ED 407 252 SE 060 160

AUTHOR Wallace, Raven; Kupperman, Jeff
TITLE On-Line Search in the Science Classroom: Benefits and Possibilities.
PUB DATE 97
NOTE 31p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, March 24-28, 1997).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Computer Uses in Education; Educational Resources; Elementary Secondary Education; Hypermedia; *Information Systems; Internet; *Navigation (Information Systems); *Online Systems; *Science Instruction; Technology; *World Wide Web

ABSTRACT

This study addresses some basic questions about students' strategies for seeking and using information from the World Wide Web. The questions pertain to the effects of environment and attitudes on students' use of online resources to find information and the development of a typology of strategies. The focus of the study was on describing student activities as they use online resources to find information. The goals of the study were to investigate how students use features of Web-based technology, and how they approach the task of finding and using information in the Web environment. From the results, it is clear that the students are not engaged and thoughtful by virtue of being online; rather, the students are constrained, which may be connected to their consistent reduction of the task to finding an answer or a perfect source. The subjects in this study adapted easily to using the Web through standard Web tools, but none of the focus students progressed beyond naive use of the tools available. Contains 31 references. (DDR)

On-Line Search in the Science Classroom: Benefits and Possibilities

Raven Wallace and Jeff Kupperman

The University of Michigan

In E. Soloway (symposium chair) Using On-line Digital Resources to Support Sustained Inquiry Learning in K-12 Science

AERA 1997

Wallace, R., & Kupperman, J. (1997,). On-line search in the science classroom: Benefits and possibilities. Paper presented at the AERA, Chicago, 1997.

The rapidly increasing quantity of free, readily available information on the Internet and in on-line[1] digital libraries creates unexplored possibilities for education. One of the most compelling of these possibilities, we believe, is the chance to allow students to engage more easily in inquiry-based learning. (Wallace, Krajcik & Soloway, 1996) Inquiry-based learning in the classroom has been difficult to accomplish for many reasons, one of which is the limited amount of materials available to the teacher. Now, students can find information on-line that would be difficult to obtain any other way: current scientific data, new ideas from research, and information from obscure sources not normally available in schools. With the enormous resources of the World Wide Web and digital libraries, students can have more opportunities to select topics in which they are truly interested. Students may become more engaged as they gain more control over their own learning. Inquiry-based learning has many components, and it is not the purpose of this paper to go into a thorough discussion of inquiry. One strength of on-line resources is information, and this paper will focus on the related component of inquiry, information seeking.

In much popular literature and in the political arena, there seems to be the assumption that connecting schools to the Internet has inherent value. For example, recent proposals for educational reform include the mandate to get every classroom and library connected to the Internet by the year 2000. (Christoff & Moorlehem, 1997; Education, 1996; Riley, 1996), and periodicals addressed to teachers are filled with articles about the benefits of being connected. (Cafolla & Knee, 1997; Dixon & Falba, 1997; Owston, 1997) As schools and libraries install connections to the Internet, more and more students are spending time looking for information on-line. Although it is easy to assume that students are having engaging, stimulating experiences

[1] Throughout this paper, on-line will be used to mean synchrously connected to the Internet, or available on the Internet.

while searching and browsing the World Wide Web, we actually know very little about what students do on-line. We do not know what strategies students use for seeking, evaluating, and using information. We have little evidence of what students understand about the resources available and the information seeking process, and how these understandings affect the strategies students use on-line.

While there is a body of research about how students seek information in libraries, looking for and using information from the Web is likely to be a significantly different process from using card catalogs, books, and shelves. We do not yet understand what these differences might be, but they may result from differing characteristics of the information, of the information seeking process, of the environment in which the activity is carried out, of the task itself, and of the attitudes and beliefs students bring to the task. Because the Web is a new genre of information, we don't have a typology of strategies to look for, and we don't have even a preliminary understanding of the aspects of students' knowledge or beliefs which are important to their success on-line.

In this study we attempt to address some of these basic questions. Our focus here is on describing students' activities as they use on-line resources to find information. The goals of the study are to investigate how students use features of Web based technology, and how they approach the task of finding and using information in the Web environment.

Theoretical Background

Research on information seeking by K-12 students in small full-text databases (such as CD Rom encyclopedias), bibliographic databases (electronic card catalogs), and physical card catalogs in library settings tells us that students are not very good at using Boolean logic, and that they do not understand the information seeking or research processes.[2] (Borgman, Krieger, Gallagher & Bower, 1990; Jacobson, 1995; Kuhlthau, 1993; Kuhlthau, 1996; Pitts, 1994; Walter, Borgman & Hirsh, 1996) Although the teacher may have a learning goal in mind (that students increase their understanding of a subject or integrate information from a variety of sources), students typically approach research assignments with a process goal (finding the right answer to a question, or finding a single source from which to draw relevant information). This may be because of the separation of research skills from classroom activities, or the lack of specific instruction on research skills. (Jacobson, 1995; Kuhlthau, 1996; Moore, 1995; Pitts, 1994; Stripling, 1995) These studies have looked at information as a library activity, in situations which are isolated from classroom activities. We do not know whether using on-line resources in a way that is more tightly connected with classroom activities can result in differences in student goals.

Elementary school students in school libraries avoid using catalogs, whether on-line or physical card catalogs. (Borgman et al., 1990; Walter et al., 1996) They prefer to go to the shelves and browse or to ask a librarian for help. When they do use on-line catalogs, they formulate simple queries: if they use more than one keyword, they tend to use "AND" connectors, avoiding any complex use of Boolean logic. In experiments with a new OPAC interface for children, Borgman reports that they prefer an intuitive graphical interface based on a bookshelf metaphor over a traditional text based interface, but they still show a preference for physical browsing over either on-line method. Children are more successful searching full text systems than bibliographic databases, in part because they can recognize relevant information (which they can readily see in full text databases) more easily

---

[2] Although there are differences between research and information-seeking, the term research is often used in K-12 classrooms as a synonym for information seeking.

than they can generate keywords (which are needed to use bibliographic databases.) (Jacobson, 1995)

At the high school level, students searching full-text databases experience an unusually high cognitive load (Marchionini, 1989): they must keep the task clearly in mind to figure out whether what they find is useful, and they must continuously appraise how they are doing in satisfying their goals for the task. Models for search behavior are based on strategies of experts (research librarians) and little is known about appropriate strategies of non-expert users in a full-text environment. Evaluation of sources in full-text systems is quite different than in bibliographic databases, since the user can assess information during the information seeking process, rather than after the document itself is retrieved from the shelves. (Marchionini, 1995)

Because the Web is a hypermedia system consisting of pages (nodes) with hyperlinks to other pages, literature on the use and impact of hypermedia in education has relevance for educational uses of the Web. This literature refers repeatedly to the cognitive overload and frequent disorientation users experience in using hypermedia. (Foltz, 1996; Hammond, 1989; Heller, 1990; Lehrer, 1993; Marchionini, Liebscher & Lin, 1991; Rouet & Levonen, 1996) In a comprehensive study of electronic text, Dillon (Dillon, 1996) points out that the characteristics which influence user behavior are not well documented or understood, and that many of the claims for use of hypertext have not held up to scrutiny. We know, for example, that reading speed is slower using electronic media than print media, but we don't know if that is a function of screen resolution or other factors. (Marchionini et al., 1991) Conflicting and confusing research results make it difficult to understand the role of hypertext in learning, or to improve designs of hypertext systems. (Heller, 1990)

Putting information in the form of hypermedia does not automatically have a positive impact on learning: the outcome is dependent on the design of the materials and the task, as well as characteristics of the user including goals, use of strategies, and understanding of the task. Jones (Jones, 1990) says it clearly: "The main point seems to be this, that it is not enough to simply present a reader with hypertext, no matter how varied the features of the system, and hope that they will develop a rich knowledge structure just because one underlies the database. What is essential is that they interact with the knowledge in ways that actively practice and develop cognitive strategies." The Web is a massive hypermedia system with minimal design consistency (provided only by the requirements of the HTML mark-up language), and we cannot assume that students will benefit from the information on the Web just because it is there and accessible.

Students' evaluation of the usefulness and relevance of what they find is another area of interest in this study. In the past, librarians have used the concept of relevance to evaluate content, but relevance has been defined as a characteristic of the system's response to a specific user query. For example, if the user searches using the term "volcano," relevance is judged by whether the documents returned are about volcanoes. Whether they meet the specific needs of that user at that time (for example, amount of detail and reading level) has not been a consideration of research on relevance.

Rather than looking at whether documents retrieved are inherently relevant, we are interested in how students make decisions about what information is valuable and how to use it. In analogous work on problem solving in history, Wineberg (Wineberg, 1991) identifies characteristics of novice problem solvers which may be applicable to behavior of students in on-line information seeking. He identifies significant differences between novices and experts in the domain of history in terms of how they view and use evidence. When given a problem, expert historians make theoretical models to interpret the evidence at hand. These models are not predetermined but emerge from evaluating the set of evidence as a whole. The experts recursively use the evidence to construct the model and use the model to understand the evidence presented. Students, on the other hand, assume there is a correct interpretation of the historical documents given, and proceed to derive an answer. They pay little attention to what the information *is* (e.g., the source, date, and reliability) focusing instead on what it *says*. This strategy is similar to the "copy-delete" strategy described by Bereiter and Scardamalia in novice writers. (Bereiter & Scardamalia, 1989) Students evaluating evidence on the Web may use similar strategies, possibly influenced by beliefs that there is a correct answer to their question, and that sources are equivalent -- in essence ignoring what a document is and focusing on what it says.

Classroom Context

Students in the study were enrolled in sixth grade classrooms at a middle school in Franklinville, a medium sized Midwestern town.[3] Based on a modified middle school model in which a team of two teachers works with a group of students for their four academic subjects, the curriculum is traditional in its coverage and presentation. The school draws students primarily from middle class families, but includes a wide range of socioeconomic levels as well as diverse racial and ethnic populations. The teacher involved in this study has been working with the University of Michigan Digital Library project since 1995, having completed several projects with students during the 1995-96 school year.

Students completed an introductory activity during which they learned fundamentals of using the Web and basic concepts of searching. The activity spanned five class periods with students using the Web to complete a somewhat structured "Scavenger Hunt." (http://mydl.soe.umich.edu/myscavenger/) Students worked in pairs to complete the introductory activity, and they

[3] Pseudonyms for people and places are used throughout this paper.

continued to work with the same partner for an ecology activity, during which data was collected for this study.

The ecology activity is part of a six-week curriculum unit which includes a variety of activities from doing lab experiments about groundwater to using a laserdisk program to study sanitary landfills, to taking a field trip to the local water treatment plant. In this class, the assignment for the on-line activity was to ask a question about ecology of interest to them and look for information about the question on the Web. The on-line activity came after two weeks of other activities about ecology, using the school district's "Physical Ecology Curriculum Guide." During these two weeks the students played roles in a simulation of a town with water quality problems, with activities about the water cycle, movement of ground water, parts per million, and a decision making exercises.

The on-line activity itself was structured and scaffolded by on-line materials created by UMDL researchers (http://mydl.soe.umich.edu/myecology/). (Hoffman, 1997) The materials gave the students a place to start on the Web, including links to on-line reference materials, to selected Web sites about ecology, and to search engines; and a mechanism for posting what they found on-line. (Bos, 1997)

The final student product from the on-line activity was in two parts: in their language arts class, the students created a "mind map," a type of concept map with their driving question as the central concept and important information they found shown in relation to their question. In science class, they presented their driving question and what they found to the class, and talked about successes and difficulties they had on-line. The activity was planned to last a total of seven days in science class, with four additional days in the language arts class for preparation and presentation of the mind map. The days in science class were used as follows:

Day 1: In class introduction, brainstorming of driving questions

Day 2-3: Lab days, on-line activities

Day 4-6: Lab or Library days - complete research and writing

Day 7: Class presentations

Prior to going on-line, the students spent time in class brainstorming questions. They were required to have their question approved before or during the first on-line session, during which they were to post their question on-line using software provided as part of the on-line learning materials. Once their question was approved, students began to look for information, beginning with the on-line ecology learning materials, and moving into Web searching and library research at their own pace. All students worked on these activities in pairs.

Part of the assignment was to keep a written journal of their information seeking activities. At various times during the project, the teacher instructed students to record their activities for the day, or their plans for the next day.

Methodology

Two pairs of students in each of two classes were selected for observation, for a total of eight students in four dyads. All pairs were single gender, at the suggestion of the teacher. The selections were based on teacher nomination, using the criteria that students represent a range of achievement, gender, and race. The students were asked if they were willing to participate, and all agreed. For purposes of this paper, the pseudonyms given to the pairs of students are Al and Bill (AB), Carl and Doug (CD), Ellen and Fay (EF), and Gail and Helen (GH).

Video data were collected for each pair of students during their time on-line, using equipment which captures video output from the computer, replicating what is seen on their monitor, along with an audio recording of the students' conversation during the session. This yields a complete record of what the students do on-line. Students were asked to talk about what they were doing, but no specific think-aloud training was provided. Student journals were also collected after the end of the unit.

Once video data were collected, the tapes were used to create a running description of students activities, including names of Web pages they visited with a description of what they did on each page; descriptions of the conversations between the students and with helpers; quotations from conversations; and a time stamp showing the tape index time for each item. Descriptions included how they got to the page (type of navigation); whether and how quickly they scrolled down the page; what they talked about while on the page; whether they encountered any problems or received any help; keywords used for searching; and the number of "hits" returned from each search.

From these description files, and with reference back to the tapes for clarification, each tape was coded to characterize the students activities along two dimensions: technology and task. In the area of technology, we were specifically interested in how students navigate on the Web and in how they use search engines. This coding involved counting the occurrence of certain events as well as characterizing the nature of the students interactions. For example, every page change was counted for a total number of page changes during the class period. A second count of unique pages visited was also made, since many of the page changes were simply steps on the way to another page. An example of this is when a student repeatedly uses the "BACK" button in the Web browser to navigate to some prior location. The "BACK" moves count as page changes, but not as unique sites.

The nature of their navigation was coded based on the variety and appropriateness of tools used: a student who used only one tool (typically the BACK button) received a "minimal" rating, while a student who used the

BACK button for going back one page, the GO menu to navigate to a somewhat earlier page, and made BOOKMARKs to get back to important pages would receive a "maximal" rating for that session.

Within the area of technology, codes were also developed for searching, including counting the number of searches and the number of keywords per search, and coding the complexity of their use of Boolean logic within searches. Search conception was analyzed by identifying episodes which gave evidence of the student's understanding of the way the search mechanisms work. For instance, an episode in which a student submits the exact same search over and over indicates some misunderstanding of the technology. Another misunderstanding is evidenced by a student who enters natural language phrases or complete sentences into the search engine.

In terms of the task, we looked for evidence of students' conception of the task, the nature of their goals and engagement with the task, and the strategies they applied to find and evaluate information. We were trying to see how the students approached the task in the dimensions of systematic v. random, engaged v. disengaged, and seeking an answer v. seeking understanding. The initial coding scheme was to mark evidence from the tape narrative using codes for goals, engagement, and systematicity, then to give the session an overall rating for each aspect. Ratings were then viewed across the five days of the on-line activity to look for consistency or change for a given pair of students.

We also coded instances when students encountered a problem with the technology, the process, or the content, and we coded those based on how the student went about dealing with the problem. Finally, we coded distinct conversation about technology, process, or content, as well as off-task conversations, and described the circumstances as well as the substance of those conversations.

A recursive coding method was used, during which themes were identified and tapes reviewed iteratively as themes emerged. For example, it became evident after the initial reduction of the tapes to narrative that many students were not doing things which helped them find content relevant to their question, but rather were trying to reduce the number of "hits" returned from a search to the smallest possible number. This might be an attempt to find an answer to their question, or, similarly, to find a perfect source for information. Tapes were reviewed to look for specific evidence of what students said about their search strategies. Another instance of recursive coding occurred as it became clear that students were visiting very few content sites. By content sites, we mean Web pages which were not provided as part of the UMDL materials, and were neither search engines nor hit lists.[4] The number of content sites visited and the amount of time per content page were added to the coding.

Findings

Four themes emerged from analysis of the data, each of which will be presented in detail below. The themes are:

1. Students don't explore much.

2. Students tend to seek answers rather than aim for understanding.

3. Students don't evaluate sources: they receive content from on-line sources at face value.

4. Students learn to use Web tools easily, but use them naively.

[4] By "hit list" we mean results returned from a search. On the Web, these are in the form of a Web page, with hot links to the sites returned. The search engine used by these students, Open Text, includes a brief excerpt from each page returned.

Theme 1: Students don't explore much.

Students were much more constrained in their Web excursions than we expected. Instead of seeing students "surfing" innumerable pages on and off task, we observed that students stayed very close to home. They were very busy, submitting searches, getting back lists of sites, and making lots of navigational moves, but they rarely went more than a few links away from the search page. The largest number of links away from the search page was five: that is, none of these students went more than five links away from the latest hit list. When students visited content sites , they rarely followed links within the site, instead looking at a page or two, then returning to the familiar turf of searches and hit lists.

There were occasional off-task episodes in which students surfed several links (in particular, there were two instances of student pairs visiting local hometown sites, including the mall home page). There was only one episode in which students found a set of content pages about which they got excited and in which they browsed. Interestingly, in that episode, the students found something which they didn't think was relevant to their own question, but which they thought one of the other pairs of students could use. These two students spent over twelve minutes looking at these sites and trying to share them with the students who they thought would benefit from them.

Table 1 shows data which gives details about what the students did while they were on-line. The first two data columns show how many page changes each pair averaged per class, and the average time per page. These data include pages which are visited incidentally as a navigation move. The average across all pairs is 49 page changes per class, with an average time per page of .81 minutes, or just over 48 seconds. Since the total times include all downloading and navigation time, it is apparent from these numbers that the students are moving quickly from page to page. In many instances, they scrolled down pages too quickly to conceivably read the words presented. Often, in the case of hit lists, students went back to search without even scrolling through the list of the first ten hits. This rate of changing pages alone tells us how little engagement students have with page content. An average page time of 30 seconds, the minimum among these students, is undoubtedly too fast to extract information.

	1	2	3	4	5	6	7
Student Dyad	# URL Changes per period	Average time per page(in minutes)	Total number of content sites (all periods)	Percent of time on content page	Total number of searchers submitted	Total # of hits returned	Maximum # links away from search page
AB	38	.96	11	27%	21	4,256	5
CD	56	.82	28	23%	25	8,737	4
EF	66	.50	37	22%	37	48,685	4
GH	35	.97	22	23%	49	18,511	4
Averages	49	.81	25	24%	33	20,047	4

Table 1: On-line activity by students

The third and fourth data columns address student use of content pages. The content pages may or may not be related to the students' questions, but the students have chosen to look at these pages while they are seeking information about their question. As the table shows, during the ecology activity these students spent an average of 24% of their time on content pages. Even though the percentage of time on content sites was fairly consistent, the number of content pages visited by each pair varied widely, with a low of 11 and a high of 37. In two of the pairs, these content pages included an episode of off-task surfing, during which they visited sites of local businesses. The interesting thing about these numbers is what they were doing with the other 76% of their time: primarily searching. The fifth column shows the number of searches submitted by each pair of students, ranging from a low of 21 to a high of 55. Over a four or five day period, this represents a lot of searching. Analysis of the number of hits returned during these searches (Column 6) reveals that these students had between 4256 and 48,685 sites returned for their consideration.

Theme 2. Students tend to seek answers rather than aim for understanding.

Consistent with previous research on student behavior in library environments, many students seemed to see their assignment as finding an answer to their question, and thus they reduced the task to finding a single page, the perfect source, on which the answer could be found. Evidence of this is found in their submitting searches one after another without looking at any of the returns until they have a very small number of hits. In several instances, students made exclamation such as "I've got it" or "Yeah!" when they got the number of hits under ten.

One pair of students, Carl and Doug, rushed to be "finished," regularly calling to the teacher for permission to stop, even though they had found no content relevant to their question. This pair reacted effusively to small hit lists, singing and calling out "yes, we got it now...hey you guys, we got it!" when they saw that the number of hits from a search was 18, then reacting with equivalent disappointment when a cursory viewing of the hit list did not reveal an obviously appropriate site: "All these things stink...cause we put in *animals*...let's delete *animals*." Later, these students produced a hit list with only three pages, and Doug exclaimed, "Oh my gosh, we got it!"

Ellen and Fay were looking for information about the question "How long does it take for water to go through the water cycle in the air?" After the second day, their journal entries read:

Fay: "What we learned today is we didn't get the answers. It's hard to find the keywords of it. I Fay and Ellen couldn't find keyword of how long so we couldn't find about our information." [sic]

Ellen: "Today we were looking for the anserw but we did NOt find it. I was hopeing that we would find it." [sic]

On the third day, Fay was absent, and Ellen wrote:

"I can't find the anserw to our question! This is relly fustrating! ... Today Fay wasn't here so I had to search by my self it got relly fustrating becuase I could of found it in a book but I didn't find it on the Internet?" [sic]

On the fourth day, Ellen wrote:

"tomorrow we need to find the anserw to our question. Today we did not relly find any thing. It got relly annyoning that we didn't find anything." [sic]

Ellen and Fay	Number of content pages visited	Maximum time per content page (min:sec)	Minimum time per content page	Average time per content page
Day 1	0			
Day 2	13	:50	:05	:24
Day 3	13	5:30	:10	1:03
Day 4	11	1:50	:10	:32
Day 5	0 (library)			

Table 2: Ellen and Fay: Interaction with Content Pages

Analysis of the actual activities of Ellen and Fay on-line shows that they rarely looked at content pages at all, and when they did, they did not spend enough time on them to read what information was there. Table 2 shows the number of content pages visited during their on-line activities, along with maximum, minimum and average times per content page. On the third day, Ellen spent more time looking at content, spending over five minutes on an EPA page, and over three minutes on a USGS page. Upon further analysis, it turns out that the EPA page is a list of hot links to topics, and, although she stays on the page for a long time, much of that time is spent talking to one of the helpers in the lab (a UMDL researcher trying to figure out good keywords. Ellen does not pick any of the links from this page for further investigation, but goes back to search again. The second page on which Ellen spent some time was a collection of surface water data from USGS. The page was selected by the teacher for Ellen when she asked for help, and the teacher instructed her to record the site in her journal, which she did, even though she did not connect the information on the page to her question in any way. On the fourth day, they visit two sites for more than a minute each, but again spend much of that time talking to the helper about not finding their answer.

On the last day of the on-line activities, Ellen and Fay went to the library and used books to answer their question. Their journal entries for that day read:

> Fay: "In the book called the Biosphere it tells about how long it take water to through air. The average days is 9 days in the air." [sic]
>
> Ellen: "Today sience we haven't found anything on the Internet we are looking in books. in a book called the Bio Sphere it said that the avrege time in most parts of the world is about 9 days but in other parts of the world like a desert it takes longer. and in the tropics it takes a shorter period of time." [sic]

Several of the students treated searches as if they were methods of finding indexes or tables of contents to the Web. When they found a good site, Al and Bill gave their neighbors the search words they used to find the site rather than the name of the site or the URL of the site. This occurred on several occasions. In fact, during the days of their on-line activity, the easiest way for some students to get back to a site they found previously was to reproduce the search and look for links they had previously followed.

Much of this evidence points to students understanding of the task or the nature of the resources available on the Web. Their task was reduced to finding an answer or a perfect source for an answer; their conception of the resource, at least initially, was that the answer or perfect page was there, if they could come up with the right way to find it. In many instances, they treated hit lists as a table of contents, and used them much the way they might use a textbook.

Theme 3. Students don't evaluate sources.

Students method of evaluating sources seemed to be to look for the words they expected to find in an answer to their question, and to accept the source as valuable if they found those words. For example, Carl and Doug posed a question about water pollution: Is your drinking water safe? They eventually decided to look for information about how cars affect water pollution, and they found a site called "Jones Act Examples." (http://www.lexitech.com/jarc/examples.html), which is a site advocating a political stance on reform of a shipping law which regulates international shipping. The page included sections about agriculture and automobiles and had many references to water. Although Carl and Doug visited the page repeatedly during their days using the computer, they never read enough of the page to understand that its content had nothing to do with their questions, and they used it as evidence that they had "finished" their assignment. In their journal after the third day, they wrote:

> Doug: "We found how cars effect water in midwest. the author is Jones act example. URL http://www.lexited.com/jarc/examples.htm/ The Jones act find damage on industries in America. ex. water, coal and agriculture."

> Carl: "We searched for Franklinville most pollution. We got 5 pages. Then we went and looked up water effect on water and Midwest. The URL is http://mydl.soe.umich.edu/myecology/keyfram.htm. The Jones act's pervasive and damaging Impact on American Industry. The act produces around $635.6 million in profits annually for the cabotage. So it gets money to help the American Industries not polute as much. The author is Huston Regional Group. The date was made is May of 1996." [sic] [5]

[5] This journal entry also reflects a misunderstanding of URLs on Carl's part, and a problem with using frame pages with naive users.

Their entries reflect an understanding of the Web as a place where you "look things up" with keywords. Carl drew the conclusion that the Jones Act page answered his question, and in his notes, he tried to make it fit.

In a related episode, Carl and Doug searched for information about the most polluted places in their town, Franklinville. They searched using the keywords "Franklinville, most polluted," and got back a hit list with five hits. Without visiting any of the sites on the hit list, they leapt to the conclusion that the five sites which appeared on the hit list were the most polluted "places" in the town. Doug told the teacher, "We want to know if we can go on with the search [start a new question] because we found what is the most polluted area in Franklinville." When the teacher pushed him about how he knew they were the most polluted places, he backed off, replying that there really was no information on polluted areas in Franklinville.

Theme 4. Students learn to use Web tools easily, but use them naively.

Students learned to use the Web in a previous introductory project, an on-line Scavenger Hunt. Few students experienced problems with learning the basics of starting up the computer, opening the Web browser, using search engines, and navigating with hyperlinks. However, closer analysis of their patterns of use indicates that many students may not have progressed beyond a very basic ability to use these tools.

Some of the patterns of use which indicate a relatively naive use of the tools are:

• Use of the BACK button as the primary means of navigating. All four pairs of students used the BACK button almost exclusively to navigate among Web pages. This regularly led to extended trips back through pages. At one point, Al and Bill used the BACK button twenty five consecutive times to find the page they were looking for. Although this was extreme, the other dyads had maximum numbers of consecutive BACKs of 9, 8, and 5 respectively. In all there were fewer than ten instances of using the GO buttons among all the pairs for all sessions; and only four bookmarks were created among all the groups.

• Incorrect use of the GO TO box for locating a page by entering the URL. Two of the dyads tried to use the GO TO box, but they incorrectly thought it was connected to the SEARCH button on the search engine page. Thus, they typed the desired URL into the GO TO box, then clicked on the SEARCH button. The search engine returned an error message, SORRY, no keywords entered. Neither group ever figured out what they were doing wrong.

Dyad	Searches	Unique Keywords	Total Keywords	Average Keywords per search
Al & Bill	21	11	58	2.8
Carl & Doug	25	16	110	4.4
Ellen & Fay	37	15	112	3.0
Gail & Helen	49	31	166	3.4

Table 3: Search Activity

• Simple, repetitive use of keywords for searching. Although students submitted many searches, they did not use feedback from the search engine to systematically improve their searches. Their search behavior was often incomprehensibly random. Table 3 shows data about searches. From this data, we see that students regularly use more than one keyword. However, there was only one instance of students using a connector other than the default AND: Ellen and Fay searched for "water cycle FOLLOWED BY time". Gail and Helen submitted a large number of searches - 49 - using many different keywords - 31 different words in all - but as described above, their analysis of the results of their searches, and their use of content pages, was unsystematic at best.

A normal pattern of search behavior was for the students to enter several keywords, then embellish those words with additional, changing keywords, looking at the hit list only if the number of hits became very low; and sometimes not looking at hits for a particular sequence of keywords at all. Table 4 is an illustration of a cycle of searching by Gail and Helen, and shows a typical pattern of behavior. These searches occurred on the first day of on-line activities, within a six minute time frame, and none of the pages on the hit lists were visited during this class period. At the end of this sequence of searches, as the class period ended, Gail and Helen copied down the URL for the hit list of the last search. In the search engine they were using, the URL does not reproduce the search results. The next day, they attempted to go to the URL, and after several failures, began to search again using similar terms to those in the table. It was thirty minutes and ten more searches into the period before they visited a content site from one of the hit lists, and they visited only one content site that day.

Search #	Time	Keywords	Hits
1	22:18	pollution in Franklinville	0
2	24:07	pollution, Illinois	743
3	24:20	pollution, Illinois Franklen County [sic. misspelled Franklin]	0
4	25:00	pollution, Illinois chicugo [sic. misspelled Chicago]	0
5	25:30	pollution, Illinois, midwest,	55
6	27:40	pollution, Illinois, midwest, countys [sic]	0
7	28:00	pollution, Illinois, midwest, Franklinville	8

Table 4: Search example - Gail & Helen

- Students sometimes consecutively enter the exact same search to the same search engine. For example, Al and Bill search on the terms "bad, water" two times in a row without taking action on the hit list either time. Ellen and Fay enter an identical search twice within half a minute and they search on "water cycle" several times during the five day period, getting the same list of 366 hits. At one point they told the teacher that "water cycle has 366 pages." On the day following the sequence shown in Table 4, Gail and Helen submitted the same search, "pollution, Franklinville, most polluted, areas," three times.

- Another problem with searching is seen when students use natural language. Several of the pairs entered phrases rather than keywords for some searches: Gail and Helen's first search used the phrase "pollution in Franklinville." Al and Bill used the phrases "effect on people" and "deaths by water" during their second day of searching. Carl and Doug inserted the preposition "on" in many of their searches on the second day, not noticing that the same hits are returned with and without the word "on" in the keyword list. Ellen and Fay repeatedly used the phrase "most polluted" in conjunction with other keywords on their second, third and fourth days of searching.

Search #	Elapsed Time	Keywords	Hits
1	2:33	most, pollution, Franklinville	0
2	5:25	Franklinville, most, polluted	5
3	6:35	Franklinville, most, polluted, areas	4
4	6:47	Franklinville, most, polluted, areas, on	4
5	7:17	Franklinville, most, polluted, areas, on, Main St.	0
6	7:43	Franklinville, most, polluted, areas, on, animals	0
7	7:59	Franklinville, most, polluted, areas	4

Table 5: Search example, Carl and Doug

• Sometimes students added keywords to a search which had obtained very few hits. This may indicate some confusion about "broadening" and "narrowing" searches. For example, Carl and Doug made the sequence of searches shown in Table 5. They may have been unclear about how the search engine worked, or they may have been trying to narrow down the number of hits to one. In this instance, the students did not look at any of the hits returned by these searches, and at the end of this sequence, they changed topics and searched using an entirely different set of keywords. This example, and others mentioned above, make it hard to understand the students' conception of searching.

Discussion

The findings of this study have focused on the details of students activity and behavior as they look for information on-line. From these results, it is clear that these students are not becoming engaged and thoughtful by virtue of being on-line, even though there is plenty of anecdotal evidence that they enjoy these on-line activities. Although they are free to explore the Web to find interesting information, their behavior is constrained, and very little excitement about the Web as a content resource is in evidence. Their unwillingness to explore may be connected to their consistent reduction of the task to finding an answer or a perfect source, and, in some cases, to an apparent conception of the Web as a place where they can find the answer. In accord with this conception, most of these students accepted what they found on the Web as true, with no consideration of the source or purpose of the information. Another possible interpretation of their willingness to use whatever they found is that they were so happy to find anything – to get their task done – that they didn't care whether it was good information or not. They adapted easily to using the Web through standard Web tools, but none of the focus students progressed beyond naive use of the tools available.

The limited scope of this study makes it impossible to draw general conclusions about larger populations. However, the study does give one view of student activity on the Web, in what might be a typical use of Web resources. Although UMDL materials were carefully prepared to get students started on-line, the activity did not include an extensive focus on teaching students the research process, and it was a short, self-contained project, embedded in their usual curriculum. This is an activity which could occur in any classroom with Internet access, requiring neither special pedagogy nor access to special technologies. The ecology on-line activity was a step toward using on-line resources as part of inquiry, but doing this activity did not require a long term commitment to inquiry based learning. For researchers it provided a chance to investigate the skills and strategies students bring to the task.

The results of the study say a great deal about some of the assumptions underlying the current rush to get all schools and libraries connected to the Internet. On the one hand, we observed that, in spite of a carefully constructed set of activities and Web pages, these students were not very successful at finding useful information or at using the Web thoughtfully. On the other hand, for most of the students in the classrooms we have worked in, the Web was a relatively new and unfamiliar resource, and they were quite successful at getting started with these new tools. The problems we observed were rooted in the students' conception of the task, and possibly in the nature of materials available on the Web. We are in the process of collecting longitudinal data on the students in this study, as well as students in high school classes, to analyze whether their strategies and understandings change over time.

Even though the results are neither what we wanted to see, nor what one would expect based on the popular talk about the Internet, we take the results to be a solid starting point both for further research and for improving students' skills in and understanding of how to seek and use information in scientific investigations. For example, among the many issues which merit further research is the question of why students don't read on-line. Is it simply because most of the materials they find are at too high a reading level, a fact which they have learned in prior experiences with Web materials? Or is it because it is hard to read on-line even for adept and motivated readers? (Dillon, 1994) Another hypothesis is that the Web environment itself distracts students from the more mundane task of reading.

Similarly, we don't know what factors are playing into students' reluctance to explore for information. Is it the nature of the task we have given them which restrains their behavior, or their reduction of the task to "school work" ? Or have we simply not given them adequate time to explore?

A problem we encountered in the course of this study was the nature of the content students were able to find. Content available on the Web for K-12 science inquiry is unpredictable, changing, and hard to find. It is almost impossible to predictably find information about a given question, especially an ill-structured, complex question. This is a bit counter-intuitive – we started this research thinking that open-ended, ill-structured questions would be the best ones to pose when using Web resources, given the nature of the Web. Instead, the opposite may be true: the Web may be better suited as a source of specific information, which answers specific, simple questions. Anecdotal evidence tells us that, when these same students have free time on the Web, they can find exactly what they want – from the MTV home page to the local mall store sites. Since this study was not focused on content issues, we have made no attempt here to address how problems with content are related to other problems identified as students use the Web. However, it is fertile ground for future research, and in upcoming activities in the UMDL research project, we will use the project's Digital Library testbed rather than the Web as the primary content base for student activity. This collection will be smaller, better organized, and searchable based on user characteristics such as level of expertise or education.

Our research will pursue several aspects of the findings of this study, and the issues raised peripherally by this study, trying both to improve the way that students interact with digital information resources and to understand better why we are seeing the particular behaviors we have seen. On the side of improving the tools and techniques, we are working on the following:

1. Our project is pursuing use of different approaches to information organization and seeking (Atkins et al., 1996), and we are designing an information seeking interface for use in the schools to reflect these differences. This interface will be piloted in the spring of 1997.

2. We are developing organized collections of Web- and non-Web based digital resources appropriate for the K-12 audience in the UMDL testbed. We will study whether a less chaotic information environment is more conducive to thoughtful approaches to information seeking.

3. We are working with teachers on improved conceptions of pedagogy, to contextualize the activities of finding and using information as a part of inquiry in science classrooms. We are looking at task definitions as well as classroom activities.

To better understand student behavior as they use on-line resources, we will continue this study with longitudinal data from several classrooms at the middle and high school levels. New tools will be deployed through UMDL, and new approaches to pedagogy and task will be implemented.

Working from the belief that on-line digital resources can be valuable in K-12 classrooms, we realize very clearly that neither the Web, nor the University of Michigan Digital Library itself, will be a panacea for solving the difficult problem of helping students do scientific inquiry. We have not yet created a context which encourages inquiry to the extent we think is possible, but with improved tools and more attention to classroom context, we are optimistic that we can make progress toward the goal of using digital resources as an integral part of scientific inquiry.

This work was supported by the NSF/ARPA/NASA Digital Library Initiative; by a grant from the NSF NIE Initiative for the University of Michigan Digital Library project (RED-9554205); and by the University of Michigan.

Bibliography

Atkins, D. E., Birmingham, W. P., Edmund H. Durfee, Glove, E. J., Mullen, T., Rundensteiner, E. A., Soloway, E., Vidal, J. M., Wallace, R., & Wellman, M. P. (1996). Toward Inquiry-Based Education Through Interacting Software Agents. Computer, 29(5), 69-76.

Bereiter, C., & Scardamalia, M. (1989). Intentional learning as a goal of instruction. In L. B. Resnick (Ed.), Knowing Learning and Instruction: Essays in Honor of Robert Glaser, (pp. 361-392). Hillsdale, NJ: Lawrence Erlbaum.

Borgman, C. L., Krieger, D., Gallagher, A. L., & Bower, J. (1990). Children's Use of an interactive science library: exploratory research. School Library Media Quarterly, 1990(Winter), 108-113.

Bos, N. (1997,). Student publishing of value-added contributions to a digital library. Paper presented at the AERA, Chicago.

Cafolla, R., & Knee, R. (1997). Creating educational web sites. Learning and Leading with Technology, 24(5), 13-16.

Christoff, C., & Moorlehem, T. V. (1997, March 7, 1997). Clinton is a big hit. Detroit Free Press, pp. A1.

Dillon, A. (1994). Designing usable electronic text. London: Taylor & Francis Ltd.

Dillon, A. (1996). Myths, misconceptions, and an alternative perspective on information usage and the electronic medium. In J.-F. Rouet, J. J. Levonen, A. Dillon, & R. J. Spiro (Eds.), Hypertext and Cognition, (pp. 25-42). Mahwah, N. J.: Lawrence Erlbaum Associates.

Dixon, J. K., & Falba, C. J. (1997). Graphing in the information age: Using data from the world wide web. Mathematics Teaching in the Middle School, 2(5), 298-304.

Education, D. o. (1996). The future of networking technologies for learning, . http://www.ed.gov/Technology/Futures/toc.html: Department of Education.

Foltz, P. W. (1996). Comprehension, coherence and strategies in hypertext and linear text. In J.-F. Rouet, J. J. Levonen, A. Dillon, & R. J. Spiro (Eds.), Hypertext and Cognition, (pp. 109-136). Mahwah, N.J.: Lawrence Erlbaum Associates.

Hammond, N. (1989, May 9-11, 1989). Hypermedia and learning: Who guides whom? Paper presented at the Computer Assisted Learning 2nd International Conference ICCAL, Dallax, TX.

Heller, R. S. (1990). The role of hypermedia in education: A look at the research issues. Journal of Research on Computing in Education, 1990(summer), 431-441.

Hoffman, J. L., Kupperman, J., Wallace, R. (1997,). On-Line learning materials for the science classroom: Design methodology and implementation. Paper presented at the AERA, Chicago.

Jacobson, F. F. (1995). From Dewey to Mosaic: Considerations in interface design for children. Internet Research: Electronic Networking Applications and Policy, 5(2), 67-73.

Jones, T. (1990, 1990). Toward a typology of educational uses of hypermedia. Paper presented at the Computer Assisted Learning 3rd International Conference, ICCAL, Hagen, FRG.

Kuhlthau, C. C. (1993). Seeking Meaning: A Process Approach to Library and Information Services. Norwood, N. J.: Ablex Publishing Corporation.

Kuhlthau, C. C. (Ed.). (1996). The virtual school library: Gateway to the information superhighway. Englewood, CO: Libraries Unlimited, Inc.

Lehrer, R. (1993). Authors of knowledge: Patterns of hypermedia design. In S. P. Lajoie & S. J. Derry (Eds.), Computers as Cognitive Tools, (pp. 197-227). Hillsdale, NJ: Lawrence Erlbaum Associates.

Marchionini, G. (1989). Information seeking in electronic encyclopedias. Machine-Mediated Learning, 3(3), 211-116.

Marchionini, G. (1995). Information seeking in electronic environments. Cambridge: Cambridge University Press.

Marchionini, G., Liebscher, P., & Lin, X. (1991). Authoring hyperdocuments: Designing for interaction. In M. Dillon (Ed.), Interfaces for information retrieval and online systems: State of the art, (pp. 119-131). New York: Greenwood Press.

Moore, P. (1995). Information problem solving: A wider view of library skills. Contemporary Educational Psychology, 20, 1-31.

Owston, R. D. (1997). The world wide web: A technology to enhance teaching and learning? Educational Researcher, 26(2), 27 - 33.

Pitts, J. M. (1994). Personal understandings and mental models of information: A qualitative study of factors associated with the information-seeking and use of adolescents. , Florida State University, Florida.

Riley, R. (1996, November 7, 1996). Statement by Richard W. Riley U.S. Secretary of Education on the Decision of the FCC/Joint Board to Provide Universal Service to Schools and Libraries.

Rouet, J.-F., & Levonen, J. J. (1996). Studying and learning with hypertext: Empirical studies and their implications. In J.-F. Rouet, J. J. Levonen, A. Dillon, & R. J. Spiro (Eds.), Hypertext and Cognition, . Mahwah, N. J.: Lawrence Erlbaum Associates.

Stripling, B. K. (1995). Learning centered libraries: implications from research. School Library Media Quarterly, 23(3), 163-170.

Wallace, R., Krajcik, J., & Soloway, E. (1996). Digital Libraries in the Science Classroom: An Opportunity for Inquiry. D-Lib Magazine(http://www.dlib.org/dlib/september96/umdl/09wallace.html).

Walter, V. A., Borgman, C. L., & Hirsh, S. G. (1996). The science library catalog: A springboard for information literacy. School Library Media Quarterly, 1996(Winter), 105-110.

Wineberg, S. S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. Journal of Educational Psychology, 83(1), 73-87.

AERA CONFERENCE 1997

SE060160

ERIC

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: On-Line Search in the Science Classroom: Benefits & Possibilities

Author(s): Raven Wallace & Jeff Kupperman

Corporate Source:

Publication Date: 3/97

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following two options and sign at the bottom of the page.

[X] **Check here For Level 1 Release:** Permitting reproduction in microfiche (4" x 6" film) or other ERIC archival media (e.g., electronic or optical) *and* paper copy.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Level 1

The sample sticker shown below will be affixed to all Level 2 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Level 2

[ ] **Check here For Level 2 Release:** Permitting reproduction in microfiche (4" x 6" film) or other ERIC archival media (e.g., electronic or optical), but *not* in paper copy.

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Sign here→ please

Signature: Raven M Wallace

Printed Name/Position/Title: Raven Wallace

Organization/Address: Univ. of Michigan, 2203 Lafayette, Ann Arbor Mi 48104

Telephone: 313 663 6030

FAX:

E-Mail Address: ravenmw@umich.edu

Date: 3/27/97