.50 / 63.3
Edited DQ (16)	0.36 / 25.1	0.36 / 23.5	0.17 / 93
Opened DQ Folder (17)	1.86 / 67.4	1.8 / 55.3	1.25 / 35.6
Saved Notes in DQ Folder (18)	0.07 / 0	0.4 / 23.7	0
Deleted a Bookmark from a DQ Folder (19)	0.07 / 0	0.12 / 19.5	0
Saved Site from the MYDL in DQ Folder (20)	2.21 / 0	1 / 61.4	1.42 / 31.7
Saved Bookmark in	0 / 69.8	0	0.08 / 0

DQ Folder. Site not MYDL (21)			
Viewed Full Website Description (22)	12.71 / 73.1	12.2 / 163.3	9.17 / 74.5
Viewed Website (23)	4.14 / 277.2	2.12 / 252.2	1.25 / 206.1
Shared Site from the MYDL with class (24)	0.14 / 320	0.76 / 105.8	0.25 / 7
Shared Site NOT from the MYDL with class (25)	0	0	0
Viewed Shared Cool Sites (26)	0.86 / 96.2	2 / 292.5	1.25 / 93.4
Viewed Shared DQ (27)	0.86 / 102.3	1.6 / 94.3	1.17 / 73.2
Viewed Website Description of Site Saved in DQ Folder (28)	1.14 / 753.3	0.24 / 438	0.08 / 0
Average Total Uses and Time in seconds	61.21 / 142.5	46.72 / 117.1	29.25 / 36.9

Table 20 shows that the A-users in Class Four conducted more searches, opened more DQ Folders, and used more of the advanced features than the other graded user groups. B and C-users created more DQs.

Examination of the above tables illustrates two perspectives of how we can measure teacher effect on the students. First, by comparing the graded groups within each individual class, we can begin to see some similarities and differences. Second, we can compare how the graded user group's scaffold use varied from one class to the next. Tables 21 –24 present a synthesis of both time and use analysis categorized using the functional categories presented above in Table 14. The figures in bold red type indicate the highest use and time across classes.

Table 21. Comparison of Graded A-Users Across the Four Classes

Scaffold Category	Class One A-users Use/Time%	Class Two A-users Use/Time%	Class Three A-users Use/Time%	Class Four A-users Use/Time%
Maintenance	**4.72 / 2.2**	3.89 / 0.6	2.4 / 0.07	**4.78 / 1.2**
Organizational	35.47 / 18.16	**39.48 / 57.12**	36.28 / 21.28	19.81 / 6.02
Searching	**50.53 /** 61.13	28.6 / 32.7	47.0 / **70.8**	37.1 / 15.6
Collaborative	0.43 / 3.24	1.2 /0.82	**14.06 / 3.88**	5.41 / 3.53
Advanced	8.79 / 15.12	6.72 / 8.51	5.53 / 3.94	**32.66 / 73.55**

Graded A-users in each class use and time figures vary. This may be an indication of the ADL features the teacher chose to emphasize. It is, however, evident that use between the A-users in each class did vary.

Table 22. Comparison of Graded B-Users Across the Four Classes

Scaffold Category	Class One B-users Use/Time%	Class Two B-users Use/Time%	Class Three B-users Use/Time%	Class Four B-users Use/Time%
Maintenance	3.12 / 0.68	3.8 / 0.96	3.37 / .27	**5.11 / 5.82**
Organizational	25.58 / 16.7	**71.08 / 91.86**	38.63 / 25.6	22.83 / 9.26
Searching	**57.2 / 74.6**	23.07 / 6.9	37.9 43.1	45.9 / 53.2
Collaborative	2.14 / 2.32	1.92 / 0.15	6.26 / 21.2	**16.41 / 23.55**
Advanced	11.87 / 5.46	0 / 0	**14.45 / 9.64**	9.61 / 8.29

Variation of scaffold use is also evident in the B-users across classes, as is shown in Table 22. It is interesting to note that each group of B-users seemed to focus on one category of activity. Class One B-users show the highest level of Searching, while B-users in Class Two used Organizational scaffolds more, Class Three B-users used more Advanced scaffolds, and Group Four shows the highest level of both Maintenance and Collaborative scaffolds.

Table 23. Comparison of Graded C-Users Across the Four Classes

Scaffold Category	Class One C-users Use/Time%	Class Two C-users Use/Time%	Class Three C-users Use/Time%	Class Four C-users Use/Time%
Maintenance	4.15 / 0.49	**10.46 / 3.69**	4.14 / .61	7.14 / 3.63
Organizational	23.03 / 18.36	**65.17 / 67.76**	33.95 / 16.01	28.56 / 15.89
Searching	**55. 82 / 72.6**	17.60 / 23.62	47.2 / 24.9	46.2 / 62.4
Collaborative	1.56 / 4.08	2.41 / 2.25	6.24 / 2.76	**13.44 / 16.03**
Advanced	**15.36** / 4.27	4.23 / 2.55	8.3 / **55.45**	4.62 / 1.91

C-users in all classes do not exhibit the same tendencies as described for the A and B-users above. Class One C-users conducted more searches and use more advanced scaffolds. Class Two C-users used more maintenance and organizational scaffolds, Class Three did not show highest use in any particular category except for the highest time spent using advanced scaffolds. Class Four C-users used more collaborative scaffolds.

Table 24. Comparison of Graded Other-Users Across the Four Classes

Scaffold Category	Class One D-users Use/Time%	Class Two F-users Use/Time%	Class Three U-users Use/Time%
Maintenance	2.18 / 0.93	3.14 / **1.76**	**3.49** / .62
Organizational	30.68 / 22.25	**45.77** / 39.44	44.05 / **43.95**
Searching	**51.6 / 72.6**	38.42 / 47.3	35.74 / 37.6
Collaborative	8.77 / 0.9	2.62 / 4.69	**9.39 / 5.85**
Advanced	6.57 / 3.16	**9.96** / 6.02	7.2 / **11.67**

Graded Other-users for all classes also varied in their scaffold use. For example, Class One Other-users used the highest level of search scaffolds, while Class Two used more Organizational and Advanced features, but spent less time than Class Three Other-users in both categories.

Tables 21-24 illustrate the impact teacher effect may have on the way the user groups interact with the ADL. By comparing each graded group across classes, we can see the scaffolds in which each group focused their interactions. These differences across class again provide evidence of the different ADL scaffolds the teachers may emphasize in their students' use of the ADL.

To further help in our characterization of the ADL users, individual graded user groups from each class were combined to place all A-users, all B-users, all C-users and all D-users into groups. There are a total of 132 A-users, 111 B-users, 121 C-users, and 44 D, F, U-users for which data were examined. Their individual user group search histories were then plotted and all A-users, B-users, etc. from each class were combined to provide a picture of what each graded user groups' interaction with the system might entail. Figures 12 through 15 illustrate how each graded user group interacted with the ADL.

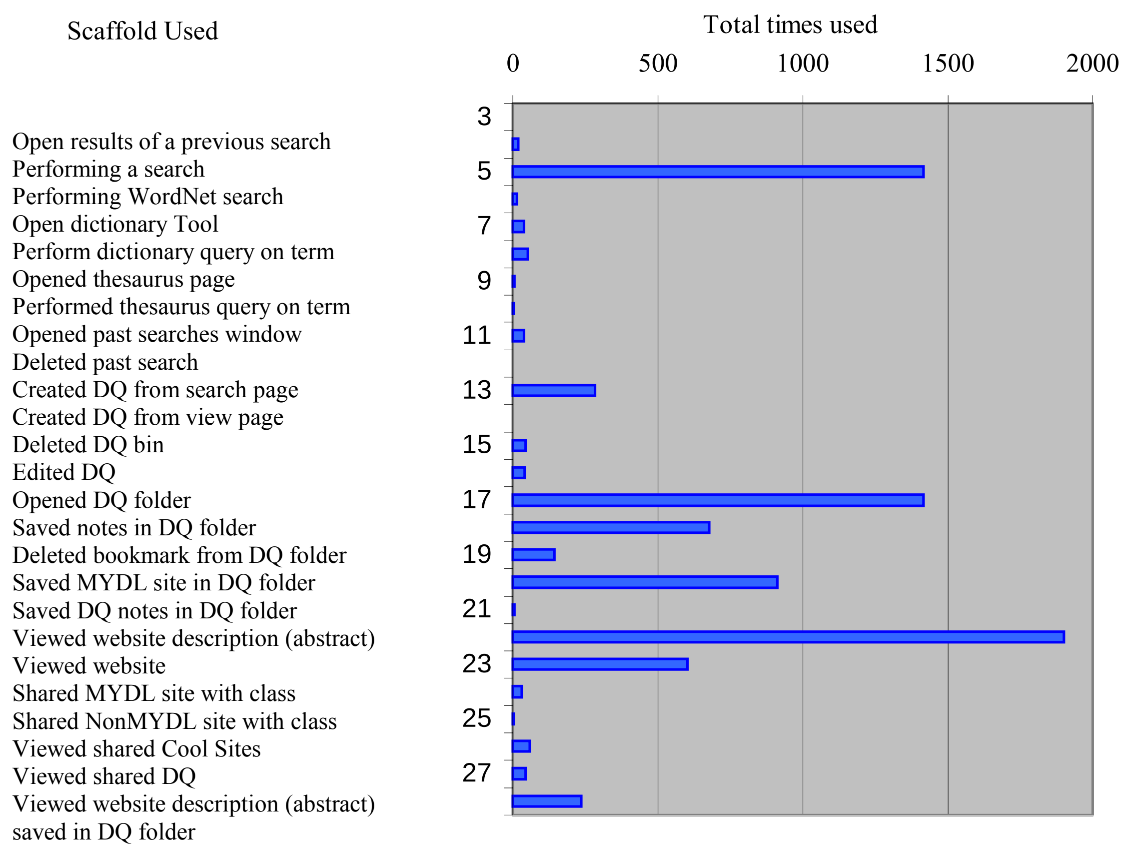


Figure 12. A-user's Scaffold Use

The total activities recorded for the A-users were 9,839. Figure 14 shows the top five scaffolds of: 1) viewing website descriptions (abstracts), 2) opening DQ folder, 3) saving MYDL sites in DQ folder, 4) saving notes in DQ folder, and 5) create DQ from search page, comprise 56.0% of the A-users' interactions with the system. It is important to note that the A-users group does not view websites as frequently as we saw evidenced in the data for the entire sample group (refer to Table 9). The A-users also use more of the organizational scaffolds such as saving notes (6.90%) and saving sites in DQ folders (9.60%) than the other graded groups. Furthermore, this group used the collaborative scaffolds (1.67%) and the more advanced scaffolds of the thesaurus, dictionary, and saved searches folders (2.60% total) less frequently than might be expected from a group that was more successful than the other students in their Artemis assignment(s).

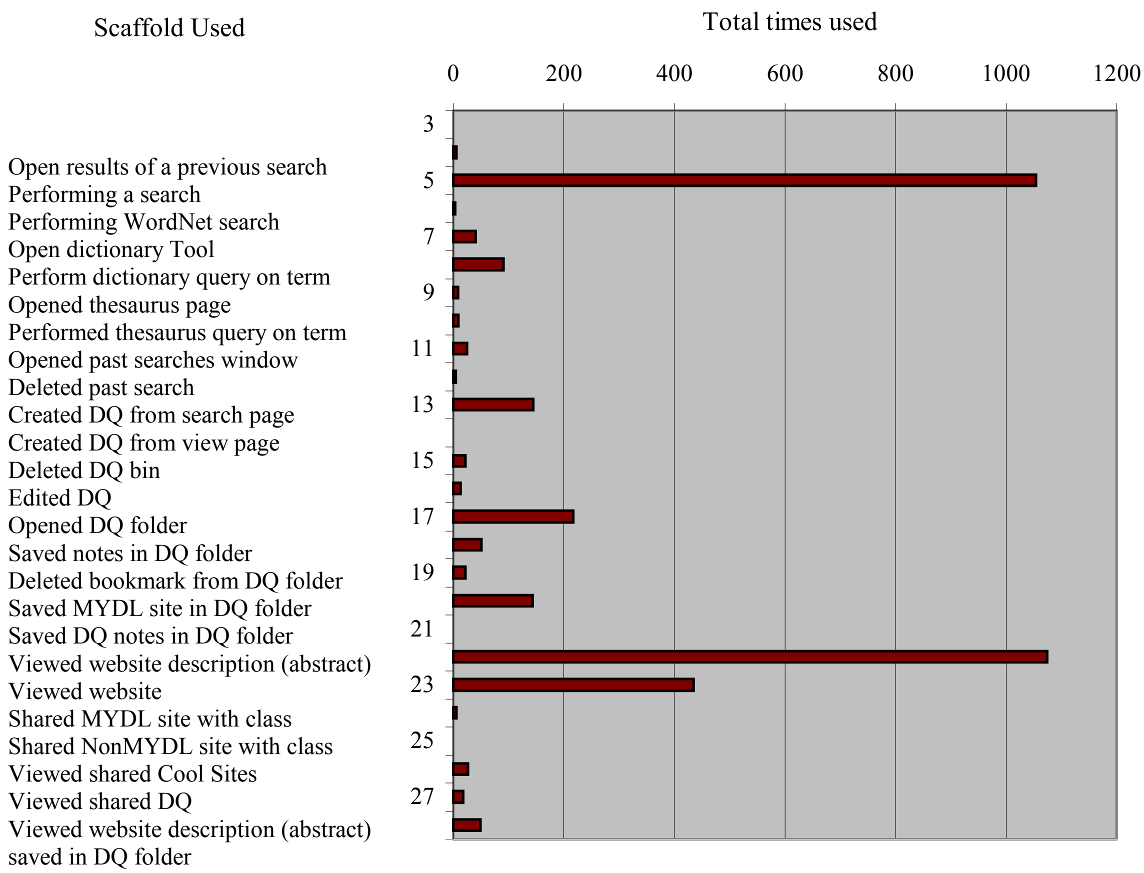


Figure 13. B-users Scaffold Use

The total activities recorded for the B-users was 5,147. Figure 13 shows the top five scaffolds of: 1) viewing website descriptions (abstracts), 2) open DQ folder, 3) create DQ from search page, 4) saved MYDL site in DQ Folder, and 5) Performed dictionary query on term. These activities comprise 42% of the B-users' interactions with the ADL. This group used very few of the collaborative scaffolds (2.36%) or the more advanced scaffolds (4.53%). It is notable, however, that the activities of searching and viewing websites comprise 36% of all B-user activity.

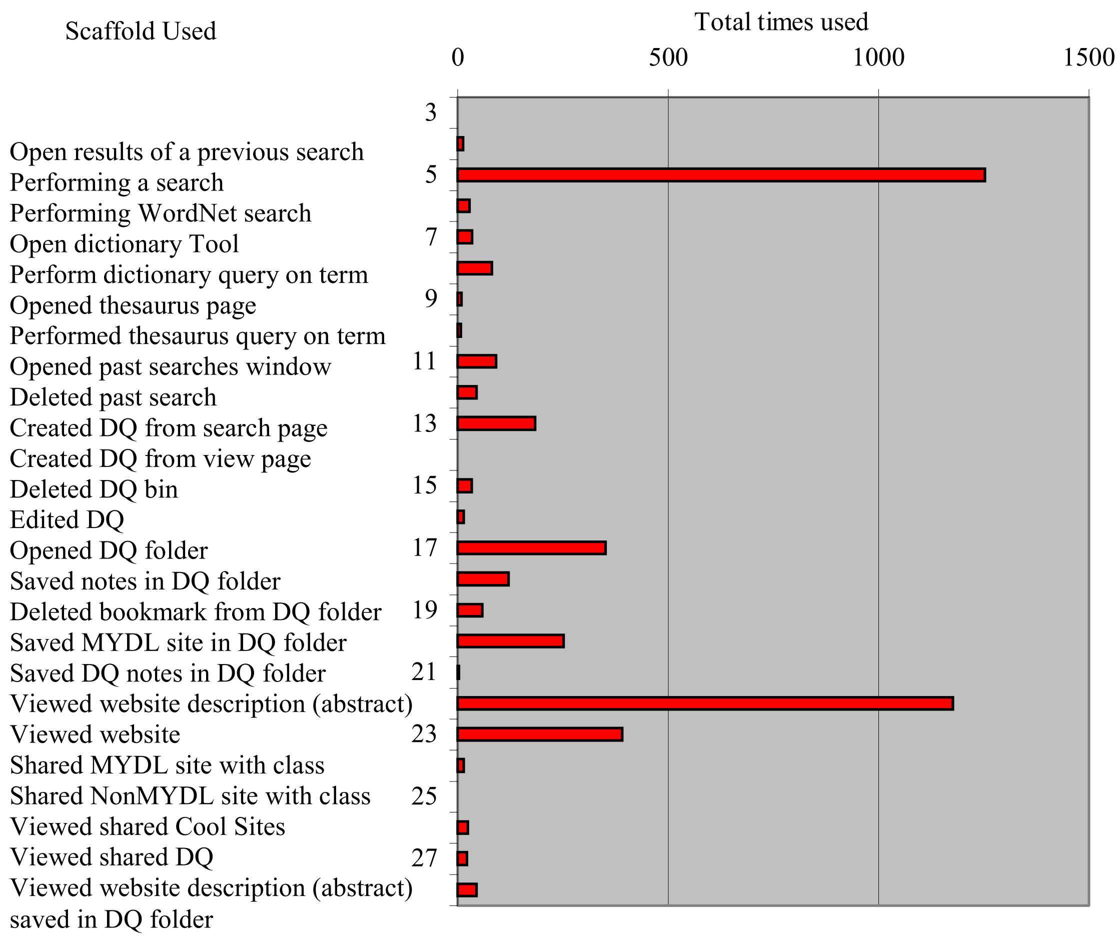


Figure 14. C-Users Scaffold Use

The total recorded activities for the C-users were 5,192. Figure 14 shows the top five scaffolds of: 1) viewing website descriptions (abstracts), 2) opening DQ folders, 3) saving MYDL sites in DQ folders, 4) create DQ from search page, and 5) saving notes in DQ folder. These activities comprise 44% of the C-users' interactions with the ADL. The activities of searching and viewing websites comprise 34% of C-user's activities. This group also did not use the collaborative scaffolds (2.15%) or the more advanced scaffolds (5.16%) very often. The C-users also used the maintenance scaffolds of creating DQ folders, Deleting and editing DQs and DQ Bins and deleting bookmarks more than the other groups.

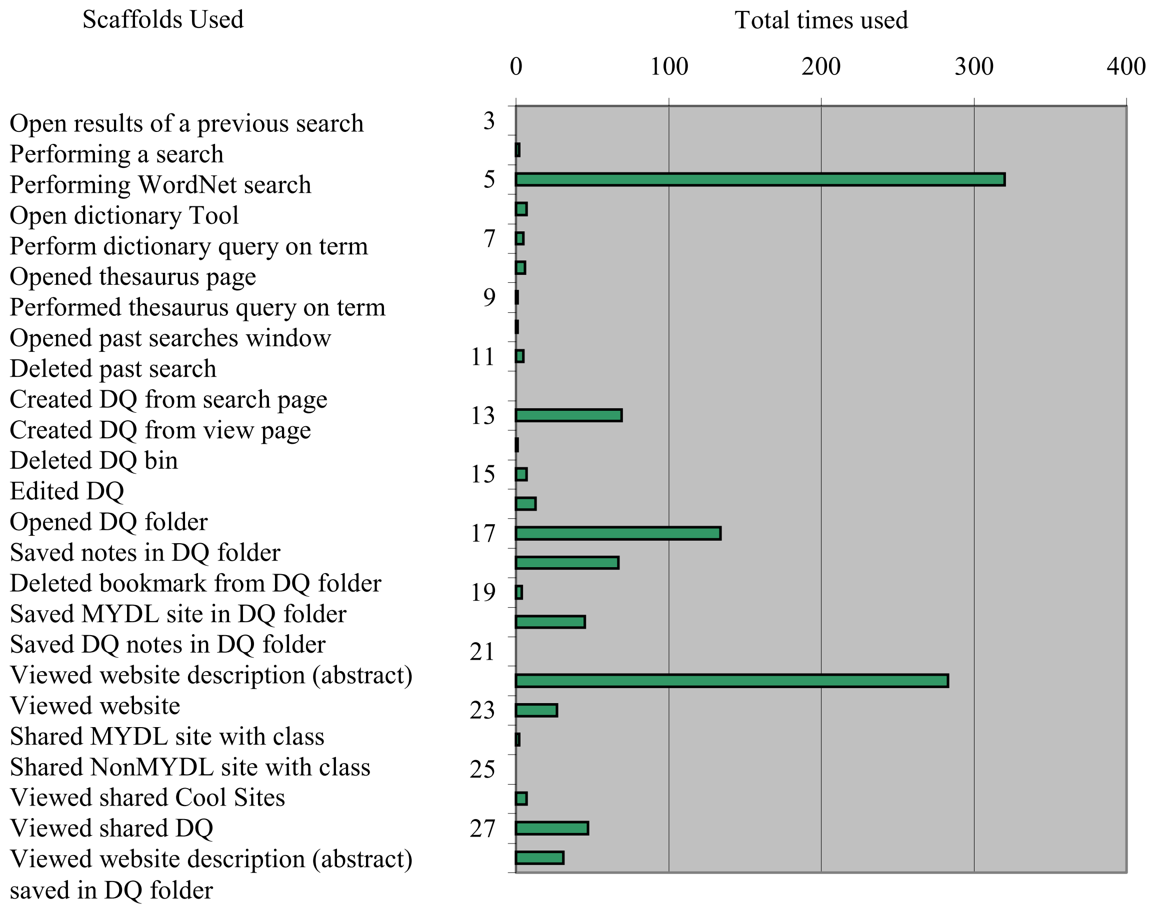


Figure 15. D, F, U-users Scaffold Use

The total activities recorded for the D, F, U-users was 1,299. Figure 15 shows the top five scaffolds of: 1) viewing website descriptions (abstracts), 2) opening DQ folders,

3) creating DQ from search page, 4) saving notes in DQ folder, and 5) viewed shared DQs. These activities comprise 46.2% of the D, F, U-users' interactions with the ADL. This group used the collaborative scaffold of viewing shared DQs more than the other groups (3.62%). This group also viewed the saved website descriptions more than the B and C-users (2.4%), with the A-users using this scaffold (2.62%). Table 17 presents each graded group's overall ADL use compared to the other groups' use.

Table 25. Categorized Percent of Use of the Scaffolds of Graded User Groups

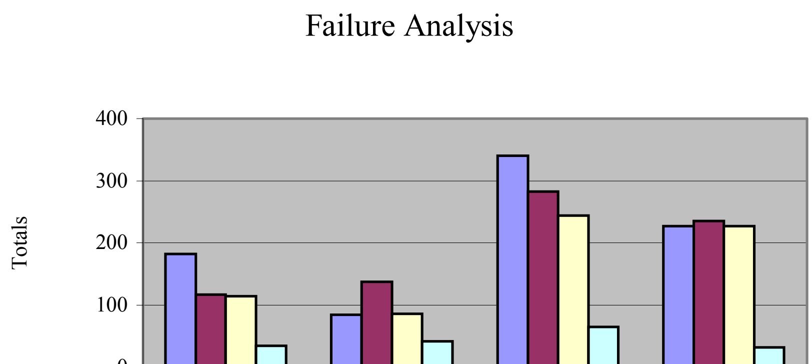
Scaffold Category	A-users	B-users	C-users	D, F, U-users
Maintenance	5.8%	6.13%	**7.50%**	7.24%
Organizational	**31.2**	9.6	14.7	19.0
Collaborative	1.67	3.13	1.78	**6.7**
Searching	45.0	**63.0**	59.0	49.0
Advanced	5.7	5.8	**6.17**	4.66

Table 25 re-emphasizes the frequency at which user groups use each category of scaffolds. Notable is that the A-users show higher use in only one of the scaffold categories, the Organizational scaffolds.

An additional means to assess system use by the graded user groups is to look at the failure analysis, search techniques and the extent of match analyses for each of the graded user groups. Each is presented in a separate section and the trends and similarities examined.

Failure Analysis for the Graded User Groups

User groups within this study often encountered obstacles that affected their information seeking activities. The obstacles of misspelled words, typographical errors, repeating searches, or repeating searches in sequence were evidenced in the user data. Figure 16 illustrates to what extent each of the obstacles were encountered by each graded user group.



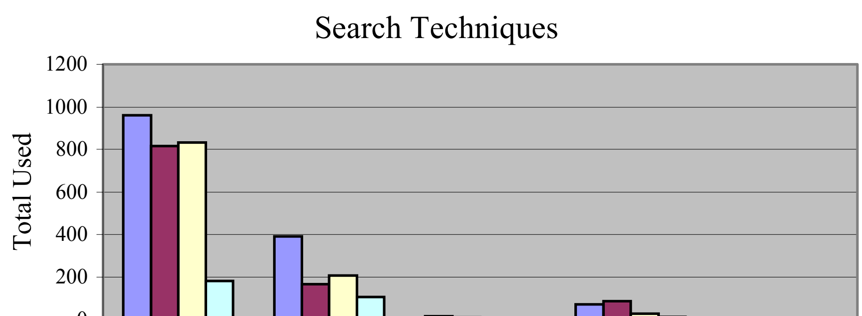
	Spelling	Typos	Repeated Search	Repeated Sequentially
A-users	182	84	340	227
B-users	117	137	282	235
C-users	114	86	244	227
D-users	34	41	64	32

Figure 16. Failure Analysis of All Graded Users

Figure 16 shows the totals for each groups' failure analysis. If we compare the groups we see some slight differences in the obstacles encountered while searching. For example, while 9.0% of the A-users' searches contained instances of misspelled terms, the other groups percent of misspelled terms were not significantly different B=8.6%, C=8.5%, and D, F, U=10.0%. However, 12.8% of the D, F, U-users' searches contained instances of typographical errors and punctuation, which is higher than the other graded users of A=5.1%, B=10.1% and C=6.5%. All groups tended to repeat their searches almost an equal percent of the time (A=20.6%, B=20.8%, C=18.4%, and D, F, U=20.0%). A total of 17.4% of the B-users' searches were repeated sequentially, compared to the other group's A=13.8%, C=17.1%, and D, F, U=10%.

Search Techniques Use for the Graded User Groups

Examining the search techniques of the graded user groups may also provide more insight into successful system use. Figure 17 shows figures for each groups' search technique use.



	Single word	Phrase	Boolean	Natural Language	Unsupported
A-users	962	389	15	71	2
B-users	815	167	9	86	8
C-users	832	207	0	28	3
D,F,U-users	181	105	2	12	3

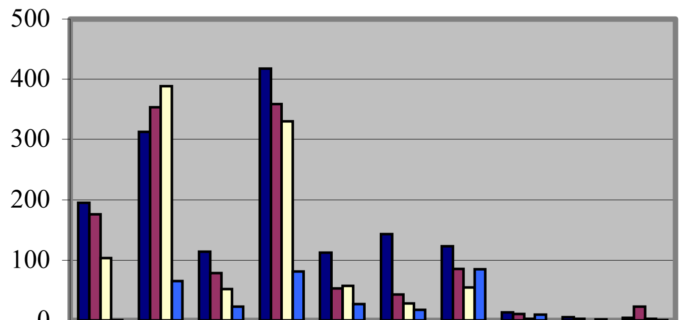
Type of Searches

Figure 17. Search Techniques of the Graded User Groups

User groups in each graded group all tended to use single word searches more frequently than the other search techniques. The C-user groups conducted slightly more single word searches (62%), as opposed to the other groups' A=58%, B=60%, and D, F, U=56% searches respectively. The D, F, U-users phrase searches comprise 33% of this group's searches, as opposed to A=23%, B=12%, and C=15%. B-users' conducted more Natural-language searches (6.4%) than the other groups, A=4.3%, C=2.1%, and D, F, U=3.8%. The A-users, however, conducted more searches than the other groups; A=1,439, B=1,085, C=1,070, and D, F, U=303.

Extent of Match Analysis for the Graded User Groups

Choosing search terms can be difficult for users. Examining the graded groups' extent of match analysis may provide insight into how often successful users' search terms match the terms in their driving questions. Figure 18 presents this analysis.



	NDQ	NM	NM-S	EM	RM	PM	PM-P	IOM	CM	GO
A	195	313	114	418	112	143	123	13	5	4
B	176	353	79	359	53	43	86	11	3	23
C	103	389	52	330	57	28	55	3	0	3
D,F,U	1	65	23	81	27	18	85	10	2	1

Figure 18. Extent Graded Users' Search Terms Match DQs

Figure 18 shows the A-users' search histories contained the highest instance of NDQs or not having DQs to compare with search terms. However, the terms in their searches more frequently exactly matched the terms in the DQs. Overall, the A-users contained the highest instance of all extent of match categories with the exception of the

No Match (NM) category. Of additional interest is the trend of the B-users searches to either not match the DQs or to match exactly with the DQs. This parallels the finding shown in Table 20 below (p. 115) for the total user groups of the ADL. Figure 20 illustrates the search terms for all graded user groups most often exactly matched the terms in their DQs.

Picture of Graded ADL Users

Reviewing the findings presented in the previous sections on the graded users, it is possible to begin to posit what each graded ADL user groups' system use might entail. This section will synthesize the above findings and present a rudimentary picture of each graded ADL user group's interaction with the system.

A-users tend to conduct more searches and engage with the system more than the other user groups. (A=9,839 activities or an average of 75/user group, B=5,147 activities or an average of 46/user group, C=5,192 activities or an average of 43/user group, and D, F, U=1,299 activities or 30/user group). This graded group tended to use the organizational scaffolds more than the other groups. However, the A-users used the collaborative scaffold less than the other groups and the advanced scaffolds less than the B-users. They also viewed websites less than either of the B-users or C-users (A=6.72%, B=9.5%, and C=7.82%).

A-users' searches had a slightly smaller instance of misspelled terms, typographical errors, but they tended to repeat searches with the same frequency of the other groups. They also had a lower tendency of sequentially repeating their searches than the B- users and C-users. Their searches tended to contain more single words than any of the other categories of phrase, Boolean, Natural-language, or Unsupported searches.

A-users' search histories contained the highest instance of NDQs, but their search terms more frequently match the terms in their DQs if a DQ was created. Compared to the other groups' extent of match totals, the A-users totals were higher than the other groups.

B-users tend to view more websites, use more collaborative scaffolds than the A-users and C-users, but less than the D, F, U-users. This graded group used more of the advanced scaffolds than the A-users and D, F, U-users, but less than the C-users. This group also used less organizational scaffolds than any of the other groups.

B-users search terms contained the second highest instance of typographical errors and the inclusion of punctuation in the searches. This group's searches were also repeated sequentially more than the other groups' searches. The B-users conducted more Natural-language searches and fewer phrase searches than the other groups.

The B-users follow the trend shown in Figure 21 (shown below on p. 116) for all ADL users. The B-users search terms either do not match their DQs (353 total) or terms that exactly match those in their DQs (359 total). It is also interesting to note that the B-users' searches contained more off task (GO) search terms than the other groups.

C-users tend to view website descriptions more than the B-users or the D, F, U=users, but less than the A-users. This graded group also viewed websites more than the A-users and D, F, U-users, but less than the B-users. These users also used the collaborative scaffolds more than the A-users, but less than the other groups. They did, however, use the more advanced scaffolds more than the other groups. The C-users used the maintenance scaffolds of creating folders, deleting folders, creating DQs, editing DQs and deleting DQ bins more than the other groups (C=7.49%, A=5.8%, B=6.13% and D, F, U=7.24%).

D, F, U-users tend to use the collaborative scaffolds more than the other groups. This group viewed more shared DQs than the other three graded groups at 3.62%, as opposed to A=.57%, B=1.13%, and C=.70%.This group also uses the organizational scaffolds of saving notes and saving sites in the DQ folders more (8.66%) than the B-users (4.5%), C-users (7.49%), but less than the A-users (16.5%). This group also viewed the website descriptions they had saved in their DQ folders more (2.39%) than the B-users (1.07%) or the C-users (.87%), but not more than the A-users (2.55%). See Appendix I for a table of graded user groups' scaffold use.

D, F, U-user's searches tended to contain more spelling errors, typographical errors and use of punctuation in the searches than the other groups. This group, however, repeated searches sequentially less than the other groups. The D, F, U-users conducted more phrase searches than the other groups but less in each of the other search technique categories.

The D, F, U-user groups' search terms had more of a tendency to exactly match the terms in their DQs than any other of the exact match categories, with the exception of the partial phrase match (PM-P) category, which was slightly higher (EM=81 and PM-P=85).

Question Two Revisited

How does children's language relate to the language used to represent the documents in the system? Can children's language be used to represent documents within the collection? Will using student language within representations affect retrieval?

Question Two provided the basis for examining one of the most problematic aspects of information seeking and retrieval, choosing the most useful words with which to represent an information need and to search for the needed documents. This question also allowed the researcher to explore issues of representation inherent in any information retrieval system, choosing the best, most representative terms with which to describe the resources so that the user may retrieve them. The ADL contains elements that give us a chance to understand more about how this age group represents their information needs and the processes they undertake to resolve that need. The ADL provides the researcher with user-side scaffolds of representation, the DQ folders and spaces where users can save DQs and past searches. These representations are valuable resources to both the user and the system. With the driving questions and stored search words it is possible for us to learn more about how this age group represents their information need, the search strategies and techniques they use, even the form of words they most often use. Examining their DQs enables us to see a little about how they think and the process they undertake to solve their information need.

An additional dimension of this research question deals with issues of retrieval. While representation is central to retrieval issues, evaluating retrieval is often separated from representation. Using the ADL scaffolds of DQs and stored search terms, it became possible to feed back into the system user-side descriptors for documents (the terms the users most frequently used to search with), thereby allowing us to evaluate any increase in functionality and successful retrieval.

This study explored both sides of the representation issues: student representation of their information need through the use of DQs and search terms (Method 3.1 A and B), and system side representation (Method 3.1 C, Method 3.2 and Method 4). Based on observations and analysis of preliminary data the following hypotheses were formulated to explore Research Question Two:

1. There will be significant differences in strategies used by the user groups.

    Sub-hypothesis A: Overall, children will use more single word searches than phrase, Boolean, or Natural-language searches.

    Sub-hypothesis B: In a significant number of searches, obstructions due to mechanical errors, such as spelling errors, typographic errors, repeated searches, and repeating searches sequentially, will occur.

2. In a significant number of searches, system representations of documents will not match children's representations of their information needs.

3. Representing documents using the language of student's searches will have a positive effect on retrieval.

In order to explore the representation issues, the data were analyzed in several ways, as defined in Chapter 3, Methods 1.1, 3.1 and 3.2 (pp.62 -63) and as described below. These methods outlined the use of various mechanisms to examine strategies used and obstacles encountered during information retrieval, and both user-side and system-side representation issues including: the extent of match between the terms in the user groups' DQs and the search terms used to retrieve the resources; and the extent of match between the most frequently used user search terms and the system's terms or the controlled vocabulary used to represent the resources. Each analysis is presented and discussed separately below. Please note that retrieval issues will be presented later in the discussion for Method 4.

Search Strategies and Techniques

Another view into the users' understanding of the information seeking process, as well as system knowledge, can be glimpsed by examining the search strategies and techniques they use to retrieve information in the ADL. Examining search strategies and techniques also provides insight into representation and system design issues. This version of the ADL search engine, from which the data were gathered, provides a truncated or stemmed search of the database fields for each resource. The fields searched are the title, description, and keyword fields. The ADL search engine allows single term and phrase searching. It also supports Boolean AND searching. However, because of the often complex nature of Boolean searching, this search technique is not encouraged by the teachers or the system designers. The ADL search engine at this time does not perform natural-language searching. Figure 19 illustrates search techniques used by the user groups.

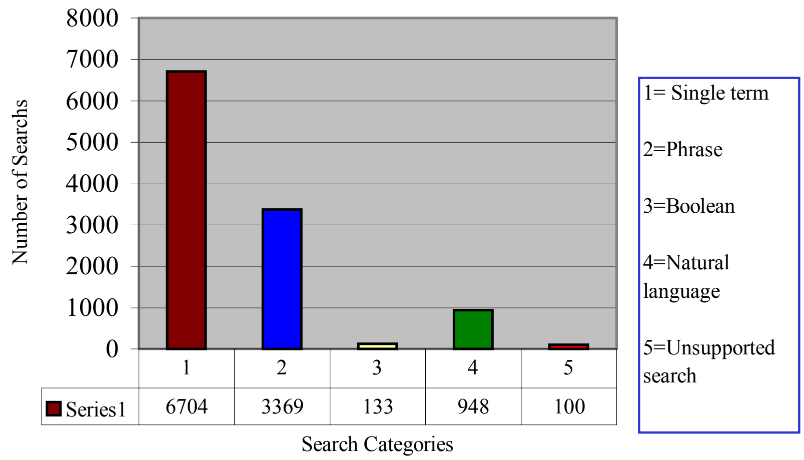


Figure 19. Search Techniques Used

Figure 19 shows that users conducted single term searches a total of 6,704 times or 57.73%, phrase searches a total of 3,369 or 29.0%, Boolean searches a total of 133 or 1.15%, natural-language searches a total of 948 or 8.16%, and unsupported search strategies and techniques a total of 100 or 0.86%. It is important to note that single term searches comprise more than half of all searches conducted by the user groups. This may suggest that representations for this age group should be comprised of single word terms instead of the current practice within the controlled vocabulary used by the ADL digital librarians to use phrases. These findings may also illustrate that further training in how to choose more complicated search terms or how to construct more complex search strings may help students achieve more desirable results. These issues comprise the heart of information seeking, representation, and retrieval. Achieving a better match between system generated index terms and user search terms requires that we understand more about search terms and strategies of searching. These findings further support the idea of using user-defined search terms to augment the system's descriptions of the resources.

Further analysis on the extent of match between student terms and the controlled vocabulary used to represent the documents, as well as the most frequently used search terms also supports this conclusion. This analysis will be presented under the section for Student Issues below. Table 26, which shows the percent of user groups that used each search technique, also supports this trend.

Table 26. Total Users With at Least One Use of the Search Technique

Search technique	Single term	Phrase	Boolean	Natural-language	Unsupported
Total users	719	490	33	217	42
Percent of users	95.4%	66.0%	4.38%	28.8%	5.6%

Table 26 illustrates that 95% of all user groups used the single term search strategy at least one time. Phrase searching was used by a total of 66% of the user groups. Notable is the lack of use of the Boolean search technique. As mentioned above, Boolean searching is not recommended by the teachers or the system designers because of its often complex nature. However, while this search technique is not encouraged, the students' searches indicate that many of their searches would benefit by connecting two words or concepts together. This was evidenced in many of the phrase searches that contained two to three word combinations that were strung together without the addition of Boolean operators. For example, the words "star" and "galaxy" appeared together as a phrase search. This search would yield any resources that contain either term in the title, abstract, and keyword field(s). However, with the addition of the Boolean operator AND the system would retrieve only the resources that contained BOTH the terms “star” and “galaxy”. It was decided early in the coding stage to code these two word combinations as phrases, as they were neither single word searches or natural-language searches. Whether or not the user groups were aware of how this form of a search strategy would work is not known. Further research into the user groups’ intentions for this type of search strategy is warranted.

It is also interesting to note from Table 26 that 28.8% of the searches contained natural-language search strings, or searches structured as partial sentences. As mentioned above, the ADL search engine does not provide natural-language searching at this time. Because of the relatively moderate level of frequency of this search technique, system designers may wish to consider including natural-language searching.

Obstacles Encountered

The pilot study uncovered obstacles encountered by the students while using the ADL. This study conducted a failure analysis of the obstacles evidenced in the pilot study. Obstacles considered include: 1) spelling errors within search terms, 2) typographic errors or use of punctuation within search terms, 3) searches repeated but not in succession, and 4) sequential repetition of the same search terms. The failure analysis revealed that obstacles encountered while using the ADL in the pilot study are also present in the searches within the present data set. As a result of filtering out the data of the user groups with fewer than three logins, a total of 11, 611 searches were analyzed. The results presented in Figures 20 and Table 27 show both the obstacles as evidenced within the total searches and the percent of users who encountered each obstacle.

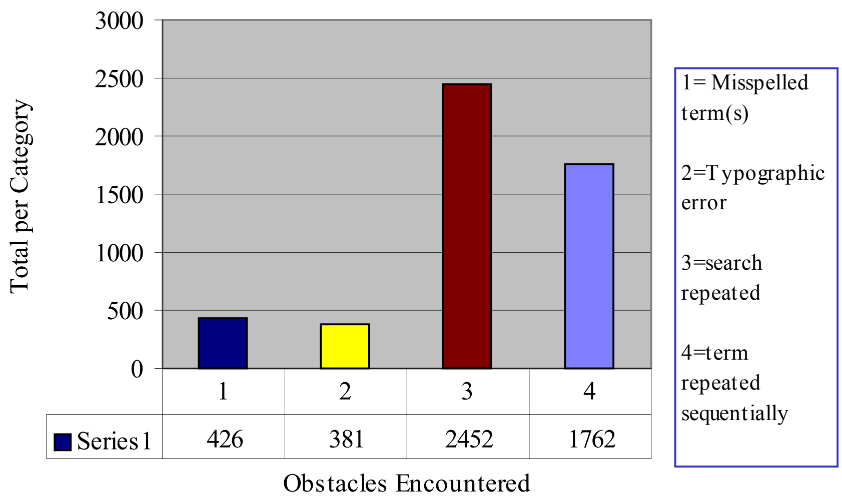


Figure 20. Failure Analysis of Total Searches Conducted by the User Groups

The chart in Figure 20 illustrates that 426 or 3.67% of the searches contained misspelled terms, 381 or 3.28% of searches contained typographic errors or contained punctuation within the search, 2,452 or 21.12% of the searches were repeated by the user groups within the course of their search history, instead of viewing and using saved past searches. Searches that were repeated sequentially totaled 1,762 or 15.18%, which could indicate a system problem or a lack of understanding of the search process by the user groups.

Table 27. Users that Encountered Obstacles while Searching

Obstacle	Spelling errors	Typographic errors	Repeated Search	Repeated Search Sequentially
Total users	311 users	196 users	598 users	434 users
Percent of users	41.25%	26.0%	79.31%	57.60%

Table 27 illustrates that at least 26% of all users encountered some obstacle while searching the ADL. These findings indicate that users might benefit by additional training on how spelling and typographical errors, and use of punctuation within a search can affect seeking and retrieval. Furthermore, the evidence of searches being repeated sequentially indicates a potential problem with the system, or a gap in the student's knowledge base regarding both the search process and the system's search functions. Training in the use of more advanced scaffolds such as opening a past search window (Action ID 11) and opening results of a previous search (Action ID 4) might reduce the number of searches that are repeated by the users.

Student Issues: Extent of Match Between DQ Terms and Search Terms

The sample for this piece of the analysis remained the same as that explained on p.68 above. It included all student user groups that had logged in to the system a minimum of three times. Further deselected were any logins that were teacher or system designer related. The total activities analyzed were 52, 781 (81.3% of the initial total). A total of 754 user groups were used in the analysis (45.8% of the initial total).

Using an Access database designed to hold the transaction log activities, the field holding the DQ Bins containing the DQs and the field holding the user groups' search terms were segmented and printed out. Using the Extent of Match Rules developed by the researcher (refer to Appendix E.) a total of 11,612 searches were compared against the User group's DQ(s). The results for the extent of match categorization and the total users evidenced using each of the categories are presented below in Figure 21 and Table 28.

Extent of Match Category	Total number of searches **N= 11,612**	Percent of all searches
NDQ	2122	18.27%
NM	3004	25.86
NM-S	1090	9.38
EM	2475	21.31
RM	657	5.65
PM	765	6.58
PM-P	1189	10.24
IOM	74	0.637
CM	24	0.206
GO	212	1.825
Totals	11,612	100%

Search terms did NOT match terms in DQs

Search terms matched exactly with DQs

Figure 21. Extent of Match between Driving Questions (DQs) and Student Search Terms

Figure 21 shows us that the user groups' representations, their search terms, in 25.86% of the searches do not match the terms in their DQs. However, the Figure also shows that in 21.31% of the searches their search terms match exactly with the terms in their DQs, or in the case of phrase searches, their search terms match in 6.58% of the searches. Also of interest in this data is that user groups' search terms were either synonymously related (NM-S= 9.38%), at least the root form of the word matched the DQ (RM= 5.65%), or in the case of phrase searches, the search terms partially matched the terms in the DQ (PM-P= 10.24%). In all, the search terms matched the DQs in some instantiation 54.0% of all search sessions. It is important to note, however, that 18.27% of the search sessions contained no DQs to compare the terms to, as evidenced by Table 9 which indicates this scaffold comprised only 3.6%. of all ADL activities.

Table 28. User Group Extent of Match

Extent of Match Category	Total of User Groups	Percent of Total Users
No DQ present (NDQ)	211	28.00%
No Match (NM)	455	60.34
No Match but Synonymous (NM-S)	267	35.41
Exact Match (EM)	505	67.00
Root Match (RM)	224	29.70
Phrase Match (PM)	202	26.80
Phrase Match Partial (PM-P)	249	33.02
Inverted Order Match (IOM)	39	5.17
Combination Match (CM)	13	1.72
Goof Off (GO)	74	9.81

Table 28 illustrates a trend similar to that evidenced in Figure 21 that 60.34% of all user groups' conducted at least one search in which the terms either contained no match to their DQs or the terms exactly matched the DQs (EM=67.0% and PM=26.80%). User groups' whose searches also contained at least one instance of synonomously related, root form, or partial phrase match terms totaled 132.2%. These findings might suggest that the user groups experienced difficulty in choosing search terms. On the other hand, the findings might suggest that the user groups do not use the DQ scaffold (developing a driving question before searching) as a means to focus their research.. Table 9, which indicates that the DQ scaffold was only used 3.6% by the total filtered ADL user groups, supports this assertion. Use of the DQ as a means by which the user groups focus their research, and the extent to which teachers require this important scaffold needs to be explored further.

These findings do, however, strongly support the assertion that we can use the ADL's user side representation scaffolds (DQs and search terms) to learn more about how this age group represents its information needs and how they then express those needs to the system in the form of search terms. The majority of all searches conducted by the user groups contained some instantiation of the terms present in the DQs, which indicates that terms with which the user groups conducted searches were directly related to the terms they originally used to express their information need.

We can approach this finding with two assertions: 1) by helping students to formulate better DQs they may understand the process of information seeking and retrieval more completely; and 2) because the terms in the user groups' DQs are directly related to their search terms, we can use these user-side representation scaffolds to enhance the system-side representations.

Driving questions rich with the terminology of the discipline or the classroom experience may help user groups to understand the science concepts they are researching. Enriching their DQs may also teach them to choose alternate search terms that may better match those used by the system's agents, the indexers, to represent the documents, thereby affecting a better match in retrieval.

While examining the DQs and search terms, the collaborative nature of the DQs was noticed. User groups who viewed the shared DQs of the other groups were seen to duplicate or use the other group's DQs with only minor modifications. This collaborative effect was an interesting but unexpected result that requires further investigation. Students may view their peer groups as the experts to follow, as opposed to modeling their teachers. However, as mentioned above on p. 72, the collaborative scaffolds comprise only 4.3% of all ADL activities. It is, however, important to note that the A-users of the system used these scaffolds less than 0.2% (0.194%). What these findings indicate is that the DQs scaffold has potential to be a very helpful scaffold. Using DQs as a means of providing focus for the user group's information seeking activities, should be encouraged both for its collaborative nature, as well as a means to further understand the iterative nature of the information seeking process.

The second assertion, using these user-side representation scaffolds (the DQs and the search terms) as a means to enhance the system-side representations, provides the basis for the further exploration of this user side scaffold as outlined in Methods 3.1B and C (refer to pp. 64 for complete description of this method) and as described below.

Student Issues: Extent of Match Between Search Terms and System's Terms

The sample for this piece of the analysis was a complete list of the entire filtered user groups' (754 total user groups) search terms. The original intention was to conduct a semantic frequency match between the terms and the DQs for which they were used. It was assumed that the terms would vary substantially to make a semantic match necessary. However, once the search terms were compiled, it became evident that there was very little variety in the terms used to search for same or similar concepts, thereby making a semantic match unnecessary at this time. The process for determining the most frequently used search terms consisted of four stages as follows.

1. Using the SPSS frequency function a complete frequency list of *all* search terms used was compiled. This list removed all deletions, but included those terms with misspellings, punctuation, and typographic errors. The list was sorted with most frequently used terms appearing first. See Appendix H. for a list of the most frequently used search terms.

2. All terms in the compiled list that were used at least five times by the user groups were extracted. A total of 856 terms that fit this criteria were examined. At this stage, a failure analysis and search technique analysis was conducted on the terms. Any terms that contained misspellings or typographic errors were filtered from the list. A total of 120 misspelled terms and 38 terms with typographic errors were filtered from the list. Any terms contained within a Boolean (S3), natural-language (S4), or Unsupported (S5) search technique were parsed out and examined separately. A total of 195 terms were filtered out. Further criteria for filtering terms included: 1) any terms that might appear off task or objectionable to the teachers or indexers such as those evidenced in the off task category (GO) or 2) any terms that contained geographic names. A total of 26 off task terms (GO) were deleted and 11 terms containing geographic names were deleted.

3. The remaining 466 terms were then compared to the system's controlled vocabulary (CV) using the Extent of Match Rules in Appendix E. The terms were coded for their extent of match, failure analysis, and the search technique used. A second Access database was developed which contained the coded analysis.

4.The fourth stage entailed the filtering out of any terms that exactly matched (EM) those in the CV, as well as any that were phrase matches (PM). Terms not present in the CV were compiled into a list of student keywords. Root matches (RM) were then compared to the CV to determine if the search term represented a concept different than the root form that appeared in the CV. If the root match represented a concept not contained in the CV, it was added to the list of student keywords. Terms that appeared as part of a phrase search that were not complete phrases, but rather a string of two or three words without the use of Boolean operators, or terms part of a partial phrase match (PM-P), were parsed out and each term compared separately to the CV terms. Terms that did not match the CV but were synonymously related (NM-S) were compared to their equivalents and individual judgments were made whether or not to include them. For example, if a term was substantially different in wording from those found in the CV or if it was found to be a related term using the Wordnet thesaurus feature of the ADL, it was added to the list of student keywords. Table 29 presents the results of the term analysis.

Table 29. Results of Controlled Vocabulary Extent of Match

Extent of Match Category	Total Terms Examined	Total Terms Added to List
EM	128	0
NM	163	116
NM-S	23	21
RM	56	16
PM	17 phrases	16 terms
PM-P	79 phrases	36 terms
Totals	466	205

A list of 205 student keywords (user side descriptors) was compiled. See Appendix I. for the complete list.

Table 29 also illustrates the extent to which the CV or the system terms match the terms most frequently used to search the database. Only 145 search terms or search phrases (31% of terms examined) exactly matched the terms on the CV; while 163 terms (35%) did not match the CV. A total of 158 search terms or search phrases (34%) contained at least a partial or synonymously related instantiation of the CV terms. These finding indicate that the terms used by the system, the CV, do not adequately represent the user groups' information needs. Using student-generated keywords (SGKs), or terms constructed as a result of the above extent of match analysis, may enable the system to provide more age appropriate representations of the resources.

Retrieval issues: Augmenting Resource Descriptions with User Side Descriptors

Method Four explored research question number two, hypothesis number three, and tested the potential advantages of using student-generated search terms (SGKs) as representations or index terms for the resources within the collection. The sample included: the list of student-generated keywords (SGKs) compiled in Method Three, the database of ADL resource descriptions, and student searches from the ADL transaction logs. It was conducted in two stages as outlined below.

1. The list of compiled student-generated keywords (SGKs) was used to add additional index terms to the ADL resources contained in a duplicated ADL database. Terms were added to the existing keywords field in the abstracts for each appropriate resource.

2. A sample of 143 original user groups' searches was re-entered by the researcher and the results analyzed to determine the effects augmenting ADL resources with student-generated keywords had on retrieval. The sample of original searches was limited to one percent of all searches conducted within the data collection period. Searches were chosen using stratified sampling. Every tenth search was chosen, and then duplicates or those containing spelling errors or punctuation were filtered out. In order to determine the effect on retrieval, results from the sample queries were then compared to the student's original results and retrieved resource descriptions were examined for their utility in answering the students' original driving questions. Each stage and the results are presented below.

Augmentation of Resource Descriptions

At the time of the study the ADL database contained a total of 4,496 resource descriptions or abstracts representing the ADL collection. A random sample of the resources was examined for potential augmentation using the list of student-generated keywords (SGKs). A total of 1,725 abstracts or 38% of all abstracts were reviewed. Of this sample, 478 or 11% of the total ADL abstracts (28% of the sample reviewed) were augmented using between one to four SGKs. The remaining 1,245 abstracts were not augmented because the SGK list did not contain terms that were appropriate to their content. (It should be noted that the terms on the SGK may be skewed by the present ADL topic assignments. Future semesters may focus on different topics and may add new terms to the SGK. It should also be noted that 449 of the abstracts reviewed contained social science related materials for another ADL system, the CLIO system, and therefore no terms on the SGK would be considered appropriate for these abstracts.)

Retrieval Results

The ADL user side scaffold of Driving Questions (DQ) and user group search terms provide us with a unique opportunity to explore retrieval issues. A stratified sample of user group searches was compiled by selecting every tenth search conducted in the ADL during the collection period. Duplicate searches and those with misspelled words or punctuation were then deleted and a list of 143 searches (one percent of the 11,611 total searches) and the results of each search was compiled. Each search was then re-executed by the researcher and the abstracts examined for: utility in answering the DQ, and augmented terms within the keyword field. Table 30 presents the retrieval results. (It should be noted at the time of the study the ADL limits the results retrieved and displayed to the user groups to a total of 25 for each search. As a result, searches with an original retrieval result of 25 were not included in the analysis. A total of 11 searches and their results were not used because their original search results totaled 25, so there is no way of knowing the actual total of results that would have been returned if this limitation was not in place.)

Table 30. Retrieval Results From Re-executed Searches

Augmentation Status	Total of Searches Conducted **N=132**	Percent
Augmented/Showed increase in results	43	32.6%
Augmented/Showed no increase in results	17	13.0%
Not Augmented	76	57.6%
Others (unexplained results)	6	4.5%

Table 30 shows that for all searches conducted 32% showed an increase in the results that were returned and contained SGKs as augmentation in the abstracts. The results also show that 13% of the searches did not show an increase, but also contained SGKs in the abstracts. Searches that returned abstracts that contained no augmentation (57.6%) show no change between the original results returned and the re-executed search, with the exception of six searches that returned unexplainable increases in the results. It may be possible that additional resources were added to the ADL after these searches were first conducted. See also Appendix J. for a complete log of the searches used, their results in both executions of the searches, and the extent of match between the search term and the user groups' DQs. These findings support the assertion that user side descriptors or student-generated keywords can have a positive effect on retrieval.

Hypotheses Revisited

This study presented the following hypotheses. Each will be examined and the results presented.

1. Children will engage with the system in various ways but common activity patterns will also be evident.

   Sub-hypothesis A: Teacher effect will have an impact on a user group's use of scaffolds

   Sub-hypothesis B: Teacher effect will have an impact on a user group's time spent using the ADL

   Sub-hypothesis C: Teacher effect will have an impact on a user group's mode of engagement (activity patterns).

2. Children will use a variety of scaffolds while engaging with the system.

   Sub-hypothesis A: Various levels of achievement will correlate with various patterns of interaction with the ADL.

3. There will be significant differences in strategies used by the user groups.

   Sub-hypothesis A: Overall, children will use more single word searches than phrase, Boolean, or Natural-language searches.

   Sub-hypothesis B: In a significant number of searches, obstructions due to mechanical errors, such as spelling errors, typographic errors, repeated searches, and repeating searches sequentially, will occur.

4. In a significant number of searches, system representations of documents will not match children's representations of their information needs.

5. Representing documents using the language of student's searches will have a positive effect on retrieval.

Hypotheses One, Two, and Three were examined within the context of Research Question One where the user groups' use of the ADL scaffolds was explored from many different dimensions. Each sub-hypotheses provided the focus to examine the issue of teacher effect on student use of and time spent using the ADL, as well as the potential activity patterns. The results of the exploration provide evidence to support the three hypotheses, as well as the sub-hypotheses under each. Children using the ADL engage with the system in various ways during their information seeking activities. There are, however, common activity patterns present in the data, that may provide us with models of how students interact with the ADL. User groups of the ADL also use a variety of the scaffolds provided within the system, with the searching and organizational scaffolds being used more than the advanced or collaborative scaffolds. Graded user group interaction also was seen to vary from class to class, indicating again teacher effect.

Hypotheses Four and Five were examined within the context of Research Question Two. This question enabled the researcher to explore both representation and retrieval issues inherent in any information retrieval system. It also provided a unique environment to examine how the use of the user side scaffold of Driving Questions and the users' search terms can be used not only to better understand the user group's information seeking activities, but also how augmenting the collection's resources with these student-generated search terms can affect retrieval. Analysis of the obstacles encountered during information seeking was also examined. Overall, children use more single word searches than phrase, Boolean, or Natural-language searches. In a significant number of searches, obstructions due to mechanical errors, such as spelling errors, typographic errors, repeated searches, and repeating searches sequentially, were evident. The extent of match analysis between the most frequently used search terms and the controlled vocabulary terms used by the system presented us with two results: a list of student-generated keywords (SGKs) which were used to augment the abstracts, and findings that indicate that the terms used by the system agents to index the resources are not adequate or age appropriate for this group of users. The findings from the re-execution of a sample of original searches conducted by ADL users indicate that using SGKs to augment the resource abstracts can have positive effects on retrieval. The findings of these explorations also provide evidence in support of the two hypotheses, as well as the sub-hypotheses under hypothesis Four.

The findings of the analysis not only provide evidence in support of the five hypotheses within the context of this system and study, but they also bring to our attention issues and areas of research that need further study. Chapter Five will explore these issues further and discuss areas that warrant further research.

CHAPTER 5

DISCUSSION AND CONSIDERATIONS

This study has increased our understanding of how children interact with digital libraries. It has also furthered our understanding of their question states, choices of search terms, search strategies, interaction patterns, as well as the obstacles they encounter during information seeking. The study also allowed the exploration of the issues of representation inherent in any information retrieval system, as well as testing the effects of user-defined descriptors on retrieval. This chapter summarizes the findings of the study, discusses unexpected occurrences within the study, outlines the limitations of the study, and speculates how these findings can be used to enhance the ARTEMIS Digital Library (ADL), as well as suggest directions for future research.

Summary

Two research questions focused the study. Each question dealt with a separate dimension of representation and use of the digital library by children. Research question One provided for the study of the children's interaction with the ADL. Multiple approaches were used to form a picture of how the children interact with the system, the system features or scaffolds they use or did not use, as well as the impact teacher effect may have on children's use and time spent using ADL scaffolds. Correlation between user's grades and scaffold use was also examined. Examining the data of the entire sample group of users revealed several findings as summarized briefly below.

1. Children engage with the system in various sequences of patterns, but common activity patterns were also evident. Overall, children used more of the searching and organizational scaffolds, than the collaborative, maintenance, or advanced scaffolds.

2. The presence of teacher effect or teacher context variables was examined by a correlation of teacher assessments with interaction or activity patterns within the individual's search sessions. Analysis of teacher effect on use and time factors of the scaffold use and activity patterns, as well as a failure analysis and search strategy analysis revealed a picture of each group of users' interaction with the system. Teacher effect is a factor in both scaffold use and time spent using the ADL. Teacher effect and scaffold use may affect student success using the ADL. (See Ch. 4 pp. 83-96 for a complete discussion of scaffold use, failure and search strategy analysis. Refer to Ch. 4 pp. 83-86 for discussion of activity patterns and teacher effect.) Overall, it was evident that user groups' interactions and scaffold use are influenced by the teacher's goals or objectives for using the ADL in their classrooms. Each teacher in the sample indicated a particular use of the ADL within their classrooms and assignments. Some teachers mentioned that the ADL assignment included other elements of the middle-school curriculum, such as language arts and social science. Because each teacher is using the ADL in a different fashion, it may, therefore, become important to provide further training to both the teachers and students on the use of the scaffolds, particularly the time saving more advanced scaffolds of Past Searches and Past Results. These scaffolds were designed to reduce the cognitive load felt by the students when conducting their information seeking activities within ADL. The scaffolds have great potential, if used. Furthermore, use of the

collaborative scaffolds also needs to be encouraged. A large part of scientific inquiry includes the collaborative interaction between its participants. As there is evidence of some student modeling within the activity patterns, ADL designers and teachers could harness this learning strategy of the children to enhance their science inquiry and use of the ADL.

Research Question Two explored the issues of representation and retrieval from both the system's perspective and the users' perspective. Obstacles to information seeking, such as user's mechanical errors, were explored. User groups' Driving Questions (DQs) were correlated with their search terms, in an effort to learn more about how children represent their information needs both prior to searching an information system and within the system environment. Extent of match analysis was conducted on the DQs and search terms. The system's terms or the controlled vocabulary (CV) used to create representations of the ADL resources were also compared to the user groups' search terms using the extent of match rules. Examining the data of the entire sample group of users revealed findings as summarized briefly below.

1. The majority of all users preferred to use single word searches over Boolean, phrase or natural-language searches. Users also tended to use a strategy of repeating the same search using the same terms, instead of using the advanced scaffold of viewing and using saved past searches.

2. A high percent of all users attempted at least one search that included spelling or typographical errors, punctuation, or sequentially repeated searches.

3. Twenty six percent of the user groups' search terms did not match the terms in their DQs. Twenty one percent of their search terms match with the terms in the DQs. Overall, the search terms matched the DQs in some instantiation 54% of all search sessions.

4. Only 145 of the most frequently used search terms or search phrases (31% of terms examined) exactly matched the terms on the controlled vocabulary, while 163 terms (35%) did not match the CV. A total of 158 search terms or search phrases (34%) contained at least a partial or synonymously related instantiation of the CV terms. These findings indicate that the terms used by the system, the CV, do not adequately represent the user groups' information needs.

From the above measure, a list of student-generated keywords (SGKs) was created. These terms were then used to re-index the ADL resources. A small subset of the user groups' original searches were then re-executed and the results of the two compared. Thirty two percent of the searches' results showed an increase and contained SGKs. The results also show that 13% of the searches did not show an increase, yet the resource descriptions contained SGKs. These findings indicate that using student-generated keywords to augment resource descriptions can have a positive effect on retrieval.

These findings outline both substantive and methodological contributions to the field and to prior literature. As outlined above, further knowledge of how middle-school students interact with the ADL and the impact the variable of teacher effect has on student interaction and scaffold use was uncovered. The study also extended earlier findings concerning obstacles encountered by children during information seeking. Most importantly, it provides ground breaking new research into children's questions as a reflection of their information needs, search techniques and strategies preferred by this user group within the digital environment, the evaluation of user defined descriptors and their use in indexing, and the effect their use in indexing has on retrieval.

The study also developed new methods and schemes that could be useful in future studies with similar objectives. For example, the failure analysis and search strategy analysis schemes could be used with other information retrieval data. The extent of match schemes could be used in studies of search term correlation to either searcher's questions or to controlled vocabularies.

Unanticipated Findings

There were several findings within the study that were surprising or unanticipated. These findings do, however, add to the picture we are formulating of ADL use and it is important that we address them here. These are explained below, in no order of importance or occurrence.

The data indicated instances of user groups modeling or copying other group's DQs. This "collaborative effect" was an unexpected occurrence that was not evident in the pilot study. It does, however, illustrate the potential advantages and also potential pitfalls of using DQs as a means to focus user's searches. As was evidenced in the extent of match analysis, DQs are a good indicator of the terms the user groups will use to search. However, poorly constructed DQs can provide the other user groups with an inadequate or incorrect model to follow. ADL teachers need to emphasize the importance of constructing adequate DQs, so if this collaborative effect remains a factor, students will have the proper model to follow. Furthermore, working to create good DQs may also provide the students with more terms used within their classroom or language learning community and also by the discipline, thereby enriching their vocabularies and their searches.

The study also revealed that user groups' searches contained a low degree of variability of terms. This finding may be attributed to the fact that particular areas of the curriculum were emphasized during the data collection period. For example, in the data collection period, the ADL units in use were astronomy, communicable diseases, and air and water quality. Inclusion of a larger, long term sample of search terms may show a larger degree of variability.

Examining user groups' DQs and their search terms can provide an understanding about this age group's information seeking processes and how they represent their needs. These user side scaffolds can also be a rich source for potential keywords or controlled vocabulary terms to use when representing the resources in the collection. Utilizing the SGKs has been shown to have positive effects on retrieval, as it eliminates some of the guessing game by both users of the system and the system agents, the indexers. Children's search terms within the ADL did not contain a high degree of variety, which may indicate that the user groups use similar terms to search for the same or similar concepts. It is, therefore, a fair assumption that the terms contained on the SGK are a good representation of this age group's (at least the ADL users) domain knowledge related to this area of science inquiry, making the user groups' terms a rich resource to use for creating representations. Because children are more likely to choose a term that they are familiar with, (such as those present on the SGK list) than a term from a nonage-appropriate controlled vocabulary, retrieval will be more effective. This finding warrants future exploration.

This finding also supports the concern that DQs need to remain a central emphasis of focusing the students' information seeking activities. The terms used by the user groups in their DQs may serve teachers by providing an indication of the students' content understanding or domain knowledge. Teaching students to use more variations of terms may increase a user group's chances of guessing the terms used by the system's agents to represent the documents, as well as enhance the student's content understanding. Teaching the users the search strategies of broadening or narrowing searches and using alternate forms of words may also aid in retrieval.

The low degree of use of the collaborative scaffolds was also a somewhat surprising finding. One of the ADL designer's main goals for the system, as well as supporters of digital libraries, is the collaborative nature of these resources. Science inquiry instruction includes an element of collaboration, or teaching the students the importance of collaborating with other scientists or experts. The ADL provides two avenues for collaboration, the View DQs scaffold and the Post Cool Sites scaffold. Using these scaffolds allows students to see what the other groups are researching, as well as to pass on helpful sites the groups encounter in the course of their own information seeking activities. This finding may be attributed to the objectives each teacher has for the ADL in their individual classrooms. Increasing avenues for teacher feedback about their intended uses of the system may provide more insight into their reasons for choosing not to use the collaborative scaffolds. However, if increasing use of the collaborative scaffolds continues to be a goal of the ADL teachers, this area will require further study to determine the benefits of this practice.

A further unexpected finding was the high degree of searches that were sequentially repeated using the same search terms. In numerous instances user groups' repeated searches as many as fourteen times sequentially with no other activity or Action ID present between searches. This finding may indicate a problem with the system, or the low degree of system and searching knowledge of the user. In either case, it is an area that should be explored further.

One further finding that was unexpected was the lack of rules present to guide the digital librarians in constructing representations or the resource descriptions. Upon examination of the guides for the ADL, it was determined that these documents were used for the new version of the ADL currently under development. When the digital librarians were contacted, none remembered having or using any guidelines or manuals while indexing. In a system that has been designed for a specific user group, it was the researcher's expectation that there would be special indexing specifications in place. Cooper & O'Connor (2001) refer to indexing as "shifting quicksand". As is often the case in indexing practice, rules, when given, are usually vague and provide little guidance to the indexer or the user. Indexers rely, instead, on gut feeling or past experience with the collection and the controlled vocabularies. The ADL digital librarians are unique in that they work closely with ADL designers and teachers, select resources that match specific curriculum requirements, and index the resources for a particular audience, thereby giving them a strong knowledge of the collection, the users, and the intended uses of the collection. One of the limitations of not having guidelines in place is that inconsistencies in the representations may occur. Library and Information Science studies have shown varying degrees of interindexer inconsistency. The affect of this phenomena can be seen from two differing perspectives: inconsistency can create confusion for users trying to guess the terms the indexer used to represent the resources, *and* using more diverse terms to represent the resources may enhance the possibility of the users to guess the terms being used. It is not, however, evident from the data of this study, what effect having no guidelines has on user retrieval or representation issues. This would be another interesting area to study further.

Limitations of the Study

This study was designed to use the method of transaction log analysis (TLA) as the primary means to collect and analyze the data. However, as the study progressed, the need for external data to corroborate the TLA data became apparent. At this point the teacher data were introduced. While TLA may remain a very powerful means to collect data via nonintrusive methods, thereby not disrupting the natural processes and environment, it is not without limitations. Collection of qualitative data cannot be gathered using TLA. For example, as noted in the pilot study, an important aspect of information seeking and representation activities is the affective process or emotive stages the users may undergo. TLA data does not allow for the capture of this form of data. User relevance judgments of the resources or reasons for revising or repeating the same queries were also difficult to ascertain using TLA. Aspects of the study could be enhanced by the use of video capture techniques, surveys, and interviews. However, because of time and geographic location limitations, these techniques were not used.

Speculations

These findings can be used to rethink our ideas and practices of scaffolds, representation and retrieval in a digital library environment, especially within the ADL. Children interact with the ADL in various ways, but common activity patterns are also evident. These models of system use can help researchers to rethink the scaffolds provided within the ADL, as well as rethink how we train the users of the ADL. It would benefit ADL designers to gather more feedback from the teachers who use the ADL in their classrooms. Learning more about their intentions, objectives, and integration of the ADL may provide valuable insights into system use not possible with the TLA data gathered in this study. Taking advantage of the findings of this study concerning the collaborative nature of the DQs, the scaffold use, the search terms most frequently used, and the preferred search strategies, may enable teaching students to become more efficient users of the ADL, as well as other information retrieval systems.

Children do encounter many obstacles when engaging with an information retrieval system. This study has identified the obstacles, but has not emphasized how to clear the path for users. Teaching the ADL users more about information retrieval techniques and strategies may be one means of alleviating some of the obstacles. Basic information retrieval techniques such as how to choose alternate terms to search, how to broaden or narrow searches, will not only enhance information retrieval, but also strengthen the student's vocabulary. Explaining the effects of misspelled terms on retrieval is also a needed piece of system knowledge.

Making use of the findings on the system's representation scaffolds is also important. Continued emphasis on the use and development of good DQs is essential. If students model their peers DQs, it is important that these models be accurate. Encouraging use of the collaborative scaffolds such as View DQs and Post Cool Sites is important. Using user groups' search terms and DQs as a means to learn more about information seeking and as a means to augment the systems representations will continue to make the ADL a more valuable and age appropriate resource. While the idea of using user-defined descriptors is not entirely new to the field, (O'Connor, O'Connor & Abbas, 1999; Hastings, 1995), it is a concept that has yet to be implemented in an information retrieval or representation system. Representation is a central problem and focus of information retrieval. Providing user-centered, age-appropriate representations within the ADL is no easy task, but it is one that can be ameliorated by using the student's own words.

The users of the ADL exhibited both the naïveté and experimentation of new users, but also continued skill acquisition. As posited in Ch. 3, ADL users may engage with the system in various stages, from a novice user to a more advanced user. Designers of the ADL included scaffolds to aid the user groups in their information seeking and retrieval. The scaffolds were designed to fade, or to be used less frequently as the student became more comfortable with the system and the process of information seeking. However, the activity patterns present in the data do not support this criterion of scaffolding. Rather, the user groups appear to be following the Five Stages of Skill Acquisition model presented by Dreyfus (1986) and later adapted by Abbas (refer to Ch. 3 p. 58, Table 6). User groups' interactions with the system appear to be in the stages one goes through to become an expert in a task. These stages include: Novice, Advanced Beginner, Competence, Proficiency, and Expert. Use of the scaffolds does not seem to fade away, but rather, each stage is characterized by particular scaffold use and activity patterns as seen in Appendix H. Table 31 shows a synthesis of activity patterns conducted within each stage.

Table 31. Revised Table 6: Stages of Interaction

Stage	Activity Pattern Name	Activity Pattern(s)	Activities in Sequence	Scaffold Use
Teacher modeling (rule based learning)	**Exploration/Beginning of ADL Activities** (three or four activities)	13 1 13 2 13 (5) 1 5 (22) 5	Create DQ from Search Page Log out/Log back in Conduct First Initial Search	Scaffold of DQ used in beginning of first session. May be used in middle sessions as well. Not often used in end sessions.
Teacher modeling (rule based learning)	**Workspace Setup and Organization/Middle sessions** (two or three activities in iterative sequence)	17 18 Repeat 17 and 18 in sequence of two or more iterations, with the addition of 19 at times.	Open DQ Folder Saved DQ Notes in DQ Folder Deleted Bookmark from a DQ Folder	Scaffold of DQ Folders used throughout sessions. Folders used to organize students information seeking activities.
Exploration and Familiarization (Novice/Beginner)	Beginning Search/Middle sessions (two or three activities)	4 22 23	Open Results of Previous Search View Abstracts View Websites	Advanced scaffold of Previous Search Results used to reduce cognitive load of student while searching.
Exploration and Familiarization (Novice/Beginner)	Beginning Search/Middle	5 22 23	Conduct Search View	Advanced searching scaffolds used to reduce

	Search/Middle sessions (two or three activities)	(20) May or may not include 23 and/or 20.	Abstracts View Websites Saved MYDL Site in DQ Folder	cognitive load of student while searching. **May occur at any point in sessions.**
Exploration and Familiarization (Novice/Beginner)	**Beginning Search/Middle sessions** (two to four activities)	7 8 with the addition of 9 and 10 at times.	Open Dictionary Search Tool Performed Dictionary Query on Term Opened Thesaurus Page Performed a Thesaurus Query on Term	Advanced scaffolds of Dictionary and Thesaurus used to aid student in choosing terms to search with.
Advanced Beginner/Competence (Advanced Beginner)	**Workspace Organization/Middle and End sessions** (two to four activities in iterative sequence)	17 15 or 18 Repeat 17 and 18 in sequence of two or more iterations, with the addition of 19 at times.	Open DQ Folder Delete DQ Bin Saved DQ Notes in DQ Folder Deleted Bookmark from a DQ Folder	Scaffold of DQ Folders used throughout sessions. Folders used to organize students information seeking activities.
Advanced Beginner/Competence (Advanced Beginner)	Extended Search/Middle and End sessions **(two activities. May be repeated)**	4 6	Open Results of Previous Search Performing a Search with Wordnet	Advanced scaffolds of Previous Search Results and Wordnet to choose terms used to reduce cognitive load of student while searching.
Advanced Beginner/Competence (Advanced Beginner)	**Extended Search/Middle and End sessions** (two to four activities in iterative sequence. May not include all five activities)	5 7 8 9 10 in sequence of two or more iterations and alternating 7 and 8.	Performing Search Open Dictionary Search tool Performed Dictionary Query on Term Opened Thesaurus	Advanced scaffolds of Dictionary and Thesaurus used to aid student in choosing terms to search with.

			page Performed Thesaurus Query on Term	
Advanced Beginner/Competence (Advanced Beginner)	**Extended Search/Middle and End sessions** (two to three activities in iterative sequence)	6 7 10	Performing a Search with Wordnet on Open Dictionary Search tool Performed a Thesaurus Query on Term	Advanced scaffolds of Dictionary and Thesaurus used to aid student in choosing terms to search with.
Competence/Proficiency (Beginning Expert)	**Extended Search/Middle and End sessions** (two activities in iterative sequence)	11 4	Open Past Searches Window Opened Results of Previous Search	Advanced scaffolds of Past Searches and Previous Search Results used to reduce cognitive load of student while searching.
Competence/Proficiency (Beginning Expert)	**Extended Search/Middle and End Sessions**	20 22 in sequence of two or more iterations	Saved Site from MYDL Viewed Abstract	Organizational and Searching scaffolds used to reduce cognitive load of student while searching. May occur at any point in the sessions, but most frequently at middle and end sessions.
Competence/Proficiency (Beginning Expert)	**Extended Search/End sessions** (two to three activities in iterative sequence)	17 28 (23) May or may not include 23.	Opened DQ Folder Viewed Abstracts Saved in DQ Folder Viewed Website	Scaffolds of DQ Folder and items saved in DQ Folder used to aid student in further information seeking. May occur at any point in the sessions, but most frequently in end of sessions.

Table 31 places the activity patterns in the stages of interaction of the Dreyfus (1986) and Abbas adaptation of the Skill Acquisition model. The various activity patterns and the scaffolds within each do not fade after initial use, rather the user groups tend to use particular scaffolds at the differing levels of skill acquisition. In the beginning sessions, user groups model the activities of the teachers. They then begin to explore the system and the use and frequency of the activity patterns varies. As they become more competent and proficient users, they use more of the advanced scaffolds. Finally as they become expert system users, they make use of all scaffolds with ease. It is in this stage that scaffold use may either peak or fade, depending on the user group, as the users become comfortable with the system and information seeking. These findings can assist teachers using the ADL to provide models of natural system use. Furthermore, they can be used to help less successful users of the ADL, the C-users and the Other-users to make more efficient use of the scaffolds. Further research in this area might explore the time frame it takes for the students to progress from one stage to the next, in the hopes of providing appropriate human and system scaffolds at the appropriate junctures.

One further finding of this study suggests the need for developing guidelines for the digital librarian to use when creating resource descriptions and entries in the ADL database. The original intention was to conduct an analysis of the instruments used by the librarians to ascertain representation practices within the ADL. However, upon examination of the sparse documentation (guidelines and tips to creating abstracts) used by the librarians, it became evident that these guides are used for the newest version of the ADL, and not for the system used in the study. Upon questioning the digital librarians, it was determined that this assumption was correct and no documentation was used to create the abstracts within the study's version of the ADL.

It is fundamental within library bibliographic control and cataloging practice to use standardized rules and procedures for creating representations of a collection's resources. However, the digital environment, with the exception of computerized online public access catalogs (OPACs), has caused the library and information science community and other information communities to rethink, and in some cases to re-emphasize the need for these practices to continue in the online environment. It is now a practice within the Web environment to provide little or rudimentary indexing or description of the resources. In this new digital environment, it is often difficult to follow the guidelines or rules that are fundamental to library bibliographic representations. With new developments and the awareness that information seeking and retrieval is problematic and difficult in this environment, we may see some change in this practice. However, for the ADL, using no standardized documentation or guidelines to create representations may prove problematic for the users of the system, as well as the system agents, the digital librarians. One aspect of system knowledge is an understanding of the rules used to create representations within the system. Lacking this knowledge can impact how the system agents create the representations, thereby affecting higher levels of indexer inconsistency. It can also handicap the users as they attempt to guess the terms used to represent the resources. ADL librarians should work to create guidelines to follow when developing resource descriptions and records within the ADL database.

Thoughts to Consider…Or Paths to Explore

The ADL provided the unique opportunity to examine many issues fundamental to the use of any information retrieval system. This digital library also afforded the ability to explore features unique to this specialized digital library. While the findings suggest potential enhancements with the ADL, many of the discoveries are generalizable to other information retrieval systems and to this particular user group's information seeking activities. The findings can help to enhance not only information retrieval systems, but can also serve to improve middle-school student's use of other information retrieval systems, such as the World Wide Web. Children's interactions indicate exploration but also skill acquisition. Children naturally experiment within new environments and adapt system resources to suit their own purposes. Learning more about how users naturally interact with systems can help in the design of more learner-centered systems.

While the findings have helped to further our understanding of users and their use of information retrieval systems, an important resource that was also made available within the ADL was that of the users' questions and search terms. The ADL afforded the researcher the unique environment to learn more about this user group's questions and question states. Questions and the act of questioning and question answering are not just a means for conveying a user's information need. By taking advantage of what we can *learn* from these questions and the processes of question answering, we exploit the greatest resource of all, human potential and curiosity. Learning how to build these features into a system would truly be a great discovery.

APPENDIX A

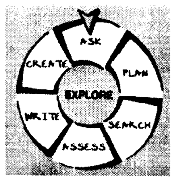
ACTIVITY SHEET 2A ASKING GOOD QUESTIONS (ASTRONOMY)

Appendix A: Activity Sheet 2A Asking Good Questions (Astronomy)

Asking Good Questions (Astronomy)

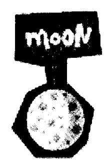
Name:____________________________

(Activity Sheet 2A)



What kind of question can you ask during the astronomy unit? First of all, ask something that is interesting to you. Start with an open-ended question -- we call this a **Driving Question.**

Think about these questions:
1) How far away is the moon from the earth?
2) Could the sun burn out one day?



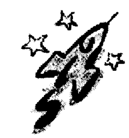
These questions are **not** really Driving Questions. They may be interesting, but they are not BIG. Good Driving Questions have much more than a sentence or paragraph answer; they are supposed to be BIG!

Driving means that it "drives" your research. A good Driving Question has a lot of depth, or might be complicated. You will need to look in a variety of places to locate information and create an answer in your own words. You will probably **never** find an answer to a good question on a single page! You will have to work at it!

Try changing the two examples into good Driving Questions:
- ____________________________________
- ____________________________________

Sometimes you might need to break your Driving Question into smaller parts so you know what to look for. Using smaller questions will help you figure out what you already know, and what you need to learn!

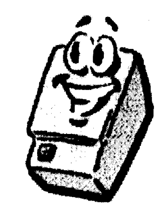
Suppose your Driving Question is:
What type of spaceship would we need to get an astronaut to Mars?
You might break it down into smaller questions like these:
- What types of spaceships are there?
- How do astronauts live in space?
- What is the surface of Mars like?
- How fast do spaceships travel?



Now you try one
Break this question down into smaller parts: What would happen to us if a large meteorite hit the earth?
- ____________________________________
- ____________________________________

Now it's your turn
Think about and write your own Driving Questions for your investigation. Try to come up with **three,** and remember good questions are interesting, BIG, and have a lot of depth.

Use Activity Sheet 2B to write and break down your Driving Questions.



Tactics and Strategies: Web Inquiry Support Unit 187

Used with permission of Hi-Ce (1998). Tactics and strategies: Leading on-line investigations. Web inquiry support manual. Ann Arbor, MI: University of Michigan.

APPENDIX B

CONTENT ANALYSIS CODING SCHEME FOR ADL ACTIVITIES

Appendix B: Content Analysis Coding Scheme for ADL Activities

Alphabetized ADL Activities	**Codes**	**ActionID**
Created a Driving Question (DQ) from the search page	CDQ-S	13
Created a Driving Question (DQ) from the View page	CDQ-V	14
Deleted a bookmark from a DQ folder	D-Bookmark	19
Deleted a DQ Bin	D-Bin	15
Deleted a past search	D-Search	12
Edited DQ	Edit-DQ	16
Login Student	LIN-S	1
Login Teacher	LIN-T	3
Logout	LOUT	2
Opened DQ folder	OPEN-DQ	17
Open Dictionary Search tool	OPEN-Dict	7
Opened Past Searches window	OPEN-PS	11
Opened results of a previous search	OPEN-RPS	4
Opened the Thesaurus page	OPEN-TH	9
Performed dictionary query on term	P-DictQuery	8
Performed a thesaurus query on term	P-TheQuery	10
Performing a search	P-Search	5
Performing a search with WordNet	P-Search-WordNet	6
Saved bookmark in DQ folder. Site not MYDL	SAVE-BK-DQ	21
Saved DQ Notes in DQ folder	SAVE-N-DQ	18

Saved site from the MYDL in DQ folder	SAVE-S-DQ	20
Shared site from the MYDL with class	SHARE-S	24
Shared site NOT from the MYDL with class	SHARE-SN	25
Viewed full website description (abstract)	VIEW-A	22
Viewed website description of site saved in DQ folder	VIEW-S	28
Viewed shared Cool Sites	VIEW-CS	26
Viewed shared DQ	VIEW-DQ	27
Viewed website	VIEW-WEB	23

APPENDIX C

SEMANTICS FOR CONTENT ANALYSIS OF SEEKING PROCESS

Appendix C: Semantics for Content Analysis for Seeking Process

Piece of Process/Code	**Semantics**
UserID/(G1, G2, G3, etc.)	number assigned/class/group
StudentID/(GoddessGirls)	alphanumeric chosen by group
Region/(R1, R2, R3, etc.)	number assigned/region
School /(S1, S2, S3, etc.)	number assigned/school
Class/(C1, C2, C3, etc.)	number assigned/class
Date	Date of session
Time	Start/end times of session
Subject	ADL subdivisions of subject content
Search Terms	student generated search terms
Result Count	system generated tabulation of results
Activity	ADL activity engaged in (refer to App. A.)
DQ Bin	Student generated Driving Question
Time Spent	System generated calculation of session time

APPENDIX D

SEARCH STRATEGIES AND FAILURE ANALYSIS CODING SCHEMES

Appendix D: Search Strategies and Failure Analysis Coding Schemes

Strategy Used	**Code**
Single term used	S1
Multiple terms used (phrase)	S2
Boolean search used	S3
Natural language (sentence structure)	S4
Other unclassified/unsupported search	S5

Failure Analysis Activity	**Code**
Misspelled term(s)	F1
Typographic error	F2
Same search repeated	F3

APPENDIX E

EXTENT OF MATCH RULES

Appendix E: Extent of Match Rules

Terms within the searches and the driving questions must follow these rules in order to be considered an exact match. Mapping can consist of any term(s) within the driving questions and the search queries.

Terms will be considered an acceptable match if at least one of the following conditions apply:

1. Root Match (RM)

If the root form of the search term(s) matches the root form of the term(s) within the driving question this term is considered a match. For example, if the driving question is: "Do electro magnetic fields cause tumors in plants?" and the search term is "electromagnetic", THEN the term is considered a match.

2. Inverted Order Match (IOM)

If the root form of the search term(s) is present but the terms are in inverse order this term is acceptable as a match. For example if the driving question is: "Do electromagnetic fields cause plants to wilt?" and the search term is "fields electromagnetic" THEN the term is considered a match.

3. Combined Match (CM)

If the root form of the word(s) is present but the term(s) are combined within a Boolean search statement, this term is acceptable as a match. For example, if the driving question is: "Can plants live when exposed to electromagnetic fields?" and the search term is "electromagnetic AND fields" THEN the term is considered a match.

4. Phrase Match (PM)

If the search term appears within a phrase (as a string of terms separated by a space on both sides) then the root form of each term must match the root form of the phrase within the driving question. For example, if the driving question is: "How long will plants live once exposed to an electromagnetic field?" and the search phrase is "electromagnetic field" THEN the phrase is considered a match.

5. No Match-Synonymous (NM-S)

If the search term(s) does not appear in the driving question, but a synonym of the search term(s) does appear, THEN the term is NOT considered a match, but will be designated as No Match-Synonymous (NM-S).

6. No Match (NM)

If the search term(s) does not appear in the driving question, and no synonym of the search term(s) is present, THEN the term is NOT considered a match.

Mapping of Search Terms to Keywords and Words within the Abstracts

Mapping of the search terms and the keywords and the terms within the abstracts will follow the same rules as outlined above. Each field of Keyword and Abstract will be compared separately and individual figures compiled.

APPENDIX F

Use Analysis for All Classes

Use Analysis for All Classes

Activity and ActionID by Class	C1AUses	C1AMeanUse	C1BUses	C1BMeanUse
Login (1)	57	5.7	327	4.19
Logout (2)	36	3.6	112	1.44
Login teacher (3)			0	0.00
Open Results of Previous Search (4)	1	0.1	3	0.09
Performing a Search (5)	241	24.1	960	12.31
Performing a Search with Wordnet (6)	3	0.3	4	0.06
Open Dictionary Search Tool (7)	11	1.1	38	0.58
Performed Dictionary query on term (8)	14	1.4	83	1.26
Opened Thesaurus page (9)	1	0.1	8	0.11
Performed thesaurus query on term (10)	1	0.1	9	0.13
Open Past Searches Window (11)	4	0.4	19	0.26
Deleted a Past Search (12)			5	0.13
Created DQ from Search Page (13)	50	5	128	1.64
Created DQ from View Page (14)			0	0.00
Deleted a DQ Bin (15)	17	1.7	20	0.30
Edited DQ (16)	3	0.3	8	0.12
Opened DQ Folder (17)	70	7	166	2.31
Saved Notes in DQ Folder (18)	13	1.3	16	0.48
Deleted a Bookmark from a DQ Folder (19)	2	0.2	21	0.32
Saved Site from the MYDL in DQ Folder (20)	32	3.2	131	1.68
Saved Bookmark in DQ Folder. Site not MYDL (21)			0	0.00
Viewed Full Website Description (22)	235	23.5	985	12.63
Viewed Website (23)	131	13.1	434	5.56
Shared Site from the MYDL with class (24)			5	0.08
Shared Site NOT from the MYDL with class (25)			0	0.00
Viewed Shared Cool Sites (26)	2	0.2	25	0.38
Viewed Shared DQ (27)			7	0.11
Viewed Website Description of Site Saved in DQ Folder (28)	6	0.6	41	0.62

C1CUses	C1CMeanUses	C1DUses	C1DMeanUses	SUMS	C2AUses	C2AMeanUses	C2BUses	C2BMeanUses
312	3.68	18	3.00	714.00	479	5.38	5	2.5
139	1.57	7	1.17	294.00	291	3.27	4	2
0		0		0.00	0	0.00	0	0
5	0.11	0		9.00	16	0.18	0	0
1009	11.21	49	8.17	2259.00	897	10.08	8	4
24	0.71	0		31.00	11	0.14	0	0
32	0.43	2	0.33	83.00	21	0.24	0	0
79	0.79	3	0.50	179.00	37	0.42	0	0
8	0.07	0		17.00	4	0.09	0	0
7	0.04	0		17.00	3	0.07	0	0
71	1.39	0		94.00	32	0.36	0	0
40	0.46	0		45.00	0	0.00	0	0
125	1.32	19	3.17	322.00	207	2.33	10	5
0		0		0.00	0	0.00	0	0
18	0.29	1	0.17	56.00	23	0.26	2	1
6	0.04	1	0.17	18.00	31	0.40	0	0
134	1.46	4	0.67	374.00	1246	14.00	10	5
10	0.14	0		39.00	611	6.87	11	5.5
5	0.04	0		28.00	139	1.56	0	0
113	0.89	5	0.83	281.00	865	9.72	6	3
0		0		0.00	5	0.08	0	0
925	10.25	47	7.83	2192.00	1411	15.85	12	6
369	4.14	9	1.50	943.00	472	5.30	1	0.5
3	0.04	1	0.17	9.00	31	0.35	1	0.5
0		0		0.00	0	0.00	0	0
16	0.39	2	0.33	45.00	21	0.27	0	0
7	0.11	5	0.83	19.00	8	0.10	0	0
29	0.07	1	0.17	77.00	209	2.35	0	0
3486	39.64	174	29.00	8145.00	7070	79.44	70	35

C2CUses	C2CMeanUses	C2FUses	C2FMeanUses	SUMS	C3AUses	C3AMeanUses	C3BUses	C3BMeanUses	C3CUses
44	2.44	27	2.70	555.00	108	5.68	41	6.83	44
34	1.89	16	1.60	345.00	56	2.95	20	3.33	20
0		0	0.00	0.00			0		0
6	0.33	2	2.00	24.00	1	0.05	3	0.50	2
136	7.56	102	6.38	1143.00	278	14.63	86	14.33	108
3	0.17	4	2.00	18.00	1	0.05	0		1
0		2	1.00	23.00	6	0.32	3	0.50	2
0		2	1.00	39.00		0.37	8	1.33	2
0		1	1.00	5.00			1	0.17	1
0		1	1.00	4.00			1	0.17	0
10	0.56	3	3.00	45.00	2	0.11	6	1.00	10
0		0	0.00	0.00			0		4
49	2.72	19	2.11	285.00	26	1.37	7	1.17	10
0		1	1.00	1.00			0		0
14	0.78	1	1.00	40.00	4	0.21	0		1
3	0.17	3	1.50	37.00	7	0.37	6	1.00	5
171	9.50	21	3.50	1448.00	101	5.32	41	6.83	47
76	4.22	23	4.60	721.00	54	2.84	24	4.00	35
52	2.89	2	2.00	193.00	2	0.11	1	0.17	2
132	7.33	23	7.67	1026.00	15	0.79	7	1.17	6
2	0.11	0	0.00	7.00			1	0.17	0
116	6.44	73	6.08	1612.00	254	13.37	77	12.83	136
22	1.22	18	2.25	513.00			0		0
11	0.61	1	1.00	44.00			0		0
0		0	0.00	0.00	4		0		0
3	0.17	0	0.00	24.00	36	0.21	2	0.33	5
2	0.11	4	2.00	14.00	36	1.89	11	1.83	13
9	0.50	4	2.00	222.00	20	1.05	8	1.33	6
895	49.72	353	19.61	8388.00	982	51.68	354	59.00	460

C3CMeanUses	C3UUses	C3UMeanUses	SUMS	C4AUses	C4AMeanUses	C4BUses	C4BMeanUses	C4CUses	C4CMeanUses
6.29	84	4.42	277.00	81	5.79	125	5	19	1.58
2.86	63	3.32	159.00	10	0.71	27	1.08	3	0.25
	0		0.00	0	0.00	0	0	0	0.00
0.29	0		6.00	8	0.57	3	0.12	0	0.00
15.43	169	8.89	641.00	229	16.36	299	11.96	76	6.33
0.14	3	0.16	5.00	18	1.29	6	0.24	3	0.25
0.29	1	0.05	12.00	14	1.00	6	0.24	1	0.08
0.29	1	0.05	11.00	89	6.36	31	1.24	3	0.25
0.14	0		2.00	1	0.07	1	0.04	0	0.00
1.43	0		1.00	1	0.07	1	0.04	0	0.00
0.57	2	0.11	20.00	10	0.71	10	0.4	3	0.25
1.43	0		4.00	3	0.21	3	0.12	9	0.75
	31	1.63	74.00	37	2.64	70	2.8	34	2.83
0.14	0		0.00	0	0.00	2	0.08	1	0.08
0.71	5	0.26	10.00	14	1.00	19	0.76	6	0.50
6.71	9	0.47	27.00	5	0.36	9	0.36	2	0.17
5.00	109	5.74	298.00	26	1.86	45	1.8	15	1.25
0.29	44	2.32	157.00	1	0.07	10	0.4	0	0.00
0.86	2	0.11	7.00	1	0.07	3	0.12	0	0.00
	17	0.89	45.00	31	2.21	25	1	17	1.42
19.43	0		1.00	0	0.00	0	0	1	0.08
	163	8.58	630.00	178	12.71	305	12.2	110	9.17
	0		0.00	58	4.14	53	2.12	15	1.25
	0		0.00	2	0.14	19	0.76	3	0.25
0.71	0		4.00	0	0.00	0	0	0	0.00
1.86	5	0.26	48.00	12	0.86	50	2	15	1.25
0.86	38	2.00	98.00	12	0.86	40	1.6	14	1.17
65.71	26	1.37	60.00	16	1.14	6	0.24	1	0.08
40.63	772	40.63	2568.00	857	61.21	1168	46.72	351	29.25

SUMS
225.00
40.00
0.00
11.00
604.00
27.00
21.00
123.00
2.00
2.00
23.00
15.00
141.00
3.00
39.00
16.00
86.00
11.00
4.00
73.00
1.00
593.00
126.00
24.00
0.00
77.00
66.00
23.00
2376.00

APPENDIX G

Time Analysis for All Classes

Appendix G: Time Analysis for All Classes

Activity and ActionID by Class	Class1TotalTime/sec	C1Mode	C1Mean	Class2TotalTime/sec	C2Mode	C2Mean
Login (1)	56935	0	84.97761	31473	1	59.8346
Logout (2)	292189	147	3075.674	112078	1	1205.14
Login teacher (3)	82719	#N/A	4353.632	100808	#N/A	4382.957
Open Results of Previous Search (4)	260	#N/A	28.88889	644	13	26.83333
Performing a Search (5)	85724	0	43.12072	54417	0	54.52605
Performing a Search with Wordnet (6)	951	8	35.22222	318	9	18.70588
Open Dictionary Search Tool (7)	3265	23	43.53333	828	22	39.42857
Performed Dictionary query on term (8)	10766	1	74.76389	6924	0	197.8286
Opened Thesaurus page (9)	676	#N/A	45.06667	141	#N/A	35.25
Performed thesaurus query on term (10)	965	1	1	89	#N/A	29.66667
Open Past Searches Window (11)	7280	6	80.88889	8426	6	168.52
Deleted a Past Search (12)	132	2	2.933333	0	0	0
Created DQ from Search Page (13)	40542	27	151.2761	13448	0	57.47009
Created DQ from View Page (14)	0	0	0	2	#N/A	0
Deleted a DQ Bin (15)	2785	18	54.60784	1928	10	49.4359
Edited DQ (16)	502	22	29.52941	653	2	21.76667
Opened DQ Folder (17)	19798	11	55.30168	110163	1	84.41609
Saved Notes in DQ Folder (18)	1691	2	49.73529	35885	1	51.70749
Deleted a Bookmark from a DQ Folder (19)	375	0	11.36364	512	0	3.065868
Saved Site from the MYDL in DQ Folder (20)	26779	9	110.2016	35147	12	41.44693
Saved Bookmark in DQ Folder. Site not MYDL (21)	0	0	0	27	3	3.857143
Viewed Full Website Description (22)	366001	4	194.8887	106720	3	72.69755
Viewed Website (23)	241708	37	325.752	85216	1	204.8462
Shared Site from the MYDL with class (24)	678	#N/A	61.63636	789	11	24.65625
Shared Site NOT from the MYDL with class (25)	0	0	0	0	0	0
Viewed Shared Cool Sites (26)	14098	#N/A	391.6111	2031	0	81.24
Viewed Shared DQ (27)	1358	26	113.1667	1240	45	59.04762
Viewed Website Description of Site Saved in DQ Folder (28)	5300	3	71.62162	15316	2	84.61878
Total activities	1263477			725223		

Class3TotalTime/sec	C3Mode	C3Mean	Class4TotalTime/sec	C4Mode	C4Mean
43003	10	338.6063	27549	8	85.82243
99515	0	6634.333	53132	6	4830.182
0	0	0	0	0	0
155	#N/A	25.83333	248	1	17.71429
18960	8	30.72934	62587	9	85.97115
88	#N/A	44	1411	19	38.13514
181	#N/A	20.11111	1268	29	43.72414
27895	#N/A	1743.438	49658	5	379.0687
19	#N/A	9.5	80	#N/A	26.66667
33	#N/A	33	49	#N/A	24.5
588	9	28	13614	1	340.35
3	1	0.75	24	2	1.6
2443	3	38.77778	7286	1	45.8239
0	0	0	3	1	1.4
201	19	22.33333	5647	4	117.6458
744	4	31	1414	60	54.38462
44642	3	157.7456	5458	7	40.13235
10878	15	74.50685	306	1	18
6	0	0.857143	68	1	5.666667
12184	2	297.1707	4128	10	59.82609
0	0	0	1	0	0.2
125115	0	213.5068	67188	3	92.41816
0	0	0	57396	170	257.3812
0	0	0	2582	#N/A	122.9524
0	0	0	0	0	0
3957	5	989.25	14286	56	207.0435
11395	42	130.977	5560	97	85.53846
14764	81	278.566	19840	10	734.8148
416769			400783		

APPENDIX H

ACTIVITY PATTERNS OF GRADED USER GROUPS

Appendix H: Activity Patterns of Graded User Groups

Table G1. Activity Patterns of Graded User Groups Divided by Class Number

Class#	Graded Group	Activity Pattern	Occurrence	Scaffold(s) Category	Frequency of Use
1	A-users	13 2 1 5	At beginning of first session.	Organizational and beginning Search	Only used by 3 user groups
1	A-users	13 5 (22)	At any point in sessions. May or may not include (22)	Organizational and beginning Searching	Used by majority of user groups
1	A-users	17 15 or 18 Repeat 17 and 18 in sequence of two or more iterations with the addition of 19 for some	At the end of search sessions, but with a few occurrences of use in middle of sessions.	Organizational	Used by all user groups, with the variance of using 19.
1	A-users	5 22 23 (20)	At any point in sessions. May or may not include (20)	Searching	Most frequent activity pattern. Used by all user groups, but not all used (20).
1	A-users	5 7 8 9 10 in sequence of two or more iterations and alternating 7 and 8.	At the end of search sessions, but with a few occurrences of use in middle of sessions.	Advanced Searching	Used by only one user group in this group.
1	B-users	13 2 1 5	At the beginning of sessions.	Organizational and beginning search	Used by all but eight groups
1	B-users	13 5 (22)	At any point in sessions. May or may not include 22.	Organizational and Searching	Used by approximately half of user groups.
1	B-users	17 15 or 18 Repeat 17 and 18 in sequence of two or more iterations with the addition of 19 for some	At the end of search sessions, but with a few occurrences of use in middle of sessions.	Organizational	Used by all user groups, with the variance of using 19.
1	B-users	17 28 23	At the end of sessions.	Organizational and Searching	Only two user groups used this pattern.

1	B-users	5 22 23 (20)	At any point in sessions. May or may not include 20.	Searching	Most frequently used activity pattern. Used by all user groups. Only eleven groups used 20.
1	B-users	7 8 9 10 in sequence of two or more iterations and alternating 7 and 8.	At the end of sessions.	Advanced Searching	Used by only five user groups in this group.
1	C-users	13 2 1 5	At the beginning of the sessions.	Organizational and beginning Search.	Used by only three user groups in this group.
1	C-users	13 5 (22)	At any point in sessions. Most included 22.	Organizational and Searching	Used by approximately half of user groups.
1	C-users	13 15	At the end of sessions.	Organizational and Maintenance	Used by only two user groups.
1	C-users	17 28	At middle or end of sessions.	Organizational	Used by only one user group, but repeated often.
1	C-users	17 19	At the middle or end of sessions.	Organizational	Used by only two user groups.
1	C-users	4 22 23	At the middle of the sessions.	Advanced Search and Search	Used by only one user group, but repeated often.
1	C-users	5 22 23 (20)	At any point in sessions. May or may not include 20.	Searching	Most frequently used activity pattern. Used by all user groups. Only six groups used 20.
1	C-users	5 26	At any point in sessions.	Searching and Collaborative	Used by only five of the user groups.
1	C-users	7 8	At the middle of sessions.	Advanced Searching	Used by only three

		9 10 in sequence of two or more iterations and alternating 7 and 8 and 9 and 10.			user groups in this group.
1	C-users	8 22 23	At the middle or end of sessions.	Advanced Searching and Searching	Used by only two user groups, but repeated often.
1	C-users	8 26	At the middle of sessions.	Advanced Searching and Collaborative	Used by only one user group.
1	D-users	13 5	At the middle or end of sessions.	Organizational and beginning Search	Used by only half of user groups, and not as starting activity.
1	D-users	17 18	At the middle of sessions.	Organizational	Only used by one user group.
1	D-users	5 22 (23)	At any point in sessions.	Searching	Used by all groups, but 23 only used by one user group.
2	A-users	13 5 (22)	Most often at beginning of sessions, but may occur at any point in sessions. May or may not include (22).	Organizational and beginning Searching	Used by majority of user groups.
2	A-users	17 18 19 in sequence of two or more iterations and alternating 17 and 18, with 19 used less frequently	At any point in sessions.	Organizational	Used by majority of user groups.
2	A-users	17 22 20	At any point in sessions.	Organizational and Searching	Used by majority of user groups.
2	A-users	17 28	At any point in sessions.	Organizational and Searching	Used by approximately half of user groups.
2	A-users	20 22 in sequence of two or more iterations.	At any point in sessions.	Searching	Used by few user groups.
2	A-users	4 6	At middle or end of sessions.	Advanced Searching	Used by few user groups.
2	A-users	5 22	At any point in sessions. May or	Searching	Most frequent

		23 (20)	may not include (20). May use 20 and not 23.		activity pattern. Used by all user groups, but not all used (20).
2	A-users	5 26	At end of sessions.	Searching and Collaborative	Used by few user groups.
2	A-users	7 8 in sequence of two or more iterations.	At the middle of the sessions.	Advanced Searching	Used by few user groups.
2	B-users	11 13	At beginning of sessions.	Advanced Searching	Used by only one user group.
2	B-users	13 5 (22)	Most often at beginning of sessions, but may occur at any point in sessions. May or may not include (22).	Organizational and beginning Searching	Used by majority of user groups.
2	B-users	17 18 19 in sequence of two or more iterations and alternating 17 and 18, with 19 used less frequently	At any point in sessions.	Organizational	Used by majority of user groups.
2	B-users	5 22 (20)	At any point in sessions. May or may not include 20. **Did not include 23 as other groups did.**	Searching	Most frequently used activity pattern. Used by all user groups. Half of user groups included 20.
2	B-users	5 26	At end of sessions.	Searching and Collaborative	Used by only one user group.
2	C-users	5 22 (20)	At any point in sessions. May or may not include 20. **Did not include 23 as other groups did.**	Searching	Most frequently used activity pattern. Used by all user groups. Half of user groups included 20.
2	C-users	5 26	At end of sessions.	Searching and Collaborative	Used by only one user group.
2	F-users	13 5 or 13 alone	At beginning of sessions.	Organizational and beginning Searching	Used by only three user groups.

2	F-users	17 18 Repeat 17 and 18 in sequence of two or more iterations with the addition of 19 for some.	At the middle of search sessions.	Organizational	Used by two user groups, with the variance of using 19.
2	F-users	17 28	At the end of sessions.	Organizational	Used by one user group.
2	F-users	5 22 (20) (23)	At any point in sessions. May or may not include 20 or 23.	Searching	Most frequently used activity pattern. Used by all user groups. Half of user groups included 20 or 23.
3	A-users	1 13 (5)	At the beginning of sessions. May or may not be followed by 5.	Organizational	Used by half of user groups.
3	A-users	17 28	At any point in the sessions.	Organizational	Used by only two user groups.
3	A-users	17 18 Repeat 17 and 18 in sequence of two or more iterations. **Did not use 19 as other groups did.**	At any point in the sessions.	Organizational	Most frequently used activity pattern in this group of users.
3	A-users	5 22	At any point in the sessions. **Did not include 23 or 20 as other groups did.**	Searching	Second most frequently used activity pattern. Used by all user groups.
3	A-users	7 8	At middle or end of sessions.	Advanced Searching	Only used by two user groups.
3	B-users	1 13	At the beginning of sessions. **Was not be followed by 5 as seen in other user groups.**	Organizational	Used by majority of user groups.
3	B-users	17 28	At any point in the sessions.	Organizational	Used by only two user groups.
3	B-users	17 18 Repeat 17 and 18 in sequence of two or more iterations, with	At any point in the sessions.	Organizational	Used by all user groups.

		the addition of 19 for some.			
3	B-users	5 22	At any point in the sessions. **Did not include 23 but some use of 20.**	Searching	Most frequently used activity pattern. Used by all user groups.
3	B-users	7 8	At middle or end of sessions.	Advanced Searching	Only used by one user group.
3	C-users	17 18 Repeat 17 and 18 in sequence of two or more iterations. **Did not use 19 as did other user groups.**	At any point in the sessions.	Organizational	Used by all user groups.
3	C-users	17 28	At any point in the sessions.	Organizational	Used by only one user group.
3	C-users	5 OR 5 22	**Sessions began with searching instead of other activities.** Also used at any other point in sessions. **Did not include 23 or 20.**	Searching	Most frequently used activity pattern. Used by all user groups.
3	U-users	1 13	At the beginning of sessions. **Was not be followed by 5 as seen in other user groups.**	Organizational	Used by majority of user groups.
3	U-users	13 26 or the inverse	At the beginning of the sessions.	Organizational and Collaborative	Used by half of the user groups.
3	U-users	17 28	At any point in the sessions.	Organizational	Used by only three user groups.
3	U-users	17 18 Repeat 17 and 18 in sequence of two or more iterations, with the addition of 19 for some.	At any point in the sessions.	Organizational	Used by majority of user groups.
3	U-users	5 22	At any point in the sessions. **Did not include 23 but some use of 20.**	Searching	Most frequently used activity pattern. Used by all user groups.
4	A-users	13 5	At any point in sessions. May or	Organizational and Searching	Used by majority of

		(22)	may not include 5 and/or 22.		user groups.
4	A-users	17 18 Repeat 17 and 18 in sequence of two or more iterations with the addition of 19 for some.	At the end of search sessions, but with a few occurrences of use in middle of sessions.	Organizational	Used by all user groups, with the variance of using 19.
4	A-users	17 28	At the end of sessions.	Organizational and Searching	Only two user groups used this pattern.
4	A-users	5 22 (23) (20)	At any point in sessions. May or may not include 20 or 23.	Searching	Most frequently used activity pattern. Used by all user groups.
4	A-users	6 8 11 7 9 4 10	At the middle or the end of sessions.	Advanced Searching	Used by half of user groups, but used frequently.
4	B-users	13 5 (22)	At any point in sessions. Majority of all user groups started with 13. May or may not include 5 and/or 22.	Organizational and Searching	Used by majority of user groups.
4	B-users	17 18 Repeat 17 and 18 in sequence of two or more iterations. Minor use of 19.	At the end of search sessions, but with a few occurrences of use in middle of sessions.	Organizational	Used by only four user groups.
4	B-users	26 27 26 in sequence with two or more iterations OR a series of 26's with or without 24's.	At any point in the sessions.	Collaborative	Used by the majority of user groups.
4	B-users	5 22 (23) (20)	At any point in sessions. May or may not include 20 or 23.	Searching	Most frequently used activity pattern. Used by all user groups.
4	B-users	7 8	At middle or end of sessions.	Advanced Searching	Used by only one user group.
4	C-users	13 5 (27)	At any point in sessions. **Majority of all user groups**	Organizational and Searching	Used by majority of user groups.

			started with 13. May or may not include 5 and/or 27.		
4	C-users	17 with 18 in few sequences. **Did not use 19 as others did.**	At any point in the sessions.	Organizational	Used by only four user groups.
4	C-users	26 27 26 in sequence with two or more iterations OR a series of 26's.	At any point in the sessions.	Collaborative	Used by only two user groups.
4	C-users	5 22 (23)	At any point in sessions. May or may not include 20 or 23.	Searching	Most frequently used activity pattern. **Only one user group included 23. No instances of 20.**

Table G2. Beginning Activity Patterns All Graded Groups Included

Group#	Graded Group	Activity Pattern	Occurence	Scaffold Category	Frequency of Use
3	A-users	1 13 (5)	At the beginning of sessions. May or may not be followed by 5.	Organizational	Used by half of user groups.
3	B-users	1 13	At the beginning of sessions. **Was not be followed by 5 as seen in other user groups.**	Organizational	Used by majority of user groups.
3	U-users	1 13	At the beginning of sessions. **Was not be followed by 5 as seen in other user groups.**	Organizational	Used by majority of user groups.
2	B-users	11 13	At beginning of sessions.	Advanced Searching	Used by only one user group.
1	A-users	13 2 1 5	At the beginning of first session.	Organizational and beginning Search	Only used by 3 user groups
1	B-users	13 2 1	At the beginning of first session.	Organizational and beginning search	Used by all but eight

		5			groups
1	C-users	13 2 1 5	At the beginning of first session.	Organizational and beginning Search.	Used by only three user groups in this group.
2	A-users	13 5 (22)	Most often at beginning of sessions, but may occur at any point in sessions. May or may not include 22.	Organizational and beginning Searching	Used by majority of user groups.
2	B-users	13 5 (22)	Most often at beginning of sessions, but may occur at any point in sessions. May or may not include 22.	Organizational and beginning Searching	Used by majority of user groups.
2	F-users	13 5 or 13 alone	At beginning of sessions.	Organizational and beginning Searching	Used by only three user groups.
3	U-users	13 26 or the inverse	At the beginning of the sessions.	Organizational and Collaborative	Used by half of the user groups.
3	C-users	5 OR 5 22	**Sessions began with searching instead of other activities.** Also used at any other point in sessions. **Did not include 23 or 20.**	Searching	Most frequently used activity pattern. Used by all user groups.

Table G2 shows the beginning activities are characterized by the creation of DQs and beginning searches.

Table G3. Middle Activity Patterns All Graded Groups Included

Group #	Graded Group	Activity Pattern	Occurrence	Scaffold Category	Frequency of Use
1	D-users	17 18	At the middle of sessions.	Organizational	Only used by one user group.
2	F-users	17 18 Repeat 17 and 18 in sequence of two or more	At the middle of search sessions.	Organizational	Used by two user groups, with the variance of using 19.

		iterations with the addition of 19 for some.			
1	C-users	4 22 23	At the middle of the sessions.	Advanced Search and Search	Used by only one user group, but repeated often.
1	C-users	7 8 9 10 in sequence of two or more iterations and alternating 7 and 8 and 9 and 10.	At the middle of sessions.	Advanced Searching	Used by only three user groups in this group.
2	A-users	7 8 in sequence of two or more iterations.	At the middle of the sessions.	Advanced Searching	Used by few user groups.
1	C-users	8 26	At the middle of sessions.	Advanced Searching and Collaborative	Used by only one user group.

Table G3 shows the middle sessions contained instances of Advanced scaffold use and Organizational scaffolds.

Table G4. Middle/End Activity Patterns All Graded Groups Included

Group #	Graded Group	Activity Pattern	Occurrence	Scaffold Category	Frequency of Use
1	D-users	13 5	At the middle or end of sessions.	Organizational and beginning Search	Used by only half of user groups, and not as starting activity.
1	A-users	17 15 or 18 Repeat 17 and 18 in sequence of two or more iterations with the addition of 19 for some	At the end of search sessions, but with a few occurrences of use in middle of sessions.	Organizational	Used by all user groups, with the variance of using 19.
1	B-users	17 15 or 18 Repeat 17 and 18 in sequence of two or more iterations with the addition of 19 for some	At the end of search sessions, but with a few occurrences of use in middle of sessions.	Organizational	Used by all user groups, with the variance of using 19.
4	A-users	17	At the end of	Organizational	Used by all

		18 Repeat 17 and 18 in sequence of two or more iterations with the addition of 19 for some.	search sessions, but with a few occurrences of use in middle of sessions.		user groups, with the variance of using 19.
4	B-users	17 18 Repeat 17 and 18 in sequence of two or more iterations. Minor use of 19.	At the end of search sessions, but with a few occurrences of use in middle of sessions.	Organizational	Used by only four user groups.
1	C-users	17 19	At the middle or end of sessions.	Organizational	Used by only two user groups.
1	C-users	17 28	At middle or end of sessions.	Organizational	Used by only one user group, but repeated often.
2	A-users	4 6	At middle or end of sessions.	Advanced Searching	Used by few user groups.
1	A-users	5 7 8 9 10 in sequence of two or more iterations and alternating 7 and 8.	At the end of search sessions, but with a few occurrences of use in middle of sessions.	Advanced Searching	Used by only one user group in this group.
4	A-users	6 8 11 7 9 4 10	At the middle or the end of sessions.	Advanced Searching	Used by half of user groups, but used frequently.
3	A-users	7 8	At middle or end of sessions.	Advanced Searching	Only used by two user groups.
3	B-users	7 8	At middle or end of sessions.	Advanced Searching	Only used by one user group.
4	B-users	7 8	At middle or end of sessions.	Advanced Searching	Used by only one user group.
1	C-users	8 22	At the middle or end of	Advanced Searching and	Used by only two user

		23	sessions.	Searching	groups, but repeated often.

Table G4 shows the middle/end sessions also contained more instances of Advanced scaffolds and Organizational scaffolds.

Table G5. End Activity Patterns All Graded Groups Included

Group #	Graded Group	Activity Pattern	Occurrence	Scaffold Category	Frequency of Use
1	C-users	13 15	At the end of sessions.	Organizational and Maintenance	Used by only two user groups.
1	B-users	17 28 23	At the end of sessions.	Organizational and Searching	Only two user groups used this pattern.
2	F-users	17 28	At the end of sessions.	Organizational	Used by one user group.
4	A-users	17 28	At the end of sessions.	Organizational and Searching	Only two user groups used this pattern.
2	A-users	5 26	At end of sessions.	Searching and Collaborative	Used by few user groups.
2	B-users	5 26	At end of sessions.	Searching and Collaborative	Used by only one user group.
2	C-users	5 26	At end of sessions.	Searching and Collaborative	Used by only one user group.
1	B-users	7 8 9 10 in sequence of two or more iterations and alternating 7 and 8.	At the end of sessions.	Advanced Searching	Used by only five user groups in this group.

Table G5 illustrates the ending sessions contained more Organizational and Searching scaffolds, more Advanced scaffolds and Collaborative scaffolds. There are also activity patterns that occurred at all stages of interaction. Table G6 shows these activity patterns.

Table G6. Any Point in Session Activity Patterns All Graded Groups Included

Group #	Graded Group	Activity Pattern	Occurrence	Scaffold Category	Frequency of Use
1	A-users	13 5 (22)	At any point in sessions. May or may not include (22)	Organizational and beginning Searching	Used by majority of user groups
1	B-users	13	At any point	Organizational	Used by

		5 (22)	in sessions. May or may not include 22.	and Searching	approximately half of user groups.
1	C-users	13 5 (22)	At any point in sessions. Most included 22.	Organizational and Searching	Used by approximately half of user groups.
2	A-users	13 5 (22)	Most often at beginning of sessions, but **may occur at any point in sessions**. May or may not include (22).	Organizational and beginning Searching	Used by majority of user groups.
2	B-users	13 5 (22)	Most often at beginning of sessions, but may occur at any point in sessions. May or may not include (22).	Organizational and beginning Searching	Used by majority of user groups.
4	A-users	13 5 (22)	At any point in sessions. May or may not include 5 and/or 22.	Organizational and Searching	Used by majority of user groups.
4	B-users	13 5 (22)	At any point in sessions. Majority of all user groups started with 13. May or may not include 5 and/or 22.	Organizational and Searching	Used by majority of user groups.
4	C-users	13 5 (27)	At any point in sessions. **Majority of all user groups started with 13.** May or may not include 5 and/or 27.	Organizational and Searching	Used by majority of user groups.
2	A-users	17 18 19 in sequence of two or more iterations and alternating 17	At any point in sessions.	Organizational	Used by majority of user groups.

		and 18, with 19 used less frequently			
2	B-users	17 18 19 in sequence of two or more iterations and alternating 17 and 18, with 19 used less frequently	At any point in sessions.	Organizational	Used by majority of user groups.
3	A-users	17 18 Repeat 17 and 18 in sequence of two or more iterations. **Did not use 19 as other groups did.**	At any point in the sessions.	Organizational	Used by majority of user groups.
3	B-users	17 18 Repeat 17 and 18 in sequence of two or more iterations, with the addition of 19 for some.	At any point in the sessions.	Organizational	Used by all user groups.
3	C-users	17 18 Repeat 17 and 18 in sequence of two or more iterations. **Did not use 19 as did other user groups.**	At any point in the sessions.	Organizational	Used by all user groups.
3	U-users	17 18 Repeat 17 and 18 in sequence of two or more iterations, with the addition of 19 for some.	At any point in the sessions.	Organizational	Used by majority of user groups.
4	C-users	17 with 18 in	At any point in the	Organizational	Used by only four user

		few sequences. **Did not use 19 as others did.**	sessions.		groups.
2	A-users	17 22 20	At any point in sessions.	Organizational and Searching	Used by majority of user groups.
2	A-users	17 28	At any point in sessions.	Organizational and Searching	Used by approximately half of user groups.
3	A-users	17 28	At any point in the sessions.		
3	B-users	17 28	At any point in the sessions.	Organizational	Used by only two user groups.
3	C-users	17 28	At any point in the sessions.	Organizational	Used by only one user group.
3	U-users	17 28	At any point in the sessions.	Organizational	Used by only three user groups.
2	A-users	20 22 in sequence of two or more iterations.	At any point in sessions.	Searching	Used by few user groups.
4	B-users	26 27 26 in sequence with two or more iterations OR a series of 26's with or without 24's.	At any point in the sessions.	Collaborative	Used by the majority of user groups.
		26 27 26 in sequence with two or more iterations OR a series of 26's.	At any point in the sessions.	Collaborative	Used by only two user groups.
1	A-users	5 22 23 (20)	At any point in sessions. May or may not include (20)	Searching	Most frequent activity pattern. Used by all user groups, but not all used (20).

1	B-users	5 22 23 (20)	At any point in sessions. May or may not include 20.	Searching	Most frequently used activity pattern. Used by all user groups. Only eleven groups used 20.
1	C-users	5 22 23 (20)	At any point in sessions. May or may not include 20.	Searching	Most frequently used activity pattern. Used by all user groups. Only six groups used 20.
2	A-users	5 22 23 (20)	At any point in sessions. May or may not include (20). May use 20 and not 23.	Searching	Most frequent activity pattern. Used by all user groups, but not all used (20).
1	D-users	5 22 (23)	At any point in sessions.	Searching	Used by all groups, but 23 only used by one user group.
2	B-users	5 22 (20)	At any point in sessions. May or may not include 20. **Did not include 23 as other groups did.**	Searching	Most frequently used activity pattern. Used by all user groups. Half of user groups included 20.
2	C-users	5 22 (20)	At any point in sessions. May or may not include 20. **Did not include 23 as other groups did.**	Searching	Most frequently used activity pattern. Used by all user groups. Half of user groups included 20.
2	F-users	5 22 (20) (23)	At any point in sessions. May or may not include 20 or 23.	Searching	Most frequently used activity pattern. Used by all user groups. Half of user groups included 20 or 23.
3	A-users	5 22	At any point in the sessions. **Did not include 23 or 20 as other groups**		

			did.		
3	B-users	5 22	At any point in the sessions. **Did not include 23 but some use of 20.**	Searching	Most frequently used activity pattern. Used by all user groups.
3	C-users	5 OR 5 22	Sessions began with searching instead of other activities. Also used at **any other point in** sessions. **Did not include 23 or 20.**	Searching	Most frequently used activity pattern. Used by all user groups.
3	U-users	5 22	At any point in the sessions. **Did not include 23 but some use of 20.**	Searching	Most frequently used activity pattern. Used by all user groups.
4	A-users	5 22 (23) (20)	At any point in sessions. May or may not include 20 or 23.	Searching	Most frequently used activity pattern. Used by all user groups.
4	B-users	5 22 (23) (20)	At any point in sessions. May or may not include 20 or 23.	Searching	Most frequently used activity pattern. Used by all user groups.
4	C-users	5 22 (23)	At any point in sessions. May or may not include 20 or 23.	Searching	Most frequently used activity pattern. **Only one user group included 23. No instances of 20.**
1	C-users	5 26	At any point in sessions.	Searching and Collaborative	Used by only five of the user groups.

Table G8 illustrates that many activity patterns were evidenced at any point of interaction with the system. These include a variety of Organizational and Searching, Advanced, and Searching and Collaborative scaffolds.

APPENDIX I

GRADED USERS SCAFFOLD USE

APPENDIX I

	SUM/A's	Percent	Average	SUM/B's	Percent	Average	SUM/C's	Percent	Average	SUM/D,F,U	Percent	Average
Total users	132 users			111 users			121 users			44 users		
ActionID												
1	725	7.37	5.50	498	9.68	4.50	419	8.07	3.46	129	10.00	2.93
2	393	4.00	3.00	163	3.17	1.50	196	3.80	1.62	86	6.62	2.00
3	0	0	0	0	0	0	0	0	0	0	0	0
4	26	0.26	0.20	9	0.17	0.08	13	0.30	0.11	2	0.15	0.05
5	1645	16.72	12.50	1353	26.29	12.20	1329	26.00	11.00	320	24.63	7.27
6	33	0.34	0.25	10	0.20	0.10	31	0.60	0.26	7	0.54	0.16
7	52	0.53	0.40	47	0.91	0.42	35	0.67	0.29	5	0.38	0.11
8	140	1.42	1.06	122	2.37	1.10	84	1.62	0.69	6	0.46	0.14
9	6	0.06	0.05	10	0.19	0.10	9	0.17	0.07	1	0.08	0.02
10	5	0.05	0.03	11	0.21	0.10	7	0.13	0.06	1	0.08	0.02
11	48	0.49	0.36	35	0.68	0.32	94	1.81	0.78	5	0.38	0.11
12	3	0.03	0.02	8	0.16	0.07	53	1.02	0.44	0	0.00	0.00
13	320	3.25	2.42	215	4.18	2.00	218	4.20	1.80	69	5.31	5.00
14	0	0.00	0.00	2	0.04	0.02	1	0.02	0.01	1	0.08	0.02
15	58	0.59	0.44	41	0.80	0.40	39	0.80	0.35	7	0.54	0.16
16	46	0.47	0.35	23	0.45	0.21	16	0.31	0.13	13	1.00	0.30
17	1443	14.70	11.00	262	5.10	2.36	367	7.10	3.03	134	10.32	3.05
18	679	6.90	5.14	61	1.20	0.55	121	2.33	1.00	67	5.16	1.52
19	144	1.46	1.10	25	0.50	0.22	59	1.14	0.49	4	0.31	0.09
20	943	9.60	7.14	169	3.30	1.52	268	5.16	2.31	45	3.50	1.00
21	5	0.05	0.03	1	0.02	0.01	3	0.06	0.02	0	0.00	0.00
22	2078	21.12	15.74	1379	26.80	12.42	1287	24.80	10.64	283	21.79	6.43
23	661	6.72	5.00	488	9.50	4.40	406	7.82	3.36	27	2.08	0.61
24	33	0.34	0.25	25	0.50	0.22	17	0.33	0.14	2	0.15	0.05
25	4	0.04	0.03	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
26	71	0.72	0.54	77	1.50	0.70	39	0.75	0.32	7	0.54	0.16

27	56	0.57	0.42	58	1.13	0.52	36	0.70	0.30	47	3.62	1.10
28	251	2.55	1.90	55	1.07	0.50	45	0.87	0.37	31	2.39	0.70
Total activities	9839			5147			5192			1299		

APPENDIX J

MOST FREQUENTLY USED SEARCH TERMS EXTENT OF MATCH

APPENDIX J

Search Term	Extent of Match	Frequency	Percent
NA		40376	76.49722438
bacteria	EM	519	0.983308387
planets	EM	236	0.447130596
virus	RM	232	0.439552112
mars	EM	203	0.384608098
hepatitis	NM	154	0.29177166
astronomy	EM	119	0.225459919
herpes	RM	118	0.223565298
cholera	EM	109	0.206513708
batteries	EM	107	0.202724465
stars	PM-P	102	0.193251359
aids	NM-S	98	0.185672875
sound	EM	95	0.179989011
gonorrhea	EM	89	0.168621284
influenza	EM	86	0.162937421
immunity	NM-S	82	0.155358936
west nile encephalitis	EM	81	0.153464315
moon	EM	79	0.149675072
cell	RM	78	0.147780451
rabies	EM	78	0.147780451
hurricanes	EM	77	0.14588583
black holes	EM	73	0.138307345
cells	EM	73	0.138307345
syphilis	EM	70	0.132623482
plague	EM	69	0.130728861
florida hurricanes	PM-P	66	0.125044997
volcanoes	EM	65	0.123150376
galaxies	NM	64	0.121255755
acid rain	EM	63	0.119361134
hiv/aids	NM-S	61	0.115571891
viruses	EM	60	0.11367727
meningitis	EM	54	0.102309543
weather	EM	52	0.098520301
universe	EM	52	0.098520301
chlamydia	EM	50	0.094731059
blizzards	NM	49	0.092836437
ebola	NM-S	46	0.087152574
clouds	EM	46	0.087152574
atoms	NM-S	44	0.083363332

black hole	RM	44	0.083363332
influenza\\	F2	43	0.08146871
comets	EM	42	0.079574089
sun	EM	40	0.075784847
ozone	EM	40	0.075784847
atmosphere	EM	38	0.071995604
snow	EM	37	0.070100983
atoms molecules	PM-P	36	0.068206362
precipitation	NM	35	0.066311741
electromagnetic radiation	EM	34	0.06441712
radiation	EM	34	0.06441712
jupiter	EM	34	0.06441712
constellations	EM	34	0.06441712
saturn	EM	34	0.06441712
antarctic snow storms	NM	34	0.06441712
genital herpes	PM-P	31	0.058733256
chicken pox	RM	31	0.058733256
tuberculosis	EM	31	0.058733256
tornado	RM	30	0.056838635
aliens	EM	30	0.056838635
immune system	PM	28	0.053049393
wind	EM	28	0.053049393
polio	RM	28	0.053049393
virus?	F2	28	0.053049393
viral meningitis	PM	26	0.04926015
sand storms	PM-P	26	0.04926015
volcano	RM	25	0.047365529
uranus	EM	24	0.045470908
microbe zoo	NM	24	0.045470908
"moon,craters"	F2	22	0.041681666
solar system	EM	22	0.041681666
sand storm	PM-P	22	0.041681666
big bang	PM-P	21	0.039787045
hurricane	RM	21	0.039787045
moons	EM	20	0.037892423
snow storms	EM	20	0.037892423
astronomy cafe	NM	20	0.037892423
gonorhea?	F2, F1	20	0.037892423
cold	RM	20	0.037892423
venus	EM	20	0.037892423
antarctic	RM	19	0.035997802
goneahea	F1	19	0.035997802

water pollution	PM	19	0.035997802
hiv	NM	19	0.035997802
planet	RM	19	0.035997802
antarctica	EM	19	0.035997802
temperature	EM	19	0.035997802
nasa	NM	18	0.034103181
ion	RM	18	0.034103181
tornados	F1	18	0.034103181
plants	EM	18	0.034103181
desert	EM	18	0.034103181
chylamdia	F1	18	0.034103181
african desert sand storm	NM	18	0.034103181
winds	RM	17	0.03220856
viruis	F1	17	0.03220856
pluto	EM	17	0.03220856
water	EM	17	0.03220856
cell phones	RM	17	0.03220856
antartica	F1	16	0.030313939
cloud	RM	16	0.030313939
africa	NM	16	0.030313939
nuclear	RM	16	0.030313939
florida	NM	16	0.030313939
tornadoes	EM	16	0.030313939
snow storm	RM	15	0.028419318
chlamydial	F1	15	0.028419318
ebola virus	RM	15	0.028419318
storms	EM	15	0.028419318
anode	NM	15	0.028419318
e.coli	RM	15	0.028419318
crickets	NM	15	0.028419318
earth	EM	14	0.026524696
galaxy	NM	14	0.026524696
cell?	F2	14	0.026524696
menigitis	F1	14	0.026524696
seeds	EM	14	0.026524696
peanuts	NM	13	0.024630075
bacteria?	F2	13	0.024630075
communicable disease	RM	13	0.024630075
learn bacteria	PM-P	13	0.024630075
causes meningitis	PM-P	13	0.024630075
flu	EM	13	0.024630075
std	RM	13	0.024630075

blizzard	NM	13	0.024630075
battery	RM	13	0.024630075
fog	NM	13	0.024630075
sun orbit	PM-P	12	0.022735454
"elements,atoms,compounds,mixtures"	ALL NM	12	0.022735454
artic temps	NM	12	0.022735454
electric current	NM-S	12	0.022735454
lightning	EM	12	0.022735454
pneumonia	NM	12	0.022735454
minerals	EM	12	0.022735454
chickenpox	EM	12	0.022735454
african dessert sand storm	NM	12	0.022735454
rocks	NM-S	12	0.022735454
life mars	PM-P	12	0.022735454
electromagnetic spectrum clips movies	NM	12	0.022735454
supernova	RM	12	0.022735454
life planets	PM-P	12	0.022735454
constelation	F1	11	0.020840833
lava	EM	11	0.020840833
meteorology	EM	11	0.020840833
tornado pictures	PM-P	11	0.020840833
disease	EM	11	0.020840833
rain forest	RM	11	0.020840833
ozone layer	PM-P	11	0.020840833
air pollution	EM	11	0.020840833
anartica	F1	11	0.020840833
microscopes	EM	11	0.020840833
syphillis	EM	11	0.020840833
greenhouse effect	PM	11	0.020840833
bacterial meningitis harm you?	F2	11	0.020840833
cathode	NM	10	0.018946212
miningitis	F1	10	0.018946212
african desert sand storms	NM	10	0.018946212
word net	GO	10	0.018946212
volcanos	F1	10	0.018946212
noise pollution	PM	10	0.018946212
causes seasons	PM-P	10	0.018946212
global warming	PM	10	0.018946212
horoscopes	NM	10	0.018946212
eruption	NM	10	0.018946212
plate tectonics	PM	10	0.018946212
mars life	PM-P	10	0.018946212

hot sun	PM-P	10	0.018946212
gravity	EM	10	0.018946212
sahara desert	NM	10	0.018946212
noise	EM	9	0.017051591
active immunity	NM	9	0.017051591
stages star	PM-P	9	0.017051591
oil spills	NM	9	0.017051591
eruptions	NM	9	0.017051591
florida weather	PM-P	9	0.017051591
ecoil virus bacteria?	F1,F2	9	0.017051591
neptune	EM	9	0.017051591
toxic waste	NM	9	0.017051591
chickenpox cdc	PM-P	9	0.017051591
cirrus	NM-S	9	0.017051591
temperature pressur	PM-P	9	0.017051591
oxidation reduction chemical reactions	PM-P	9	0.017051591
alkaline	EM	9	0.017051591
metoroligy	F1	8	0.015156969
bacteria drinking water	EM	8	0.015156969
science fair	PM-P	8	0.015156969
deserts	EM	8	0.015156969
finland	NM	8	0.015156969
humidity	EM	8	0.015156969
sound waves	EM	8	0.015156969
sex	RM	8	0.015156969
atom?	F2	8	0.015156969
moon get orbit	S4, BOTH EM	8	0.015156969
constellation	RM	8	0.015156969
african desert	NM	8	0.015156969
astronmy	F1	8	0.015156969
exploring mars	S4, EM, RM	8	0.015156969
oxidation	EM	8	0.015156969
chickenpoxcdc	RM	8	0.015156969
125jwah1	GO	8	0.015156969
sound pollution	PM-P	8	0.015156969
echo	NM	8	0.015156969
bgjgnbjgggbjhrnbjh	F1	8	0.015156969
polymers	EM	8	0.015156969
botany	EM	7	0.013262348
tell viruses bacteria?	S4	7	0.013262348
nuclear waste	NM	7	0.013262348
materials batteries with?	S4	7	0.013262348

acoustics	NM	7	0.013262348
florida climate	PM-P	7	0.013262348
kind chemicals toxic waters?	S4, F1	7	0.013262348
earth core	PM-P	7	0.013262348
snowstorms	RM	7	0.013262348
http://atrs.arc.nasa.gov/r_t/1996/scien	F1	7	0.013262348
electric current?	F2	7	0.013262348
hurricane temperatures	PM-P	7	0.013262348
violent storms	PM-P	7	0.013262348
explosive volcanoes	PM-P	7	0.013262348
desert sand storms	PM-P	7	0.013262348
hail	EM	7	0.013262348
crystals	NM	7	0.013262348
wet cell	NM	7	0.013262348
satellite	RM	7	0.013262348
living things	NM	7	0.013262348
erosion	NM	7	0.013262348
tell bacteria	S4	7	0.013262348
monocot	NM	7	0.013262348
big bang theory	PM	7	0.013262348
games	NM	7	0.013262348
mumps	NM	7	0.013262348
microbezoo	NM	7	0.013262348
food science	PM-P	7	0.013262348
constallations	F1	7	0.013262348
hoilnhj\\	GO	6	0.011367727
milky way	NM	6	0.011367727
african weather	PM-P	6	0.011367727
science fair projects	NM	6	0.011367727
turberculosis	F1	6	0.011367727
florida hurricane	PM-P	6	0.011367727
general herpes	PM-P	6	0.011367727
radiation leaks	PM-P	6	0.011367727
detroit river polluted	S4	6	0.011367727
mercury	EM	6	0.011367727
viral menigitis	PM	6	0.011367727
troposphere	NM	6	0.011367727
ear	EM	6	0.011367727
chemicals located batteries?	S4	6	0.011367727
tuberculosis?	F2	6	0.011367727
sand	RM	6	0.011367727
sound levels	PM-P	6	0.011367727

communicable disease?	F2	6	0.011367727
detroit river	NM	6	0.011367727
universe created	S4	6	0.011367727
cancer	EM	6	0.011367727
milk different experation	S4, F1	6	0.011367727
std's	EM	6	0.011367727
jupiter's storms	PM-P	6	0.011367727
phnemonia	F1	6	0.011367727
stomers	F1	6	0.011367727
blackholes	RM	6	0.011367727
dicot seed	PM-P	6	0.011367727
asteroids	EM	6	0.011367727
hepatitus	NM	6	0.011367727
rain	EM	6	0.011367727
drinking water	PM	6	0.011367727
bacteria water	ALL EM	6	0.011367727
chemicals	EM	6	0.011367727
hepititis	F1	6	0.011367727
reduction?	NM	6	0.011367727
climate	EM	6	0.011367727
plantes	F1	6	0.011367727
influneza	F1	6	0.011367727
floods	EM	6	0.011367727
electromagnetic waves	PM	6	0.011367727
viurs	F1	6	0.011367727
harmful batteries	PM-P	6	0.011367727
spiral galaxies	NM	6	0.011367727
wind speeds	NM	6	0.011367727
hepatitisb	NM	6	0.011367727
science rollercoasters	S4	6	0.011367727
chemicals batteries	ALL EM	6	0.011367727
living mars	S4	5	0.009473106
predict	NM	5	0.009473106
ckicken pox sick	F1	5	0.009473106
gas planets	ALL EM	5	0.009473106
information chichenpox	S4	5	0.009473106
african sandstorm	PM-P	5	0.009473106
acoustic	NM	5	0.009473106
bacteria vs. virus	EM, RM	5	0.009473106
vascular plants	NM	5	0.009473106
blizzards form	NM	5	0.009473106
quaser	F1	5	0.009473106

tb	EM	5	0.009473106
sahara	NM	5	0.009473106
electromagnetic spectrum	PM-P	5	0.009473106
tempatures florida	F1	5	0.009473106
erupt	NM	5	0.009473106
physics projects	S4	5	0.009473106
hurricanes ranges temperature	S4	5	0.009473106
universe's size	PM-P	5	0.009473106
temperature florida	PM-P	5	0.009473106
ecoli	RM	5	0.009473106
flordia hurricanes	F1	5	0.009473106
health risk cell phones	S4	5	0.009473106
desert sand storm	PM-P	5	0.009473106
space shuttle	PM	5	0.009473106
asteriods	F1	5	0.009473106
antartic snowstorm	PM-P	5	0.009473106
antartic snow storm	PM-P	5	0.009473106
harmful	NM	5	0.009473106
orgin moon	F1	5	0.009473106
redution?	F1	5	0.009473106
african climate	NM	5	0.009473106
bla	F1	5	0.009473106
bacterial meningitis	PM	5	0.009473106
astronomers	EM	5	0.009473106
duracell	EM	5	0.009473106
west nile virus	PM-P	5	0.009473106
earthquakes	EM	5	0.009473106
planet rings	PM-P	5	0.009473106
centripetal force	NM	5	0.009473106
microwaves	EM	5	0.009473106
black holes form	S4	5	0.009473106
fertilizer	NM	5	0.009473106
african deserts	PM-P	5	0.009473106
nasa mars	PM-P	5	0.009473106
african	NM	5	0.009473106
monkeys	PM-P	5	0.009473106
rocks minerals	NM, EM	5	0.009473106
temputer	F1	5	0.009473106
cell phone batteries	PM-P	5	0.009473106
oil pollution	PM	5	0.009473106
venus report	PM-P	5	0.009473106
blood	EM	5	0.009473106

clouds humidity	ALL EM	5	0.009473106
pollution	EM	5	0.009473106
human body	NM	5	0.009473106
oxygen	EM	5	0.009473106
hurricans	F1	5	0.009473106
genital herpies	F1	5	0.009473106
gonnorhea	F1	5	0.009473106
eletromagnetic radiation	F1	5	0.009473106
create new driving folder	S4	5	0.009473106
geology	EM	5	0.009473106
water purification	NM-S	5	0.009473106
nebula	RM	5	0.009473106
star	PM-P	5	0.009473106
temperature pressure	ALL EM	5	0.009473106
plate movement	PM-P	5	0.009473106
battery chemistry	ALL EM	5	0.009473106
ncan tell ecoil?	F1	4	0.007578485
hot	NM	4	0.007578485
sharks	NM	4	0.007578485
cold/flu	S2	4	0.007578485
african desert sandstorms	NM	4	0.007578485
milk	NM	4	0.007578485
headaches	NM	4	0.007578485
sun's gravity	NM	4	0.007578485
tornado damage	NM	4	0.007578485
ear damage repaired	S4	4	0.007578485
stephen fat bat cat mat	GO	4	0.007578485
space	EM	4	0.007578485
severe weather	PM-P	4	0.007578485
hazardous waste wells?	F2	4	0.007578485
gential herpes	F1	4	0.007578485
humans mars	S4	4	0.007578485
stars form	S4	4	0.007578485
air pressure	NM	4	0.007578485
disease?	F2	4	0.007578485
causes plague?	F2,S4	4	0.007578485
aquarius	NM	4	0.007578485
life	EM	4	0.007578485
galaxys	F1	4	0.007578485
crust	NM	4	0.007578485
non-living things	NM	4	0.007578485
waste wells?	F2, NM	4	0.007578485

dust	NM-S	4	0.007578485
cumulonimbus	NM-S	4	0.007578485
constillation	F1	4	0.007578485
passive immunity?	F1, F2, NM	4	0.007578485
oil spill	NM	4	0.007578485
vaccine available for influenza.	S4	4	0.007578485
inside tornado	S4	4	0.007578485
common cold	PM-P	4	0.007578485
chicken poxs	F1	4	0.007578485
clouds megellan	F1	4	0.007578485
stars form?	S4	4	0.007578485
astrology	NM	4	0.007578485
desert storms	NM-S	4	0.007578485
cell movie	S4	4	0.007578485
satillite	F1	4	0.007578485
travel mars	S4	4	0.007578485
communicable dieases	F1	4	0.007578485
nuclear power	F2	4	0.007578485
antarctic storms	PM-P	4	0.007578485
thunder storms	NM	4	0.007578485
florida temperatures	S4	4	0.007578485
syphills	F1	4	0.007578485
flu/colds	ALL EM	4	0.007578485
big gas planets	S4	4	0.007578485
florida tempatures	F1	4	0.007578485
astroides	F1	4	0.007578485
greenhouse gases	PM-P	4	0.007578485
meningities	F1	4	0.007578485
kind batteries cell phones use?	S4	4	0.007578485
ocean pollution	NM	4	0.007578485
smog	EM	4	0.007578485
earth's atmosphere	PM-P	4	0.007578485
menegiti	F1	4	0.007578485
science fair topics	S4	4	0.007578485
finland country	NM	4	0.007578485
kind+volcones	S5	4	0.007578485
battery chemicals	S4	4	0.007578485
forces motion roller coasters	S4	4	0.007578485
battery anatomy	PM	4	0.007578485
predict volcano	PM-P	4	0.007578485
herpies	F1	4	0.007578485
new planets	PM-P	4	0.007578485

life outside earth	S4	4	0.007578485
tell tubercuiosis syphills?	S4	4	0.007578485
tell tuberculosis syphilis?	S4	4	0.007578485
reduction	NM	4	0.007578485
ideas physics projects	S4	4	0.007578485
flu/cold	F2	4	0.007578485
temparture	F1	4	0.007578485
photosynthesis	EM	4	0.007578485
high-frequency sound	PM-P	4	0.007578485
zodiac signs	NM	4	0.007578485
wordnet	GO	4	0.007578485
precipatation hurricanes	F1	4	0.007578485
stones	EM	4	0.007578485
rollercoaster	NM	4	0.007578485
jupter	F1	4	0.007578485
ronald sutton jr	GO	4	0.007578485
pepper soap	NM	4	0.007578485
atoms molicules	F1	4	0.007578485
cells look	S4	4	0.007578485
mad cow	NM	4	0.007578485
astronauts	EM	4	0.007578485
differance betwwen atom molecule?	F1,F2,S4	4	0.007578485
stars there?	F2	4	0.007578485
sun stages	PM-P	4	0.007578485
gonorrhea effects	S4	4	0.007578485
galaxay	F1	4	0.007578485
experiments	EM	4	0.007578485
types clouds	S4	4	0.007578485
tonmdery partick nahgeramn lkasjdfo	GO	4	0.007578485
antartica snow storms	F1	4	0.007578485
batteries harmful	S4	4	0.007578485
diseases	RM	4	0.007578485
pysics projects u.s.	S4	4	0.007578485
hepattitis	F1	4	0.007578485
zora neale hurston	GO	4	0.007578485
space travel	NM-S	4	0.007578485
mercury free	PM	4	0.007578485
toxic waters	NM	4	0.007578485
biology	EM	4	0.007578485
curris cumulonimbus clouds	NM-S	4	0.007578485
electromagnetism	EM	4	0.007578485
kinds batteries	S4	4	0.007578485

{black hole}	S5	4	0.007578485
average temperature florida hurricane	S4	4	0.007578485
pictures jupiter	S4	4	0.007578485
point electromagnetic waves affective	S4	4	0.007578485
rabies?	F2	4	0.007578485
black hole star exploded?	S4,F2	4	0.007578485
food science- liquids	S4	4	0.007578485
mars + travel	S3	3	0.005683864
viral meningisit	F1	3	0.005683864
meningitus	F1	3	0.005683864
"mars,the planet"	S5	3	0.005683864
viral meningitis.	F1	3	0.005683864
cassiopeia	NM	3	0.005683864
nuclear energy	PM	3	0.005683864
equator	NM	3	0.005683864
west nile virus treatable	S4	3	0.005683864
tell virus bacteria	NM-S	3	0.005683864
stratus	S4	3	0.005683864
causes seasonal changes	S4	3	0.005683864
earthqaukes	F1	3	0.005683864
weather forcast	PM-P	3	0.005683864
living non-living things	S4	3	0.005683864
weather wildlife	S4	3	0.005683864
composition jupiter	S4	3	0.005683864
oldest stars	S4	3	0.005683864
marie	GO	3	0.005683864
astonomy	F1	3	0.005683864
animals	EM	3	0.005683864
westnileencephalitis	F2	3	0.005683864
people temperature	S4	3	0.005683864
ebole virus	F1	3	0.005683864
pesticides	NM-S	3	0.005683864
herpes.com	S5	3	0.005683864
aid	NM	3	0.005683864
origin asteroids	S4	3	0.005683864
electromagnetic radiation waves	PM-P	3	0.005683864
genitel	F1	3	0.005683864
helen	GO	3	0.005683864
precipitatiom	F1	3	0.005683864
dew	NM	3	0.005683864
sahara winds	NM	3	0.005683864
florida precipitation	S4	3	0.005683864

blackhole	RM	3	0.005683864
stars + elements	S3, RM	3	0.005683864
radioactivity	EM	3	0.005683864
hydrochloric acid	NM	3	0.005683864
e. coli	EM	3	0.005683864
african desert storms	S4	3	0.005683864
lead comtamination	F1	3	0.005683864
nuclear acid	NM	3	0.005683864
lung cancer	NM	3	0.005683864
water molecules	NM-S	3	0.005683864
peanut facts	S4	3	0.005683864
pluto/coldness?	S4,F2	3	0.005683864
average temperatures	S4	3	0.005683864
amusement parks	NM	3	0.005683864
physical changes water	S4	3	0.005683864
phneumonia	F1	3	0.005683864
unhirversity	F1	3	0.005683864
science fair for technology.	S4	3	0.005683864
giant gas planets	S4	3	0.005683864
living non living things	S4	3	0.005683864
oozing volcanoes	PM-P	3	0.005683864
damage	NM	3	0.005683864
toronto news amusment park	S4	3	0.005683864
pictures turberculosis	S4	3	0.005683864
magnet	RM	3	0.005683864
ashes	NM	3	0.005683864
crustal rocks	NM	3	0.005683864
desert temperatures	S4	3	0.005683864
viruse	F1	3	0.005683864
phases moon determine day month?	S4,F2	3	0.005683864
windtempure	F1	3	0.005683864
alto	NM-S	3	0.005683864
fog formed	S4	3	0.005683864
antarctic snowstorm	NM	3	0.005683864
living	RM	3	0.005683864
tubercluosis t.b.	F1	3	0.005683864
death stars	NM	3	0.005683864
tell virus	S4	3	0.005683864
asteroid belt	PM-P	3	0.005683864
bacteria tap water	S4	3	0.005683864
earth science	NM	3	0.005683864
lighting	EM	3	0.005683864

atom	PM-P	3	0.005683864
brown water	NM	3	0.005683864
flood	RM	3	0.005683864
plutonium	NM	3	0.005683864
air pressure affect pressure?	S4	3	0.005683864
exobiology	NM	3	0.005683864
national geographic	GO	3	0.005683864
stars stay place	S4	3	0.005683864
blood coagulation	PM-P	3	0.005683864
ear damage	PM-P	3	0.005683864
eboli	RM	3	0.005683864
find climographs	S4	3	0.005683864
orgin th moon	S4, F1	3	0.005683864
cell structure	NM-S	3	0.005683864
luna	NM	3	0.005683864
cold&flu	S3	3	0.005683864
electromagnetic field	RM	3	0.005683864
cathode ?	F2	3	0.005683864
thermal pollution	NM	3	0.005683864
west nile encepalitis	EM	3	0.005683864
amune system	F1	3	0.005683864
hepatitis hiv/aids	S4	3	0.005683864
magneteism	F1	3	0.005683864
acid	RM	3	0.005683864
exploding volcanoes	S4	3	0.005683864
lunar orbit	NM	3	0.005683864
asia's weather	NM	3	0.005683864
hurricane data	PM-P	3	0.005683864
sexually-transmitted disease	EM	3	0.005683864
cells?	F2	3	0.005683864
percipitation	F1	3	0.005683864
hepaitis	F1	3	0.005683864
average afriacan desert temperatures	S4,F1	3	0.005683864
big bang start?	S4	3	0.005683864
microscopes improve world	S4	3	0.005683864
radaition	F1	3	0.005683864
seasons changes	S4	3	0.005683864
wind desert	S4	3	0.005683864
birds weather	S4, NM, EM	3	0.005683864
cells look like?	S4	3	0.005683864
wet	NM	3	0.005683864
cotyledon	NM	3	0.005683864

comets come	S4	3	0.005683864
soundlevels	F2	3	0.005683864
science fair project dealing technology	S4	3	0.005683864
garlic	NM	3	0.005683864
birth stars	S4	3	0.005683864
co2	NM	3	0.005683864
creat new dq	GO	3	0.005683864
weather africa	S4	3	0.005683864
making cell phones safer	S4	3	0.005683864
mars madness	S4	3	0.005683864
size galaxy (types)	S4	3	0.005683864
oil spill projects	S4	3	0.005683864
happens solar exclispe?	S4	3	0.005683864
areoxidation reduction chemical	S4	3	0.005683864
humans	NM	3	0.005683864
causes seasonal changes?	S4	3	0.005683864
science	EM	3	0.005683864
menegitis	F1	3	0.005683864
nitrogen	EM	3	0.005683864
microwave	RM	3	0.005683864
size galaxy	S4	3	0.005683864
tubercluosis	F1	3	0.005683864
sunblock	NM	3	0.005683864
gonorhea	F1	3	0.005683864
anarctic	F1	3	0.005683864
ice	EM	3	0.005683864
supernovas	RM	3	0.005683864
sucrose	NM-S	3	0.005683864
locations battery	S4	3	0.005683864
rats mazes	S4	3	0.005683864
germination	NM	3	0.005683864
botany experiments	S4	3	0.005683864
orbit moon	S4	3	0.005683864
gonnerhea	F1	3	0.005683864
explosions	EM	3	0.005683864
wet get africa	S4	3	0.005683864
types volcanoes	S4	3	0.005683864
microbe	RM	3	0.005683864
"atoms, elements, compounds, molecules,	S5	3	0.005683864
star constallations	F1	3	0.005683864
coriolis force	F1	3	0.005683864
zodiacs	NM	3	0.005683864

lucy	GO	3	0.005683864
desert wind	S4	3	0.005683864
super nova	F1	3	0.005683864
south america	NM	3	0.005683864
mars lander	NM	3	0.005683864
astroid burn reaches earths	S4	3	0.005683864
learn syphilis	S4	3	0.005683864
weather cloud	ALL EM	3	0.005683864
thunder	NM	3	0.005683864
average hurricane temperature	S4	3	0.005683864
problems humans living mars	S4	3	0.005683864
effects human body space	S4	3	0.005683864
temparture pressure	F1	3	0.005683864
earthqauke	F1	3	0.005683864
dew drops	NM	3	0.005683864
lightning storms	PM-P	3	0.005683864
sattern	F1	3	0.005683864
heavy metals	NM	3	0.005683864
tropical climate regoin	S4	3	0.005683864
floods efect world	NM	3	0.005683864
sandstorms	S4	3	0.005683864
health risks phones	S4	3	0.005683864
frogs	EM	3	0.005683864
new planet life	S4	3	0.005683864
dq	GO	3	0.005683864
detroit river water	S4	3	0.005683864
sun's stages	S4	3	0.005683864
uranus tilt	S4	3	0.005683864
science fair project	S4	3	0.005683864
water poulltion	F1	3	0.005683864
meteorites	F1	3	0.005683864
monocot seed	NM	3	0.005683864
sky	NM	3	0.005683864
electromagnetic	NM-S	3	0.005683864
bateria	F1	3	0.005683864
sound frequency	PM-P	3	0.005683864
vaccines	EM	3	0.005683864
hiv255	NM	3	0.005683864
electromagnetic radiation?	F2	3	0.005683864
life cycle star?	S4	3	0.005683864
"food molds fastest placed place.	S5	3	0.005683864
learn tubercluosis	S4	3	0.005683864

temperatures	RM	3	0.005683864
mineral	RM	3	0.005683864
find cholera??	S4	3	0.005683864
decibels	F1	3	0.005683864
decibals	NM	3	0.005683864
trees	EM	3	0.005683864
zeldie jennifer	GO	3	0.005683864
american football	GO	3	0.005683864
glucose	EM	3	0.005683864
recycle	RM	3	0.005683864
west nile encephalities	F1	3	0.005683864
plague?	F2	3	0.005683864
food science- milk	S4	3	0.005683864
hurricane conditions	S4	3	0.005683864
force motion	S4	3	0.005683864
electromagnetic radiaiton	F1	3	0.005683864
cell phones radiation	S4	3	0.005683864
chylamydia	F1	3	0.005683864
causes clustise	F1	3	0.005683864
bactria	F1	3	0.005683864
comet origin	S4	3	0.005683864
chlaymdia	F1	3	0.005683864
average temperature for florida	S4	3	0.005683864
cathode?	F2	3	0.005683864
chicken poxs diease	F1	3	0.005683864
snow formation	S4	3	0.005683864
comets formed	S4	3	0.005683864
really bad lightning storms	S4	3	0.005683864
projects	NM	3	0.005683864
cell phone radiation	EM	3	0.005683864
technology	EM	3	0.005683864
cnn mars	NM	3	0.005683864
chlmydia	F1	3	0.005683864
e.coli bacteria	PM-P	3	0.005683864
dieases	F1	3	0.005683864
different types batteries?	S4, F2	3	0.005683864
chicken pox&gonerhea	F1,F2	3	0.005683864
antartica climate	S4	3	0.005683864
volcanic eruption	PM-P, RM	3	0.005683864
health risk radiation prevented	S4	3	0.005683864
seasons latitude	S4	3	0.005683864
dust created?	S4,F2	3	0.005683864

frozen	NM	3	0.005683864
wet cell?	F2	3	0.005683864
voyager	NM	3	0.005683864
core	NM	3	0.005683864
sun block	NM	3	0.005683864
birds	EM	3	0.005683864
electromagnetic radiation parts	S4	3	0.005683864
eathquake	F1	3	0.005683864
turgor pressure	NM	3	0.005683864
energy used?	S4	3	0.005683864
african wind force	S4	3	0.005683864
snowstorm	RM	3	0.005683864

APPENDIX K

STUDENT GENERATED KEYWORDS

App. K: Student Generated Keywords (SGKs)

acoustics
active immunity
aids
air pressure
alto
amusement parks
anode
aquarius
ashes
asteroid belt
astrology
atoms
battery
big bang
blizzards
blood clotting
blood coagulation
cassiopeia
cathode
cell phone batteries
cell structure
centripetal force
cirrus
climographs
CO2
common cold
compounds
core
cotyledon
crickets
crust
crustal rocks
crystals
cumulonimbus
death of stars
decibels
desert sand storm
desert sand storm
desert storms
dew
dew drops
dewpoint
dicot seed
dust
e.coli bacteria
ear damage
earth science
earth's atmosphere
ebola
eboli
echo
ecoli
electric current
electromagnetic
electromagnetic radiation waves
electromagnetic spectrum
elements
equator
erosion
eruption
exobiology
fertilizer
fog
food science
food science- liquids
food science- milk
force
fusion
galaxies
games
garlic
genital herpes
genital warts
geotropism
germination
greenhouse gases
headaches
heavy metals
hepatitis b
herpes
high-frequency sound
hiv
hornets
horoscopes
human body
humans
hydrochloric acid
immunity
life in space
lightning storms
living
living things
lunar orbit
lung cancer
mad cow
magnet
mars lander
milk
milky way
mixtures
mold
monkeys
monocot
monocot seed
motion
mumps
nasa
nature
non-living things
northern lights
nova
nuclear
nuclear acid
nuclear waste
ocean pollution
oil spills
ozone layer
parrots
peanuts
pesticides
physical science
physics projects
planetary alignment

plate movement
plutonium
pneumonia
polio
precipitation
prediction
radiation leaks
radon gas
recycle
reduction
rocks
sandstorms
science fair
science fair projects
sedimentary rocks
sex
severe weather
sharks
sky
small pox
sound frequency
sound levels
sound pollution
space probes
space travel
spiral galaxies
stars
std
stratus
sucrose
sunblock
sun's gravity
sun's stages
supernova
tap water
thermal pollution
thunder
thunderstorms
tornado
toxic dumping
toxic waste
toxic waters
transient
troposphere
turgor pressure
ufo
vascular plants
volcanic eruption
voyager
waste wells
water molecules
water purification
weather forcast
wet cell
wind speeds
zodiac signs

APPENDIX L

RESULTS OF ALL RE-EXECUTED SEARCHES

Appendix L: Results of All Re-executed Searches

ID	Search_Term	Result_Count	Augmented	DQ	Comments
54	ozone layer	12	19	matches DQ	
144	atoms molecules	17	74	matches DQ	Term broken out from full phrase
310	hepatitis	10	10	No DQ	Not augmented
440	herpes	14	16	No DQ	Augmented
828	andramada	0	0	Not same as DQ" Where do stars come from?"	Misspelled, not augmented
902	black hole	19	35	If you go into a black hole where do you come out?	Augmented
1474	plague	6	6	No DQ	Not augmented
1548	hiv/aids	1	8	No DQ	aids and hiv broken out into two words
1586	bacteria	25	36	matches DQ	Most frequently used term to search with, not augmented in CV
2530	virus	25	36	Does not match DQ	One of the most frequent terms searched, not augmented in CV
2552	std	11	17	Synonymous match to DQ	Term broken out from full phrase
3212	cells	25	105	matches DQ	Because of 25 total search result threshhold, it is difficult to see degree of change
3306	aids	10	14	matches DQ	Term broken out from full phrase
3928	influenza	12	12	matches DQ	Not augmented
4072	syphilis	12	12	No DQ	Not augmented
4198	gonorrhea	9	9	matches DQ	Not augmented
4554	chickenpox	5	5	matches DQ	Not augmented
5264	pneumonia	3	6	No DQ	Augmented term
5330	ebola	5	5	Does not match DQ	Augmented term
5622	genital herpes	7	7	Does not match DQ	Not augmented
6150	e.coli	2	2	matches DQ	Not augmented
6296	hepatitis b	0	3	No DQ	Augmented term
7412	cholera	6	6	No DQ	Not augmented
8114	tuberculosis	10	10	matches DQ	Not augmented
8784	meningitis	6	3	Partially matches DQ	

9568	disease	25	37	matches DQ	Because of 25 total search result threshhold, it is difficult to see degree of change
9626	viral meningitis	0		No DQ	
9780	chlamydia	12	12	matches DQ	Not augmented
11726	west nile encephalitis	5	5	matches DQ	Not augmented
14686	death rate	0	0	Does not match DQ	Not augmented
14980	sexually-transmitted disease	12	16	Synonymous match to DQ	Not augmented
15618	smallpox	2	2	Does not match DQ	Synonymous form augmented
16588	water	25	209	Partially matches DQ	Because of 25 total search result threshhold, it is difficult to see degree of change
16658	detroit water	1	1	Partially matches DQ	Not augmented
16912	bacterial meningitis	1	1	matches DQ	Not augmented
17176	sea sick	0	0	No DQ	Not augmented
17902	precipitation	25	38	matches DQ	Because of 25 total search result threshhold, it is difficult to see degree of change
18064	winds	25	25	matches DQ	Not augmented
18478	temperature	25	58	Partially matches DQ	Because of 25 total search result threshhold, it is difficult to see degree of change
18580	clouds	25	28	Partially matches DQ	Because of 25 total search result threshhold, it is difficult to see degree of change
18694	temperature pressure	15	15	Partially matches DQ	Not augmented
18782	air pressure	6	8	Partially matches DQ	Augmented term
19140	mad cow	0	0	No DQ	Not augmented
19600	immune system	14	14	matches DQ	Not augmented
21716	bug	2	2	No DQ	Not augmented
22272	genital warts	3	3	No DQ	Augmented term
22786	active immunity	0	0	No DQ	Augmented term
23116	common cold	5	11	Does not match DQ	Augmented term
23834	samonella	0	0	No DQ	Not augmented
27642	transient	2	2	No DQ	Not augmented
28596	polio	1	5	matches DQ	Augmented term

31324	innate	0	0	No DQ	Not augmented
32416	dengay fever	0	0	Synonymous match to DQ	Not augmented
33314	games	25	25	no DQ Goof Off	Augmented term
37604	super novas	0	2	matches DQ	Synonymous form augmented
37976	blizzards	5	5	no DQ	Not augmented
38068	weather forcast	0	28	No DQ	Augmented term
38226	tornado damage	2	2	Partially matches DQ	Not augmented
38810	echo	3	3	matches DQ	Augmented term
39224	sahara desert	3	3	no DQ	Not augmented
40472	astrology	1	1	no DQ	Not augmented
40576	general herpes	0	0	No DQ	Not augmented
40944	science fair projects	0	1	No DQ	Augmented term
40970	oil spills	4	5	matches DQ	Augmented term
41044	jupiter	22	22	matches DQ	Not augmented
41186	solar system	10	56	No DQ	Unexplained result. Could be that more sites were added after search conducted.
41268	nasa sites jupiter	2	5	No DQ	Term nasa added as augmented term
41336	orbit moon	8	8	Does not match DQ	Not augmented
41476	orbit earths moon	3	3	No DQ	Not augmented
41528	storms jupiter	0	0	Partially matches DQ	Not augmented
41820	fungus	23	23	Does not match DQ	Not augmented
42210	wind chill	9	12	Partially matches DQ	Unexplained result. Could be that more sites were added after search conducted.
42958	weather clouds	2	2	Partially matches DQ	Not augmented
43078	sun orbit	8	8	matches DQ	Not augmented
43182	disinfecting hands	0	0	matches DQ	Not augmented
43494	cancer	7	7	No DQ	Not augmented
43752	phases moon	3	3	No DQ	Not augmented
43854	navigation	3	3	Does not match DQ	Not augmented
44316	galaxies	25	71	No DQ Term broken out from phrase "Stars & Galaxies"	Because of 25 total search result threshhold, it is difficult to see degree of change
44424	spiral galaxies	1	1	No DQ	Augmented term
44962	meteorology	16	16	Synonymous match to DQ	Not augmented
44982	cirrus cumulonimbus	0	2	Synonymous match to DQ	Each term added as augmented term
45140	battery	19	19	matches DQ	Augmented term
45202	battery disposal	0	0	Partially matches DQ	Not augmented

45232	battery chemicals	2	2	Partially matches DQ	Not augmented
45308	electric current	6	9	No DQ	Augmented term
45420	oxidation reduction	7	17	Does not match DQ	Term "reduction" added as augmented term
45522	biotech grain	0	0	Does not match DQ	Not augmented
45774	milky way	3	3	No DQ	Augmented term
45946	big bang	2	6	matches DQ	Term shortened from phrase "big bang theory" which was used less frequently
46698	sun's stages	0	5	Partially matches DQ	Augmented term
46970	radiation leaks	0	2	No DQ	Augmented term
47054	electromagnetic radiation	13	14	matches DQ	Augmented term
47142	power plant melt	0	0	Partially matches DQ	Not augmented
47210	food science	2	8	Synonymous match to DQ	Augmented term
47252	food science- milk	0	0	Synonymous match to DQ	Augmented term
47876	snow fall	0	0	matches DQ	Not augmented
48130	architecture	7	13	No DQ	Unexplained result. Could be that more sites were added after search conducted.
48320	vascular plants	2	2	Does not match DQ	Augmented term
48592	volcano erupt	5	28	matches DQ	Augmented term "volcanic eruption" added is root form
48710	travel star	5	5	Does not match DQ	Not augmented
49352	finland	5	5	No DQ	Not augmented
49444	doldrums	0	0	matches DQ	Not augmented
49662	mold	7	7	No DQ	Augmented term
49732	sandstorms	0	2	No DQ	Augmented term
49896	waste wells	1	1	No DQ	Augmented term
50032	fertilizer	0	0	Does not match DQ	Augmented term
50504	life mars	11	11	matches DQ	Not augmented
50674	saturn rings	3	3	Partially matches DQ	Not augmented
51010	acoustics	0	0	matches DQ	Augmented term
51178	life planets	21	21	matches DQ	Not augmented
51314	dog kiss	0	0	No DQ	Not augmented
51374	birds weather	0	0	matches DQ	Not augmented
51408	weather birds	0	0	matches DQ	Not augmented
51648	explosive volcanoes	2	2	matches DQ	Not augmented
52486	nasa	13	24	No DQ	Augmented term
52690	dust	12	15	matches DQ	Augmented term
53004	holography	0	0	Does not match DQ	Not augmented
53062	alignment planets	0	0	Partially matches DQ	Synonymous form augmented

53122	sun's gravity	2	2	matches DQ	Not augmented
53246	humidity	5	6	partially matches DQ	Not augmented
53384	dew point	1	9	Partially matches DQ	Augmented term
54028	cell phones	21	21	Partially matches DQ	Not augmented
54186	quartz	1	1	Does not match DQ	Not augmented
54580	antarctic snow storms	0	0	Does not match DQ	Not augmented
54616	florida cyclones	0	0	No DQ	Not augmented
54956	electromagnetic radiation	13	14	matches DQ	Augmented term
55238	african desert sand storm	0	0	No DQ	Not augmented
55296	nuclear waste	2	2	matches DQ	Augmented term
55386	radon gas	2	5	Synonymous match to DQ	Augmented term
55454	nuclear power plant	8	8	Partially matches DQ	Not augmented
55536	radioactivity	15	15	Partially matches DQ	Not augmented
56168	rotation	4	7	matches DQ	Not augmented
56512	oxidation	10	11	Partially matches DQ	Not augmented
56544	troposphere	0	0	Does not match DQ	Augmented term
56626	crust	12	16	matches DQ	Augmented term
56666	plate movement	3	24	Partially matches DQ	Augmented term
58424	alkaline	5	5	Partially matches DQ	Not augmented
58658	water plants	23	23	No DQ	Not augmented
59582	rock crystals	1	21	matches DQ	Augmented terms
60234	new planets	6	6	Partially matches DQ	Not augmented
60942	water irrigation	2	3	Does not match DQ	Unexplained result. Could be that more sites were added after search conducted.
62668	water contamination	10	14	matches DQ	Unexplained result. Could be that more sites were added after search conducted.

REFERENCES

Allen, B. (1991). Cognitive research in information science: Implications for design. *Annual Review of Information Science and Technology, 26*, 3 - 37.

Bellardo, T. (1985). An investigation of online searcher traits and their relationship to search outcomes. *Journal of the American Society for Information Science, 36*(4), 241-250.

Bilal, D. (2000a). Children's use of the Yahooligans! web search engine: I. Cognitive, physical, and affective behaviors on fact-based search tasks. *Journal of the American Society for Information Science, 51*(7): 646-665.

Bilal, D. (2000b). Children's use of the Yahooligans! web search engine: II. Cognitive, physical, and affective behaviors on research tasks. *Journal of the American Society for Information Science, 52*(2): 118-136.

Bishop, A. & Starr, S. (1996). Social informatics for digital library use and infrastructure. In: M.E. Williams, (Ed.), *Annual Review of Information Science and Technology, 31*, (pp. 301-401). Medford, NJ: Information Today.

Blair, D. (1990) *Language and representation in information retrieval*. Amsterdam: The Netherlands: Elsevier Science.

Borgman, A. (1999). Holding on to reality: The nature of information at the turn of the millennium. Chicago: IL: University of Chicago Press.

Borgman, C. Hirsh, S., Walter, V., & Gallagher, A. (1995). Children's search behavior on browsing and keyword online catalogs: The Science Library Catalog Project. *Journal of the American Society for Information Science, 46*(6), 663-684.

Borgman, C. (1999). What are digital libraries? Competing visions. *Information Processing and Management* 35: 227-243.

Brown, I. (2000). Kid's catalogs: Help or hindrance for information retrieval. Available at http://home.earthlink.net/~mikirmbrown/KidsFinal.htm .

Buckland, M. (1991). Information. Information-as-thing. In Information and information systems (pp. 3-13 and 43-54). New York: Praeger.

Carlyle, A. (1989). Matching LCSH and user vocabulary in the library catalog. *Cataloging and Classification Quarterly, 10*(1/2), 37-63.

Churchland, P. (personal communication, Spring 1996).

Cooper, W. & O'Connor, B.C. (personal communication with O'Connor, 2001).

Doyen, S. & Wheeler, D. (1989). Use of a controlled vocabulary index in information retrieval tasks. In Salvendy, G. & M. Smith , (Eds.), *Designing and Using Human-Computer Interfaces and Knowledge Based Systems*. Elsevier: Amsterdam.

Drabenstott, K. & Vizine-Goetz, D. (1990). Search trees for subject searching in online catalogs. *Library Hi Tech, 8*(3), 7-20.

Dresang, E. (1999). More research needed: Informal information-seeking behavior of youth on the Internet. *Journal of the American Society for Information Science, 50*(2): 1123-1124.

Edmonds, L., Moore, P. & Balcom, K. (1990). The effectiveness of an online catalog. *School Library Journal* (October): 28-32.

Fenichel, C. (1981). Online searching measures that discriminate among users with different types of experiences. *Journal of the American Society for Information Science,32*(1), 23-32.

Fidel, R. (1987). What is missing in research about online searching behavior? *The Canadian Journal of Information Science, 12*, 54-61.

Fidel, R., Davies, R., Douglass, M., Holder, J., Hopkins, C., Kusher, E., Miyagishma, B, & Toney, C. (1999). A visit to the information mall: Web searching behavior of high school students. *Journal of the American Association for Information Science, 50*(1), 24-37.

Goodrum, A. (1997). Evaluation of text-based and image-based representations for moving image documents. Ph.D. Dissertation, University of North Texas, Denton, TX.

Graesser, A., Person, N. & Huber, J. (1992). Mechanisms that generate questions. In T.W. Lauer, E. Peacock, & A. Graesser, (Eds.), *Questions and information systems* (pp. 167-187). Hillsdale, NJ: Erlbaum.

Greenberg, J. (2001). Automatic query expansion via lexical-semantic relationships. *Journal of the American Association for Information Science, 52*(6), 487-498.

Hastings, S. (1995). An exploratory study of intellectual access to digitized art images. *Proceedings of the Fifty-ninth Annual Meeting of the American Society for Information Science*, 3-8.

Highly Interactive Computing in Education (Hi-Ce) (1998). Tactics and strategies: Leading on-line investigations. Web inquiry support manual. Ann Arbor, MI: University of Michigan.

Hirsh, S. (1997). How do children find information on different types of tasks? Children's use of the Science Library Catalog. Special Issue: Children and Digital Libraries. *Library Trends, 45*(4), 725-745.

Hirsh, S.G. (1998). Relevance determinations in children's use of electronic resources: A case study. In *Proceedings of the Sixty-First ASIS Annual Meeting*, Pittsburg, PA , 35, 63-72.

Hirsh, S. (1999). Children's relevance criteria and information seeking on electronic resources. *Journal of the American Society for Information Science, 50* (14), 1265-1283.

Hoffman, J. (1997). On-line learning materials for the science classroom: Design methodology and implementation. Unpublished manuscript. The University of Michigan.

Hoffman, J. (1999). Information seeking strategies and science content understandings of sixth grade students using on-line learning environments. Unpublished dissertation. University of Michigan.

Ingwersen, P. (1996). Cognitive perspectives of information retrieval interaction: Elements of a cognitive IR theory. *Journal of Documentation, 52*(1), 3-50.

Jacob, E. & Shaw,D. (1998). Sociocognitive perspectives on representation. In Williams, M.E. (Ed.), *Annual Review of Information Science and Technology, 33*, (pp. 131-185). Medford, NJ: Information Today.

Jansen, B., Spink, A. & Saracevic, T. (2000). Real life, real users, and real needs: A study and analysis of user queries on the web. *Information Processing and Management, 36*(2), 207-227.

Jorgensen, C. (1996). Indexing images: Testing an image description template. *Proceedings of the Fifty-ninth Annual Meeting of the American Society for Information Science*, Pittsburg, PA, 33, 209-213.

Kafai, Y. & Bates, M.J. (1997). Internet web-searching instruction in the elementary classroom: Building a foundation for information literacy. *School Library Media Quarterly,* 25(2), 103-111.

Kuhlthau, C. C. (1988). Developing a model of the library search process: Cognitive and affective aspects. *RQ, 28*(2), 232-42.

Kuhlthau, C. (1991). Inside the search process: Information seeking from the user's perspective. *Journal of the American Society for Information Science, 42*(3), 361-371.

Kuhlthau, C. C. (1995). The process of learning from information. *School Libraries Worldwide*, 1(1), 1-12.

Kulper, U., Schultz, U., & Will, G. (1997). *Information Services and Use, 17*, 201-214.

Large, A. , Beheshti, J. & Breuleux, A. (1998). Information seeking in a multimedia environment by primary school students. *Library and Information Science Research*, *20*(4), 343-376.

Large, A. , Beheshti, J. & Moukdad, H. (1994). A comparison of information retrieval from print and CD-Rom versions of an encyclopedia by elementary school children. *Information Processing and Management* 30(4), 499-513.

Large, A. , Beheshti, J. & Moukdad, H. (1999). Information seeking on the Web: Navigational skills of grade-six primary school students. Proceedings of the 62nd Annual Meeting of the American Society for Information Science, Washington, D.C., 36, 84-97.

Large, A., Tedd, L. & Hartley, R. (1999). Information seeking in the online age: Principles and practice. New Providence, N.J: Bowker Saur.

Lester, M. (1989). Coincidence of user vocabulary and Library of Congress subject headings: Experiments to improve subject access. Ph.D. dissertation, University of Illinois at Urbana-Champaign.

Lundgren, L. (1998, February). Helping children help themselves. *Scandinavian Public Library Quarterly*, 10-12.

Lynch, C. & Garcia-Molina, H. (1995). Interoperability, scaling, and the digital libraries research agenda. Available online at: http://www.hpcc.gov/reports/reports-nco/iita-dlw/main.html .

Lyons, D., Hoffman, J., Krajcik, J., & Soloway, E. (1997, March). An investigation of the use of the World Wide Web for sustained on-line inquiry in a science classroom. Paper presented at the meeting of the National Association for Research in Science Teaching, Chicago, IL.

Marchionini, G. & Teague, J. (1987). Elementary students' use of electronic information services: An exploratory study. *Journal of Research on Computing in Education, 20*, (Winter),139-155.

Marchionini, G. (1988). Hypermedia and learning: Freedom and chaos. Educational Technology, 28(11), 8-12.

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.

Marchionini, G. (1995). *Information seeking in electronic environments*. Cambridge:MA.

Markey, K. (1984). *Subject searching in library catalogs: Before and after the introduction of online catalogs*. Online Computer Library Center: Dublin, OH.

Maron, M.., Cooper, W., & Robertson (1982). Probability of relevance: A unification of two competing models for document retrieval. *Information Technology: Research and Development, 1*, 1-21.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. W.H. Freeman and Company: San Francisco.

McNally, M. J. & Kuhlthau, C. C. (1994). Information search process in science education. *Reference Librarian*, 44(1), 53-60. New Providence, NJ: Bowker Saur.

Meister, D. & Sullivan, D. (1967). *Evaluation of user reactions to a protoype on-line information retrieval system*. Report to NASA by the Bunker-Ramo Corporation under Contract No. NASA-1369, Report No. NASA CR-918, 1967. ERIC ED 019 094.

Miralpeix, C. (1994). Children's catalogs: A comparative study of catalogue use in two children's libraries. *ICBC, 23*(3), 43-46.

O'Connor, B. (1996). *Explorations in indexing and abstracting: Pointing, virtue and power*. Englewood, CO: Libraries Unlimited, Inc.

O'Connor, B. (personal communication, February 2001).

O'Connor, B., O'Connor, M. & Abbas, J. (1999).

Olson, G., Duffy, S. & Mack, R. (1985). Question-asking as a component of text comprehension. In A.Graesser & Black, J., (Eds.), *The psychology of questions*. Hillsdale, NJ: Erlbaum.

Ornager, S. (1997). Image retrieval: Theoretical analysis and empirical user studies on accessing information in images. Digital collections: Implication for users, funders, developers, and maintainers. *Proceedings of the 60th Annual Meeting of the American Society for Information Science*. Washington, D.C., 34, 202-211.

Peters, T. (1993). The history and development of transaction log analysis. *Library Hi Tech, 42*(11:2), 41-66.

Quintana, C. (2001). Symphony: A case study for exploring and describing design methods and guidelines for Learner-Centered Design. Unpublished dissertation. The University of Michigan.

Schacter, J., Chung, G., & Dorr, A. (1998). Children's Internet searching on complex problems: Performance and process analyses. *Journal of the American Society for Information Science, 49*(9), 840-849.

Scherer, M. (Ed.). (1997). How Children Learn. *Educational Leadership*, 54, (6).

Shera, J. (1965). Putting knowledge to work. In: Shera, J., (Ed.), *Libraries and the organization of knowledge*. Hamden,CT: Archon Books.

Solomon, P. (1993). Children's information retrieval behavior: A case analysis of an OPAC. *Journal of the American Society for Information Science, 44*(5), 245 - 264.

Sutcliffe, A. & Ennis, M. (1998). Towards a cognitive theory of information retrieval. *Interacting with Computers, 10,* 321 - 351.

Taylor, A. (1984). Authority files in online catalogs: An investigation of their value. *Cataloging & Classification Quarterly, 4*(3), 1-17.

Wallace, R. & Kupperman, J. (1997). On-line search in the science classroom: Benefits and possibilities. In E. Soloway (symposium chair), *Using on-line digital resources to support sustained inquiry learning in K-12 science*. Symposium conducted at the meeting of the American Educational Research Association, Chicago, IL.

Wallace, R., Kupperman, J., Krajcik, J. & Soloway, E. (1997). On-line information seeking in a sixth grade science classroom. Unpublished paper. University of Michigan.

Wallace, R. (1998). ARTEMIS: Learner-Centered Design of an information seeking environment for K-12 education" *Proceedings of ACM CHI 1998*, Los Angeles, 1998, 195-202.

Walter, V.A. (1994). The information needs of children. *Advances in Librarianship*, 18, 111-129.

Watson, J. (1998). “If you don’t have it, you can’t find it.” A close look at students’ perceptions of using technology. *Journal of the American Society for Information Science, 49*(11), 1024-1036.

Waters, D. (1998). What are digital libraries? *CLIR (Council on Library and Information Resources) Issues,* No. 4. Available online at: http://www.clir.org/pubs/issues/issues04.html .