

## Introduction

There is ongoing debate about the ways in which college students approach information gathering and which tools they use to find information. Since the onset of online searching, librarians have advocated databases with complex search mechanisms that employ controlled vocabulary and require highly stylized queries. However, as public search engines now dominate the market, younger students have grown up with simpler interfaces and the more natural language of search engines.

There is a growing movement to develop technologies designed to harness the power of metadata and taxonomies of controlled vocabulary while providing the end-user a simple and seamless search interface. Database developers and librarians are at a crossroads in information retrieval. Should students learn to use more complex search strategies based on print-based research tools or should designers develop interfaces that more closely relate to millennials' mental models of Internet-based information retrieval with engines that more accurately parse a simpler, more natural language query? If developers choose to do the latter, what are millennials' mental models of information retrieval? How do today's college students differ from the generations who preceded them and how can developers of design interfaces with these mental models in mind?

## Literature Review

A recent study found that 96% of college students surveyed use search engines to find information for at least some of their assignments; 37% use them for most assignments and 42% use them for all assignments (OCLC, 2002). Most (70-75%) believe they are successful with Internet search engines (OCLC, 2002). However, 77% express at least some degree of frustration while searching, and as many as 29% were frustrated with Internet searching (Sullivan, 2001).

These frustrations are even higher with library resources. Although OCLC (2002) found that 70% of students use the library's website for their research, 43% thought that other sites had more valuable information than the library, and a majority of students (55%) cited difficulty finding full-text articles. They said that they have difficulty navigating and searching library tools and suggested that libraries make their resources easier to access and use (OCLC, 2002).

This research attempts to examine the types of queries students create and to understand their mental models of search in order to determine how tools can be designed to better meet their needs.

## Participants

This study observed 21 students in their first semester at the University of Baltimore (UB). UB's first-year curriculum revolves around the concept of learning communities with three interrelated courses—one humanities, one social science and either an information literacy or speech communications course. I recruited students from three of the four learning communities in Fall 2008.

Although small in number, the 21 students demographically represent the entering class at UB: 52% women and 48% men; 48% white, 43% African American and 9% Asian; 29% 19 years old, 48% 18, 14% 17, and 5% 16 years old.

All participants had at least one computer at home: 57% had two or more. Ninety-one percent had a profile on a social networking site. Figure 1 shows their daily Internet use.

	At least 1 hour/day	2+ hours / day
Hours of Internet use	95%	62%
Hours of Web surfing	71%	33%
Hours of Social Networking	51%	24%

Figure 1. Internet use

## Methodology

The study used a contextual inquiry methodology to examine students' approach to research and their mental models of information retrieval. Contextual inquiry is a field research method whereby the researcher observes and interviews the participant while he/she conducts a real-world task. This methodology has three primary elements: 1) the researcher gathers data within the context of the participant's work, 2) the researcher and participant become partners in the exploration of the task and solutions, and 3) the research is focused on a defined set of issues rather than on a list of questions (Raven & Flanders, 1996). This methodology is appropriate for study of a current product or process in completing routine tasks (Kantner, Sova, & Rosenbaum, 2003).

Each student had a research assignment for which they chose to use online resources. During a three-week period in September-October 2008, I met with each student for approximately 60 – 80 minutes. After an initial introduction, the students researched their assignments; although there were four distinct assignments, 15 of the 21 researched two political parties' plans for a social or economic issue. I observed students use their natural choice of search engines over article databases. In addition to personal observation and note taking, I videotaped each session and captured the search paths of each participant using Morae software. I encouraged each student to walk me through their research process, thinking aloud about his/her search strategies and anticipated search results. After the observation I asked each participant to explain and/or diagram his/her concept of the search system and how the search is processed.

## Results

The students performed a total of 209 discrete searches; 129 were in search engines, and 80 were in library databases. An overwhelming majority of searches (73%) were in Google. Most database searches (74%) were in Academic Search Premier (See Figure 2).

Search Engine	Searches	%	Database	Searches	%
Google	93	72.1	Acad. Search Premier	59	73.8
Alta Vista	7	5.4	JSTOR	6	7.5
Yahoo	5	3.9	Opposing Viewpoints	4	5.0
Ask	5	3.9	InfoTrac	4	5.0
Google Scholar	4	3.1	LexisNexis	3	3.8
Google Gov	3	2.3	Lib. Catalog	1	1.3
Google Images	3	2.3	Journal Finder	1	1.3
PolyCola	3	2.3	Project Muse	1	1.3
MSN	2	1.6	Gen. Business File	1	1.3
Wikipedia	2	1.6	Baltimore Sun	0	0.0
Clusty	1	0.8			
Encarta	1	0.8			
<b>Total</b>	<b>129</b>	<b>100</b>	<b>Total</b>	<b>80</b>	<b>100</b>

Figure 2. Search engines/databases used

Students generally conducted very simple searches, using personal names or two- or three-word phrases that conveyed a single concept. Those who tried to combine two or more distinct concepts often linked them with inappropriate Boolean operators, particularly in the databases. Figure 3 indicates the types of searches students conducted and the additional actions (both positive limits or truncation and incorrect Boolean logic or search mechanisms).

## Results

Of the seven students who did use Boolean logic and truncation, five used only the AND operator, and two used it incorrectly. Of the two who used both AND and OR, only one used them correctly. These mistakes paired with frequent misspellings led to few if any relevant results in the databases. In fact, more than half (52%) made one or more spelling error. Students rarely saw those mistakes; only in two instances did the participants catch their own mistakes (in Google they were alerted with a "Did you mean?" set of results). A majority of students (57%) did attempt to narrow a search by adding terms, but they did not consistently use this strategy (11.5%) and often returned to an original overly broad search in a different tool. Only four students seemed to be deliberate in their attempts to focus their searches, but they, too, made mistakes in these strategies. When asked afterwards about alternative search terms, some stated they might look for synonyms, but only one looked for additional terms during the search.

Method	No.	%	Method	No.	%
Simple Searches	80	38.3	Broaden / Narrow	24	11.5
Phrase w/ Multiple concepts	79	37.8	Limits	2	1.0
Boolean search	38	18.2	Truncation	5	2.4
Subject search	3	1.4	Boolean mistakes	18	8.6
Use of subject headings	9	4.3	Incorrect punctuation	4	1.9
	209	100	Spelling errors	18	8.6

Figure 3. Searches and use of standard operators

Four students went directly to sites they knew to find information. Almost half (48%) searched within sites: some searched in news sites for their topics, for those researching political parties three searched for their issue within the party site. Additionally, three students used "find within page" features in the browser or Adobe to find search terms within a document.

Several students relied heavily on one or two sites in their research, browsing more than searching. Twelve (57%) used sites' global navigation to find information, particularly in the case of the political assignment, using party sites to find information on their issues. Another 52% (11) used contextual links to move from a source to its references or related links to acquire additional information.

When asked to explain how a search engine worked and how it used search terms, most had a vague sense of keyword matching and the mechanics of spiders searching and collecting sites. Several stated that one had to be very specific in searching (and when using simple search phrases would indicate that they weren't being specific enough), but few could articulate what they meant by specific; only one indicated that a search could be narrowed by adding additional keywords. When asked how to narrow a search, one participant stated that she thought there was a way and suggested putting a phrase in quotes. Another stated, "Google doesn't know exactly what I'm looking for, but I think based on what I search for, they seem to give an idea of what the public is looking for in that topic range."

Few students seemed to have a clear mental model of the work of a search engine. When asked to give an analogy for a search engine, 13 students (61%) compared it to a comparable print tool (library - 3; phone book - 2; card catalog - 1; filing cabinet - 1; indexable of contents - 1; book/encyclopedia - 3; librarian - 2), but all of them described them more as storage spaces and did not indicate how such tools would pull out information for the user. Only one of the students articulated any classification system or organizational structure that is used to retrieve sites.

## Discussion

Although small in sample size, this study demonstrates some prevailing thoughts about students' search skills and their performance with both search engines and library databases. While students stated that search engines retrieve sites by matching keywords, most did not demonstrate a strong conceptual model of search such that they could effectively narrow or focus a search to retrieve relevant materials. Furthermore, they were often unable to recognize a problem (incorrect Boolean logic, spelling errors, etc) and resolve it for results. Although they considered themselves successful, their skills were rudimentary at best.

The observations uncovered several other areas for further research. One such area is the way in which students read and process information. Nine students (43%) scrolled through material so quickly that they were not reading it for comprehension but were simply scanning text for keywords. Much like the tools they were using, they seemed to assume relevancy by the frequency of their keywords. Conversely, only seven (33%) read sites or articles in depth. Also, 13 students (62%) either highlighted text or followed with the mouse as they read, and six students (29%) actually read material out loud.

It was interesting to see how students multitasked in these observations. Seven students opened new tabs/windows as they selected sites, and 12 students frequently toggled between various sites and tools such as Word, Notepad or email applications to copy and paste material and even compare material.

Lastly, more study is warranted in exploring students' methods of evaluating sources and their opinions about certain sites, such as Wikipedia, and search engines. Two thirds of the students made some comment about the quality of sites: half of those rated sites favorably based solely on design and layout; others accepted .edu sites as valuable even if they had no relevance to the topic. Nine students talked specifically about Wikipedia: five were adamant against it, decriing its accuracy; four regularly used its entries or its references, and one stated she never used Wikipedia as she viewed two of its entries.

This research only begins to discover how students perceive and understand search tools and how they use them in research. As developers and librarians design new search interfaces and retrieval systems for this and future generations of students, continued research in the search habits of students is vital.

## References

- Kantner, L., Sova, D. H., & Rosenbaum, S. (2003). Alternative methods for field usability research. *In SIGDOC '03: Proceedings of the 21st annual international conference on Documentation*. Retrieved February 28, 2009 from [ACM Digital Library](#).
- Online Computer Library Center (2002). *OCLC white paper on the information habits of college students: How academic librarians can influence student's web-based information choices*. Retrieved February 28, 2007 from: <http://www.oclc.org/community/informationhabits.pdf>
- Raven, M. E., & Flanders, A. (1996) Using contextual inquiry to learn about your audience. *SIGDOC Journal of Computer Documentation*, 20(1), 1-13. Retrieved February 28, 2009 from [ACM Digital Library](#).
- Sullivan, D. (2001). WebTop search range study. *The Search Engine Report*. Retrieved March 20, 2007 from Sullivan, D. (2001). WebTop search range study. *The Search Engine Report*. Retrieved March 20, 2007 from: <http://searchenginewatch.com/showPage.html?page=2163451>

## Contact information

Lucy Holman  
Langsdale Library, University of Baltimore  
1420 Maryland Ave., Baltimore  
Baltimore, MD 21201

T: 410.837.4333  
F: 410.837.4319  
E: [holman@uball.edu](mailto:holman@uball.edu)