Millennial Students' Mental Models of Information Retrieval

by
Lucy Holman

October 2009

Presented to the
School of Information Arts and Technologies
University of Baltimore

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Communications Design

Approved by: _____
Advisor: Dr. Nancy Kaplan

_____
Committee member: Dr. Kathryn Summers

_____
Committee member: Dr. Evelyn Daniel

Abstract

This qualitative study examines first-year college students' online search habits in order to identify patterns in millennials' mental models of information retrieval. The study employed a combination of modified contextual inquiry and concept mapping methodologies to elicit students' mental models. The researcher confirmed previously observed millennial behavior including preference for searching rather than browsing and hypertext reading. The study found three basic types of mental models; those with a network view conducted more searches overall and more complex (Boolean and topic + focus) searches. However, none of the participants demonstrated strong mental models that increased their effectiveness in searching.

Acknowledgements

I would like to thank my advisors Dr. Nancy Kaplan and Dr. Kathryn Summers for their encouragement and support throughout my doctoral program. Thanks also go to Dr. Evelyn Daniel for her guidance on my dissertation; she challenged me to excel.

I would also like to thank those who assisted with this project: Dr. Daniel Gerlowski, Dr. Daniel Page, and Ms. Catherine Johnson who allowed me to recruit first-year students from their courses; I could not have completed the study without their cooperation.

Lucy Holman
November 2009

Table of Contents

List of Figures and Tables

Figure

Table

Introduction

There is ongoing debate in academic libraries and in the information retrieval community about the ways in which college students, particularly the millennial generation, approach information gathering and which tools they use to find materials suitable for research in their college courses. Since the development of electronic databases and the onset of computerized searching, librarians and information professionals have encouraged students to use well-indexed collections with complex search mechanisms that employ controlled vocabulary and require highly stylized queries. However, as the amount of material available on the World Wide Web and the popularity of public search engines grow, younger students gain experiences with simpler interfaces and the more natural language of public search engines. Because students can retrieve long lists of results on a topic with the simplest of search syntax or natural language queries, they are likely to feel a certain level of success, even if the results are not totally relevant to their initial search or are not scholarly in nature. Students, then, as they grow accustomed to these search engines, often eschew online catalogs and subscription databases with their complicated Boolean logic and controlled vocabulary in favor of public search engines such as Google. They believe that they can find everything they need on the public Web and see little reason to learn to use more difficult article databases.

Information literacy, which is defined as "the set of skills needed to find, retrieve, analyze, and use information" (ACRL 2007), has become a required student learning outcome at many colleges and universities across the country. For years librarians have offered undergraduate students instruction with research tools in order

to assist them in their scholarly research. Beginning with print indexes, continuing with CD-ROM-based indexes, and moving more recently to Internet databases, academic libraries have developed instruction programs to teach students search strategies and evaluation skills to improve the quality and efficacy of their use of such tools for their academic research. As more and more information has become available and retrievable from Internet-based tools, librarians have taught students appropriate searching techniques for online indexes and library catalogs, including subject headings and controlled vocabularies, which differ from more general searching in public Internet search engines. In many cases students question the need for such instruction, believing that they have the requisite skills to find and retrieve adequate information from these new tools. Yet when they use online indexes, their searches often yield few, if any, results, and those are generally of questionable relevance, based on the poor quality of their queries.

There is, however, a growing movement among database producers and libraries to develop technologies designed to harness the power of metadata and taxonomies of controlled vocabulary in order to harvest materials from a wide variety of collections while providing the end user a simpler and more seamless search interface. Some database designers see a need to become more "Google like," and to develop more powerful algorithms and employ more well-developed synonym rings and variant spelling detection so that even ineffective searches will provide better and more relevant results.

Database developers and librarians appear to be at a crossroads in information retrieval and information literacy instruction. Should students (and young faculty)

learn to use complex interfaces with less intuitive search strategies based on mental models of print-based research, or should designers develop interfaces that more closely relate to millennials' mental models of Internet-based information retrieval with engines that more accurately and effectively parse a simpler, more natural language query? If developers choose to do the latter, do they understand the new mental models of millennial students? How do today's (and tomorrow's) college students differ from the generations who preceded them? What are their mental models of search, and how can developers of commercial databases and online catalogs design interfaces with these mental models in mind?

Research does show significant differences between the current and subsequent generations of college students and their predecessors in the tools they use to locate information and the way in which they find, evaluate, and use that information. If users' prior experience and existing mental models of search systems have an impact on their efficacy with those systems, it is critical that librarians and designers conduct research and determine what the mental models of the millennial generation are in order to expand upon those existing frameworks. This research will ascertain the nature of millennial students' mental models of search engines, library catalogs and online article indexes and how they retrieve information in order to provide a roadmap for database developers and librarians alike to design interfaces, search mechanisms and information literacy instruction sessions that more closely align with current and future student skills and needs.

These findings can be used to compare identified mental models with the assumptions librarians make about undergraduates and the models used as examples

in information literacy sessions. In addition, these research findings may have

implications for developers of online information retrieval tools and their

documentation.

Literature Review

Millennials

Societal and cultural changes are seen in every generation. Some of these shifts are cyclical; one generation of students returns to the values of their grandparents or great grandparents. Other shifts, however, reflect changes in technology and innovation and are unique to the latest generation. In our lifetime we have experienced the age of the baby boomers, Gen Xers, and now Generation Y or the Net Generation. While each of these generations has particular traits and characteristics influenced by the times in which they live, the current group of traditional college-age students, in particular, have grown up in a period of great technological shifts and changes in information sources that have significant impacts on the way in which they communicate, learn, and approach academic research.

It is important to identify a number of general characteristics of the new generation of college students in order to better understand the way in which they approach research and learning. These students, often dubbed "millennials," came of age using a variety of digital devices. They grew up in homes with computers and Internet access; most own MP3 players, cell phones and other mobile handheld devices (Salaway, Caruso & Nelson, 2008; Rainie, 2006). Instant messaging (IM) has become a major means of communication. As of 2006, 75% of teens and 66% of adults 18-28 years old use instant messaging as compared to 52% of Gen Xers and 47% of all online adults (Fox & Madden, 2006). Today, the number of college students using IM has increased to almost 84% (Salaway, Caruso, & Nelson, 2008).

As of 2002, 86% of college students used the Internet; almost half of them began using it in high school or earlier (Jones, 2002). More recent studies show that 93% of teens are online, 65% own a cell phone, and more than half (51%) own an iPod or MP3 player (Macgill, 2007). Today, college students spend an average of almost 20 hours a week on the Internet; 7.4% are online 40 hours a week (Salaway, Caruso & Nelson, 2008). Today's students are constantly wired and continually connected.

This online activity is increasingly participatory; students now are creators as well as viewers. A recent study (Salaway, Caruso, & Nelson, 2008) noted that more than 40% of undergraduates report posting to a blog, wiki or video site such as YouTube. A large majority (85.2%) engage in social networking through such sites as Facebook and MySpace (Salaway, Caruso, & Nelson, 2008). They enjoy communicating online and value each other's work as much as they do that of "experts."

As a result of technological and social shifts in their lifetimes, millennials exhibit generational learning characteristics that may impact their research habits and the way they approach information retrieval. This population tends to be more social than previous generations, and students look to learn from their peers as well as their instructors; their learning tends to be more group-oriented (McGlynn, 2005; Rainie, 2006). They are much more comfortable with multitasking (Rainie, 2006; Sweeney 2006); they can simultaneously search for information, email or text a friend, write a paper, and listen to music. This multitasking has both positive and negative effects; it may lead users to new material through concurrent searches (Spink, 2004), but it may reduce performance in reading and concentration in reviewing results (Fox, Rosen &

Crawford, 2009; Levine, Waite, & Bowman, 2007) and may increase users' distractibility (Ophir, Nass, & Wagner (2009). They seem to need more guidance in their academics (Howe and Strauss, 2000); students appreciate clear and well-defined parameters and guidelines for assignments. This may lead to searches for the specific topic as assigned and less creative use of vocabulary. Additionally, students express frustration with faculty who fail to provide clear instruction on assignments, stating that the lack of information hampers their research (Head, 2007).

And, with their experience in "creating" content, whether it be videos on YouTube, their own mixes on their iPods, or personal tags on del.icio.us, students are more likely to develop their own categorizations, metadata and taxonomies rather than using Library of Congress subject headings or other controlled vocabulary used in today's online catalogs and commercial databases.

Millennial students are experiential learners; they value learning through discovery as much, if not more, than learning through the more traditional lectures or readings (Sweeney, 2006). College students prefer courses that are at least 50% interactive (Roberts, 2005). They learn by creating material, simulating an experiment in a lab or testing business theory in a new entrepreneurial venture. They learn by solving problems in their classes and following a project through all stages of development. This is generally done through social interaction with their peers as much as their professors; unlike some previous generations, millennials thrive on group work and collaboration (Windham, 2005). They enjoy reaching a conclusion from exploring and discovering information on their own and appreciate a learner-centered approach to material, where the learning is of personal relevance to the

student (Windham, 2005). And while many prefer to complete projects with the use of technology, they do so only if the technology is relevant and useful for the task at hand, not simply because it is the latest technology (McNeely, 2005). Even though they may be adept at using a particular technology, they may not fully understand how it works. They do not necessarily have conceptualizations of the technology then.

Closely related to millennials' preference for experiential learning is their extensive participation in online games. The Net generation has often been dubbed the "gamer" generation; as they have grown up with computers in the home they have also grown up with a generation of video and online games. Although many play games simply as pastimes, numerous educational games have appeared, and more elementary, secondary, and higher education institutions are incorporating games into their curriculum in order to engage these students. Seeing this generation as interactive learners, a number of libraries including the University of North Carolina at Greensboro, George Washington University, and the New Jersey Institute of Technology have developed game-based information literacy tutorials to teach students how to access, evaluate, and use information.

In their exploration and search for information, millennials often do not discriminate between information that is found online and that discovered in traditional print and broadcast media. In fact, they may not even consider more traditional print resources as they determine their information need and think of sources for answers to that need. They do not necessarily value peer-reviewed or professionally-edited material online and thus may not start a search with refereed

material (Abram & Luther, 2004). Nor do they limit themselves to text; they are likely to use audio- or video-based information streamed to them on their computer or portable device (Abram & Luther, 2004).

Students entering colleges today, who have literally grown up with computers, also demonstrate different cognitive skills than their predecessors. They seem to be "hypertext" learners; their brains are trained to move around and retrieve bits of information from a variety of places rather than from a single narrative (Prensky, 2001). They can move easily and quickly from one source to another and may lose interest with information that does not engage them immediately (Oblinger & Oblinger, 2005). The University College of London (UCL) CIBER study confirms this characteristic; it found that younger students scan sites quickly and click other links often. Boys, in particular, rarely read pages in a linear fashion (UCL CIBER, 2008). Some studies (Bilal & Kirby, 2002), in comparing middle school and graduate students' Internet search habits, noted that children search as much as they browse through information. Furthermore, the younger students tend to backtrack and jump around in searching more than older students; they are less systematic or linear in their approach to the results – only one third of the middle school students' moves were in the right direction as compared with 53% of the graduate students' moves. Children also seem to be stymied in their searches more than adults (Bilal & Kirby, 2002). Williams (1999), in his study of ten-year olds who are now 18-19, found that many do not fully read sites or online documents; rather, they scan materials for facts or small chunks of information within the text and do not spend time reading the entire document. This is consistent with what the UCL CIBER (2008) project

describes as "horizontal information seeking;" this study found that students skim material, as many as 60% viewing just a few pages of an e-journal article and then "bouncing" to another.

Additionally, millennial students are more comfortable learning from visual cues than their predecessors. They are accustomed to digital material loaded with images and tend to ignore large blocks of text, whether they are text-based documents or verbose instructions. They may even evaluate the material based on its visual relevancy rather than the textual content (Williams, 1999). If databases or search tools are not intuitive, lack strong visual cues, and require text-rich guides or manuals, millennials are less likely to use them (Oblinger & Oblinger, 2005).

When faced with an information need, students of the Net Generation seem to build searches as they go; they often let the information they find lead them to more. Millennials learn through inductive reasoning and discovery rather than through deductive processes (Oblinger & Oblinger, 2005). This is partly due to millennials' preference for active learning – once again, they tend to be experiential and immediate (Brown, 2005). Another factor is their focus on achievement; they are less interested in how something works or what something means than how can they use it or what can they build with it (Brown, 2005). This constructivist approach to learning may then lead to a similarly constructivist method of searching for and gathering information..

Several researchers have specifically applied some of these constructivist models to Generation Y students to look at generational characteristics. Holliday and Li (2004) found that millennial students do follow some stages of Carol Kuhlthau's

constructivist Information Search Process (ISP); however, the researchers discovered

several unexpected differences from Kuhlthau's studies. Many students seem more

uncomfortable with the uncertainly in the topic identification phase and quickly make

topic choices (Holliday & Li, 2004). This reflects the characterization of millennials

needing structure. Some make decisions on their topic simply based on the ease with

which they can find information on it; they seem to have little tolerance with topics

for which information is not immediately and readily available (Holliday & Li, 2004).

This is consistent with the characterization that millennials, who have grown up with

information at their fingertips, expect immediate feedback, and are impatient with

delay (Oblinger & Oblinger, 2005). Students do not seem to conscientiously and

constructively form a topic; rather the first items they find drive the process.

Moreover, they seem to spend less time on topic formulation/focus than Kuhlthau's

model suggests (Holliday & Li, 2004). Furthermore, Holliday and Li (2004) note that

in their analysis, students in Kuhlthau's fifth stage, information collection, also seem

to stop at a point when they find the requisite number of sources, not when they feel

that they have completed a comprehensive search or when they begin to see

redundancy in their sources. This research seems to point to less of a constructivist

approach than the idea of information foraging or simply "bumping into" information

and using what they immediately find, regardless of relevance (Holliday & Li, 2004).

Millennials do not seem to delve deeply into their subject matter or become more

focused as the process continues. This lack of drive in searching can certainly impact

the use of proprietary databases. Students who never receive formal library

instruction or tune out such instruction if it is not engaging may find online library

resources too cumbersome to use and prefer a simpler interface such as Google or Google Scholar with more immediate results (Lippincott, 2005).

Millennial students have been dubbed the "Google Generation" (UCL CIBER, 2008); they have grown up with search engines, and although they may not retrieve the "best" materials using public Internet search engines, they deem them "good enough" resources to satisfy their needs. They value ease and convenience over quality. Thus, they may not feel they need more specialized library resources; in fact, they may shy away from those tools that require more expertise or skill. They rarely use more advanced features in the search engines, expecting the engine itself to know what they need or want and to employ those features on its own (UCL CIBER, 2008). A 2000 study conducted by Roper Starch Worldwide found that 47% of those questioned are looking for academic research on the Internet (Sullivan, 2001). And although the survey did not explain how they define time spent on search, 60% said they spend more than one hour searching for information (Sullivan, 2001). Unfortunately, libraries have not yet made widespread efforts to integrate their services and resources with more public search tools, nor have they developed a presence in campuses' course management systems for simple, seamless use of materials related to courses (Lippincott, 2005). Several researchers (OCLC, 2002; Lippincott, 2005; Thomas & McDonald, 2005) suggest that libraries have not adapted to these student needs and continue to organize sources by format (book, journal, audio-visual) or by library collection rather than a more student-centric organization by subject or discipline. Consequently, traditional library resources require students to search in a variety of places or tools and use different search syntax in order to find

all that the library offers on a topic. Students then may prefer a one-stop shop such as Google Scholar that does not force students into separate searches for related materials (Lippincott, 2005). Libraries may lose this generation without some sort of discovery tool that will harvest data of varying formats from a variety of locations.

Recently, research from the Pew Internet and American Life project (Jones, 2002) found that 73% of college students use the Internet more than college libraries for *academic* research, while only 9% use the library more than the Internet. Academic libraries across the country are noting declines in the number of questions that students ask librarians or the frequency with which they ask for assistance in finding information (Association of Research Libraries as cited in Jones, 2002). The Pew researchers observed students using the library for purposes other than specific library research; they check email or use the Internet but are not using library websites or subscription databases. A similar Pew study on teenagers found that the trend continues with younger students, where 94% state they have used the Internet for their school assignments, and 71% consider the Internet as the primary source of research material. (Lenhart, Rainie & Lewis, 2001). A study commissioned in 2001-2002 by the Online Computer Library Center (OCLC), a national library consortium, found that 96% of those surveyed use search engines to find information for at least some of their assignments; 37% use for most assignments and 42% say they use search engines for all of their assignments (OCLC, 2002). More recently, a similar OCLC study (DeRosa, C., Cantrell, J. Hawk, J., & Wilson, A., 2006) found that 89% of college student start their research with a search engine as compared to 2% starting with the library's website. Responses to a similar question on the survey indicated

that for future information needs, 90% of college students would consider using a search engine compared to 66% who would consider a physical library; while 72% would go first to a search engine, only 14% would first try a library (DeRosa, C., Cantrell, J. Hawk, J., & Wilson, A., 2006).

These statistics are obviously of concern to academic libraries, and many of the rationales indicate that students find search engines easier to use and effective (enough). For example, the 2006 OCLC study found that 96% of the college students surveyed find Google a valuable source for information, while 84% found the library website worthwhile. These studies tend to rate search engines higher than libraries in terms of ease, convenience, speed and reliability; library resources rank higher in accuracy and credibility (DeRosa, C., Cantrell, J. Hawk, J., & Wilson, A., 2006). It would appear, however that students value the first four qualities more than the latter two.

Students seem to be satisfied with the quality and quantity of information retrieved from search engines. In the 2006 OCLC study, 91% of students were satisfied with the amount of material they found through Internet searching. Most believe (between 70-75%) they are successful searchers with Internet search engines (DeRosa, C., Cantrell, J. Hawk, J., & Wilson, A., 2006). Although this and smaller studies (Head, 2007) have found that students do use library resources and find them useful, they rarely feel they need to go beyond the Web unless required to do so.

A key point here is that millennial students often lack the sophistication to understand exactly what information they need and frequently have difficulty in developing effective search strategies and then later judging the relevance of their

search results. Furthermore, the way in which they rapidly scan content on sites and online documents does not allow for a deliberate evaluation of material (UCL CIBER, 2008).

This preponderance of Internet searching is not without its complaints, however. Although 78% said that they find what they want on the Internet, 29% note that they find Internet searching very frustrating, and 77 % experience at least some degree of frustration while searching (Sullivan, 2001). These complaints arise from too much information, particularly seemingly completely irrelevant information, to finding little or no information on a narrow topic (Head & Eisenberg, 2009). Furthermore, 86% believe that searching can be more efficient (Sullivan, 2001). Students do question the accuracy of Internet sources (OCLC, 2002; Lenhart, Rainie, & Lewis, 2001); in one study they rated the accuracy of Internet resources as 6.2 on a 10-point scale (OCLC, 2002).

A number of studies also examine millennials' use of library resources. The OCLC study (2002) did find that 70% of students use the library's website for at least some of their research, and 20% use it for most of their research. Almost half (43%) think that other sites have more valuable information than the library. As mentioned previously, the more recent OCLC study indicated that students rate search engines more highly than library resources in terms of ease, convenience and speed. When asked about the library resources they do use, two-thirds (67%) stated that they use full-text journals, followed by the library catalog (57%), subscription databases (51%), periodicals (44%), and reference tools (32%) (DeRosa, C., Cantrell, J. Hawk, J., & Wilson, A., 2006).

There are differing observations on exactly how much students are using the library. In 2006 a large study (DeRosa, C., Cantrell, J. Hawk, J., & Wilson, A., 2006) revealed that 57% of college students were visiting their library's website (and 87% were visiting the physical library). More recent studies (Salaway, Caruso, & Nelson, 2008) actually saw higher student use of library resources; 93.4% of students reported using their library's website with a median frequency of once a week.

Most students (55%) cite difficulty finding the full text of articles and suggest that libraries make their resources easier to access and use. Students say they have difficulty navigating and searching library products (OCLC, 2002). They are frustrated when they find a citation to a relevant source and realize that they can't access the full text of the article (Head & Eisenberg, 2009). And they are lost when they need to use the physical library to find primary sources in archival material or even finding a book on the shelf (Windham, 2005).

Other research cites similar concerns with library databases and millennials. Stein, Bright, Hurlbert, Linke & St. Clair (2006) found that students prefer searching the Internet to library resources because they often are confused about which databases to use for their topics, and they are unaware of or unsuccessful at cross searching in the databases. Students appreciate the simplicity of a simple, single interface to search a variety of sources (Stein, Bright, Hurlbert, Linke & St. Clair, 2006) and find that search engines are good sources for general background information on a topic. Some skills (or perception of skills) increase throughout college, however; almost 55% of seniors believed they are very skilled or expert in

using library resources compared to 37% of first-year students (Salaway, Caruso, &

Nelson, 2008).

Having grown up with the Internet and the immediacy of information, this

generation wants information at the point of need (Thomas & McDonald, 2005) and

does not want to wait for books or articles through interlibrary loan or other library

services. Millennial researchers suggest that libraries embed their own resources into

search engines such as Google Scholar (Thomas & McDonald, Lippincott, 2005) or

integrate these search tools into their own systems.

With their preference for experiential and group learning, their reliance on the

Internet and search engines for information, and their desire for multimedia content

and presentation of results, millennial students require a new approach to information

retrieval. They want materials that are easy to find and access, and they prefer simple

search interfaces with systems that draw from a number of sources. Libraries and

developers of online catalogs and library databases need to consider the

characteristics of this generation of secondary and higher education students as well

as the characteristics of generations to come as they design the interfaces and search

functionality of their own retrieval systems.

Differences between Search Engines and Article Databases

While all share similarities in their design, online catalogs, databases, and

Internet search engines differ in a number of ways. Some of the differences grow out

of their initial purpose and the audiences they originally intended to reach. In the

early days of electronic information retrieval, developers created integrated library

systems (ILS) or electronic catalogs and abstracting and indexing (A&I) databases for professional librarians and researchers who were familiar with standard cataloging practices and indexing rules and who often paid for searches by the number of results retrieved or the time the system spent processing the search and retrieving results. These professional searchers were willing and able to use complex command languages and strictly defined controlled vocabulary to retrieve a highly precise and comprehensive set of results. Public search engines, on the other hand, developed from the need of consumers and non-professional end users to search for documents within the vast universe of web sites that exploded as the World Wide Web grew. These documents, placed on the Web by individuals with varying levels of experience, knowledge and background in the subjects, had no uniform system of metadata to retrieve or categorize them. Online catalogs and A& I databases contain records of articles and books that were published in a large but manageable universe of the commercial publishing world. These materials were created in somewhat similar print formats (initially, at least), were inventoried in some fashion by national libraries and journal indexes, and thus were able to be categorized in clear and standard ways. Pages and documents on the web, on the other hand, were produced and posted in a variety of formats with few coordinated standards or guidelines. Database and catalog producers generally control the metadata and indexed content at the point of input into the system; conversely, web site creators and authors typically generate their own sites' metadata, which search engines simply collect and categorize as they retrieve and store them. These very different universes of content and systems of production gave way to major differences in the way metadata was

developed for websites as opposed to that for article or book records and thus differences in the way individuals search and retrieve that content.

In the early days of the World Wide Web, subject directories employed human indexers to sort and classify pages into a variety of categories (Green, 2000). Automated search engines grew in popularity and soon overshadowed directories as the information retrieval systems of choice for most users. These search engines employ spiders, which are computer programs that troll the universe of Internet sites/pages for certain keyword matches (Green, 2000). Once a spider has matched a site to its keyword, the site is saved into its database, into which the user queries.

Search engines typically match users' queries against their database of sites and site content (Rosenfeld & Morville, 2002). Generally, search engines' algorithms match query search strings against their index of site key terms, retrieved from both the metadata (tags, contributors, page descriptions, etc), and the content of the online document itself. The metadata, that is, the data used to describe the organization or content of the page, may conform to a standardized format such as the Dublin Core, or use a subject-specific ontology (Baeza-Yates & Ribeiro-Neto, 1999). These standards, however, are generally more common in article databases and online catalogs than in websites. More often websites will employ descriptive information coded in extensible markup language (XML) designed specifically for the pages in the site or site database. Typically, the algorithms applied in search engines search for exact matches to the terms, and search engines compensate for common user errors such as misspellings and problematic phrasing. They routinely filter out common "stop words" and allow for automatic stemming to include variant spellings (Baeza-

Yates & Ribeiro-Neto, 1999).   They also employ synonym rings to expand the possibilities of a given search term to include common synonyms (Rosenberg & Morville, 2002). Some algorithms also match documents on key criteria and will then offer the user "more like this" or "similar pages" on a particular retrieved item (Rosenfeld & Morville, 2002).

The early search engines classified and indexed pages found by spiders by analyzing the number of times a word or phrase was used in a page and the prominence of the word(s) in the title or first paragraph (Green, 2000). The next generation of search engines incorporated existing search engine indexes of sites and further analyzed matching pages with level of popularity among those matching pages (Green, 2000). The third generation of search engines, exemplified by Google, uses links-based analysis to focus not only on the content of a site but also the relationship between the page and the pages that link to and from the original page. The positioning of search results is based on the number and strength of connecting links (Green, 2000). Therefore, relevance may be based on the number of times the keywords appears in particular parts of a page, such as the title or the contents of the page, or may be measured in terms of a page's relationship with other sites; a site that is linked from a large number of sites will be retrieved as a more relevant result (Green, 2000). But here relevance is computer-based; it is not a function of manually indexed terms or fields that are categorized with specific criteria in mind.

Article databases, on the other hand, typically build records for each article or item and index words or characters within record fields such as title, author, abstract, or subject. When a user queries the database, the system matches the keywords in the

query against the same keywords in the designated field (e.g., author) or in multiple fields and retrieves results matching those words in those fields. This field-level searching allows users to include parameters or limits on their search. Without changing their subject or keywords, students may choose to limit by date of publication, type of document (scholarly journal, magazine, book, or government report), language, or audience. Several well-structured databases with robust indexes and extensive metadata, such as *PsycInfo*, allow the user to further focus their search to specific the study population by age or sex or to limit by methodology (clinical trial, longitudinal study, case study, etc).

Online databases often employ controlled vocabulary systems or authority files to map user search terms to equivalent terms or values in the database index (Rosenfeld & Morville, 2002). Many of these databases, notably *Medline* and *PsycInfo*, also provide users with a thesaurus or list of the system's controlled vocabulary with a hierarchical structure of broader, narrower, and related terms for potential subject browsing. These thesauri and controlled vocabularies can provide relevance feedback to users and allow for more focused searching; however, students quickly may become frustrated when challenged to find appropriate subject terminology. Often the thesauri or subject headings do not change rapidly to accommodate language changes in society; commonly used terms for phenomena may take years to appear as subject headings. For example, *PsycInfo* still uses the somewhat antiquated phrase of *drug abuse* rather than *substance abuse*. If a student conducts a subject search on the commonly used term substance abuse without first consulting the thesaurus to identify the appropriate subject for that concept, he/she

will retrieve no results (the interface does provide a message suggesting that the student should use its SmartText Search feature, but the student may or may not choose to do so).

As specialized subject databases proliferated on the web, librarians and professional searchers demanded that consistent standards be implemented to enable users to build platforms that could search across these databases. The Z39.50 protocol was developed to establish standardized fields and definitions so that Z39.50-compliant databases could be searched simultaneously through a third-party interface (Chu, 2003; Baeza-Yates & Ribeiro-Neto, 1999). Now libraries incorporate the concept of federated searching, enabled by the Z39.50 protocol, and employ link resolvers which rely on the OpenURL standard. OpenURL, like Z39.50, establishes standards for metadata, so that an online article can be defined by metadata such as the publication's unique International Standard Serial Number (ISSN), and the article's volume, issue, pages, etc. The article's static URL is substituted with an OpenURL that includes the article's metadata (Ferguson & Grogg, 2004). The authors here use include a number of OpenURLs as examples (Ferguson & Grogg, 2004). OpenURL, then allows libraries to link a citation indexed on one database to a full-text article in another database or a book in the library's online catalog.

Today, online databases are moving to use XML (eXtensible Markup Language) to provide descriptive information about the website, online document, or database item. XML allows web content creators to develop a hierarchy of tags to describe an item within a sub category within a category within a super category, etc.

A specialized search engine like Google Scholar attempts to bridge the divide between public search engines and the proprietary article databases and citation indexes such as the *Science, Social Science, and Arts and Humanities Citation Indexes* found in *Web of Science*. While *Web of Science* only indexes the most highly respected peer-reviewed journals, Google Scholar also indexes books, conference papers and proceedings, as well as articles in institutional repositories and open-access databases (Schroeder, 2007). Ten studies conducted since Google Scholar's launch in 2004 (Schroeder, 2007) found its indexing to be inferior to *Web of Science* but also found its search box simplicity and its lack of advanced functionality as more familiar and less intimidating to younger students who have greater experience with search engines than commercial databases . A number of universities, including the University of Baltimore, are linking book and journal holdings to Google Scholar so that the more popular search engine will point students to full text material in the libraries' own subscription databases.

Unlike search engines, which developed more rapidly and somewhat haphazardly, online catalogs grew from more than one hundred years of search tradition. The first online catalogs were simply automated versions of the physical card catalog; their databases maintained the record structure of the catalog card and used the same search fields of title, author, subject, etc. (Borgman, 1996). Subsequent generations of online catalogs have evolved in interface design and have borrowed search features from other information retrieval systems, but their underlying functionality has remained the same. Most conform to a standardized format, such as the Machine-Readable Cataloging Record (MARC), which orders and

organizes the fields according to national cataloging standards (Baeza-Yates & Ribeiro-Neto, 1999), based on many of the cataloging rules developed in the days of the printed catalog card and authority file. After more than ten years of research in online catalog searching, Borgman (1996) identified three sets of skills or knowledge that online catalogs require of their users: 1) a basic understanding of the process of information retrieval, 2) an ability to formulate an appropriate search strategy for a given system and how to use various search features, and 3) a knowledge of search syntax and an ability to enter a query correctly using that syntax. Unlike many search engines that allow for more exploratory seeking, online catalogs (and to a certain extent commercial databases) assume that users know something specific about the information to be found – a title, an author or a specific controlled vocabulary subject. The catalog record is very limited in the amount of text or information included; only occasionally does a record contain more than cursory notes on the content of the book. Even more general keyword searching still matches against these specific fields; a search using a recently adopted phrase may not match the prescribed subject headings, which take years to change. This assumption, however, does not reflect the way many users approach an information need and conduct searches without a clear or specific understanding of the topic. Borgman further states that catalogs assume "that users formulate a query that represents a fixed goal for the search and that each search session is independent" (Borgman, 1996, p. 496). As seen throughout this review of the literature and in other research concerning millennial students, young users are typically more constructivist in their information seeking, and results found in their initial queries often shape subsequent queries. Rarely does a user stop at a

single search. Unfortunately, sometimes the initial query in an online catalog retrieves few if any items useful for a subsequent search.

Another problematic difference between commercial databases and search engines is the search syntax that databases require users to master. Some, for example, use automatic phrase searching while others require users to place phrases in quotations. Other tools accept varying levels of nesting and Boolean operators. Each database vendor's interface requires varying symbols such as *, ?, or ! for truncation and wildcard searching and employ different rules for retrieving plurals or synonyms. They may or may not offer proximity searching. As mentioned earlier, search engines are more understanding of users' misspellings and natural language searches and create advanced synonym rings to retrieve close but not exact matches of user queries (Baeza-Yates & Ribeiro-Neto, 1999). Othman and Halim (2004) compared 25 products from twelve database providers and found a wide range of features and application of those features. For example, there are six different approaches to Boolean operators among the twelve vendors, and there are variations even with the same database for phrase searching. Database producers employ five different truncation symbols and four different methods for applying Boolean logic. Lastly, there are 19 distinct features offered by fewer than five of the twelve providers (Othman & Halim, 2004). These differences often confuse users who search several subject-specific databases across vendors and platforms. Users may lack the understanding of the requirements of a given system, such as use of Boolean operators and truncation and thus pose an inaccurate or less than effective search query. Students may not even realize they are searching different systems or may not

note the difference between a vendor's native interface and his/her library's federated search interface. They may continue to use syntax learned from their first database and experience poor results in another without understanding why or recognizing their error. Moreover, most online catalogs do not allow for more complicated or nested Boolean logic and thus may retrieve much different results set than would an article database for the same search, further confusing and frustrating users.

One study (Xie, 2004) compared databases, search engines, meta-search engines and subject directories based on user-defined precision and ease of use in searching for business and health-related topics. She found that Google scores highest in precision on both types of searches; in the case of business topics, a subject-specific database scores second in precision, while the subject-based database is the least highly rated for precision in the health-related searches (Xie, 2004). In terms of ease of use, some users appreciate the browsability of a directory such as Yahoo, while others like the simplicity of Google (Xie, 2004). Others prefer the search limiters in the database *Factiva* and its ease of use. On the negative side, students often are frustrated by the vast numbers of results retrieved in Google and the arcane query language used by some of the commercial databases such as *Dialog* (Xie, 2004). Overall, however, users seem to prefer the simpler interfaces of search engines like Google and the ability to use more natural language. They also like being offered additional subject terms to use for subsequent searches. The author concluded that IR systems need to develop easy-to-use interfaces that support both the novice and expert use (Xie, 2004).

The preference for simplicity of search is also evidenced in a number of studies that analyze search engine queries. Spink, Wolfram, Jansen & Saracevic (2001) compared their own study of more than one million searches to a number of related studies with data from a combined total of 153,695,523 queries and found that most search queries are quite short and simple. The number of keywords average between 2.3 and 2.4 in the three studies; only 18% in each study use three or more terms (Spink, Wolfram, Jansen, & Saracevic, 2001). Users rarely take advantage of advanced search features or complicated search syntax. Only from five to ten percent of queries studied included Boolean operators, and only five to seven percent used + or – signs to convey Boolean concepts. In one study, only five percent used quotation marks (Spink, Wolfram, Jansen, & Saracevic, 2001). Moreover, many of the queries using Boolean logic used it incorrectly. For example, of the seven percent of queries that incorporated a plus or minus sign, a majority (5%) used them incorrectly. In addition, these users also failed to take advantage of relevance feedback; only five to eleven percent of those observed clicked on "for similar results" links (Spink, Wolfram, Jansen, & Saracevic, 2001). Of course, these studies are several years old and were conducted when the World Wide Web was relatively young; users today may now use more of these features and use them correctly.

And interestingly, most web searching studies to date have limited their analysis to search logs (Jansen & Pooch, 2001). Unlike studies of databases and online catalog, few actually have observed users in search tasks and questioned them on their reasoning during or immediately following a search. While the literature on more traditional IR systems attempts to support theories of mental model

development or the information seeking process, studies of searches on the World Wide Web have simply focused on the metrics of the searches in terms of query length, duration of sessions, or other specific metrics (Spink, Wolfram, Jansen, & Saracevic, 2001; Jansen & Pooch, 2001). More study is sorely needed in the cognitive models students have of the web and search engines and distinctions users make between searching the three major types of information retrieval systems.

Still, users seem to be much more familiar with the simplified interfaces of search engines and lack an understanding of the more complex syntax required in article databases. While users can greatly benefit from the robust indexing and rich metadata in databases and retrieve both more precise and more comprehensive search results, they may also continue to face difficulties and become frustrated with poor search results. Here, librarians must reevaluate their training in the use of these systems in order to teach students the intricacies of the search syntax and designers must reconsider the complexity of the database interface.

Research in Information Retrieval

Electronic information retrieval system development and information seeking research have gone hand in hand for the last four decades. The first mechanized information retrieval systems arrived in the early 1960s (Bourne & Hahn, 2003), and research into the way users approach searching followed soon afterwards. In the 1970s information retrieval research focused on professional searchers, including librarians. It was not until personal computers were in widespread use that researchers turned their attention to skilled (and not-so-skilled) end users. This lag in

end-user information-seeking studies was in part due to the fact that trained librarians were the first to use online systems and conducted mediated searches for users the first decade or more of IR systems. Although there was discussion of end-user searching in the 1970s, many saw the information retrieval market as made up of exclusively librarians (Bourne & Hahn, 2003). It was not that developers thought that users were unable to learn to operate the terminal; they were concerned that users would not be able to build effective searches and soon would become frustrated. Developers saw librarians who were well trained in taxonomies and controlled vocabulary as the ones to develop the search strategies necessary for effective searching.

This focus on professional searchers changed with the advance of the Internet and distributed databases. As electronic databases emerged first on CD-ROMs and later on the Internet, IR research widened to include all categories and classifications of end users (Marchionini & Komlodi, 1996) and has exploded in recent decades. The relatively new field of Human Computer Interaction (HCI) has grown drastically in the last forty years, analyzing and evaluating operators' use of interfaces, software applications and information retrieval systems. As one branch of HCI, user searching and IR interface design research appears in library science, computer science, and cognitive science literature (Marchionini & Komlodi, 1996). This research constitutes a massive body of literature; a search on "information retrieval" as a subject (rather than simple keyword phrase) in five prominent library and information science databases (*Library, Information Science & Technology Abstracts; Library Literature; Inspec; ACM Digital Library and IEEE*) retrieves more than 78,000

results. These articles include research in three broad directions: user search behaviors, the IR system designs, and the interaction between the two.

Chowdhury (1999) in his review of information retrieval (IR) research identifies ten major areas of study: 1) analysis of retrieval, 2) search output, 3) search engines, 4) organization of information, 5) bibliographic control or metadata of the Internet, 6) intelligent agents in IR, 7) information reliability 8) user search behavior, 9) interface design, and 10) comparisons between traditional databases and the Internet. The literature in each of these areas is extensive; therefore, this study will focus on the final three. Comparisons between databases and the search engines are addressed elsewhere in this text; this section focuses on the literature of user search behavior and interface design.

The literature on user search behaviors focuses both on the information seeking process and the stages individuals move through as they identify an information need and formulate and modify search queries. Early research showed that when users searched a system their information needs became more focused and clear and that they learned to articulate those needs much more precisely (Gerrie, 1983). The initial query would start as fairly broad, and results users found to be relevant would help to narrow the scope of the query. Sometimes the query would express only a part of the information need and did not retrieve all that the user desired (Gerrie, 1983).

Many theorists have since proposed constructivist approaches of information retrieval and have studied end users to support these models. Three major approaches include Brenda Dervin's (1998, 1999) sense-making model, Nicholas Belkin's (1980)

concept of anomalous states of knowledge and Carol Kuhlthau's (1999, 2004)

information seeking process (ISP). Dervin (1998, 1999) describes her "sense-making"

approach to information studies as a journey of knowledge acquisition where the user

finds him/herself in a new situation, realizes that there are gaps in his/her knowledge,

seeks information to fill those gaps, assesses and evaluates that information, and

finally adds knowledge to face the next situation (Dervin, 1998). In her review of

more than 15 years of user studies she notes that users rarely can be easily

categorized; their personal attributes alone do not define the information need.

Rather, the need encompasses the situation at hand, the previous knowledge the users

bring to that new situation, the criteria that they use to make sense of the new

situation, and the information they find as a result. Individuals change through

knowledge and the context within which they live, and thus it is difficult to make

generalizations about the search process (Dervin, 1998). Still, studies do tend to

place user reactions to information and information system results in three broad

categories; they narrow the focus, they change the perspective or direction of

information seeking, or they find relevant information in the search (Dervin, 1998).

Belkin (1980) sees information retrieval as a communication process between

an individual and an electronic information storage and retrieval system. The

individual user begins with a recognition that his/her knowledge of an area is lacking;

Belkin (1980) terms this information gap an anomalous state of knowledge. This lack

of understanding or awareness leads to an initial query of the system which is, in turn,

modified as the system provides information that the individual uses to learn more

about the subject at hand. The communication between the user and the system

creates an iterative process in which the information need is clarified and the search is refined as the user gains knowledge. Initial searches tend to be less precise and less relevant until the seeker fully understand his/her information need.

Carol Kuhlthau (1999, 2004) offers a six-stage information search process (ISP) model that incorporates cognition, emotion and behavior. In this model, the first stage is that of task initiation, where a student recognizes a need for information. During stage two, topic selection, the student identifies and selects a general topic and may conduct a preliminary search or scan material for other topics. The student then embarks on stage three, prefocus exploration, where the student strives for a general understanding of the topic and seeks focus. In stage four, focus formulation, the student finds a focus among information he/she has found and then has a clearer direction to proceed with stage five, information collection. The process ends with stage six, search closure, where the student completes his/her search, either with a sense of accomplishment or disappointment (Kuhlthau, 2004). Kuhlthau, unlike other researchers, also addresses students' feelings of confusion, anxiety, elation and satisfaction and the impact of emotion on the search process. Like Belkin, her research highlights the fact that students in the search initiation and topic exploration phases often do not know exactly how to articulate their search needs and are unfamiliar with topic terminology (Kuhlthau, 1999, 2004); this makes searching more difficult, particularly in systems that rely heavily on controlled vocabulary and results in less relevant results sets. Searchers gain precision as they move through the six stages of development. A number of researchers (Vakkari, 2001; Holliday & Li,

2004) have used Kuhlthau's ISP model in studying the research behaviors of students and their interaction with information retrieval systems.

Vakkari (2001) specifically focuses on how students select and combine search terms in the different phases of the search process and how they evaluate the relevance of retrieved results. He finds that students use more synonyms and related terms for narrower topics and drop the number of broad concepts as they move through the process. They also increase the use of Boolean "ors" to expand their searches. Students focus on documents that include empirical research and methodology as they move forward and find the more theoretical background information less relevant to their research (Vakkari, 2001).

Other researchers have portrayed information seeking as less directed and more opportunistic. Bates (1989) postulates a "berry-picking" model, whereby users follow paths between clusters of relevant information, much like choosing heavily-berried parts of a bush. Erdelez (1999) describes the phenomenon of information encountering or incidental information acquisition whereby users happen upon information accidentally in the pursuit of other material. This may occur while browsing a library shelf, but users particularly experience this while searching on the Internet. Thus they appear to switch topics or to jump from one information need to another (Erdelez, 1999). And other studies focus on users' cognitive representations or mental models of information systems in their search behavior (Borgman, 1996; Pitts, 1994; Savage-Knepshield, 2001; Westbrook, 2006).

In his overview of information retrieval, Marchionini (1996) outlines six key factors in information seeking: the setting within which the information seeking takes

place, the information seeker, the task at hand or information need, the domain or

discipline in which the information resides, the information system and the search

outcomes.   The setting involves both the social and cognitive situation of the search

– the cultural or organizational milieu under which the search takes place - and the

physical conditions (time, location, accessibility, cost) under which the user operates

(Marchionini, 1996). The information seeker brings his/her own unique

characteristics, skills and abilities; his/her preexisting knowledge of and experience in

the domain; his/her mental models of the search system; and his/her understanding of

the task (Marchionini, 1996). Certainly differences among users in skills,

understanding of the subject matter and clarity of the information need greatly impact

the interaction with the system. The task or information need varies in specificity,

volume and time required (Marchionini, 1996). The user may need or expect a very

specific "answer" or may conduct a comprehensive search of a field expecting a large

quantity of information. Users may also have a preconceived idea of how long it

should take to successfully complete the task. The domain of information seeking

represents the discipline or knowledge base. Again, domains may vary in volume,

specificity, and types of representations (text, formulas or equations, visual images or

graphics, sounds) (Marchionini, 1996). Domains may require differing skills and

abilities on the part of the information seeker and different search systems and

interfaces.

The search system generally comprises the database or repository of

information and the interface that interacts with the user. The systems differ in terms

of their content –in terms of the subject matter, the type of data (citations, statistics,

text, images, etc), and the quality and quantity of information they contain (Marchionini, 1996). They also vary in the way in which they collect, store and organize the information. The interface of the system is the means of communication with the user. Variations in the way the interface presents information and provides feedback to the user influence the success of the search and the development of user mental models (Marchionini, 1996).

This last factor in information seeking, the system interface, is the focus of the second primary direction in information retrieval research. Much of this body of research focuses on the analysis of individuals' use of search features and query language and the quality of resulting searches in terms of recall (comprehensiveness of the search) and precision (percentage of relevant results). Jansen and Pooch (2001) reviewed the major Web search studies to analyze common elements. Three of the studies that focused on search engine searching (Hoelscher, 1998 as cited in Jansen & Pooch, 2001; Jansen, Spink & Saracevic, 2000 as cited in Jansen & Pooch, 2001; Silverstein, Henzinger, Marias & Moricz, 1999 as cited in Jansen & Pooch, 2001) analyzed transaction logs to determine typical query language and length, retrieval results. Unfortunately all three studies had flaws in the way they defined query terms and "user" – several studied logs from a single public machine that included multiple users, for example.

Today, much of the research in online information retrieval compares proprietary databases against search engines. Xie (2004) evaluated users' searches of article databases and search engines for precision and ease of use and found that although the databases generally have higher precision rates, users prefer the search

engines for their interface and general usability. Markey (2007a) reviewed 25 years

of research studies on information retrieval systems, including online library catalogs,

digital libraries, online article databases and web search engines. Again, this research

examines measures such as the average number of queries per user, the length of

query or number of search terms, and length of each search session. Markey (2007a)

also explored study data to report the use of Boolean or other search syntax

(truncation, phrase searching in quotes, etc). She found that few users employ

Boolean operators (less than 15% of queries include AND, less than 3% include OR,

and less than 2% include NOT), although more searchers do use the equivalent

symbols + and –. Less than 15% of queries include quotes around phrases, and less

than 5% employ truncation. Interestingly, when users try to employ advanced search

features they use them incorrectly approximately 37% of the time. Users rarely (less

than 10%) choose to take advantage of the relevance feedback such as "find more like

this" (Markey, 2007a). Markey (2007a) also discovered that users tend to experience

overload at more than 200 results and often do not choose to move to any of the

results links displayed.

Although the searches analyzed were not sophisticated or focused, Markey

found that most users are at least somewhat satisfied with their results. This finding

tends to negate claims from librarians and other researchers that users would greatly

benefit from training in various online IR systems (Markey, 2007a).

Efthimiadis (2000) studied interactive query expansion whereby users selected

related controlled vocabulary after their initial keyword searches and found that users

see roughly 30% of the terms presented as useful to expand their search. In almost

half of the searches (42%) users select terms with a single item that they have already retrieved, which frustrates them. However, the 66% of the terms the systems suggest do have relationships to other query terms and lead users to generally triple the number of relevant materials (Efthimiadis, 2000).

Markey (2007b), in the second part of her evaluation of 25 years of information retrieval research, suggests that IR system designers can benefit from studies that combine these two paths of research. Future studies should compare users' tactics between initial and subsequent searchers to see what actions they tend to repeat and how they adapt query language or search strategies over the course of their quest for information (Markey, 2007b). She also recommends using that combined research to design IR systems that aid users in their attempts to articulate query language or to refine/limit searches with more advance features. Lastly, she warns researchers against the temptation of using as subjects the eager population of graduate students in library and information science who have a better understanding than most of the more complex search strategies and devising research studies that ask participants to perform specific tasks (Markey, 2007b). Researchers should employ contextual inquiry or more exploratory research designs that observe what students actually do when searching for information rather than noting how they perform a more artificial request for information.

This leads to the third, and perhaps smallest, branch of IR literature, which looks at the common element between the user and the system – that of relevance. Saracevic (2007a, 2007b) in a two-part review of 30 years of relevance research by himself and others, offers a summary of what he considers the "invisible hand of

relevance" over information science studies and how relevance research bridges the

gap between information retrieval systems research and studies of users' information

seeking behaviors as both are concerned with the concepts of recall and precision (in

other words, relevance). The former focuses on algorithmic retrieval, and the latter

focuses on the users' cognitive and affective processes, and each either ignores or

criticizes the other. Relevance research, however, attempts to bring the two sides

together to determine ways to benefit each side of the equation (Saracevic, 2007a).

Although Saracevic admits there is no true relevance theory tested in the research, he

outlines a number of descriptive models including his own stratified relevance model

(Saracevic, 1997 as cited in Saracevic, 2007a; Saracevic & Kantor, 1997 as cited in

Saracevic, 2007a; and Cosijn & Ingwersen, 2000 as cited in Saracevic, 2007a). This

model enumerates the system's strata of content, processing and engineering meeting

the user's strata of context, situation, affective, and cognitive at the surface level of

the system interface. He considers that most of the more than sixty studies he

reviewed use one of two approaches: those that observe and analyze criteria users

employ to assess relevance (Saracevic, 2007b) and those that classify the type of

relevance (topical/subject, cognitive, affective or situational) assigned to retrieved

information (Saracevic, 2007a). Some of the studies (Smithson, 1994 as cited in

Saracevic, 2007b; Bruce, 1994 as cited in Saracevic, 2007b; Tang & Solomon, 1998,

2001 as cited in Saracevic, 2007b) observed changes in relevance as the users

progress through the information seeking process; others (Koenenmann & Belkin

1996 as cited in Saracevic, 2007b; Spink & Saracevic, 1997 as cited in Saracevic,

2007b; and Ruthven, Lalmas, & van Rijsbergen, 2003 as cited in Saracevic, 2007b)

analyzed the effectiveness of relevance feedback from the IR system and its impact on searcher behavior.

While the literature on information retrieval is abundant and rich, changes in user characteristics, IR systems, and the nature of online publishing warrant new studies on the relationship between user and system, user expectations and understanding of search and the importance of relevance in search results. As an entire new generation of information seekers arrives in colleges and universities, researchers should continue to hypothesize new models of information retrieval and explore the interactions between searcher and system.

Mental Models

Mental models are internal cognitive representations of systems and processes that assist in explaining and predicting one's interaction with such systems (Norman, 1983). In other words, a mental model is the mental image a user creates of a system and its functions to aid in his/her operation of that system. Such mental visualizations are generally incomplete and unstable (they can easily fade without use) with blurred boundaries between models of similar systems. They reflect users' belief systems, and as such, are subject to illogical thinking. They are typically simplistic and often will demand more physical effort in order to reduce cognitive efforts (Norman, 1983). Mental models are dynamic; they evolve as individuals gain experience and add to their understanding and conceptualization of the system or process (Borgman, 1986 as cited in Westbrook, 2006; Doyle, J. & Ford, D., 1998 as cited in Westbrook, 2006).

Mental models of search systems impact individuals' learning and ultimate success with research tools. These models vary in accuracy and depth depending on

the users' knowledge and ability to visualize the system/process. But, as the adage goes, perception is reality, and to the user, his/her mental model of the system or process *is* the system or process (Fischbein, E., Tirush, D., Stuvy, R. & Oster, A., 1990 as cited in Westbrook, 2006). Because novice and even intermediate users' mental models are incomplete and often inadequate for successful use of a system, designers and educators must understand the role that mental models play in system use and take users' prior experience in information retrieval into consideration when developing interfaces, instructional guides or training for new or complex systems. If users have no direct experience with the system at hand, designers will do well to develop interfaces to which users can adapt more easily from their existing mental models of other tools or search products.

Young (1983) focuses on three types of mental models, two of which are pertinent to information retrieval. Analogy models are those in which users employ a model of a familiar system to apply to a less familiar one; for example, certain properties of text-messaging on a mobile phone could be likened to a computer keyboard. These analogies, however, focus simply on user actions. Surrogate models provide a procedural or mechanical representation of how a system works. Mapping models emphasize the relationship between a user's actions and the actions of the system itself in accomplishing a given task (Young, 1983) and thus can enable the user to predict certain results or consequences.

Additionally, Borgman (1986) distinguishes between a conceptual model, provided to the user by a trainer or system designer, and the user's own mental model. The conceptual model is the way in which the designer wants the user to

visualize the system – again, the designer may present it as an analogy to a similar, more familiar system or a simple visualization. The mental model, on the other hand, is solely the user's representation – he or she incorporates the designers' conceptual model and his/her own observations and experiences into a working mental model.

Generally the literature on mental models of information retrieval have followed two related but distinctly different paths – 1) research on users' conceptualization of the search *process* and study of users' visualization of the information retrieval *system(s)*. Some leading researchers (Belkin, 1980; Kuhlthau, 2004; Cole, Lin, Leide, Large & Beheshti, 2007; Cole & Leide, 2003; Vakkari, 2001) have developed models and theories concerning the way in which researchers approach the research process and their conceptualization of how they satisfy an information need. Others (Borgman, 1986; Pitts, 1994; Savage-Knepshield, 2001; Brandt & Uden, 2003) focus more on users' visualizations of computerized search systems, whether they are library catalogs, electronic databases or Internet search engines. They compare students' representations of what happens between their input of search terms or phrases and their receipt of a list of results. As mentioned earlier, Young (1983) in his description of task action mapping, highlights a model that conceptualizes the connection between the user's action (e.g., query language), the system's response (parsing the query), and the overall task (information retrieval for a particular information need). Although research on mental models of the search process certainly sheds light on students' overall conception of research and how the conceptualization impacts their search behaviors, this research focuses more on the literature concerning mental models of search systems. The research question here

asks if students of the millennial general who have more experience with computers and computerized searching have different mental models or conceptualizations of such search systems than those of earlier generations or if their mental models are more fully formed than those of older students.

These mental models may be either conceptual or visual in nature. As previously mentioned, a model may simply be an analogy (Ramsey & Grimes, 1983; Young, 1983), providing the individual a means to apply knowledge of one system or process to another but without mental visualization of the full system and its function. For visual mental models, users create mental "pictures" of the system and their interactions with it (Rouse & Morris, 1986 as cited in Savage-Knepshield, 2001; Rasmussen, 1986 as cited in Savage-Knepshield, 2001). A strong mental model facilitates better understanding of a system and leads to greater success with its use; a user can more accurately predict the system's response to his/her actions and more effectively react to problems in the system (Savage-Knepshield, 2001). Users may operate systems with models that are incomplete or mistaken, but various studies (Kerr, 1990; Kaiser, Jonides, & Alexander, 1986 as cited in Savage-Knepshield, 2001) have shown that weak or incomplete mental models negatively impact user performance on information retrieval tasks. Students with inaccurate models of search engines, for example, often do not review results beyond the first page and either give up quickly or attempt a new search (Brandt & Uden, 2003). Without a clear understanding of the system, students cannot try alternative paths or troubleshoot a particular poor result. Explanations of a system or training may build or strengthen mental models (Borgman, 1986), but without follow-up use of the

system, the models will quickly weaken. However, users' experiences with a variety of search tools and resources certainly have an impact on their search performance and ability to improve their search strategies.

Mental models typically begin with the user's observation of a system and his/her common sense predictions for how the system works based on that observation (diSessa, 1983). Students create visualizations through direct experience with a given system or through explanation or instruction on that system. The individuals then, in turn, apply the model or mental construct to new experiences with different systems, and their model develops with each experience. They believe in the efficacy of the model whether or not that belief is warranted, and the model influences their use of the system. These models can be in part intuitive and implicit (Pitts, 1994); individuals may not consciously be aware of forming these models. Nevertheless, the models impact their use of information retrieval systems. Mental models are equally persistent (Pitts, 1994); students continually test their model against their current experience. They retain and use an existing model until further exposure and interaction with the system alter their understanding of it and modify their model of it.

When individuals have not had direct experience with a system they may often transfer their model of a similar system with which they have had experience to the one at hand (Stevens & Gentner, 1983). An example of this might be applying one's experience at an ATM with a new experience at a different electronic kiosk. However, when faced with differences between the two systems, users may not readily see deficiencies in their conceptualization as they transfer to the new system

and thus may build faulty models and may experience a level of failure and frustration. For example, millennials who have extensive experience with Internet search engines may transfer their understanding of those systems to their use of online databases and apply their same search strategies (i.e., natural language searches) with less than satisfactory results. Faced with a sense of great complexity in the databases and the perception of poor quality or quantity, millennial students may reject an unfamiliar database in favor of a simpler and more successful search engine.

Several researchers have examined the effect of mental models on individuals' use of more traditional electronic retrieval systems and search efficacy. Dimitroff (1992) found that university students with more robust mental models are more effective searchers in an online library catalog; those without some conceptualization of the system are not as successful at completing searches. Kerr (1990) also discovered that faster searchers have more complete and detailed visualizations of the system than slower searchers. Marchionini (1989) reported that high school students' conceptual models of print encyclopedias serve as useful frameworks and facilitate their searching an electronic version; users easily incorporate their prior models into the development of new models. And although systems have transferability and allow searchers to build models based on previous experiences with similar systems, sometimes working with a number of similar systems may actually confuse users rather than strengthening their models (Saxon, 1997). In her work with seventh graders, Saxon discovered that many students mistake a feature of one system in another, which negatively impacts their search efficacy.

Others have researched students' mental models of other aspects of their

search. Returning to the interaction between user and system, Vigil (1988, p. 165)

describes a "mental matrix" of combinations of keywords that users must develop and

manipulate as they evaluate search results. This matrix, while made up of words, is

seen as a spatial unit, much like a mental image of a chess board (or an image of a

Venn diagram in one's head). While this visualization is not specifically of the

system itself, it does evaluate how students approach the system and examines

searchers mental model of their relationship to the system and its processing.

Katzeff (1990) emphasizes the importance of system feedback in the

development of mental models. Users often become confused and make mistakes

because the system does not provide clues to complete the next step in the search

process. Borgman (1986) found that training that uses conceptual models is not fully

successful in establishing or strengthening students' mental models; the lessons one

learns in training, if not quickly reinforced, fade and are not incorporated into long-

term mental models.

Pitts (1994) found that because students have weak or no mental models for

use of library systems, they often fail to find library material. Rather than looking for

information in more appropriate library resources, they move to find material outside

of the library. Unfortunately, the librarians may overestimate the students' mental

models and do not teach or coach students in their use of the available systems. In

fact, librarians' own mental models of information retrieval systems inform their

explanations and instruction to students and the suggestions they make to database

designers. Westbrook (2006) then, studied the mental models of library school

students to better understand budding librarian models and found that generally the conceptualizations include four separate components: 1) the searcher, 2) the Internet, 3) the library, and 4) other people (librarian, professor, etc). Even library school students (most of whom Westbrook studied were in their 20s) see the library and the Internet as two completely separate entities and see online databases as more connected to the Internet than to the library that subscribes to them. The students perceive the information search process as an iterative one, generally starting with keywords that can be refined to locate information through search engines or online databases (Westbrook, 2006).

Differences in search efficacy, then, may be related to age-related differences in mental models of search. Slone (2002) suggests that individuals with mental models of information retrieval as an online catalog may have difficulty forming an accurate mental model of Internet searching and may not easily adapt from one system to the other. Conversely, younger students with no direct experience with a physical card catalog may have difficulty forming a mental model of an online catalog retrieval system. Some researchers (Brandt & Uden, 2003) argue that providing conceptual knowledge about a tool or system is not enough to form a new mental model. Librarians have traditionally worked with students to connect existing mental models (for example, use of a telephone directory to that of an online index) (Brandt, 1997). However, as young students may have fewer existing mental models of print sources or hierarchical taxonomies and structures, librarians may have difficulties in facilitating such connections and helping to develop new mental models.

Studies of younger students (Guinee, Eagleton, & Hall, 2003) have witnessed middle-schoolers using techniques such as adding .com to their search topic (e.g., www.JamesBond.com); although this is an unsophisticated approach, it often results in location of a relevant site. They also observed students going to familiar sites such as Amazon to find information. When using search engines most students searched with a single term or phrase rather than combining concepts.

Some research (Stronge, Rogers, & Fisk, 2006) indicates that efficacy in Internet searching is not simply a matter of age but is more related to experience and exposure to Internet search strategy and techniques. For example, older adults rely more on site-provided categories and controlled vocabulary than do young adults, and are less knowledgeable of web-specific search strategies (Stronge, Rogers & Fisk, 2006). On the other hand, studies such as Zhang (2008) do not see a significant link between experience with web resources and strength or sophistication of mental models.

A few studies have compared mental models of different user groups. Zhang & Chignell (2001) examined the impact of educational status and computer experience on mental models and found significant differences between the mental models of librarians and students, with graduate students more closely related to librarians than either undergraduates or high school students. Particularly, librarians view initial search queries as untargeted (possibly seeing a query as a way to find relevant subject headings or other controlled vocabulary to develop further, more targeted searches) where as students see them as targeted (Zhang & Chignell, 2001).

Early studies of mental models (Borgman, 1986) were concerned with the effect of system training on model development. These studies have had great implications in the area of library instruction. Now that a generation has grown up with Internet-based tools, it is time to reexamine students' mental models in order to determine if such training is necessary. The proprietary databases may still require specific training on their use for maximal effectiveness, but are students' search habits on the public Internet sufficient for a reasonable level of success? Kerr (1990) notes that a training session that offers a new conceptual model to an inexperienced user may actually confuse and detract from his/her current level of understanding.

Furthermore, there is a question as to whether library instruction expands on existing mental models or rather refers to examples outside young users' framework. Considering the disappearance of the card catalog, the decline of print indexes and the contrasting growth of Internet-based search in millennials' childhood, very few, if any, of today's college students have actually seen or used physical catalogs or print tools. Therefore, they have no visual or physical memory of them and may have limited visual models. And interestingly, even research conducted in the 1980s (Borgman, 1986) when library card catalogs were still actively used, showed that using analogies to card catalogs was not particularly effective in helping to develop appropriate mental models of online catalogs.

The study of mental models has implications for the design of online catalogs and proprietary databases. Studies (Savage-Knepshield, 2001) have found that users do not understand how databases or search engines process their searches. Without an accurate mental model that helps the user predict a system response to a user

action, an individual may experience difficulty using interfaces that demand more complex Boolean logic, and, as a result, may lose confidence or be easily frustrated. As studies have shown only a loose, if any, connection between databases and the libraries that subscribe to them (Westbrook, 2006), libraries may want to strengthen those connections by offering linear, step-by-step paths to assist students in tracking their keywords and the databases they've used and offering suggestions for alternative searches if their searches lack accuracy or effectiveness (Westbrook, 2006). Library websites or the databases or search engines themselves might assist searchers by walking them through a more effective search or offering assistance when a search retrieves too few or too many hits (Brandt & Uden, 2003).

This work on mental models, particularly age-related differences in mental models, has profound implications for database designers as they develop new interfaces for their article databases and online library catalogs and as they design systems to search across databases of different content and formats and harvest data from a variety of library holdings. It also may radically impact the nature of information literacy instruction, which has a long tradition in the explanation of controlled vocabularies, subject taxonomies and the use of traditional forms of information retrieval. Librarians who present analogies of print tools that are beyond the experiences of younger students may be lost on millennials and will have no lasting impact on the development of their mental models of search. More research on the mental models of today's students may very well transform the library and online publishing world.

Methodology

The methodologies used in studying individuals' mental models of information retrieval are quite varied; scholars disagree on how one can derive a user's mental model of a search system. Some approach mental models as analogies and elicit metaphors, drawings, or diagrams from users. Others take the stance that mental models grow from extensive use of a system and thus can be measured through transactional logs (Johnson & Crudge, 2007). However, few of these studies demonstrate a clear link between such logs and mental models. Researchers who have analyzed literature in the area of information retrieval (Markey, 2007b) have called for more studies employing contextual inquiry or other qualitative methodologies to more closely examine users actual search behaviors. Such studies perhaps can offer greater insight into students' understanding of search tool and better connect mental models or visual conceptualizations to students search strategies.

Recent research approaches the notion of students' conceptualizations of search and gathers data in a variety of ways. Several studies, both large (OCLC, 2002; DeRosa, C., Cantrell, J. Hawk, J., & Wilson, A., 2006) and small (He & Jacobsen, 1996; Tolppanen, 1999) have surveyed students on their information seeking habits. Others (Fister, 1992; Seiden, Szyborski & Norelli, 1997) have conducted individual and group interviews to gain insight into student research behavior. Still others (Valentine 1993; Head & Eisenberg, 2009) have employed focus or discussion groups or a combination of discussion groups and surveys (Head, 2007). All of these techniques rely on students' memories and perceptions of their behaviors; the research is gathered through the prism of the students' perceptions and

experiences. Watson (1998) has used a variation on the traditional interview, encouraging her eighth-grade subjects to develop stories around their information seeking and technology use.

Other researchers have taken a more qualitative, ethnographic approach to understanding students' information-seeking behaviors. Qualitative research strives to study people in their natural environment, in order to interpret behavior from the perspective of those observed (Denzin & Lincoln, 2005 as cited in Creswell, 2007). Those employing qualitative methodologies inductively analyze data to identify patterns of thought or behavior in their subjects. A number of studies have used variations of contextual inquiry as they observed college students in their research efforts. Law (2008) studied 30 undergraduate and graduate students as they researched an actual assignment. Makri, Blandford, Gow, Rimmer, Warwick and Buchanan (2007) observed graduate students in Library and Information Studies and Human-Computer Interaction as they researched their thesis topics. In both cases, the participants used a thinking-aloud protocol. This protocol requires that the user talk through his/her thought processes, explaining to the researcher what he/she is seeing and considering as he/she works and what is involved in his/her decision-making process. Kerr (1990) studied 99 education graduate students and Guinee, Eagleton & Hall (2003) observed middle- and high-school students as they completed a number of assigned search tasks. Zhang (2008) analyzed and characterized undergraduate student drawings and compared them to students' effectiveness in performing two given search tasks. Cole and Leide (2003) employed concept mapping to determine undergraduate mental models of information space. Each methodology has its

advantages; surveys offer a means to collect data from large samples of students, while contextual inquiry and concept mapping allow researchers to observe students in actual searches and gain a deeper understanding of their conceptualizations of the search system.

This study used a combination of contextual inquiry and concept mapping methodologies with first-year students at the University of Baltimore in order to examine millennial students' approach to research and their mental models of information retrieval. Such an understanding of millennials' mental models can guide librarians in preparing information literacy instruction aimed at younger generations and assist database developers in designing more effective interfaces and search algorithms.

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 to the task, and 3) the research is focused on a defined set of issues rather than on a list of questions (Holtzblatt, Wendell, & Wood, 2005; Raven & Flanders, 1996). This methodology is appropriate for study of a current product or process in completing routine tasks (Kantner, Sova, & Rosenbaum, 2003); it enables the researcher to observe manifest behavior (that what the subject actually does) rather than rely on subjects' reporting of their ideal behavior (that what he/she thinks he should do) (Blomberg, Giacomi, Mosher and

Swenton-Wall, 1993). It has been used in several recent studies of graduate and undergraduate students' research behaviors (Makri et al., 2007; Law, 2008).

These observations generally require that a student participant use a "thinking aloud" protocol, whereby he/she talks through not only his/her search behaviors but also the thoughts and feelings directing those actions. This in-the-moment, play-by-play commentary tends to be somewhat stream of consciousness and thus is much less filtered or measured than survey responses or answers to interview questions. Observers employing contextual inquiry do not have a predefined script or list of questions but rather follow up on actions or statements made in order to probe the participant's decision-making processes.

According to the online *Dictionary of Information and Library Management* (2006), a concept map is "a way of representing knowledge in the form of a diagram, with links indicating the relationships between concepts." Concept mapping offers students a visual means to demonstrate their understanding. Novak (1998) recognizes that concept mapping is a way in which learners construct meaning and integrate new concepts into their existing knowledge base and considers it a key component in his theory of meaningful learning. He concludes that concept mapping is an effective method of evaluating students' understanding of a concept or set of concepts. "Moreover, as students gained skill and experience in constructing concept maps, they began to report that they were learning how to learn. They were becoming better at meaningful learning and found they could reduce or eliminate the need for rote learning. Concept maps were helping to empower them as learners" (Novak, 1998, p. 27).   Novak and others have promoted concept mapping as both a pedagogical tool

and as an assessment measure. When they draw a concept map, students show both their correct and mistaken notions of the concept (Novak, 1998).

In order to identify students' mental models of search systems and their applications of those mental models in the actual search process, this study used contextual inquiry and concept mapping methodologies in tandem. The researcher observed 21 students and followed the observation with a series of questions eliciting students' conceptualizations of their searches, both in terms of verbal responses and drawing concept maps of the relationship between search terms and results.

Pilot Project

Prior to the observations in the fall, the researcher met with three University of Baltimore students – two male incoming freshmen and one female continuing student. She presented them with a required assignment from one the first-year students' required courses – IDIS 101, First Year Seminar. The IDIS 101 course is a typical first-year introduction to college course in which students develop study skills and are exposed to campus support services and co-curricular opportunities. The identified assignment was a career exploration paper that required Internet research. The researcher observed each of the three students explore possibilities in business administration, criminal justice and jurisprudence and then asked a series of follow-up questions. These pilot tests uncovered problems with the assignment; the students generally found all the information they needed to complete the assignment on a single site, and none of them expanded their research beyond what was necessary.

The research requirement was neither comprehensive nor complex, and it did not afford the researcher the opportunity to explore students' research behaviors in depth.

Upon seeing the lack of depth in the IDIS 101 assignment, the researcher reconsidered the project design. She met with faculty teaching in University of Baltimore's First and Second Year Program (FSP) and identified three courses within the curriculum that included more robust research assignments. The faculty of those courses agreed to invite her to one of their class sessions early in the fall semester to recruit students for the project.

The students' responses to the post-observation questionnaire in the pilot study and subsequent discussion about the wording of the questions pointed out flaws in the questions and confusing wording. The researcher modified the interview questions for the fall participants as a result of the pilot project.

Participants

The University of Baltimore's curriculum for first-year students revolves around a model of learning communities. These learning communities, required in the students' first two semesters, are comprised of three interrelated courses – a humanities course, a social science course, and a "skills" course – either information literacy or speech communications. The skills courses use the subject matter from the two discipline-specific courses in skills-based assignments; the three faculty tie discussions and assignments together. The underlying assumption in this model is that students can learn information-searching techniques and oral communication skills more effectively within the context of subject-specific information needs.

The researcher recruited students from three of the four learning communities with an information literacy skills course in the Fall 2008 semester. These three clusters were: *Citizenship in America* (including Economics 100 - Economics of Contemporary Issues**,** Philosophy 101 - Introduction to Philosophy, and Interdisciplinary Studies 110 - Introduction to Information Literacy); *The Art of Work and the Work of Art* (including Arts 101 - Music & Arts as Craft, Entrepreneurship 101 - Imagination, Creativity and Entrepreneurship, and Interdisciplinary Studies 110 - Introduction to Information Literacy); and *Understanding Modern America* (including History 112 - Modern America, Management 112 - Business in a Changing World, & Interdisciplinary Studies 110 - Introduction to Information Literacy).

The researcher visited one class in each of the three learning communities in early September, in which she explained the nature of the study and asked for volunteers. Approximately 26 students volunteered for the study, giving the researcher their telephone and email contact information.

The researcher contacted the 26 individuals and scheduled sessions with 23 of them. Several students missed appointments and were unable to be reached for rescheduling. Fortunately, two students outside the original recruitment group learned of the study through classmates and volunteered to participate. Between September 23, 2008 and October 15, 2008 the researcher observed and interviewed 21 students. All except three were enrolled in an information literacy course; two were in learning communities that included speech communications as the skills course, and one student used an assignment for a course outside her learning

communication and did not identify herself as being in a particular learning community. Interestingly, only one participant indicated that she had attended a library instruction section even though 18 were in information literacy courses taught by reference librarians.

Prior to each observation, either when she scheduled appointments with the participants or when she first met them, the researcher asked each student a series of eleven screener questions related to his/her Internet use, including time spent searching the Web, participation in social networking activities, and use of email and text-messaging). The researcher used the screener questionnaire to obtain basic demographic information and to ascertain each student's level of technology use. No participant was eliminated because of lack of general Internet use or familiarity with usability testing and database design (see Appendix A for screener questionnaire).

Although relatively small in number, the 21 students demographically represented the entering class at UB: 11 (52%) in the study were women and 10 (48%) were men, compared to the full cohort of 49% women and 51% men (University of Baltimore Office of Institutional Research, 2009). Ten students (48%) in the study were white, 9 (43%) were African American and 2 (9%) were Asian. This compares with a class ethnic makeup of 40% African-American, 29% white, 7% Asian, 3% Latino, 2% International students, and 19% unknown (University of Baltimore Office of Institutional Research, 2009). All participants were under 20 years old; six (29%) were 19 years old, ten (48%) were 18, three (14%) were 17, and one (5%) was 16 years old. One student (5%) did not share his age. The average age was slightly over 18, comparable to the cohort average age of 18.5 (University of

Baltimore Office of Institutional Research, 2009).   See Figures 1 – 4 for participants'
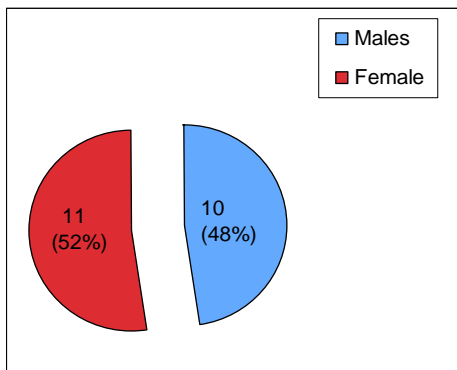demographic information.
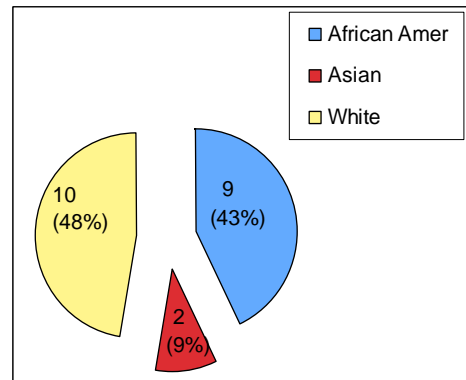

Figure 1. Participants by Sex
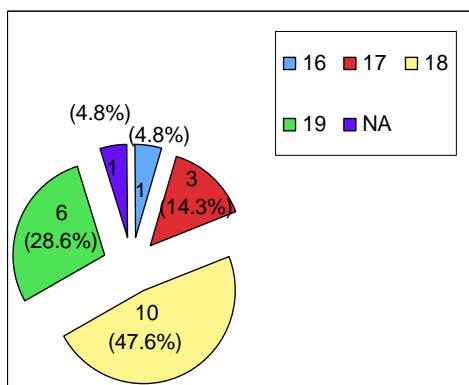

Figure 2. Participants by Race
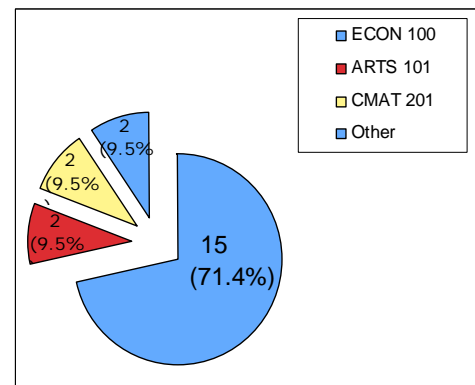

Figure 3. Participants by Age

Figure 4. Participants by Course

The screening questionnaire also addressed students' general computer and
Internet use. All participants had at least one computer at home; 57% had two or
more. Almost all of the students (86%) checked email daily; however, less than half
(43%) checked more than once a day. Almost ten percent checked their email at least
four times per week; only one student (5%) checked email only one to three times per
week. Almost 62% of the students sent text messages at least once a day; more than
half (52%) sent at least 20 messages a day. Only 5 students (24%) sent less than one

text message a week. Approximately 91% (19 students) had a profile on a social

networking site.

Almost all of the students (95%) used the Internet at least once a day; more

than three-fourths (76%) used it more than twice daily. Only one student (5%) did

not use the Internet daily, but he reported using it at least four times a week. The

questionnaire asked students to distinguish between searching the Web and Internet

activity such as social networking. Table 1 shows the students' daily Internet use.

Table 1

Participants' Daily Internet Use

n=21	At least 1 hour/day	2+ hours / day
Hours of Internet use	95%	62%
Hours of Web surfing	71%	33%
Hours of Social Networking	52%	24%

Each student had a current research assignment for which he/she chose to use

online resources. After an initial introduction to the study and a few preliminary

questions to describe their assignment and what they expected to find, the

observations began as the students researched their assignments. Although there were

four distinct assignments, 15 of the 21 (71%) researched two political parties' plans

for a social or economic issue.   These issues included energy and the environment,

the wars in Iraq and Afghanistan, the economy, poverty and taxes, and education.

Other assignments included exploratory papers on the lives and art of both a classic

and a contemporary artist, a mid-term exam that applied urban planning concepts to

local economic development projects, a bibliography on the three-strike law and its

effect on deterring crime, and a psychology career exploration paper (See Figure 4 for

the breakdown of participants' courses and Table 2 for courses and specific

assignments).

Table 2

Participants' Course and Assignment

Student	Course	Assignment
P1	ARTS 101: Music and Art as Craft	Leonardo daVinci
P2	ECON 100: Economics of Contemporary Issues	Republicans and Socialists: Global warming
P3	ARTS 101: Music and Art as Craft	6 local artists
P4	ECON 100: Economics of Contemporary Issues	Democrats and Libertarians: Iraq & Afghanistan
P5	ECON 100: Economics of Contemporary Issues	Democrats and Republicans: Poverty
P6	CMAT 201: Communicating Effectively	Forgotten or unknown cultures
P7	IDIS 110: Introduction to Information Literacy	Three-strikes law
P8	ECON 100: Economics of Contemporary Issues	Democrats and Party for Socialism & Liberation: Teaching virtues in education
P9	ECON 100: Economics of Contemporary Issues	Democrats and Party for Socialism & Liberation: Iraq and Afghanistan
P10	ECON 100: Economics of Contemporary Issues	Democrats and Libertarians: Banking regulation
P11	ECON 100: Economics of Contemporary Issues	Republicans and Democrats: Economic Downturn
P12	ECON 100: Economics of Contemporary Issues	Democrats and Party for Socialism & Liberation: Iraq & Afghanistan
P13	IDIS 101 – First Year Seminar	I-Search: Careers in Psychology
P14	ECON 100: Economics of Contemporary Issues	Democrats and Libertarians: Environment and energy
P15	ECON 100: Economics of Contemporary Issues	Democrats and Libertarians: Funding for education

P16	ECON 100: Economics of Contemporary Issues	Democrats and Party for Socialism & Liberation: Oil independence
P17	ENTR 101: Imagination, Creativity and Entrepreneurship	Midterm: Baltimore economic development projects
P18	ECON 100: Economics of Contemporary Issues	Democrats and Libertarians: Economic policy
P19	CMAT 200: Communicating Effectively	Pop culture figure: Pharrell Williams
P20	ECON 100: Economics of Contemporary Issues	Democrats and Libertarians: Poverty and taxation
P21	ECON 100: Economics of Contemporary Issues	Democrats and Libertarians: Environment (air pollution)

Observations

The researcher met with each student for approximately 60 to 80 minutes in the School of Information Arts and Technology usability lab at the University of Baltimore. After offering a brief explanation of the study, the researcher asked each participant to describe his/her assignment as he/she understood it and asked what generally was his/her first choice of search engine or article database. She then encouraged the student to start with that tool. The researcher modeled the think-aloud protocol and then asked each participant to verbally walk through his/her thought process. She observed students as they used their natural choices of Internet search engines, article databases or library catalogs for their given research assignment. In addition to personal observation and note taking, the researcher also videotaped each session and captured the search paths of each participant using Morae software. During the observation of the research task, the researcher occasionally engaged in a conversation with the participant, asking follow-up or probing questions concerning

the student's actions or decisions. She occasionally reminded more quiet students to use the think-aloud protocol.

After the researcher observed the student for approximately an hour, she concluded each session with a post-observation interview that included a series of approximately ten questions (See Appendix C). These questions included a review of the keywords the student used in the session, his/her reflections on the success of his/her searches and possible alternative search terms, and several questions designed to ascertain the student's conceptual understanding of search tools and how the tools used search terms. The researcher also asked each student to draw or diagram the relationship among the terms used and their relationship to items retrieved in the search. Several studies (Kerr, 1990; Thatcher & Greyling, 1998; Zhang, 2008) have used such drawings to identify users' mental models of the Internet.

Students did not receive directions or guidance as to how to draw their diagrams. This lack of structure led to various types of drawings or diagrams, several of which depicted the entire search process rather than the representation of the terms.

Data Coding and Analysis

The researcher reviewed each tape and transcribed the dialogue and actions taken (see Appendix E for selected transcripts). From those transcriptions, she identified actions and quotes to be included on the accompanying CD of video clips. With close review of the transcripts, she identified several areas for analysis – the number of searches and the search tools used, the types of searches conducted (single word or concept vs. phrase searches vs. multi-concept searches), additional search

techniques used (Boolean operators, truncation, quotations, etc), problematic search techniques (misspellings and incorrect usage of search techniques), searching within an individual site or page, and strategies used to refine (broaden or narrow) searches. Additionally, the researcher identified patterns in non-search behavior such as subject browsing, use of site navigation, and behaviors related to item management, collection of material, reading, organization, and evaluation and citation of sources. Using a modified affinity mapping procedure, whereby all participants' actions and comments were written on sticky notes and posted on a wall to be placed with related actions and statements, the researcher grouped all similar behavior and comments together and categorized them into the areas mentioned above.

Several researchers have used drawings to elicit mental models. Thatcher & Greyling (1998) had 51 university faculty, staff and students complete a questionnaire that required a drawing of their conceptualization of the Internet. In his study of 99 graduate students' mental models of information retrieval, Kerr (1990), after observing them use an online catalog, asked participants to either draw a diagram of the database they used or describe it in writing. Cole and his colleagues observed 80 undergraduates as they diagrammed keywords derived from their essay topics using circles and lines to connect and depict relationships among them (Cole, Lin, Leide, Large, & Beheshti, 2007). Similarly, Zhang (2008) reviewed undergraduates' drawings of their conceptualizations of the Internet to identify four categories of mental models, such as technical view, functional view, process view and connection view. In this research, students' concept maps were organized into one of categories similar to Zhang's for the purpose of comparison and analysis.

Results

The 21 students performed 210 discrete searches in 21 distinct search tools; 130 were in search engines, and 80 were in library databases (see Table 3). Sixteen students (76%) started with a web search engine, three (14%) began their research with online databases, and two (10%) started with online encyclopedias (Wikipedia and Encarta). An overwhelming majority of search engine queries (72%) were in Google. This almost mirrors a large international report (DeRosa, C., Cantrell, J. Hawk, J., & Wilson, A., 2006) where 68% of the 396 college students surveyed used Google as their search engine of choice, followed by Yahoo and MSN. A large majority of the database searches (almost 74%) were in Academic Search Premier. This seemed to be the database that most students recognized, possibly because it is the first alphabetically in the list of multidisciplinary databases.

Table 3

Searches Conducted in Search Engines and Databases

Search Engine	No. of Searches	% of Engine Searches	% of Total Searches	Database	No. of Searches	% of Database Searches	% of Total
Google	94	72.3	44.8	Acad. Search Premier	59	73.8	28.1
Alta Vista	7	5.4	3.3	JSTOR	6	7.5	2.8
Yahoo	5	3.8	2.4	Opposing Viewpoints	4	5.0	1.9
Ask	5	3.8	2.4	InfoTrac	4	5.0	1.9
Google Scholar	4	3.1	1.9	LexisNexis	3	3.8	1.4
Google Gov	3	2.3	1.4	Lib. Catalog	1	1.3	0.5
Google Images	3	2.3	1.4	Journal Finder	1	1.3	0.5
PolyCola	3	2.3	1.4	Project Muse	1	1.3	0.5
MSN	2	1.5	0.9	General Business File	1	1.3	0.5
Wikipedia	2	1.5	0.9				
Clusty	1	0.8	0.9				
Encarta	1	0.8	0.5				
Total	**130**	**10.00**	**61.9**	**Total**	**80**	**100.0**	**38.1**

In addition to queries in search engines and databases, eight students (5%) conducted another 14 searches in 11 individual sites, bringing the number of total searches to 224. These searches in individual sites included news sites such as CNN.com (2), the *New York Times*, *National Geographic* and *Education Week* and organizational and political sites such as GOP.gov (the Republican Congressional caucus), democrats.org (the Democratic Party site) (2), politico.com, and a Libertarian blog. One student also searched the University of Baltimore web site for psychology information.

Search Behavior

Using a categorization scheme similar to that of Guinee, Eagleton and Hall (2003), the researcher characterized searches into four types: 1) simple, single-term searches, which consisted of a personal name or a one- to three-word concept such as Republican party, public schools or three strikes law; 2) topic plus focus searches, which consisted of a string of words containing two related concepts such as Libertarian economic policy or republican taxes or a single Boolean AND search such as democrats and environment, 3) phrase searches, which consisted of a multiple-word descriptive phrase or sentence such as Libertarian policy on pollution or forgotten cultures in South Africa, and 4) advanced searches, which employed multiple Boolean operators such as political parties and education or schools and 2008 election or (Libertarian or democratic)+(economic or fiscal)+(policy or theory or belief).

Most students conducted very simple searches (34%), using personal names or two- or three-word phrases that conveyed a single concept or performed topic plus focus searches (30%) using a Boolean AND, such as Democrats and taxes. Students attempted 31 more complex Boolean searches (13.8%), combining two or more distinct concepts; however, for a majority of those searches (22), students crafted the queries with inappropriate Boolean operators, particularly in the databases. Students who employed both search engines and databases generally used the same keywords and syntax in both types of search tools. Table 4 indicates the types of searches students conducted.

Table 4

Type of Searches Performed

Search Type	Number	Percentage
Simple searches	77	34.4
Topic plus focus searches	68	30.4
Phrase searches	39	17.4
Boolean searches	31	13.8
Use of subject headings	9	4.0
Total (includes 210 searches in search engines and 14 searches within individual sites)	224	100.00

Approximately one sixth of the queries (17%) used phrases; again, phrase searching was similar in the search engine and databases. Students did not radically

change their syntax between the two; the major difference was in the built-in use of

Boolean operators provided in the databases. Three of the phrases were in the form

of a question: "Does the three strikes law deter crime?," What does the Democratic

party stand for?" and "Why did EBDI have a rehab delay?" The remaining 36 phrase

searches were strings of words such as "financial corruption leading to current

economic downturn," "history of the democratic party," or "democratic views on

Afghanistan/Iraq." A few of the phrase searches included the element of the source

or type of source such as "Edward Rutkowski in the Baltimore Sun" or "articles on

U.S. Democratic Party and air pollution."

In both phrase and topic plus focus searches students used a variety of

punctuation marks including quotation marks, slash marks, parentheses and plus

signs, as seen in "poverty and taxes," "democratic views on Afghanistan/Iraq,"

"(Libertarin[sp] or democratic) + (economic or fiscal) + (policy or theory or belief)"

and "energy independence + party for socialism and liberation." Frequently, the use

of such punctuation did not have the intended effect of focusing or clustering results;

incorrect punctuation, particularly in the databases, led to failed or less than

successful searches with limited or no results. Students generally assumed that few

results indicated a problem with their choice of keywords rather than a problem with

punctuation or syntax, so they would continue to experiment with new search terms

rather than review spelling or syntax.

Again, students did not seem to differentiate punctuation strategies between

the search engines and databases. Several participants, although they did not use such

punctuation in their own searches, talked about using punctuation in answer to the

researcher's probing questions or in the post-observation interviews. Several said that one could use parentheses or quotations around phrases. None seemed to recognize differences in phrase search capabilities between search engines and databases.

Six students used Boolean logic (other than a single AND) in 32 searches. Of those six, two used only multiple AND operators. Of the remaining four who used both AND and OR, only two used them correctly, and only one with successful results (one of them misspelled words resulting in unsuccessful searches). These mistakes, paired with frequent misspellings, led to few, if any, relevant results in the databases. As they did in the cases of incorrect punctuation, students failed to recognize that the lack of results was from mistakes in logic rather than their choice of terms. Each of these cases illustrate a statement made in the UCL CIBER (2008, p. 24) report that students "simply do not recognize that they have a problem; there is a big gap between their actual performance in information literacy test and their self-estimates of information skill and library anxiety."

Spelling errors had a significant impact on students' search success and highlighted a major distinction between search engines and database algorithms. Eleven students (52%) made one or more spelling errors. Students rarely saw those mistakes; only in two instances did the participants catch their own misspellings. Fortunately, in Google they were alerted with a "Did you mean?" set of results, and those students who made mistakes and were presented with a "Did you mean," clicked the revised search immediately, appearing to be very familiar with the alternative spelling link. When misspellings led to few or no results in the databases,

students assumed there were simply no results on that concept; they did not question the accuracy of the search string.

Students seemed somewhat haphazard in their strategy to focus their searches or to expand or narrow their 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 (10.7%) 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.

For the few students who correctly applied limits to their searches such as scholarly/peer-reviewed sources or those who effectively used truncation or wildcard symbols, many more students made mistakes in spelling, punctuation or logic, as shown in Table 5.

Table 5

Application of advanced techniques and common mistakes

Technique / Error	Number of Attempts	Percentage of total searches (n=224)
Broaden / Narrow	24	10.7
Limits	2	0.9
Truncation	5	2.2
Spelling errors	18	8.0
Boolean mistakes	21	9.3
Unnecessary words	6	2.7
Incorrect punctuation	4	1.8
Total	**80**	**35.6**

Four students went directly to sites they knew to find information on their topic rather that use a search tool. Once at a specific site, almost half (48%) searched within that site rather than browsing it; some searched in news sites for their topics, for those researching political parties, three searched for their issue within the party site. This preference for searching over browsing is consistent with other studies (Bilal & Kirby, 2002; UCL CIBER, 2008). Additionally, three students used "find within page" features in the browser or Adobe Acrobat Reader to find search terms within a document.

Six students used subject headings a total of nine times; five of the six only used subject headings a single time. The sixth participant used subject headings as a topic focus in a Google search on Leonardo DaVinci, using such headings as DaVinci inventions and DaVinci quotations. She seemed familiar with subject headings and indicated that she often relied on them to narrow her search; the other five seemed to happen upon them in the databases. They did not indicate that they understood the impact of subject headings on their searches, nor did they mention subject headings as a search strategy they had used in the past or typically used.

Three students attempted to perform a subject search in databases; however each of them structured the search in a way that subject searching was not appropriate. For example, one used the phrase, "3 strikes law deters crime" as the subject rather than just "three strikes law;" one included a Boolean string "banking and regulat* and United and states "and the third used "freedom or education and Obama." The students' use of phrases or Boolean operators within a concept sabotaged the effectiveness of subject searching. However, none of the three seemed

to recognize this nor did they consider retrying the subject search; they simply modified their keywords or changed the query altogether.

Several students relied heavily on one or two sites in their research, and some did browse more than search within the site. Twelve (57%) used sites' global navigation in at least one site 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.

Students could not easily articulate their understanding of why a search may have failed or why they retrieved the results they did. Some relevant comments included, "I don't know how I typed in global warming and got the entire website," and "I don't know exactly why they do it, but if you switch…..if you're looking for education and virtues and then you switch to virtues and education for some reason it brings up completely different websites." One student commented, "It gives me results because I only put one party. I don't really know [why], I think because the parties are…..have different things." Another noted, "I just took one word out and got a whole other page of information." Still another admitted that she didn't get results "because I was a little too vague," but she could not articulate what specifically was "too vague" about her search terms. One student who misspelled words and incorrectly applied Boolean logic stated that it didn't "look like they use it [Boolean] - I guess I'll just go one [term] at a time."

Others took actions to narrow or broaden the search without confidence that those actions would work. For example, one student, when asked why she deleted

certain phrases, commented, "I honestly don't know, and I don't see why it would have anything like political policies, education, and schools."

A few students did seem to have a vague notion of adding terms to narrow a search and reducing terms to broaden. One stated, "I think when I take out other words, it makes it a little bit broader, that way more stuff can come up." Another commented, "I'm going to try to change the words around because you can get a lot more…..you get more hits every time you change it around….they words that they kind of find relevant." Another noted, "If I type a synonym I may get more results." However, participants' errors in the use of Boolean operators often negated their use of synonyms or alternative terms.

Only a few students attempted to use more advanced techniques or features in their searches. Three students used truncation symbols in *Academic Search Premier* and *JSTOR*; unfortunately, two of the three had misspelled words or logic errors in their searches and thus were unsuccessful. One participant stated that she thought there was a way to search a phrase and suggested putting the 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."

Mental Models

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.

A number of information literacy advocates stress the need for students to have a well-formed mental model in order to be successful Internet searchers. (Large, 2006, as cited in UCL CIBER, 2008). "We need not only a broad understanding of how retrieval systems work and how information is represented within bibliographic or full text databases, but also some appreciation of the nature of the information space and how spelling grammar and sentence structure contribute to effective searches and evaluate the results" (UCL CIBER, 2008, p. 22). This research attempted to gain a better understanding of millennial students' mental models of search engines and how they function.

After observing the students search for materials the researcher conducted a brief post-observation interview to ascertain students' conceptualization of search. Several questions in the interview directly related to mental models. The researcher asked each student to explain, with as much detail as possible, how a search engine works – what it does with his/her search terms. She also asked how the search engine knew what sites the student needed. In answering these questions, students seemed to know more about how search engines collect sites than how their algorithms match various keywords in a search.

Five students noted that search engines send out spiders (one student further explained that a spider was a computer program developed to search for sites) to find and collect sites. Two others stated that search engines search for sites based on

page tags or on their source code. Five students described search engines as databases, indexes, or archives of sites and said that the engines match keywords based on their own index rather than the entire universe of sites. Four discussed Google's ranking of relevance by popularity.

Several students said that a search engine looks for what one types in the search box, but they seemed to understand it in terms of the keyword(s) as the concept of what one typed in rather than the literal string of letters. For example, one student said, "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 usually is looking for in that topic range. So in a way I'm siding with society and what they're generally looking for, and I have to try to find a way to surpass that and go into more detail in what I'm looking for." Several participants said that users have to be specific in their searches, but they rarely could articulate what they meant by specific. One student used an example of narrowing in a hierarchy; he suggested that the word dog would result in more hits than the term Shih Tzu. Still, in this explanation he did not articulate how the search engine read Shih Tzu differently from dog; he seemed to indicate that it was because dog was a broader category than a specific breed. Students also varied in their answers about phrase searching; some believed that the search engine searches for each word used, including what they called "little words" like "the" or "of," while others believed that the engines only use the primary terms. One participant stated that the search engine "takes the most basic words from my search and tries to bring up as many searches as it can." He continued by saying," I don't know what it does because I've never thought of this; I just type my words in.

Another student stated, "it takes out the meaningless words like 'does' and 'how' [his search was, "how does three strikes law deter crime] and it basically looks for three strikes, deter and crime." Still another said, "None of [the search engines] connect all the words together….they just kind of pull out whatever they feel is the most important word."

A number of students seemed to understand that they could narrow a search by adding terms; one stated, for example, that for a previous occasion she typed "pig wearing lipstick Obama speech" to find a video of that particular speech. But others, when asked how to narrow a search, mentioned techniques such as quotes or advanced searches. Interestingly, though, no student accurately used quotes in any of the observed searches. Others incorrectly thought that using parentheses (rather than quotations) directed the search engine to search for words as a phrase. Again, each of these cases confirms the UCL CIBER (2008) study that notes students don't recognize problems with specific search syntax, logic or spelling.

The researcher also prompted students to articulate their mental models by asking them how they would describe a search engine to someone who had never seen a computer or to provide a non-computer analogy. 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, 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. In fact,

one student suggested in another answer that Library of Congress classification and call numbers impede access to books rather that facilitate it.

Four compared a search engine to a library full of information; two gave the analogy of a phone book. Two compared search engines to card catalogs, although one added that he meant like a file cabinet. Neither elaborated on the card catalog's arrangement of cards or information. Four suggested that search engines were like encyclopedias, books or magazines; one of them went on to compare Google to a table of contents in a magazine, and one added that the databases were like books, but Google was like a documentary – "it gives you information on various things based on one search, based on one topic." Two students described search engines as librarians who help one search and find relevant sources from a much vaster universe of resources. One described Google (and the Internet) as being like water; both are accessible, free and reusable. The comparisons all related to what the sources held rather than how they organized or dispensed information.

Lastly, the researcher asked the students to describe how the search system (either database or search engine) selects and presents the list of results. More than half (57 %) suggested that searches were ranked by the number and proximity of the search terms used, while at least six students mentioned the popularity of the site or the number of sites that link to it; only one single participant suggested that both of these factors played into the list of results. When asked how the results were presented, some of the comments made were: "whatever has the most words that you used," and "it just presents anything that has the words you put up on the search bar....I believe [Google] orders it from the file that has most of the words to the ones

that may have only one or two or may not exactly be in that order." Another added, "either [the search engine] starts with the ones that have the most of your words together, maybe stated the most in the articles or the websites and it goes…it dwindles down to the least amount used. A fourth student commented, "I think it's how close the words are together…it will have to have one or both of them [terms], but the ones with both on them in it will be close to the top, and the ones with both of them, like right next to each other like you typed it are even closer to the top."

Others mentioned Google's popularity rankings, although they had difficulty articulating the concept. One stated, "What it does …. it basically goes on the number of hits that that page would generally…the number of hits the page generally gets in a day." Another commented, "It finds the top links that are connected to a certain website, and from the number of links that are in…that that website has, it sorts them out to which one is the top one and then so on and so on." A third stated, "I guess it's based on what other people have searched and what other people….I don't know."

In addition to the questions posed to students to elicit mental models, the researcher asked participants to draw or diagram the relationships between or among their keywords and the relationship between the keywords and the results. Appendix F includes all those diagrams.

Zhang (2008) in her study of students' mental models of the Internet as a whole identified four basic categories of drawings: 1) technical view, 2) functional view, 3) process view and 4) connection view. Students with a technical viewed characterized the Internet as a network of hardware and communication devices.

Participants with a functional view described the Internet as a place to accomplish

tasks or functions such as paying bills, shopping of finding information. Those with a

process view focused on search engines and information found through them, and,

lastly, participants with a connection view emphasized the connections between

computers, websites and the information they contained (Zhang, 2008). All of these

conceptualizations (with the possible exception of the process view) visualized the

Internet or World Wide Web as a whole. This study, by contrast, focused on

students' models of search engines or databases and how they process user input to

generate search results. Here, the mental models were narrower in scope; students

drew a representation of how their key words related to each other and the search

results the search engine or database returned. From the 21 drawings, three primary

categories emerged:

1) process view. Students in this category outlined the entire search as a

linear path or task flow diagram that included the search strategies and the results that

followed. Figures 5 and 6 depict examples of this category.
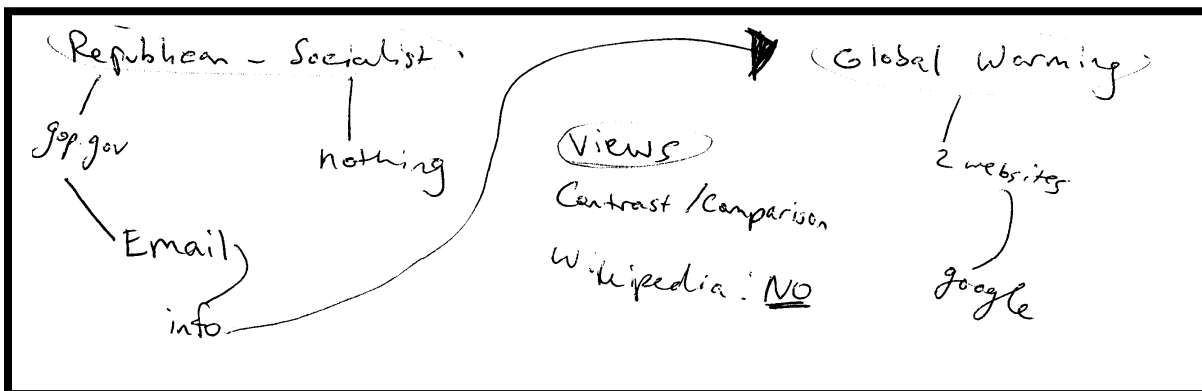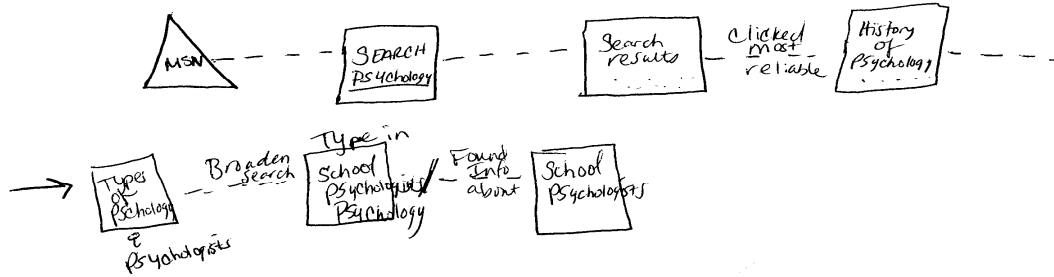

Figure 5. Process View Mental Model

Figure 6. Process view mental model

2) hierarchical view. Students here diagrammed a major topic with subtopics

or results highlighting aspects of the topic. Figures 7 and 8 depict examples of this
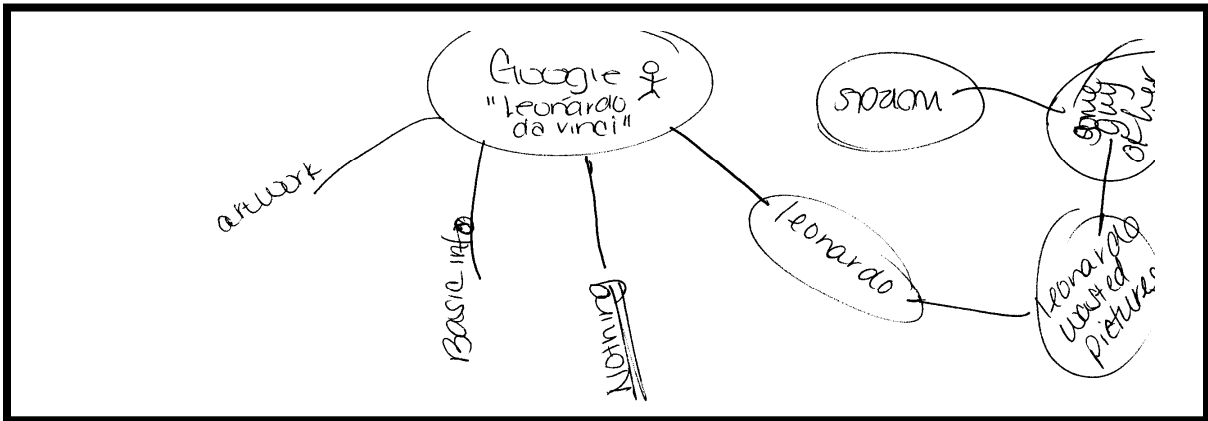
category.

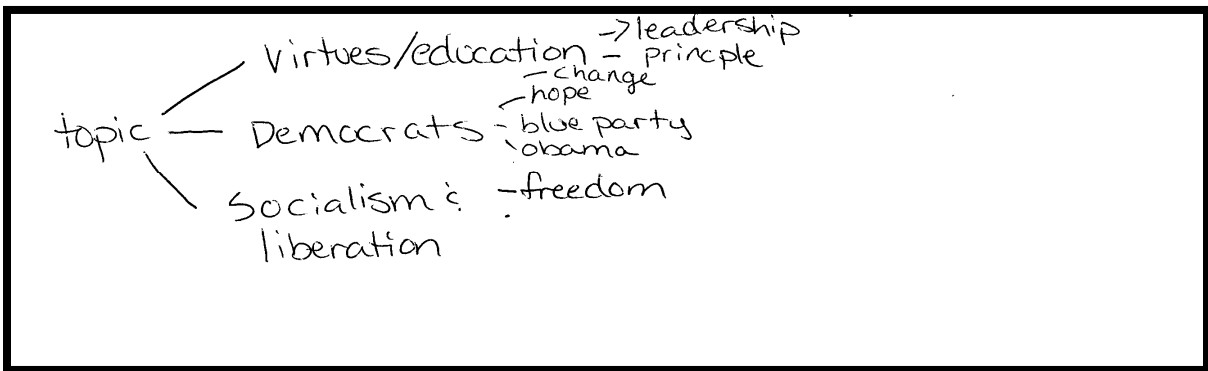Figure 7. Hierarchical View Mental Model

Figure 8. Hierarchical View Mental Model

3) network view. Students here created webs of interconnected terms.

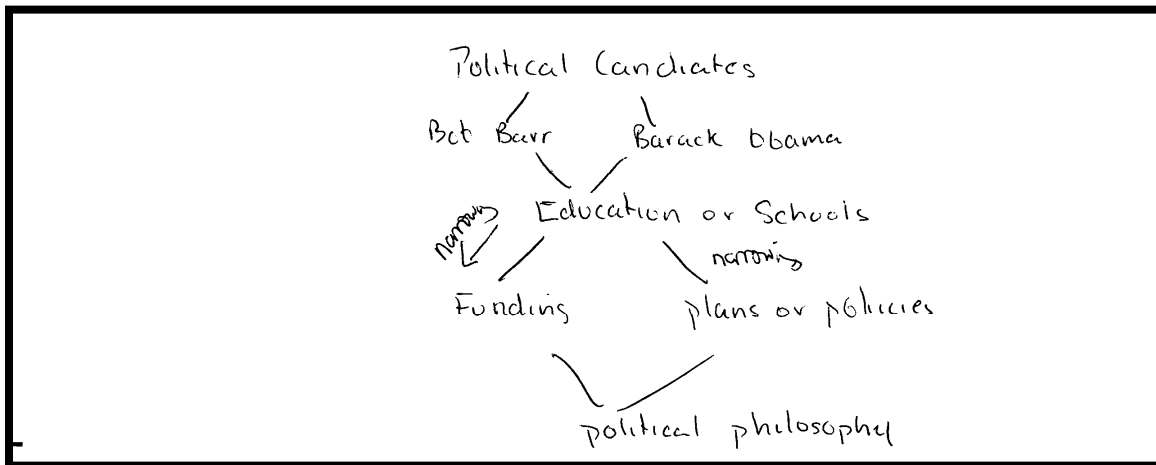Figures 9 and 10 depict examples of this category.

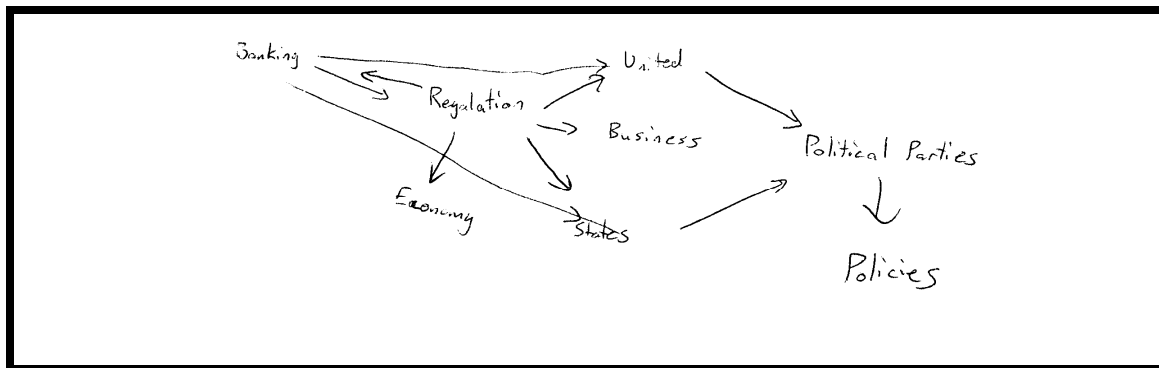Figure 9. Network View Mental Model

Figure 10. Network View Mental Model

In addition to these three views, two students simply listed all the terms they had used in the search in a chart-like form. These two responses were not included in the following observations.

As mentioned previously, the researcher gave no explanation or instruction prior to the students drawing their diagrams. Perhaps because of this lack of direction, none of the diagrams demonstrated a solid conceptualization of the type of relationship between the terms. For example, few illustrated what terms or combinations would narrow or broaden the search or if terms reflected the same concept or multiple concepts. Students showed connections between multiple words

but did not indicate how the words were connected or if certain subsets of words had stronger connections than others.

Those students with a network mental model conducted the majority of the searches (58.7%) as compared with those with a process view (11.7%) and a hierarchical view (29.6%). They also performed the most sophisticated searches; for example, they constructed 83% of the Boolean searches. Students with a hierarchical view composed the other 17%; those with a process view did not use Boolean operators (See Table 6). By comparison, 46% of the searches performed by those with a process view were simple searches; 79% were either simple or phrase searches. Among those with network views only 25% were simple searches while 65% were either topic + focus or Boolean searches.

Table 6.

Type of Search by Mental Model

Type of Search	Process	Pct	Hierarchy	Pct	Network	Pct	Total (n=206)
Simple	11	15%	32	44%	30	41%	73
Topic+Focus	5	8%	10	17%	45	75%	60
Phrase	8	20%	10	24%	23	56%	31
Boolean	0	0%	4	17%	20	83%	24
Subject Head.	0	0%	5	63%	3	38%	8
Total	24	11.6	61	29.6	121	58.7	206

Table 7

Number of Searches by Mental Model

Mental Model	Simple	Pct	Topic + Focus	Pct	Phrase	Pct	Boolean	Pct	Subject Heading	Pct
Process	11	46	5	21	8	33	0	0	0	0
Hierarchy	32	52	10	16	10	16	4	7	5	8
Network	30	25	45	37	23	19	20	17	3	2

Although the differences between the three mental model views are small, one may argue that those students with a process view see the system as more of a "black box" that gives searchers results with little reference made to the quality of search query. They may have the least sophisticated understanding of search tools and therefore craft the most simplistic queries. By contrast, those with a network view show the greatest emphasis on the searcher and the strength of the query. However, none of the three views indicate a solid mental model of the relationship between search terms and search results retrieved by either search engines or article databases.

Discussion

Although somewhat small in sample size, this study supports several ideas 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, none demonstrated a conceptual mental model of search such that they could effectively narrow or focus a search to retrieve relevant materials. Furthermore, they were rarely able recognize a problem (incorrect Boolean logic, spelling errors, etc) and resolve it for better results. Most of the students' searching strategies were rudimentary at best, and their ability to troubleshoot problems with queries in order to improve them was extremely limited.   Ultimately, this research found that millennial students do not have a working mental model of information retrieval upon which to build more complex search strategies. Despite this lack of a well-defined mental model, the students believed that they were competent and successful searchers. When their search strategies failed to retrieve their desired results, they questioned the search tool rather than reviewing their own logic or search strategy.

Confidence and Satisfaction

When asked about their ability to locate information online, most of the students considered themselves successful in their searches that day and believed themselves to be competent searchers overall.   Based on the searches they conducted during the observation sessions, this success generally meant locating five to fifteen sites or articles that students would save for future reading or review. They were

satisfied with using very simple search strings (single concept or topic plus focus); if

they retrieved at least ten items they believed that they had found enough material.

Participants demonstrated the concept of satisficing, a term coined by Herbert Simon

(1956, 1979) and recently popularized by Barry Schwartz (2004). Simon posits that

in the absence of the knowledge of the perfect choice (e.g., the "best" article), two

factors come into play – searching and satisficing. One can search indefinitely for the

perfect item, but his/her aspirations for the optimum outcome will adjust over time

and effort to what is reasonably attainable. At that point, once he/she finds an item

that meets the adjusted level of aspiration, he/she is satisfied with the result. The

longer the search takes, the more likely the person is to satisfice (Simon, 1956, 1979).

Schwartz proposes that some people tend to be satisficers, that is, they do not explore

all the choices available but settle quickly on a satisfactory choice rather than

spending greater time finding the optimal choice (Schwartz, 2004). If they are not

passionate about or personally engaged in a subject, students are generally are

satisfied with whatever information they find so long as it appears to be relevant to

their need and reasonably reliable (and some students may be satisfied without the

reliability). They do not attempt to craft a better search string and continue the search

for the best article or site or *all* the articles or sites. In this study all 21 students,

regardless of the quality of materials located, were somewhat, if not entirely, satisfied

with their results, and did not plan to look for additional material unless they later

found that they were missing certain information. Although none did more than

scratch the surface of their topic, and few had performed truly well-crafted searches,

all expressed satisfaction. In this study no student was concerned with the millions

of results he/she retrieved from Google or other search engines; only a few reviewed more than the first two pages of results. Typically participants found at least five to ten sites that they deemed relevant and useable on the first two pages of search results and went no further to focus their search.

Search strategies

Although several students had some understanding of Boolean logic and proper search punctuation, few chose to use such techniques and gave up quickly if their searches did not retrieve the expected results. Most students preferred simple searches even if they required the students to scan a larger list of results. When participants encountered a perceived problem with a search, they typically crafted a new query rather than refine their existing one. None seemed either aware of or interested in developing more complex searches to retrieve more targeted results.

The large number of spelling errors observed and students' failure to notice their own spelling mistakes call attention to the difference between search engines and databases in their ability to recognize misspelled words and to offer searchers alternative spellings. All of the students who made mistakes in the search engines seemed very familiar with the "Did you mean?" feature in Google and quickly clicked on a revised search with the correct spelling. Database developers who design algorithms that make allowances for spelling errors or alternative spellings will facilitate student search success and greater student satisfaction with their products. It is unclear if these errors result from poor spelling skills or just mistakes made in typing haste; however the cause here is irrelevant. Searchers will continue to

make spelling or typographical mistakes, and search tools must make the necessary accommodations for those mistakes. Some database producers such as EBSCO have recently released new products with more robust search mechanisms, and others have such products in development, but all the database vendors must make these algorithms for spelling variants and synonyms a priority in order to retain younger users.

Although most participants did express an understanding that the search engines and databases match their indexes for keywords used in a student's search, that understanding often did not transfer into specific searches, and students were often ill-equipped to troubleshoot particular problems. Some appeared to see the keywords more as concepts rather than as strings of letters to be matched; they would perhaps broaden the concept without broadening their search strategy. Participants did not consider that sites or articles might only include one political party's stance and therefore unduly limited their results by building queries that required two or more parties in a single site or article. Furthermore, their choice of synonyms or related terms, sometimes automatically generated by Microsoft Word or an online thesaurus, were often off the mark. Students rarely could effectively refine or improve an existing search; rather they simply typed a new query. In fact, this was the strategy that most students employed. In only 24 cases (slightly more than 10% of the total searches) did participants effectively broaden or narrow the focus of an existing search. More commonly, when a student was stymied by incorrect spelling or logic, he/she continued to generate alternative terms. This was done very quickly without the student paying much attention to the search terms used, and several

students repeated failed searches by retyping already used keywords. As they changed search tool they routinely would fail to remember which terms had been successful in the previous tool and would sometimes start with an unsuccessful keyword.

As mentioned previously, many of the students repeated failed searches and made a number of mistakes in spelling and search logic. Perhaps the students' rapid pace of searching, scanning and evaluation led to some, if not many, of their mistakes. The participants rarely took the time to look at materials for relevance or stop to consider an appropriate keyword for a search; rather, they would try a search and quickly modify it again and again. Perhaps due to their multitasking or their haste, students did not fully attend to their searches. This lack of attention often led to repetition of mistakes in spelling or returning to a problematic search strategy in a new tool. This behavior and their speed of processing seem consistent with observations of researchers (Prensky, 2001; Oblinger & Oblinger, 2005; UCL CIBER, 2008), who note that millennials are hypertext readers and tend to move quickly from one bit of information to the next. This rapid pace may result in less than careful queries and students' impatience may lead to a hasty and incomplete review of results.

Selecting and Evaluating Sources

The notions of hypertext reading and multitasking certainly significantly impact students' evaluation of the sites and articles they find. Although critical thinking skills were not the focus of this study, the researcher observed methods and

criteria students employed and knowledge they applied in evaluating Internet sites

and articles for their academic work. Most of the students reviewed retrieved sites

and articles very quickly and made decisions to keep sites (generally by saving

URLs) within seconds. Almost all students at least clicked on a site in the results list

to quickly review its content rather than simply relying on the snippet of description

from the search engine or the citation in the database, but rarely did they take more

than a cursory look at the site before deciding if it was worth reviewing in more detail

or selecting for later review. Two thirds of the students made some comment about

the quality of the sites they encountered; 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. For example, two students located university websites that

simply described (or even advertised) undergraduate and graduate programs that they

deemed relevant for their topic. One student who was searching for information on

party policies on pollution suggested that a university's description of its biology

program was not only credible but appropriate for his research. In other cases

students made evaluations based on misinformation or lack of knowledge. One

student, finding an article from *USA Today,* wondered if the source was a newspaper.

Another who found a *Baltimore Sun* article did not believe it to be an actual article

from the paper, despite the full citation and the *Sun*'s URL. One student researching

alternative energy sources located Chevron's website. He stated that he did not know

who or what Chevron was, but after reviewing the site, he thought that the company's

information on energy and the environment might provide all he needed for

background on the issue. While some students appreciated and trusted sites such as

CNN and the *Economist*, others disregarded those very sites. One student dismissed a report issued by the Cato Institute because the Institute's home page did not seem informative based on its appearance. These findings are similar to those demonstrated in the UCL CIBER (2008) study that indicated students make judgments on articles sites on relevance based on keywords only. Here participants often based their evaluation of site on visual appearance and/or frequency of keywords and did not always judge based on accuracy or authority. Students lacked reasonable criteria for evaluation sites and often applied inappropriate criteria to accept a site as authoritative or reliable.

Additionally, students sometimes purposefully looked for blogs or other more personal postings, saying that they wanted to get "average" people's opinions as well as using more authoritative sources. Participants did not seem to particularly value the credibility of established sources or seek information from professional journalists or researchers.

Wikipedia as a Reliable Source

Nine of the 21 students talked specifically about the use of Wikipedia in their research. Recent studies (Salaway, Caruso, & Nelson, 2008; Head & Eisenberg, 2009) have that found that many students use Wikipedia and its references as a starting point in their research. Here the comments fell into one of three categories: those who would never use Wikipedia, those who use only its external references or links and those who use it for background or for research. Five students were adamant against Wikipedia, decrying its lack of accuracy; four regularly used its

entries or its references. At least three students used the external links in Wikipedia entries to find material. Several students said that they would use Wikipedia for their own information needs, but they would not use for research projects. On participant stated, "I was told that for my project that Wikipedia is not a good enough source to use, so I would skip it, but I would normally go to Wikipedia because it's really informative."

Those who denounced Wikipedia did so from a very dualistic framework, and their views appeared to be informed by high school teachers or librarians. Most of the negative comments seemed to parrot admonitions that Wikipedia should be avoided at all costs because it is inaccurate and unreliable. In fact, one student stated, "The ones [sic] I usually stay away from is Wikipedia just because it's been so repeated since high school, 'don't use Wikipedia because anyone can say anything they want about that artist, and it doesn't have to be true.'" When asked why they did not use the site, most answered that it was because it could be changed at any time by anyone. These students' avoidance of the online encyclopedia does not seem to reflect other researchers (Head & Eisenberg, 2009) who see large numbers of students using the source regularly even if only for background on a topic. Some of the students here may protest the source too much, however; in fact, one stated she never used Wikipedia as she scanned the contents of two of its entries.

Reading and Multitasking

Although this study was not designed to deeply explore students' reading habits, the researcher observed ways in which students consume information and

process and manage their sources. At least 18 participants quickly skimmed or scan materials searching for keywords. This is consistent with studies (UCL CIBER, 2008; Williams, 1999) that found that younger students scan materials for small chunks of information within the text and do not spend time reading the entire document. Similarly, Watson (1998) finds middle-schoolers who differentiate between reading a "resource," which they skim, and reading a "treasure," which they read more fully (Watson, 1998, p. 1032).

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%) used a number of supportive actions to keep their place or remain focused, such as highlighting text or following with the pointer as they read. Six students (29%) actually read material out loud; however, it is unclear if they were doing so only because the think-aloud protocol or if it was a natural behavior. Still, a majority of students needed some mechanism (reading out loud or following text with pointer and eye) to focus their attention on the text. A number of studies have researched students' scanning habits (Williams, 1999; Prensky, 2001; Oblinger & Oblinger, 2005; UCL CIBER, 2008); few have discussed their reading or comprehension skills. This research questions students' ability to focus their attention on the text at hand.

As mentioned previously, millennials are master multitaskers (Rainie, 2006); they are comfortable quickly moving back and forth between multiple projects. This

research noted participants' extensive use of tabs and the move to cut/copy and paste URLs or content from sites into documents or emails and their superficial scanning of information received. Recent studies have examined college students' multitasking, particularly their use of instant messaging (IM) while reading, studying or searching for information. Although some studies (Spink, 2003) show that multitasking information behavior, that is, searching for results on several topics at once, can be iterative and effective; several other studies (Cummings, 2004; Fox, Rosen, & Crawford, 2009; Levine, Waite & Bowman, 2007) found that users engaged in multitasking, specifically instant messaging, demonstrated negative impacts on other task behaviors. Cummings (2004), in his study of Navy missile operators, saw a lower performance score among participants engaged in instant messaging. Fox, Rosen, & Crawford (2009) found that participants who used instant messaging on a daily basis experienced lower reading comprehension scores. Levine, Waite & Bowman, (2007) noticed increased distractibility among those using instant messaging in their reading, and distractibility was inversely related to the amount read. They posit that such multitasking can lead to shorter attention spans. A recent study (Ophir, Nass, & Wagner, 2009) found that frequent multitaskers were more easily distracted than those who rarely multitask, and they were less able to stop task-switching and focus on one task.

Participants in this study demonstrated a strong tendency to multitask in their research process. Seven students opened new tabs/windows for each new site/article they reviewed or selected, and twelve students frequently toggled between various sites and tools such as Word, Notepad or email applications to copy and paste

material and even compare information between sites. While many of the students simply copied and pasted URLs or citations for later review, a few students copied large amount of text (sometimes the entirety of an article or page), in some cases with no citation or attribution. Although some of these students mentioned that they were careful about plagiarizing, it is easy to see how students could even inadvertently restate a thought or idea verbatim from an online source.

Certainly here students multitasked to capture material for later use. In the absence of a particular tool designed to automatically save material or citations that students found of interest, participants had to either copy and paste or make written notes about the materials they found for later use in writing papers. While the participants all were comfortable toggling between tabs and windows and suggested that they often multitasked while doing work at home, the observed multitasking – researching, note-taking, emailing – was all directly related to the research task. However, it is unclear as to whether or not the continued toggling between applications had an effect on students' ability to attend to the materials they were reviewing and evaluating.

Conclusion

As colleges and universities enroll younger users with more experience with publicly-available Internet resources, librarians and faculty will continue to see generational shifts in their information retrieval practices and students' search skills and preferences for search tools. Academic librarians who teach information literacy will face the need to reexamine how college students today search for and retrieve information and rethink the way they teach millennials (and beyond) online search strategies and resource evaluation. Furthermore, librarians who purchase proprietary databases and the producers who design them also must consider younger students' conceptualizations of search and their preferred search methods and strategies. One of the original questions this research posed is whether students (and young faculty) should learn to use complex interfaces with less intuitive search strategies based on mental models of print-based research tools or should designers develop interfaces that more closely relate to millennials' own mental models of Internet-based information retrieval with engines that more accurately and effectively parse a simpler, more natural language query. Some educators believe that to do so would "dumb down" information literacy instruction and information retrieval to the lowest common denominator. However, if today's students do not operate in a linear fashion and learn (and search) by discovery, the focus of information literacy instruction should be on evaluating search results and developing more effective strategies for refining a search rather than on constructing a near-perfect initial search. More attention should be placed on evaluation and critical thinking skills that will endure beyond the specific techniques and syntax necessary for the current search tools.

It is likely that the tools will change, but the need for evaluation and critical review of sources will only increase as the volume of information on the web expands. If students do not have preexisting mental models of online information retrieval, librarians providing information literacy instruction may do well to step back and help students conceptualize the relationship between the keywords in their queries and the result sets they retrieve. Likewise, database developers would do well to build more robust synonym rings that will capture results with related terminology, given that young searchers do not seem concerned with the number of hits in a particular results set and are less likely to include related terms or spelling variants in their initial searches.

Although students may not be strong searchers, employing complex and effective search strategies, neither will they be likely to want or appreciate instruction that simply teaches more advanced search techniques. If students are indeed satisfied with results from simpler searches and often do not employ more advanced techniques, such as quotations around phrases, Boolean operators or search limiters, even when they know how to apply them, database developers and librarians must ask themselves how much effort should be placed on training students in these advanced features. Why should students be taught to craft complex search strategies if they purposefully choose not to use them or fail to use them correctly? Rather than teaching students more effective search syntax, should more attention be placed on explaining the general nature of search systems so that students can revise their searches, or on developing critical thinking and evaluative skills?

Therefore, librarians may be forced to accept that users conduct less than perfect searches and are generally satisfied with materials based on visual cues and readability rather than on accuracy, reliability or caliber of scholarship. Where librarians may have an opportunity to have a greater impact is in the area of evaluating results. Still, it may be idealistic for librarians (and faculty) to expect millennial students to dig beyond a surface level to find the "ideal" source. Unless faculty explicitly require more scholarly sources in undergraduate research and stress the importance of authority and credibility of sources, students will continue to value sites and articles that appeal to them. Perhaps librarians stand alone in their emphasis on scholarly sources for academic research, particularly at the undergraduate level. Although these findings stray from the original intent of the research to focus solely on students' mental models of search, this study points to a critical need to address students' evaluation skills and to provide a stronger foundation for students to detect both blatant and more subtle biases in print and online publications and sites. Librarians need to consider a shift in focus from search to evaluation and critical thinking skills. Likewise, faculty may need to reexamine the parameters they set for the sources students used for research assignments and may need to model critical evaluation of material for their students.

Database designers also need to consider changes in students' conceptualizations of search and their information retrieval habits. Some leading producers such as Ebsco, Proquest, and OCLC are already incorporating synonym rings and spelling variants in their search algorithms, offering students results when they misspell common terms or use related terms. Still these engines are not as robust

as Google or other public search engines or updated as frequently with new terms. For example, on May 28, 2009 a search in Ebsco's *Academic Search Premier* database on a misspelled Barac Obama retrieved two hits compared with 18,401 hits for a correctly spelled Barack Obama. The database suggests Bara Obama, Baras Obama, and Baram Obama as possible alternatives, and the suggestions are not displayed prominently for searchers to see. In Google, the same search brings up a "Did you mean" suggestion of the correct Barack Obama. Databases need to improve their algorithms for spelling variants and related terms. This has not appeared to be an immediate priority to information professionals; while producers are developing new algorithms, the library community does not seem to be voicing strong advocacy of more robust tools. A few years ago, Edward Proctor (2002) noted how few articles in library and information science literature focused on robust search mechanisms, particularly those that employ spelling variants or synonym rings, and how few in the field were demanding more of the database providers. He points to not only spelling errors made by users but also frequent spelling mistakes in online content itself. Even more recently, Péter Jascó (2007) pointed out that most commercial database providers don't focus on spelling enough and a message of "no results found" as opposed to "Did you mean" appropriately drives users away from those commercial tools. While some key vendors are employing thesauri and spelling variants, many major producers are not and thus may lose younger users who believe their searches have failed in their systems without understanding why.

Another issue that librarians and database providers face is the number of databases any given academic library offers, the slow development of discovery tools

and the inadequacies of currently used federated search systems. Students using a single search engine such as Google or even Google Scholar need not first make a choice as to which database is an appropriate domain for their search; they simply type their query, regardless of the discipline in which it is. Students using their library's online databases often must choose a database with a name that provides little or no clue as to its content. For example, in this study several students inadvertently searched in *JSTOR* and *Project Muse*, not understanding what materials either of those databases contained and not knowing if either was appropriate for their topic. Students frequently make their first mistake by selecting an inappropriate database. Federated search tools such as Ex Libris' *MetaLib* or Serials Solutions' *360 Search* provide users a common interface through which to search a number of databases, but, unfortunately, as these tools search across platforms with differing degrees of metadata, they may not provide users with as targeted a search as they may expect. Haya, Nygren, and Widmark (2007) in their study comparing Google Scholar to Ex Libris' *Metalib* federated search tool, found that undergraduates found Google Scholar to be more intuitive and less difficult to use. They also found more than twice as many relevant documents (as noted by the number they saved) in the Google Scholar results. Although students were not particularly satisfied with either tool, they preferred Google Scholar in almost all aspects of the evaluation (Haya, Nygren & Widmark (2007). If students choose not to use commercial databases, and there seems to be little incentive or advantage for them to use such products, there will be little opportunity to teach them more advanced search syntax that is less necessary in Google or other search engines. Although they may still allow for complex Boolean

or other advanced search syntax, databases may need to redesign interfaces and algorithms to appeal to younger users' preference for simplistic interfaces and natural language searches. Furthermore, as information professionals build data harvesting and discovery tools designed to retrieve items in a variety of formats from a variety of sources, the need for more robust algorithms becomes even greater.

Because students do not seem to have solid mental models of the search mechanisms and thus do not craft well-honed queries nor employ subject headings, database producers should continue to build more robust discovery tools designed to harvest materials in a variety of formats from a variety of sources based on less refined searches. The leading producers must adopt and adhere to metadata standards that will facilitate cross-platform searching. Students and young faculty demand more seamless searches in a single interface regardless of the location or format of the material. They expect the system to correct spelling or grammar errors and want to use more natural language in their queries.

Librarians may still continue to develop or strengthen students' mental models of search mechanisms but more in an approach that appeals to millennials' sense of discovery. If today's students prefer to craft simple queries, information literacy instruction should emphasize problem-solving strategies when searches prove ineffective and concentrate on ways in which students can broaden or narrow lists of results. Certainly, information literacy should focus on evaluation of sites and materials, again drawing on millennials' preference for constructivist learning and teach students to refine searches more effectively.

The implications of this research can be far ranging and may impact pedagogy beyond information literacy instruction. As academics experience students' limited critical thinking skills and willingness to satisfice, they may need to rethink their expectations of the sources students use and reevaluate the way they craft research-oriented assignments.

Lastly, additional research in millennials' mental models of information retrieval and their understanding of scholarly and non-scholarly material will clarify what librarians and database producers must do to develop more effective search tools for students and capture the interest of millennials so that they will become better searchers, and more importantly, better consumers of information able to discern the quality of the material they obtain online.

Although this research focused on students' mental models of information retrieval, certain observations uncovered several other areas that call for further research. These behaviors include the ways in which students process, evaluate, and select information.

Reading and Processing Information

As mentioned previously, this study observed students quickly scanning materials; very few students carefully read materials, and those who did read content, often did so by reading out loud, highlighting the text or following the text with the mouse and pointer. Additional research is needed to determine how much information is obtained by scanning and how students focus their attention on the material they read. Have younger students become so accustomed to hypertext

reading that they have difficulty reading longer passages for comprehension? Are supportive techniques such as reading out loud or physically following the text necessary for focused attention among readers today?

Evaluating Sources

This research calls for more study of millennials' critical thinking skills and the criteria they apply to evaluate sites and articles they retrieve. Participants here judged sites as much for their visual appeal as they did for their accuracy, objectivity and authoritativeness. More research is necessary to identify the factors students value and use most to judge material for academic work. Without implied or explicit encouragement from faculty or librarians to judge information on its content rather than design and to consider issues of authorship, purpose and bias, do students use these criteria in their evaluation?

While this study indicates the need for librarians to reevaluate and restructure information literacy instruction and for database developers to invest more into the development of more robust algorithms for spelling variants and synonyms, it also calls for additional research on students' mental models of information retrieval. Frequent use of these tools does not necessarily equate to a deep understanding, and as the volume of online material expands at a rapid pace users will continue to either retrieve the most convenient items or struggle to focus a search and evaluate material.

Although this study calls for additional research in areas of students' search strategies and critical thinking skills, it suggests actions that university librarians, faculty and database developers must consider in their service to this generation of

users. Librarians need to consider millennials' levels of self-confidence and preferences for simple searches and focus on troubleshooting problematic searches rather than designing an initial query. Database producers need to develop more robust algorithms for alternative spelling and synonyms that offer students suggestions when searches fail to achieve the intended results.

References

Abram, S. & Luther, J. (2004). Born with the chip. *Library Journal, 129* (8), 34-37. Retrieved November 1, 2007 from *Academic Search Premier* database.

Association of College and Research Libraries (ACRL). (2007). *Introduction to Information Literacy.* Retrieved February 11, 2007 from: http://www.ala.org/ala/acrl/acrlissues/acrlinfolit/infolitoverview/introtoinfolit/ introinfolit.htm

Baeza-Yates, R., & Ribeiro-Neto, B. (1999). *Modern information retrieval.* New York: ACM Press.

Bates, M. (1989). The design of browsing and berry picking techniques for the online search interface. *Online Review, 13* (5), 407-424.

Belkin, N. J. (1980). Anomalous states of knowledge as a basis for information retrieval. *Canadian Journal of Information Science, 5*, 133-143.

Bilal, D., & Kirby, J. (2002). Differences and similarities in information seeking: Children and adults as Web users. *Information Processing and Management, 38* (5), 649-670.

Blomberg, J., Giacomi, J., Mosher, A., & Swenton-Wall, P. (1993). Ethnographic field methods and their relation to design. In D. Schuler, & A. Namioka, (Eds.). *Participatory design: Principles and practices.* Hillsdale, NJ: Lawrence Erlbaum.

Borgman, C. L. (1984). Psychological research in human-computer interaction. In M. E. Williams (Ed.). *Annual Review of Information Science and Technology* (Vol. 19) (pp. 33-64). White Plains, NY: Knowledge Industry Publications.

Borgman, C. L. (1986). The user's mental model of an information retrieval system: An experiment on a prototype online catalog. *International Journal of Man-Machine Studies, 24,* 47-64.

Borgman, C. L. (1996). Why are online catalogs *still* hard to use? *Journal of the American Society for Information Science, 47* (7), 495-503. Retrieved November 10, 2007 from *ABI Inform* database.

Bourne, C. P. & Hahn, T. B. (2003). *A history of online information services, 1964-1976.* Cambridge, MA: MIT Press.

Brandt, D. S. (1997). Constructivism: Teaching for understanding of the Internet. *Communications of the ACM, 40* (10), 112 – 117. Retrieved April 6, 2007 from *ACM Digital Library* database.

Brandt, D. S., & Uden, L. (2003). Insight into mental models of novice Internet searchers. *Communications of the ACM, 46* (7), 133-136. Retrieved April 6, 2007 from *ACM Digital Library* database.

Brown, M. (2005). Learning spaces. In D. G. Oblinger & J. L. Oblinger (Eds.), *Educating the net generation* (pp. 12.1 – 12.22). Boulder, CO: Educause. Retrieved February 11, 2007 from http://www.educause.edu/educatingthenetgen/

Chowdhury, G. G. (1999). The Internet and information retrieval research: A brief review. *Journal of Documentation, 55* (2), 209-255.

Chu, H. (2003). *Information representation and retrieval in the digital age*. Medford, NJ: Information Today.

Cole, C., & Leide, J. E. (2003). Using the user's mental model to guide the integration of information space into information need. *Journal of the American Society for Information Science & Technology, 54* (1), 39-46.

Cole, C., Lin, Y., Leide, J., Large, A., & Beheshti, J. (2007). A classification of mental models of undergraduate seeking information for a course essay in history and psychology: Preliminary investigations into aligning their mental models with online thesauri. *Journal of the American Society for Information Science & Technology, 58* (13), 2092-2104.

Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches* (2nd ed.). Thousand Oaks, CA: Sage.

Cummings, M. L. (2004). The need for command and control instant message adaptive interfaces: Lessons learned from tactical Tomahawk human-in-the-loop simulations. *Cyberpsychology & Behavior, 7* (6), 653-661. Retrieved September 13, 2009 from *Academic Search Premier* database.

DeRosa, C., Cantrell, J. Hawk, J., & Wilson, A. (2006). *College students' perceptions of libraries and information resources: A report to the OCLC membership.* Dublin, OH: OCLC Online Computer Library Center. Retrieved April 3, 2008 from: http://www.oclc.org/reports/pdfs/studentperceptions.pdf.

Dervin, B. (1998). Sense-making theory and practice: An overview of user interests in knowledge seeking and use. *Journal of Knowledge Management, 2* (2), 36-46.

Dervin, B. (1999). On studying information seeking methodologically: The implications of connecting metatheory to method. *Information Processing and Management, 35*, 727-750.

Dictionary of information and library management (2006). London: A.C. Black. Retrieved March 23, 2009 from Credo Reference: http://www.credoreference/acbinfomanage/concept_map.

Dimitroff, A. (1992). Mental models theory and search outcome in a bibliographic retrieval system. *Library & information science research, 14* (2), 141-156.

diSessa, A. A. (1983). Phenomenology and the evolution of intuition. In D. Genter & A. Stevens (Eds.), *Mental models* (pp. 15-33). Hillsdale, NJ: Lawrence Erlbaum.

Erdelez, S. (1999). Information encountering: It's more than just bumping into information. *Bulletin of the American Society for Information Science, 25* (3). Retrieved February 15, 2007 from http://www.asis.org/Bulletin/Feb-99/erdelez.html.

Efthimiadis, E. N. (2000). Interactive query expansion: A user-based evaluation in a relevance feedback environment. *Journal of the American Society for Information Science & Technology, 51* (11), 989-1001.

Ferguson, C. L. & Grogg, J. E. (2004). OpenURL link resolvers. Computers in Libraries, 24 (9), 17-24. Retrieved November 29, 2007 from *Academic Search Premier* database.

Fister, B. (1992). The research processes of undergraduate students. *Journal of Academic Librarianship, 18* (3), 163-169.

Fox, A. B., Rosen, J., & Crawford, M. (2009). Distractions, distractions: Does instant messaging affect college students' performance on a concurrent reading comprehension task? *Cyberpsychology & Behavior, 12* (1), 51- 53. Retrieved September 12, 2009 from *Academic Search Premier* database.

Fox, S. & Madden, M. (2006). *Generations online*. Washington, DC: Pew Internet and American Life Project. Retrieved November 11, 2007 from http://www.pewinternet.org/pdfs/PIP_Generations_Memo.pdf.

Gerrie, B. (1983). *Online information systems: Use and operating characteristics, limitations, and design alternatives*. Arlington, VA: Information Resources Press.

Green, D. (2000). The evolution of Web searching. *Online Information Review, 24* (2), 124. Retrieved October 22, 2007 from *ABI Inform* database.

Guinee, K., Eagleton, M. B., & Hall, T. E. (2003). Adolescents' Internet search strategies: Drawing upon familiar cognitive paradigms when accessing electronic information sources. *Journal of Education Computing Research, 29* (3), 363-374.

Haya, G., Nygren, E., & Widmark, W. (2007). Metalib and Google Scholar: A user study. *Online Information Review, 31* (3), 365-375. Retrieved May 27, 2009 from *ABI Inform* database.

He, P. W., & Jacobsen, T. E. (1996). What are they doing with the Internet: A study of user information seeking behaviors. *Internet Reference Services Quarterly, 1* (1), 31-51.

Head, A. J. (2007). Beyond Google: How do students conduct academic research? *First Monday: Peer-reviewed Journal on the Internet, 12* (8). Retrieved August 13, 2008 from http://firstmonday.org/issues/issue12_8/head/index.html.

Head, A. J., & Eisenberg, M. B. (2009). Finding context: What today's college students say about conducting research in the digital age. Project Information Literacy Progress Report, Information School. University of Washington. Retrieved March 23, 2009 from http://projectinfolit.org/pdfs/PIL_ProgressReport_2_2009.pdf.

Holliday, W., & Li, Q. (2004). Understanding the millennials: Updating our knowledge about students. *Reference Services Review, 32* (4), 356-366. Retrieved February 20, 2007 from *Emerald* database.

Holtzblatt, K., Wendell, J. B., & Wood, S. (2005). *Rapid contextual design: A how-to guide to key techniques for user-centered design.* San Francisco: Morgan Kaufman Publishers.

Howe, N., & Strauss, W. (2000). *Millennials rising: The next great generation.* New York: Vintage Books.

Jacsó, P. (2007, May). *Trends in professional and academic online information services.* Presentation at INFORUM 2007, 13th Conference on Professional Information Resources, Prague, Czech Republic. Retrieved May 27, 2009 from http://www.inforum.cz/pdf/2007/jacso-peter.pdf.

Jansen, B. J., & Pooch, U. (2001). A review of web searching studies and a framework for future research. *Journal of the American Society for Information Science & Technology, 52* (3), 235-246.

Johnson, F. C., & Crudge, S. E. (2007). Using the repertory grid and laddering technique to determine the user's evaluative model of search engines. *Journal of Documentation, 63* (2), 259-280. Retrieved March 28. 2009 from *ABI Inform* database.

Jones, S. (2002). *The Internet goes to college: How students are living in the future with today's technology*. Washington, DC: Pew Internet and American Life Project. Retrieved February 10, 2007 from http://www.pewinternet.org/pdfs/PIP_College_Report.pdf.

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*.

Katzeff, C. (1990). System demands on mental models for a full text database. *International Journal of Man-machine Studies, 32* (5), 483-509.

Kerr, S. T. (1990). Wayfinding in an electronic database: the relative importance of Navigational cues vs. mental models. *Information Processing & Management, 26* (4), 511-523.

Kuhlthau, C. C. (1999). Accommodating the user's information search process: Challenges for information retrieval system designers. *Bulletin of the American Society of Information Science, 25* (3),   Retrieved February 15, 2007 from *ABI Inform* database.

Kuhlthau, C. C. (2004). *Seeking meaning: A process approach to library and information services* (2nd ed.). Westport, CT: Libraries Unlimited.

Law, J. (2008, February). *Observing student researchers in their native habitat*. Presentation at the VALA - Libraries, Technology and the Future Biennial Conference, Melbourne, Australia. Received from author through personal email.

Lenhart, A., Rainie, L., & Lewis, O. (2001). *Teen life online: The rise of the instant-message generation and the Internet's impact on friendships and family relationships*. Retrieved March 20, 2007 from http://www.pewinternet.org/pdfs/PIP_Teens_Report.pdf

Levine, L. E., Waite, B. M., & Bowman, L. L. (2007). Electronic media use, reading, and academic distractibility in college youth. *Cyberpsychology & Behavior, 10* (4), 560-566. Retrieved September 13, 2009 from *Academic Search Premier* database.

Lippincott, J. K. (2005). Net generation students and libraries. In D. G. Oblinger & J. L. Oblinger (Eds.), *Educating the net generation* (pp. 13.1-13.15). Boulder, CO: Educause. Retrieved February 11, 2007 from http://www.educause.edu/educatingthenetgen/

Macgill, A. R. (2007). *Parent and Teen Internet Use*. Washington, DC: Pew Internet and American Life Project. Retrieved November 11, 2007 from http://www.pewinternet.org/pdfs/PIP_Teen_Parents_data_memo_Oct2007.pdf

Makri, S., Blandford, A., Gow, J., Rimmer, J., Warwick, C., & Buchanan, G. (2007). A library or just another information resource? A case study of users' mental models of traditional and digital libraries. *Journal of the American Society for Information Science & Technology, 58* (3), 433-445.

Marchionini, G. (1989). Making the transition from print to electronic encyclopaedias: Adaptation of mental models. *International Journal of Man-Machine Studies, 30* (6), 591- 618.

Marchionini, G. (1996). *Information seeking in electronic environments*. New York: Cambridge University Press.

Marchionini, G. & Komlodi, A. (1996). Design of interfaces for information seeking. In M. E. Williams (Ed.). *Annual Review of Information Science and Technology* (Vol. 33) (pp.89-130). Medford, NJ: Information Today.

Markey, K. (2007a). Twenty-five years of end-user searching, part 1: Research findings. *Journal of the American Society for Information Science and Technology, 58* (8), 1071-1081. Retrieved November 3, 2007 from *Wiley InterScience*.

Markey, K. (2007b). Twenty-five years of end-user searching, part 2: Future research directions. *Journal of the American Society for Information Science and Technology, 58* (8), 1123-1130. Retrieved November 18, 2007 from *Wiley InterScience*.

McGlynn, A. P. (2005). Teaching millennials, our newest cultural cohort. *Education Digest, 71* (4), 12-16. Retrieved October 27, 2007 from *Academic Search Premier* database.

McNeely, B. (2005). Using technology as a learning tool, not just the cool new thing. In D. G. Oblinger & J. L. Oblinger (Eds.), *Educating the net generation* (pp. 4.1-4.10). Boulder, CO: Educause. Retrieved February 11, 2007 from http://www.educause.edu/educatingthenetgen/.

Norman, D. (1983). Some observations on mental models. In D. Genter & A. Stevens (Eds.), *Mental models* (pp. 7-14). Hillsdale, NJ: Lawrence Erlbaum.

Novak, J.D. (1998). *Learning, creating, and using knowledge: Concept maps as facilitative tools in schools and corporations*. Mahway, NJ: Lawrence Erlbaum.

Oblinger, D., & Oblinger, J. L. (2005). Is it age or IT: First steps toward understanding the net generation. In D. G. Oblinger & J. L. Oblinger (Eds.), *Educating the net generation* (pp. 2.1- 2.20). Boulder, CO: Educause. Retrieved February 11, 2007 from http://www.educause.edu/educatingthenetgen/.

Online Computer Library Center (2002*). OCLC white paper on the information habits of collegestudents: How academic librarians can influence student's web-based information choices.* Retrieved February 28, 2007 from http://www.oclc.org/community/informationhabits.pdf.

Ophir, E., Nass, C., & Wagner, A. D. Cognitive control in media multitaskers. In *Proceedings of the National Academy of Sciences.* Retrieved September 14, 2009 from http://www.pnas.org/cgi/doi/10.1073/pnas0903620106

Othman, R. & Halim, N. S. (2004). Retrieval features for online databases: Common, unique and expected. *Online Information Review, 28* (3), 200-210. Retrieved December 8, 2007 from *ABI Inform* database.

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 (Doctoral dissertation, Florida State University, 1994). Retrieved April 6, 2007 from *Digital Dissertations* database.

Prenksy, M. (2001). Digital natives, digital immigrants, part II: Do they really think differently? *On the Horizon, 9* (6), 15-24. Retrieved April 6, 2007 from http://www.marcprensky.com/writing/.

Proctor, E. (2002). Spelling and searching the Internet: An overlooked problem. *Journal of Academic Librarianship, 28* (5), 297-305. Retrieved May 27, 2009 from *Academic Search Premier* database.

Rainie, L. (2006). *Life online: Teens and technology and the world to come*. Speech given at the 2006 Public Library Association Conference, Boston, MA. Retrieved February 9, 2007 from http://www.pewinternet.org/ppt/ Teens%20and%20technology.pdf.

Ramsey, H. R., & Grimes, J. D. (1983). Human factors in interactive computer dialog. In M. E. Williams (Ed.). *Annual Review of Information Science and Technology* (Vol. 18) (pp. 29-59). White Plains, NY: Knowledge Industry Publications.

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.*

Roberts, G. R. (2005). Technology and learning expectations of the net generation. In D. G. Oblinger & J. L. Oblinger (Eds.), *Educating the net generation* (pp. 3.1-3.7). Boulder, CO: Educause. Retrieved February 11, 2007 from http://www.educause.edu/educatingthenetgen/.

Rosenfeld, L. & Morville, P. (2002). *Information architecture for the World Wide Web* (2nd ed.). Sebastopol, CA: O'Reilly.

Salaway, G., Caruso, J. B., & Nelson, M.R. (2008). *The ECAR Study of Undergraduate Students and Information Technology, 2008* (Research Study, Vol. 8). Boulder, CO: EDUCAUSE Center for Applied Research. Retrieved February 28, 2009 from http://www.educause.edu/ecar.

Saracevic, T. (2007a). Relevance: A review of the literature and a framework for thinking on the notion in information science. Part II: Nature and manifestations of relevance. *Journal of the American Society for Information Science, 58* (13), 1915- 1933. Retrieved November 10, 2007 from *Wiley Interscience*.

Saracevic, T. (2007b). Relevance: A review of the literature and a framework for thinking on the notion in information science. Part III: Behavior and effects of relevance. *Journal of the American Society for Information Science, 58* (13), 2126-2144. Retrieved November 10, 2007 from *Wiley Interscience*.

Savage-Knepshield, P. A. (2001). Mental models: Issues in construction, congruency, and cognition (Doctoral dissertation, Rutgers University, 2001). Retrieved April 6, 2007 from *Digital Dissertations* database.

Saxon, S. (1997). Seventh grade students and electronic information retrieval system: An exploratory study of mental model formation, completeness and change (Doctoral dissertation, Florida State University, 1997). Retrieved November 4, 2007 from *Digital Dissertations* database.

Schroeder, R. (2007). Pointing users toward citation searching: Using Google Scholar and Web of Science. *Portal, 7* (2), 243-248. Retrieved November 18, 2007 from *Project Muse* database.

Schwartz, B. (2004). *The paradox of choice: why more is less*. New York: ECCO.

Seiden, P., Szymborski, K., & Norelli, B. (1997). Undergraduate students in the digital library: Information seeking behavior in a heterogeneous environment. Retrieved June 29, 2005 from the Association of College and Research Libraries (ACRL) website at http://ala.org/ala/acrlbucket/nashville1997pap/ seidenszymborski.htm.

Simon, H. A. (1956). Theories of decision-making in economics and behavioral science. *American Economic Review, 49* (3), 253-283. Retrieved October 7, 2009 from *JSTOR* database.

Simon, H. A. (1979). Rational decision-making in business organizations. *American Economic Review, 69* (4), 493-513. Retrieved October 7, 2009 from *JSTOR* database.

Slone, D. J. (2002). The influence of mental models and goals on search patterns during web interactions. *Journal of the American Society for Information Science and Technology, 53* (3), 1152-1169. Retrieved April 6, 2007 from *ABI Inform* database.

Spink, A. (2004). Multitasking information behavior and information task switching: An exploratory study. *Journal of Documentation, 60* (4), 336-351. Retrieved September 12, 2009 from *ABI Inform* database.

Spink, A., Wolfram, D., Jansen, B. J., & Saracevic, T. (2001). Searching the Web: The public and their queries. *Journal of the American Society for Information Science and Technology, 52* (3), 226-234. Retrieved November 26, 2007 from *ABI Inform* database.

Stein, J., Bright, A., George, C., Hurlbert, T., Linke, E., and St. Clair, G. (2006). In their own words: A preliminary report on the value of the internet and library in graduate student research. *Performance Measurement and Metrics, 7* (2), 107-115.

Stevens, A. & Gentner, D. (1983). Introduction. In D. Genter & A. Stevens (Eds.), *Mental models* (pp. 1-6). Hillsdale, NJ: Lawrence Erlbaum.

Stronge, A. J., Rogers, W. A., & Fisk, A. D. (2006). Web-based information search and retrieval: Effects of strategy use and age on search success. *Human Factors, 48* (3), 434-446.

Sullivan, D. (2001). WebTop search rage study. *The Search Engine Report*. Retrieved March 20, 2007 from http://searchenginewatch.com/showPage.html?page=2163451.

Sweeney, R. (2006). Millennial behaviors & demographics. Retrieved August 8, 2008 from http://library1.njit.edu/staff-folders/**sweeney**/Millennials/Article-Millennial-Behaviors.doc.

Thatcher, A., & Greyling M. (1998). Mental models of the Internet. *International Journal of Industrial Ergonomics, 22*, 299-305.

Thomas, C., & McDonald, R. H. (2005). Millennial net value(s): Disconnects between libraries and the information age mindset. *Florida State University D-Scholarship Repository*, Article #4. Retrieved February 24, 2007 from http://dscholarship.lib.fsu.edu/general/4.

Tolppanen, B. P. (1999). A survey of World Wide Web use by freshman English Students: Results and implications for bibliographic instruction. *Internet Reference Services Quarterly, 4* (4), 43-53.

University College London (UCL) CIBER. (2008). *Information behaviour of the researcher of the future: A CIBER briefing paper*. Retrieved February 28, 2009 from http://www.jisc.ac.uk/media/documents/programmes/reppres/ gg_final_keynote_11012008.pdf.

University of Baltimore Office of Institutional Research (2009). *Enrollment data for first-year students, Fall 2008*. Received from Office through personal email.

Vakkari, P. (2001). Changes in search tactics and relevance judgments when preparing a research proposal: A summary of the findings of a longitudinal study. *Information Retrieval, 4*, 295-310. Retrieved November 3, 2007 from *ABI Inform* database.

Valentine, B. (1993). Undergraduate research behavior: Using focus groups to generate theory. *Journal of Academic Librarianship, 19* (5), 300-304.

Vigil, P. J. (1988). *Online retrieval: Analysis and strategy*. New York: John Wiley & Sons.

Watson, J. S. (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.

Westbrook, L. (2006). Mental models: A theoretical overview and preliminary study. *Journal of Information Science, 32* (6), 563-579. Retrieved November 3, 2007 from *Wiley Interscience*.

Williams, P. (1999). The net generation: The experiences, attitudes and behaviour of children using the Internet for their own purposes. *Aslib Proceedings, 51* (9), 315-322. Retrieved November 3, 2007 from *ABI Inform* database.

Windham, C. (2005). The student's perspective. In D.G. Oblinger & J. L. Oblinger (Eds.), *Educating the net generation* (pp. 5.1- 5.16). Boulder, CO: Educause. Retrieved February 11, 2007 from http://www.educause.edu/educatingthenetgen/

Xie, H. (2004). Online IR system evaluation: Online databases versus Web search
    engines. *Online Information Review, 28* (3), 211-219. Retrieved October 22,
    2007 from *ABI Inform* database.

Young, R.M. (1983). Surrogates and mappings: Two kinds of conceptual models for
    interactive devices. In D. Genter & A. Stevens (Eds.), *Mental models* (pp. 35-52).
    Hillsdale, NJ: Lawrence Erlbaum.

Zhang, X., & Chignell, M. (2001). Assessment of the effects of user characteristics
    on mental models of information retrieval systems. *Journal of the American
     Society for Information Science and Technology, 52* (6), 445-459. Retrieved
    April 6, 2007 from *ABI Inform* database.

Zhang, Y. (2008). The influence of mental models on undergraduate students'
    searching behavior on the Web. *Information Processing and Management, 44* (3),
    1330 – 1345.

Appendix A

Millennials Research Screener

How many computers (desktop or laptop) do you have at home?
 [   ] none   – TERMINATE
 [   ] 1-2 – CONTINUE
 [   ] 3 or more – CONTINUE

Is anyone in your family a librarian or database developer?
[ ]   Yes   TERMINATE
[ ]   No   CONTINUE

How often do you use the Internet?
[   ] 2+/day   - CONTINUE
[   ] 1/day   - CONTINUE
[   ] 4+/wk   - CONTINUE
[   ] 1-3/wk   - CONTINUE
[   ] less than 1/week - CONTINUE
[   ] never – TERMINATE

How many hours do you spend on the Internet?

[   ] More than 2 hours/day   - CONTINUE
[   ] 1-2 hours/day   - CONTINUE
[   ] 4 or more hours/week   - CONTINUE
[   ] 1-3 hours/week   - CONTINUE
[   ] less than 1 hour/week - TERMINATE

How many of those are spent surfing the Web (excluding time emailing, playing video games, social networking, etc)?

[   ] More than 2 hours/day   - CONTINUE
[   ] 1-2 hours/day   - CONTINUE
[   ] 4 or more hours/week   - CONTINUE
[   ] 1-3 hours/week   - CONTINUE
[   ] less than 1 hour/week - TERMINATE

Have you ever participated in a class on library resources, such as the library catalog or online databases within the last year?
[   ] Yes - CONTINUE
[   ] No – CONTINUE

How often do you check your email?
[   ] 2+/day   - CONTINUE
[   ] 1/day   - CONTINUE
[   ] 4+/wk   - CONTINUE
[   ] 1-3/wk   - CONTINUE
[   ] less than 1/week – CONTINUE
[   ] n/a– CONTINUE

How often do you text-message people?
[   ] 2+/day   - CONTINUE
[   ] 1/day   - CONTINUE
[   ] 4+/wk   - CONTINUE
[   ] 1-3/wk   - CONTINUE
[   ] less than 1/week – CONTINUE
[   ] n/a– CONTINUE

Do you have a profile on MySpace, Facebook or another social networking site?
[   ] Yes – CONTINUE
[   ] No – CONTINUE

How much time do you spend on social networking sites?
[   ] More than 2 hours/day   - CONTINUE
[   ] 1-2 hours/day   - CONTINUE
[   ] 4 or more hours/week   - CONTINUE
[   ] 1-3 hours/week   - CONTINUE
[   ] less than 1 hour/week - TERMINATE

In what category does your age fall?
[   ] Under 18   - CONTINUE
[   ] 18 – 19   - CONTINUE
[   ] 20 – 22   - CONTINUE
[   ] 23 – 25   - CONTINUE
[   ] over 25 - TERMINATE

Have you participated in a marketing research focus group, an in-depth interview, or
a usability test in the past six months?

[   ] Yes - TERMINATE
[   ] No – CONTINUE

Appendix B

Participation Agreement

University of Baltimore School of Information Arts & Technologies
_____

THE STUDY:
The researcher is observing students under 25 conduct research in order to determine millennial students' understanding of information retrieval systems and processes.

PROCESS:
During this session, which will last approximately 90 - 120 minutes, the researcher will ask to research a topic for an assignment in IDIS 101. She will take notes and record your comments on videotape. She will ask also ask you a series of questions after you have completed your search.

Your participation in this study is entirely voluntary: You choose whether to participate. If you choose to participate in the study, you can stop your participation at any time. If you want to withdraw from the study, please tell the researcher that you want to withdraw.

CONFIDENTIALITY:
The researcher and her advisors will have access to the observations and notes made by the researcher as well as the videotape of this session. However, the researcher will not use your name when quoting your comments in her dissertation or any subsequent publications. Only the researcher will have access to the videotape and session-related documents that identify you by name.

COMPENSATION:
If you choose to participate, you will receive a gift certificate of your choice valued at $15.00.

IF YOU HAVE QUESTIONS OR CONCERNS:
You can ask questions about this research project now or at any time during the session or you may contact Lucy Holman at 410-836-4333 or lholman@ubalt.edu.

I have read the information above and consent to participate in this study as outlined.

_____
Participant's Signature   Date

Appendix C

Post Observation Interview

Participant: _____   Date: _____

What information were you trying to find in your search today?

What key words or search terms did you use to find that information?

What kinds of sites did you expect to find using those key words?

Are there other keywords you would use to find more information?

What is more important to you - finding as many articles/sites as you can on a topic or finding a few sources that closely match what you are looking for? Why?

How does a search engine know what you're looking for?

With as much detail as possible, explain how your search tool works. In other words, what does the system "DO" with your search terms?

If you were trying to explain this system to someone who had never seen it, how would you describe how it works? Is there another system you could use as an analogy? How would you describe their similarities and differences?

Draw a diagram to show how your key words relate to each other and materials you found.

How does the system select and present the list of results?

Appendix D

Student Search Script

Participant No:_____   Date: _____

The Morae Script uses the following format:
R = indicates researcher comments   P = participant response expected

What the Researcher Says and Does	*What the Respondent Says and Does*
Introduction R: Good (morning/afternoon) _____; thank you for making time in your schedule to participate today. I have asked you here today is to help me better understand your thought process when you search for information. I understand that you are enrolled in_____ and for one of your assignments you are to research _____. I would like to observe you as you search for material on your topic. I am not judging what you do - There is no right or wrong way to search. I just want to watch how you go about finding what you're looking for. Is that ok with you? P: Participant acknowledges R: Very good, if you will please, just sign my release form here. It says …….. P: Participant signs release form	
System Introduction and Calibration R: The system we will be using to record our session today is called Morae. Before we begin, we will need to set up a session in the software. Would it be OK if I used your first name to identify this session? P: Participant Agrees R: Very good, … We will need to set up the recording. Morae will be capturing each screen that you select and will be videotaping you while you work. As we go though this session, although I may ask you to talk about your thought processes, I would like you to keep your focus on the screen, rather than looking at me. I will also be writing notes about where you go and what you say. Does that sound all right? P: Participant responds R: OK, here we go… (start and begin recording…) Now we will begin the session.	

Familiarization R: For _____ your assignment is to _____. Generally when you need to find material for a class assignment where do you tend to go?	P: Participant responds
Think-Aloud Protocol R: As you search for information here today, I would like you to "think aloud." Tell me why you are doing what you are doing and what you anticipate the system will do. This helps me see what you're thinking when you take certain actions. It may seem awkward at first, but it will help me understand your thought processes. As we go along, I may remind you to continue to talk about what you are thinking. Do you think you can "think aloud?"	P: Participant responds
Getting Started R: Here is the assignment for your course.	P: Participant may read assignment.
R: What kind of information do you need for the assignment; what do you need to look for?	P: Participant responds.
R: Based on that assignment, where would you start and why?	P: Participant responds
R: What kind of sites do you expect to find?	P: Participant responds
R: OK, well let's go there and see what you find. Remember to talk out loud and walk me through your thought process.	

Actions	Comments

Actions	Comments

Conclusion

That should do it for the day. I appreciate your time completing them for us.

Thank you for allowing me to observe you as you research your topic. Your participation will help me better understand how students today look for information. Hopefully this research will impact the way librarians and database producers design their information to make it easier for students to use.

Appendix E

**Selected Transcripts – Responses to Mental Models questions
(researcher questions in bold italics)**

6. ***How does a search engine know what you're looking for?***

    P1: From my knowledge I perceive that Google knows through tags. I don't really know what a tag is; I just know that it researches through things that have his name in it or a person that is famous. It's like you can Google my name or your name. So someone has tagged that name into the databases of information in computer land and it pops up on Google or anywhere.

    P2: I know that Google has its own archive of previous searches or generic searches, and I gave them some generic searches today. I'm assuming that they give me the top hits; I don't think that they just randomly mix up the websites every time. If they keep on searching every day maybe it might change a little bit, but it would be the same exact websites that would appear on the first 10 pages or so. 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 public usually is looking for in that topic range. So in a way I'm siding with society and what they're generally looking for and I have to try to find a way to surpass that and go into more detail into what I'm looking for.

    P3: Well I actually found this out in my one class the other day. They do it by popularity, how often site visited, that's why I like going to Yahoo answers, and sometimes that's helpful if you're looking for a specific question because sometimes people will give you other links you're probably not going to get that aspect.

    P4: Um, it...basically…after you type in something to search for, it basically…it sends out a spider that picks up all the links are connected to it from the key word that you typed in, and then afterwards it just tries to match up what might be the purpose of the article, so if it's something more political then it's going to show articles that deal with elections and stuff like that, but if it's something more….something fun or anything like that, it's not going to show really serious articles and stuff and then afterwards it shows all of top sites that have been clicked on and have links to it.

    P5: Well…we learned that Google picks up on basically your key words.
So you can't be too vague, and you can't be too narrow.
You have to pretty much get it in between because it just picks up on certain ones.

What do you mean it picks up on certain ones?
Well like if I type in Republican and democratic on taxes it would probably give me stuff on Republican, stuff on Democratic and maybe some on taxes. It wouldn't give me something that had both all together, all three of them in one area.

P6: By what you type in…Specifically whatever you type in that's what it looks for. If you type in unknown culture, then they'll all search out unknown cultures; if you don't make it specific to what you need, like unknown human cultures, they won't know. Specifically whatever you type in that is what they search.

P7: I'm guessing it uses web spiders to search throughout the databases, and when the spiders come across that word, it attaches itself and brings it back to the search engine site.

P8: We actually talked about that in class the other day…it tries to pull what it thinks is most relevant to what you type in to try and….it puts that at the top but it also feeds off of who goes to those websites and like how many people go to those websites the most, and then it rates them. But each website, each search engine rates them differently, so that's why you don't get the same web pages on Google as you would get on Yahoo….

P9: Ummm….I think it just matches the words that I put in to any of the words that are in any part of the articles, which isn't always good, but that's what it does….or to the website….

So it matches the words to any words in the site?
Yeah….

Is there something you can do to limit that or narrow that?
I think you can ….like, if I wanted this whole phrase, I could put it in quotes or if I wanted um…..say,…or I can do an advanced search also ….

Do you ever do that?
Only when I 'm not getting anything from the regular search.

P10: Google will search the source code of websites for what you're entering into the search bar, and it will use that to compose a list starting from most recent …. most commonly accessed websites going down…not of all websites, though, only search certain websites for their source code.

P11: I don't know.   I know how Google does it – I guess it has to do with links and how many people have links to it – but that may be wrong. That's what I thought it was.

P12: From what I know, Google just finds every single word that you look for. They don't have to be in order. They don't even need to be related as long as it sees those words, and it pulls it up for you. The database is more restricted, I'm guessing, because if you put two words next to each other, it's going to search for those two words next to each other, it's not…so it will be more related, more specific.

P13: I wanted to say that since you're typing in whatever term that you're looking for …I want to say that it pulls different parts out of it…maybe it runs it through the system where it can figure out, ok….maybe I can find…you can use this…the more specific you can get it can fish out the ones that you don't need and that you do need. So I think MSN would probably be one of the best search ones for me, search engines, and images as well – it has good images.

P14: It searches pages. It's like an index of pages that has all the words, and if the words I type in match up with the word on the page it brings it up.

P15: From the word I put in, the actual keywords that I put in and then sometimes… in the first time I had did it, I put the quotations on each side because he said when you do that it brings up that together, so if I had education policies in quotations, it would bring up all the websites where these words were listed together. That way I can find what I need through that keyword.

P16: The words, the keywords that I use.

P17: I know that there are far better search engines in the databases that I can go to, but because I'm learning about the rest of those databases now as I didn't know before I'm still prone to use Google because that's what I grew up on using. What Google does is take words that you're looking for, it doesn't really….it's really vague, so you get a lot of hits. So that's why you've got to be very specific when you're using Google, and sometimes you even have to think as if you were writing the article, what the article would say, so that a line could pop up that you are looking for and you get that information.

You would have to be specific when using Google because Google can give you ANYTHING that involved in the littlest words that you use.

So what does specific mean to you, how would you get more specific?
I would get more specific by trying to possibly match what I'm looking for or trying to guess what would be written and what I'm looking for, that's how you've got to use Google.

P18: It uses this up here (pointing to keywords). This is one of the better ones from what I've seen and from what Professor Salvetti recommended I guess. It's better because it offers this over here, the and, or, or not, all of this right here – it accepts Boolean searches so when you put that in (pointing to problem or situation) you put your synonyms down here like problem or situation. It's a faster search, and it gets journals, newspapers and magazines, it gets you the better sources to quote instead of…the only other one you might want, the only other one might be a book, but that's usually if you have a lot of time because you've got to, I mean, books can give you a lot, a lot of solid support but it takes too long usually to find the quote or the part that you want to paraphrase unless you're writing a book yourself sometimes.

P19: Usually, when I use Google, I'll try to make it specific; I won't try to make it broad, or type in something broad because it will give you so many different things that you're not looking for like, what was I looking up? I was looking up something….I think it was political….something political, and we just typed Obama, and it would pop up all this stuff on blogs, who likes Obama and speeches from Obama, and you have go through and you have to take off….you actually have to type in what you're looking for instead of just typing in his name.

   How would you make it more specific?
   Like when I was looking up the speech he had made something with a pig and lipstick. I was looking up that, I was trying to use that on one of my papers and I thought "how in the world am I going to find this?" and I just went straight to YouTube and typed in pig wearing lipstick Obama speech and it popped up. Or if you want to look up the debates, you don't just type presidential debates cause then it will show up more than just one…usually it will pop up the one that you're looking for, but you don't want to use that (so what else would you type) You'll put in a dash and then Obama vs. McCain or Sarah Palin vs. Biden.

P20: Something about a spider, spider sends information to memory thing and then the memory sends it to Google.

P21: It's basically based on what I write, like if I write Democratic Party – Google would give me search results, like history, policies and all this stuff. But if I write Democratic party and air pollution, for example, it gives me sometimes links to newspapers and magazines on air pollution, and it sometimes gives me results on totally different things from what I'm looking for –like it gives me issues on air pollution on other countries other than the US, so I normally specify U.S. Democratic party. But in the Academic Search Premier it gives me results from news, magazines and journals. I don't think it gives results from the webpage.

7. *With as much detail as possible, explain how your search tool works. In other words, what does the system "DO" with your search terms?*

P1: [Laughs] Google with my term Leonardo da Vinci I believe searches through online information like information in the world that I can't see on my own. It goes through and it sorts out information with his name in it and that's basically it. It gives me…it doesn't do anything for me but it gives me resources. Google gives me resources to find information. So Google is just like the home base to branching out other places with Leonardo da Vinci in it.

P2: Kind of back to my old answer…..They search in their archives, I'm not sure where, but that's how I think of it. . .they search in their archives for as much information as they can provide on that topic. Sometimes, if the topic is rarely searched, they'll get maybe the word you typed in and make a quote or highlight the terms of another source that might not even have anything to do with your topic and it will appear in the top most searched listings, how can I go more in depth with that….[laughs] I think that's it.

P3: They look for the websites that have mentioned the most word and that are viewed the most. It goes according to that. And if you type in the person's name it would bring up both the first name and the last name separately and bring up those. Image searches aren't so good with looking for a particular artist -- that's why I prefer to try to find their actual websites using their name versus trying to find their images because the images will just bring up things like "Sarah" versus Sarah whoever you're looking for.

P4: Um, it...basically…after you type in something to search for, it basically…it sends out a spider that picks up all the links are connected to it from the key word that you typed in, and then afterwards it just tries to match up what might be the purpose of the article, so if it's something more political then it's going to show articles that deal with elections and stuff like that, but if it's something more….something fun or anything like that, it's not going to show really serious articles and stuff and then afterwards it shows all of top sites that have been clicked on and have links to it.

P6: When you put your search phrase in, Google sends little spiders out to find specific links or specific topics or blogs or whatever you get the information, little spiders to pull that word out. If you type in apple pie recipes, Google looks for apple, then pie, then recipes. If they have all of them – apple pie recipe - it will be their #1. They're based on what's more general to your site, from what is more to what not so more, but is dealing around the same range of what you're looking for.

P7: Didn't I just answer that? It's looking for key words, it takes out the meaningless words like does and how. It basically looks for 3 strikes, deter and crime.

Does it look at them all the same?
Well I guess unless you put in AND and OR and NOT, I guess it does.

P8: What Google does is it uses a spider, which is a type of web program where it actually pulls up each site and searches for those words within the site, and then it puts it down into an index, so you're not seeing the entire page, it just gives you the links and key words in it and then it goes out and it rates them.

P9: I think it just takes it and goes through the archives of….it doesn't have all the information but it has a lot of information, so it goes through the archives and matches it to websites that even mention it, whether it's together or not, and brings them up, and I think it shows the most relevant first.

What do you mean by most relevant?

Like with… the whole phrase in it, and then like towards the end it gets just some of the words split up in it.

Anything else about how you search?

Umm….all I can say is that sometimes it seems like you're typing in the right thing, and you could get four results, but you could just change one of the words like "the" or "on" or "in" and can get 20 more results.

P10: Google will use my search terms in order to check for links in the web code sourcing of a website to see if it holds relevant information to what I'm asking, and it will start with more recently accessed sites and move down to lesser accessed ones….It does not mean that it's looking up scholarly sources or accurate sources, it's instead looking to see what just holds at least mentioning of the topic and which is most accessed of that [topic].

P11: Looking through information it already has, it's hitting the keywords – I don't know how …I guess it's already done search, accessing a search that's done on the keywords.

P12: It just pulls any article or any information that's on the web with the words you typed in…unless, I believe, you put them in quotations, then it searches for that certain phrase. But if you don't, it pretty much just gathers all the words and then brings you random sites that might not be related or might be so you have to watch out.

P13: Let's see…..MSN simply gives you more information, like the background about it and it gives you a lot of pages, well, according maybe to

what you're looking for, since I was looking for school psychology, it highlighted the bold, in bold, like that actual word, school psychology within different articles. It gives you a quick synopsis of whatever the article was talking about, and it highlighted school psychology. That's one thing that got my attention with that, and it basically gives you a quick introduction before you click on it so you can be able to tell what a reliable source is and what's not a reliable source.

P14: I type in my terms, it has a lot of pages cached – it has all the text from those pages. When I search it matches what I type in with words on those pages and then it gives me those links.

P15: Well I guess when you type it in, it'll just bring up a list of everything that's in connection to what you said, and unless you use some of the tools, which are the quotations and the and/or and this little line thingy I don't know what that is, it can narrow it down because it can see what you're talking about because it already know those codes. But with Academic Search Premier, I think that works a little bit better because you can say if it's an academic journal and then it has a list of those sources. So I think that works a little bit better because it gives you everything that you need in the paper from the actual article to what's going to be in a citation.

P16: I guess it has, I want to say an arsenal, but it has like a bunch of websites that it has, and from those websites they try to find what they think I'm looking for based on the key words that I put in….whatever key words I put in, they try to find as many sites that match those keywords or have those keywords in them or they jump out.

P17: What Google does is, Google finds each hit, each link with…like if you have a word that's not …if you have a tag instead of a hierarchy, like a hierarchy would mean dog or something, but if you have Shih tzu or something like that – what it would do is find each link that has Shih tzu in it versus if you had a broad keyword it would probably give you a link that was specifically about dogs, meaning if you typed in dogs, it wouldn't give you a link about animals. If you type something as broad as dogs it would give you a link about dogs, but if you typed in Shih tzu, it would be prone to give you a link that was about anything but it has the word shiatsu in it. Somehow that's how it works.

P18: Well, you put it in like this with the asterisk [econom*] - it looks up any word, all the forms of the word instead of just economy, economics, economies, it will look for all the branches that come from that word. When you put in or [demonstrates] you can set up another word or synonym that lets you find…that finds the other parts that it would miss by just searching for this [points to econom*].

Whereas JSTOR where you just search for the words that you looking for it might skip over just because the letters are off like sometimes it's just a matter of it being a plural or not and you can't really be sure if it's getting everything because again sometimes they just don't explain how it searches, so you're forced to use another database sometimes just because of that. That and the fact that you usually have different sources.

If it doesn't seem to work, you just go to another database?

Even if it does work, you still have to move on to another one because there's…instead of going all the way to …this search [points to JSTOR] only came up with a couple of sources, so it wasn't like a lot, but if you do out on Google you get like thousands and thousands of searches. You're not going to go through all of them, sometimes you're just going to check like the top 4 or 5 pages and hope those are actually close and not the sponsored ones. So you usually try to …you switch anyway just because they offer different topic choices so those first 5 or 10 pages turn out search results pages if actually going in depth 'cause if you're just looking up something casual you might just go off the first page maybe the second.

P19: They will pop up with YouTube where it actually shows the whole debate. They might pop up some blogs where people have written out what they like about debates or what they didn't like or parts of the debates they didn't understand or wanted to ask questions about. They might pop up news articles maybe that have said whatever issues.

How does it know which things?

By the words, I suppose – like with the Pharrell example, pick anything that has to do with words Pharrell and just pop it up to let you pick and choose what you want.

P20: Well if it were in Google, if I put it in parentheses (did quotes in air), then it would know not to separate those words...then the spider and all that stuff. If I don't put it in parentheses it can break it all up, and then I get a whole other group of information.

P21: It takes the most basic words from my search, and tries to brings as many searches as it can, and actually I don't know what it does, because I've never thought of this, I just type my words in. But Academic Search Premier – I'm kind of new to it. I think it's a…it's very…it's a straightforward search machine. If you type something that is not straightforward, it doesn't give you results. It needs something …you have to type…it has to be straight forward, meaning it gives you a straight answer. If you type Democratic Party, it gives you strictly the Democratic Party, it doesn't give you other searches. For example, Google gives you similar; it gives you a similar search

sometimes. If you type Democratic Party, it might bring views that are similar to the views of the Democratic Party, but might not necessarily be the Democratic Party. That's all I can say.

8. ***If you were trying to explain this system to someone who had never seen it, how would you describe how it works? Is there another system you could use as an analogy? How would you describe their similarities and differences?***

P1: How do I compare Google? Google is one of a kind. If I met someone who had no idea of Google is I would say it's …I would first say search engine. If I knew that they knew what a search engine was because that's basically what it is I would say it's a search engine or a place where you can get valuable information either in a written form, picture format and or other things. For example, Google can be a source of driving directions, it can be a source of phone book directions. Google has a lot of other things that aren't just about search. So it's pretty broad like Google maps and dictionaries and things like that.

P2: List other search websites such as Alta Vista, Yahoo, AOL has a search, Ask.com.

And what if someone had never seen a computer?
First I would explain what a computer is and then explain what technology has brought us today …it's pretty much a library in your hands – I wouldn't really explain how the sources are unreliable, because usually when you're talking with your friends you just say, "it's great, blah, blah, blah….I would say that pretty much anything you want to find out about, Google can provide the information for you, as long as it's in English. Any topic, any discussion, Google will be there to find it for you. That's how I would give a view of Google.

P3: Kind of a library in a sense, just more of a thinned out library because it won't always give you exactly what you need or as many possibilities in a sense, but at the same time it's like a library because you're not always going to find specifically what you need – like if you still just search for creative artists on Google and in the library you probably have difficulties finding exactly what you need because of the categories of how they set it up and the search ways of going about it.

P4: I don't think I could compare to something, but I might describe it as…it's basically just a search engine that….it doesn't have everything out there on the Internet, but it does show what are some of its top…from the index that it has, what are some of the top links that it has. So maybe you don't want to use just Google, but you want to use other search engines like

Yahoo or MSN or anything else. I can't think of anything right now, but that's basically how I would explain it.

P6: Water – water is accessible to everybody and it's not necessarily free, but generally the thought is that water is free. Google is accessible to anybody and it's free. So when you turn the faucet on there goes the water.
When you type in Google, whatever you want, it comes on. If you think about it, water and Google play hand to hand. Google is world wide, water is world wide. Google is reusable; water is reusable in so many ways.

P7: I'd describe it as….basically you know I'd ask them I guess…
something you'd look for and you'd say like dogs well if you go to Google and type in dogs and press enter and Google will search throughout the Internet and fins what sites, articles and stuff on, for example, dogs.

P8: I don't really have any analogies… you could kind of say it's like a magazine, I guess….Because you start with your ideas of what you want to get so that would be like your cover page, your front of your magazine, and then when you type in what you want and it pulls it up, it's kind of like an index or different searches, like different websites that you can go to and then from there as you click on, you actually get that article, the actual information that you're looking for.

When you see the index, like a table of contents?
Yes

P9: I'd say it's like flipping through an encyclopedia, or a book full of articles and information, it takes *a lot* less time because pretty much you would find your words, the words that you were looking for and um….stop and read through the stuff, but you don't have to do that with a search engine.

Several times you searched within a search, why were you doing that?
To get more specific results….'cause maybe like the links that I saw before I clicked on the one that I was in didn't look very promising, but I clicked one just to check, the one that seemed like it the most relevant…by its little caption, but um….I guess the more specific you get in the search, hopefully the more specific the results will be.

P10 : It's like a very experienced librarian or bibliographer who has access to a set of files and their classification, and she could quickly go through those; she would hand them out instead on the most commonly viewed basis, not though on which is most scholarly, so you may ask her for, say, government information, and she would hand you booklets on which …the most commonly handed out booklets not the set that has the most information.

Did you learn that in class?
(Couldn't hear answer)

P11: Sort of like library catalog or just a filing cabinet – it holds lots of information, you just slide it open and you find it, so you just type it in and press search and you don't even have to find it – it just pops up.

P12: I might say it's an experienced librarian with info….with stuff that's available to much more stuff than the library has. It's an experienced person trying to find something in a library or a place that's bigger than a library so they just gather anything that might look like it's what you're looking for.

P13: Let's see….I would say since I use it faithfully, I would think of it as…and this is for someone who's never seen it and they're trying to get to it or something….let's see…I would probably ask the person what was the last thing they've seen, if it's a Martian it's space [laughs]; if it's a person on a deserted island I would ask them maybe do they know how to look up different things in an encyclopedia – do they know how to use other sources, like encyclopedia or a newspaper and then in some encyclopedias and newspaper they may have different topics, and it may have the definition of MSN and hopefully they could try to get it from there, I guess. That's a good question -- that had me stumped.

P14: I'd tell them to type in whatever you're looking for and it will bring up websites related to that.

P15: Oh….hmm…mumbles, I'm sure there's probably a lot; I just can't think of any. Yeah, I can't think of any.

P16: Ummm. it's like a library with a whole bunch of information inside of it, that, ummm, I guess, the way that a library has certain things and certain categories, like when we type our keywords in, it's like the categories that we…that it looks for, I guess.

P17: I want to say….yes, yes, yes yellow book – isn't that what all books are called? Or what do you call what you look up - (*telephone book*) – yeah, telephone book. Basically the links would be the businesses, and the information that you're looking for would be the telephone numbers. And the keywords would be the hierarchy that you're looking for, the keywords would be like the topic …the topic in the phone book.

So using the example of the yellow book I would type in pets and you would go to that section in the yellow book those pages about pets?

right, right

P18: It works like a library – it's just faster. And I guess you see lots of spaces, but like a library it stores information, but instead of having it hidden with call numbers and rows and letters and all, that it's just by keywords they put on a document or the title so usually that's enough to find at least what you vaguely looking for.

P19: It's like phone book ….if you compare it with a phonebook, just like when were looking up movie or you're looking for someone to fix floors, it will just pop up anything, everything that shows flooring or every movie theatre you could possibly think of. It can help you out in narrowing down whatever you're trying to look for.

P20: Um….Google… how did those little paper catalogs work in libraries? If I were looking for poverty could I go into the little p and look for poverty? It's kind of like that, just look for a keyword.

Have you ever seen one?
In elementary school we had to learn how to do it.

P21: I would say that Google is like … Ok, Academic Search Premier is like a book, a book on a specific topic, if you buy a book on air pollution, it doesn't give you something different from air pollution. It gives you things on air pollution. But Google would be like a documentary – it gives you information on various things based on one search, based on one topic.

9. *Draw a diagram to show how your key words relate to each other and materials you found.*

P1: There's Google and here's you and here are your friends. There's this bubble. And I type in Leonardo and I'll search and it will say 'here is this.' It's artwork …it's basic information. This is nothing. Google sometimes provides sites with nothing on them so I call it nothing. It provides other words …it will give you …takes his name, Leonardo, and this will become Leonardo wasted pictures, which will become this, which is like some guy who has some information on a photograph …For example I was looking up Leonardo da Vinci the other day and I can't find the website now but it was a guy who apparently had writing from Leonardo da Vinci because he wrote upside down and to the left and that was like from 1452 like I don't know….people who have other information who can sometimes lead to like words that I have never heard of. Sometimes it's simple, and sometimes it just becomes… I get lost. It's almost like YouTube. You can have one video that has nothing to do with the next video.

So it takes you on a path?
…it takes you on a path, on an entire journey, and that's how I find out.

It takes me to interesting stuff, so I learn new words to me. Like I didn't know what a fable was, and it may seem like a small word, but I don't, and it's probably a word I won't forget.

And the idea is that all these areas have some connection to Leonardo?
Yes all of them are connected.

P2: OK…let's see….so you have Republican and Socialist party….those go hand and hand in my paper. And then you have global warming, which is ALL the way over here. I'm looking for the views…views, contrasts, comparisons, views because I used it as a search, I'll circle…and then websites….Republican GOP.gov….I found 2 websites on global warming, which are related, which I found off of Google and I searched with a .pdf…did I find anything? No…..and then socialist….nothing. That's going to be hard for me in the future. Gop.gov I found, let's see what else. I used email to find information….what else did I do? And I unfortunately I did not get to connect this to this. But my email was in relation to global warming….I didn't find any website that said issues global warming issues Republican Party.

P3: OK, so it's just based off the keywords. I'm just going to do it like a little spider web or a branching tree thing. I want you to know that in case it looks like a big old mess you can try to make some sense of it. From there I guess that would be the start of the search and you can ….from there ….and this is just me as I did both of these. I don't know if it makes that much of a difference it's just a one-word difference …mixed art versus mixed media ….I believe that's mainly the ones I used because they're related I guess.

P4: That's what I type in….and I'm guessing….I really don't know…so basically, these are different articles, blogs, and they basically they just try to match up keywords that you might find in one of the articles and everything just comes up….I don't really understand…that's as best as I can explain it.

P8: This would be the topic…and then you can draw those together...and the materials I found…

P9: I think that Bob Barr is related and global warming is related…..

P10: Regulation….United…States….banking could be combined to all three of those…they can talk about united banking firms, regulations themselves or state banking…the regulations could go to United States regulation, not specifically bank itself, or bank regulations, along with that, they could talk about regulations on things beyond these topics, such as, say, regulation of certain economy, economy regulation…if it's not using 'and' it can just find common information, it can find regulations on certain businesses beyond banking…the United States, it could lead to democratic parties, just by

searching on United States it could find political parties…which could in turn lead to policies….

P11: So if I wrote 'economic downturn' the bottom would be real estate corruption. I kind of think of that separately.

P14: First I type in Libertarian and then that sent me to a page with Libertarian ideologies and positions. I type in Libertarian and environmentalism or energy policy….that would give me libertarian opinions on environment and energy policies. Libertarian/democrat.
That was basically what I did….This is bigger (Libertarian ideologies and positions) because it's not just the environment.

P15: So basically, political candidates were kind of what we started off in 2008 elections, so I know mine…well, I could put Libertarian and democratic because those are my two political parties and the topic is education. So I'm going to have to talk about what they are saying about education and schools and then I narrow that down to funding, which coincides with their plans for funding and I have to include there political philosophy. So this is narrowing from education to funding or plans.

P16: I don't know how I would draw a diagram….I don't know if this is what you're looking for….Is that what you were looking for? Well, the…basically the topic was the environment, and then I broadened [narrowed?] it down to the problem, the problem is oil, the problem of oil effects the oil costs, the oil costs effects the consumers, so we need a solution. The broad solution for both of the parties was energy independence and alternatives to oil and then their policies.

P17: Keyword box….and I type in EBDI project I'm looking for how EBDI project relates to the concept about businesses being trapped…this is my search this is step 1 , step 2 would be clicking on search and step 3 would be the link….then I click on the article…this is the article…EBDI project is vision about relocating families so on…..I would think I can't use this one…it doesn't relate to the topic and this is the topic…so I would get another search downfall of EBDI split screen and search 3 http…..click….the downfall which would be in the link….the downfall of EBDI was due to great vision or investment plans ...what I'm looking for and this relates…..

And how are these two sets different?
Because this one is more vague…I mean this one is broad and vague
And this one is more – I don't want to use the word more specific – it's narrower.

P18: That's one I should have used – platform
These two might connect like this. This might just search separately though.

Sometimes so you can try to get an understanding because the minute you add Democrat to this whole situation you open up a lot of those articles about politics. I don't know how demonstrate that but…any time you look up too many words with Libertarian you get nothing

So the more you put in the less you get?
The narrower you get …especially when you include this [libertarian??] nobody really, well I won't say nobody, but it's significantly less likely you find anything. That's another thing about your keywords, like if I was explaining to that person, the more specific or the more like niche the word you use like. the more, the less often it is used the less likely you'll find what you're looking for if you're looking for something broad, or the more likely you'll find exactly what you're looking for at the same time. It's like a tradeoff – you either go the broad way and get a lot or you look up the technical term to find, to pinpoint it. It's usually a balancing act between that.. .That's the thing usually especially like when it comes to economic or fiscal, this [fiscal?] you can find more technical like papers on, this is the one you're going to get results, especially now because I've been using this [fiscal?] but I don't get anything usually Libertarian and philosophy I don't tend to get anything ... philosophy gets some things, but theory unless you're using it with political, like political theory doesn't really pick up much on its own but it might get you sometimes exactly what you want.

You'd use these clusters to get at the same thing?
Yes.

And every now and then it will try to connect libertarians and democrats, which is good for me in comparing them but it doesn't really go into a lot of detail before it tries to lump Libertarian into Democrats so that kind of messes with….Especially if you use both of them at once that why you should make sure to keep them separate.

P19: I put down, actually I put down Obama ….trails off…..and it will pop up something like another one of his speeches…..I basically can't remember what it came up as…I forgot where it was…..I just say…..I don't remember that….

P20: Politics and taxes are linked especially in the upcoming election. Poverty is directly related to rising in taxes. Therefore, all three topics are essentially intertwined.

P21: I put air pollution, contamination…they are more narrower [sic] so they give me better results on Google, but the more broader [sic] topics like the Democratic and the Libertarian Parties, the oil companies, the environment as a whole, and oil companies as a whole, they are more broader [sic]so they give me better results in Academic Search Premier.

10. How does the system select and present the list of results?

P1: Just website by website just like a list, a long list, easier than going to the library, more complex than reading the book. In Google all of it just comes up I don't have to find the call number and go into the library and search through all that, but it's just as hard as reading the book. Just as time consuming because if you want the good information you depending on what the essay is about if this.

P2: It picks the top ten, the top searches….list in top 10 and so forth, based on the public's general searches.

P3: Whichever has in a sense more hits, views, and the keywords that whatever word you typed in as a keyword if they have a certain thing that indicates that that word or topic is presented so many times. So the more times it's presented it will affect if it's first or not.

P4: It finds the top links that are connected to a certain website, and from the number of links that are in…that that website has, it sorts them out to which one is the top one and then so on and so on. It finds the top links that are connected to a certain website, and from the number of links that are in…that that website has, it sorts them out to which one is the top one and then so on and so on.

P6: Ordering from the highest source – source that is most related to what you are looking for and then it goes down, the less….the more it's driving away from your topic but it still has a source of your topic. If I type in apple pie recipes, the first thing will be like applepie.com or something and you'll have exact recipes – exactly what you need for apple pie, and the second will be maybe like a blogger typing about their experience of baking an apple pie, and then the third source will be like a definition of an apple pie or pictures of apple pies or something like that.

P7: It uses the subject line I think or the subjects what we use as like, oh, what is it called, I know that Wikipedia does it where it just adds a bunch of links and where Google will search on dog, and the subject line will be dog, dog, dog with a bunch of links that's how Google finds it.

P8: What it does …. it basically goes on the number of hits that that page would generally…the number of hits the page generally gets in a day. The example that I've heard the most is the miserable failure, where the kid decided that he wanted for every time someone typed in miserable failure to Google that they would get stuff on President Bush, and then he just worked with his friends to get more and more hits on the website and it did become number 1 on Google. [Google bombing]

Why search within site?

I was on the Democratic website, actually I figured that since I was in the site, and it was the political party, if I typed in an extra word it would…it would be guaranteeing me that there would be a connection to that party, but not necessarily relevant, but it would still hopefully get me somewhere close to my topic.

P9: I would say probably like ….either it starts with the ones that have the most out of your words together, maybe stated the most in the articles or the websites and it goes…it dwindles down to the least amount used, like maybe one of the terms will be used in the last one.

P10: (Already answered)

Anything else you want to say?

Normally, if I was in hurry I would use much less sites [sic] to search… chances are I would just check maybe two big sites, which in this case would be Google and the UB site, and when one's looking for sources, they tend not to switch constantly between.

What does AND do?

It's making me look for…it makes sure that an article will use the words banking, regulation, United and States, so I will not just get an article that will say the banks of Scotland or United Omaha or something; it will make sure the article specifically mentions what I am looking for …not exactly what I'm looking for, but closer to what I'm looking for than if I did not include the AND.

P11: I guess the most people that have, not just visited but...you can have people visit it, but Google doesn't know they're there they need to make a link to it, so I guess the most links made to something with the keyword in it is the top one or puts it in front of another one.

I don't usually…there's no real reason I like Alta Vista. I do it just for research stuff but I usually use Google, I think Yahoo uses Google. I was using Alta Vista, but I use Google for convenience at home, but when I do a research thing I think Alta Vista, but there's no actual logic behind that.

P12: It just presents anything that has the words you put up on the search bar; I mean, it doesn't exclude unless you put it in quotations, which will look for that exact phrase, but other than that, it really doesn't separate anything...I believe it orders it from the file that has most of the words to the ones that may have only 1 or 2 or 1 or may not exactly be in that order.

P13: It basically presents it in a list form and gives like a quick introduction about some of them and maybe what you're actually searching for are in bold, like it may stand out to you more and that's what gets your attention cause it's bold letters.

I guess it would be ordered from maybe the dates maybe like when the information was found or put on the actual web itself and how useful it is because at the bottom of some of them like the actual search, MSN, it says is this information useful and you click yes or no and if it's no it will try to re-search it again and make it a little bit better for you and you can always get more specific for what you're searching for as well. Since psychology is a broad search, you can always get down into it, like how I did with school psychologists.

How do you get a smaller number of results?

I would probably broaden, like maybe narrow the search down to maybe school psychologists in Maryland since I'm interested in school psychologists in Maryland, so that would narrow it down since it's a specific state or maybe adding more words like school psychologists' salary or something like that; you can always change it and narrow it down versus having a broad search which is school psychologists itself, like I can give specifics like in Maryland, salary, maybe in 2008 like school psychologists in 2008 or in the millennial [sic].

P14: I think it's how close the words are together…it will have to have one or both of them [terms], but the ones with both on them in it will be close to the top and the ones with both of them, like right next to each other like you typed it are even closer to the top.

Anything else about how you search?

To narrow it, I would add another term to it, like I had Libertarian and I added environmentalism and it narrowed the topic. To broaden it, I would take something off, like if I type in Libertarian and energy policies, or Libertarian energy policy vs. Democratic energy policy, I would not have gotten much, but if you simplify you get more.

P15: I think they do it by the top ones from the keywords that you put in, the top ones are the ones that are, would probably be the most valuable, and then it just goes along from there because as you go along you find some stuff that's like entertainment wise, and then things that would be on television and then maybe jokes along the topic, not necessarily facts about the topic, which is what you want to get.

What do you mean by most valuable?
Ones that the computer thinks would be most factual. But not all the time do you find those to begin with, though. Sometimes you find crappy ones.

What is the computer using?
I think they're using the keywords, so if they notice that your keywords are recognized the most in this particular article or the most in this particular website they'll probably bring it up like that.

Anything else about how you choose the words you use? How do you broaden or narrow the search?
Well, mainly since I don't like to search the computer, I try to get it over with as much as possible so I can come up with keywords like that, which normally gets me where I want to go, somehow I can get around it. But I prefer stuff like Academic Search Premier, where it gives you not only the actual article, but information about the article because that makes it faster, and the faster you get it off the computer, the less time you have to spend on the computer and actually work on the paper or work on the assignment.

P16: I guess ….. I guess it's based on what other people have searched and what other people….I don't know (laughs).

Is there any order?

It looks like in the headings, the …when umm….I can show you rather than explain…..Goes to Google…. Like when I type in energy independence (types in and gets results) like the title of the webpage I guess comes up and whichever one they think is….since I'm looking for energy independence, it's obviously going to give me the one that it thinks is the best, what I'm really looking for and I'm looking for energy independence so it wouldn't bring up anything that says, like anything that has to do with food, or anything, it's going to bring up what I'm looking for and what I searched for and then as the pages go down …well actually they're the same ones…but I thought as they got farther away they would go off the topic a little bit.

Generally how many pages deep do you go down in a list?
Probably 2-3, and then I use another search engine. I've been using….I forgot what the name of the website was, but the one that Professor Salvetti told us about …it's Poly something [PolyCola], where it splits the computer screen into 2 search engines, and then I just see which one both of them have and obviously that one is more effective. So I usually use the top 5 that come up on both of them.

P17: Based off words I type in, based off of the broadness of the word …..

Order of the list?
Links with most words that you type in the keyword search are at the top
From greatest words to least.

P18: From what I know about Google …Google will just throw up whatever
that has the most links so it's always going to be a news website or maybe a
popular blog site so I really don't want to use Google because in-depth articles
and journals aren't going to have a lot of links to them because usually they
are only read by like experts or in the company or something and those again
aren't very popular especially with news because they're trying to get
everybody. They don't want to throw something technical and throw off the
large majority of people.

What's it linking on?
The most links to the site or the article itself…so if a lot of other sites might
say, hey, look over at this… From what Google has stored of the Internet it
will usually go by what has the most links to it and what page so it will
probably pull up like….since a lot of people don't like Fox they do blogs that
try to link articles to Fox to kind of bash it or whatever but by linking to it
they make it more popular, so you're more likely to find this Fox article that
everybody is saying is so bad…it kind of gets in the way, but I guess what
Google is doing is trying to find the most efficient way to casually search. It's
not really meant to be for in-depth searches, because you only ever get exactly
what you want as far as complex articles if you already know the title that's
the only way to really get it.

P19: Basically…um….It just shows you anything with the words in it like
with Pharrell if it has anything with the name Pharrell in it pops in… coming
down as a list.

Any order to it?
Whatever has the most …I'll go back to Obama with the speech with Sarah
Palin – if you say Obama calls Palin a pig wearing lipstick then it will pop up
whatever has the most words that you used. It'll pop up like it would try to
make it exact…that's about it.

P20: The most hits go first (*what do you mean by most*) – like the most
viewed, most popular or what they think is relevant, but you can't always go
by that. Sometimes I just randomly click on page 10 and see; maybe it just
hasn't been viewed a lot.

P21: On this one…it's easier on Academic Search Premier because it gives
you all, everything. It gives you, you see a section that says all results , you
find all results – journals, you click, it gives you just journals, so you don't
need to look for journals in case you're trying to cite it, you use journals , and
it gives you magazines, news and the rest. You click on each of those, and

you get specific on those, so it's way more easier [sic], so if you want to cite, it's very good to use Academic Search, a very useful site.

Within journals, any order, how do they appear?
I don't know how they are listed, but I don't think they have a specific order, but if you click on any topic on journal, it gives you where the journal is from the specific place it's from, the publisher, the date, the page, all the specifics.
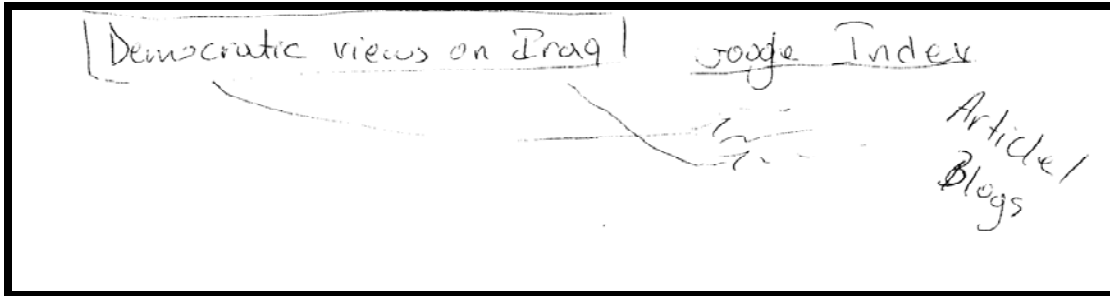
Appendix F

Concept Maps
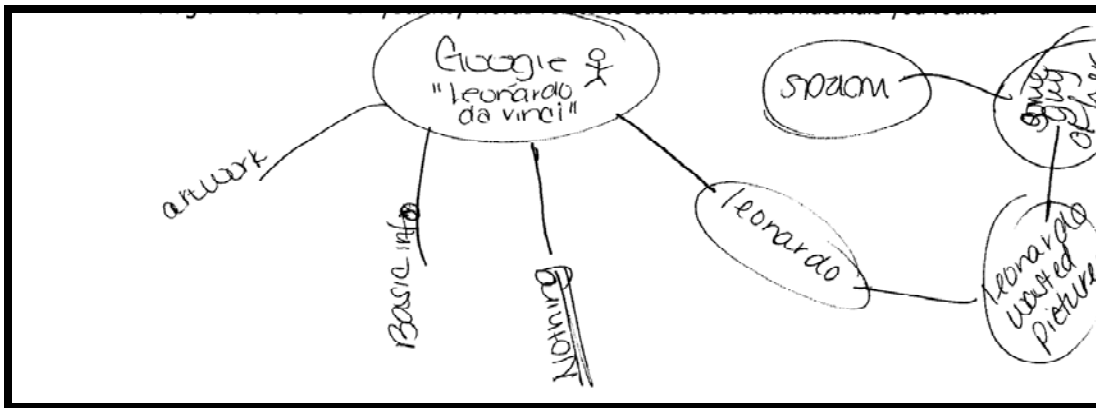
Participant 1 – Hierarchical view
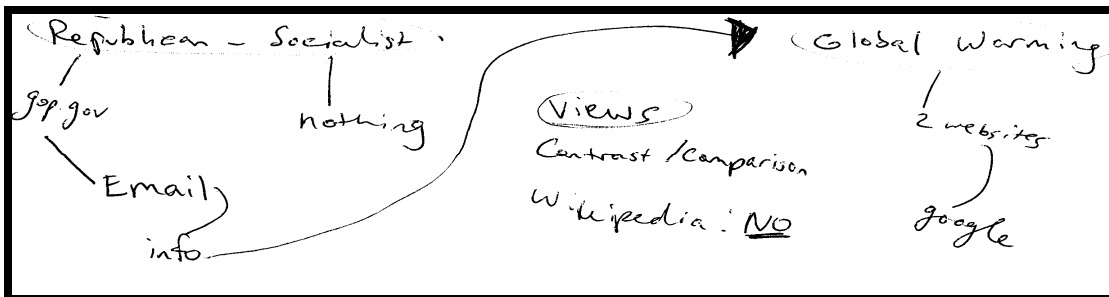
Participant 2 – Process view
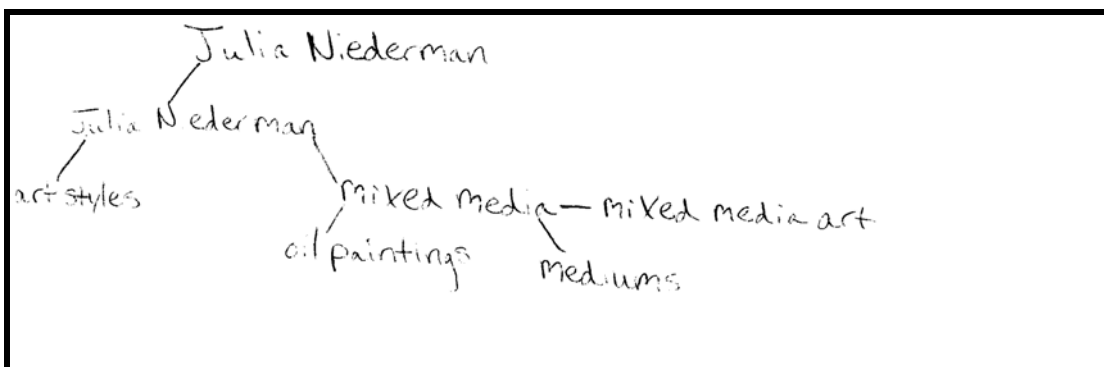
Participant 3 – Hierarchical view

Participant 4 – Process view
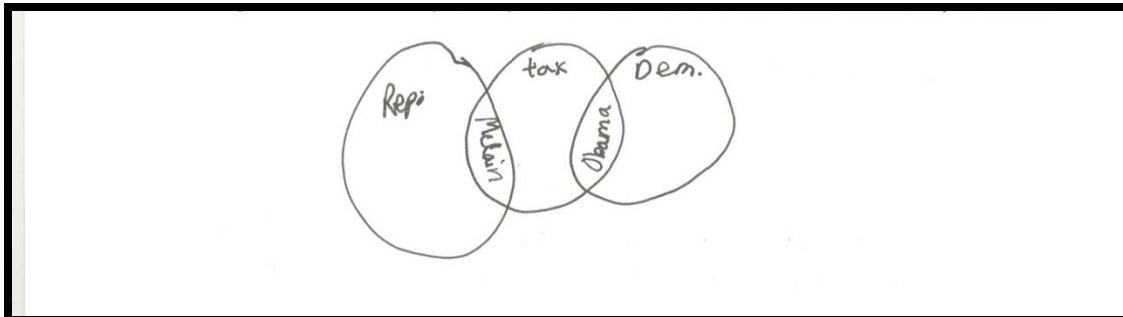
Participant 5 – Network view
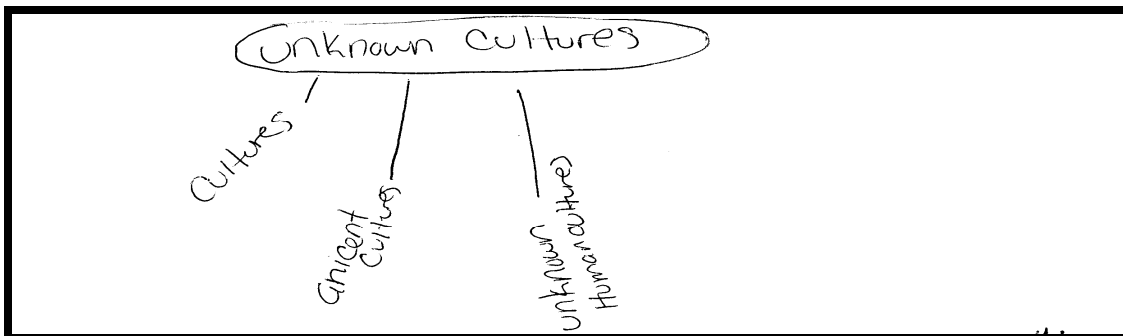
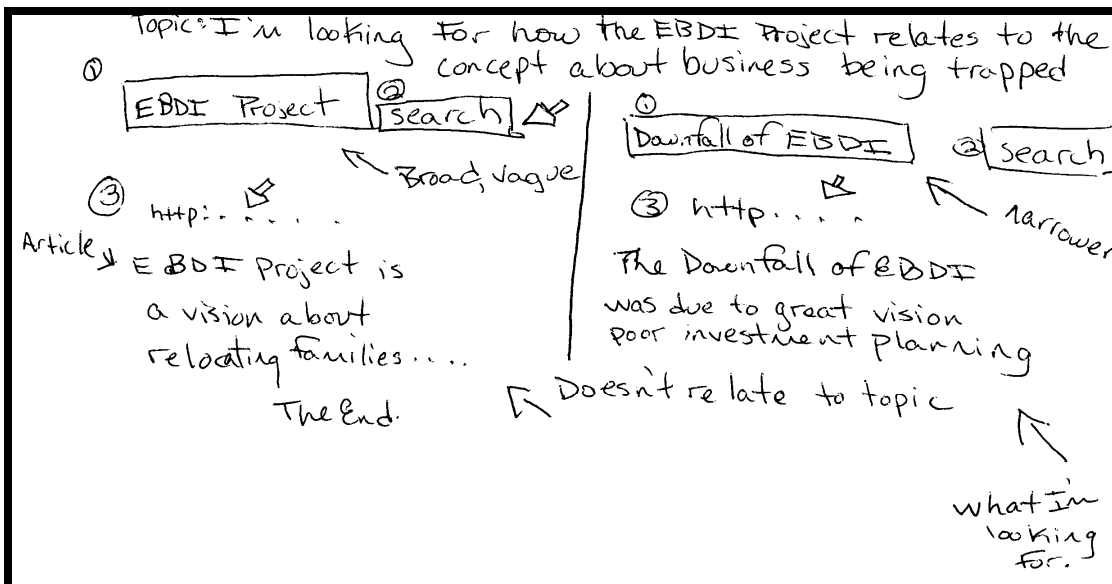Participant 6 – Hierarchical view
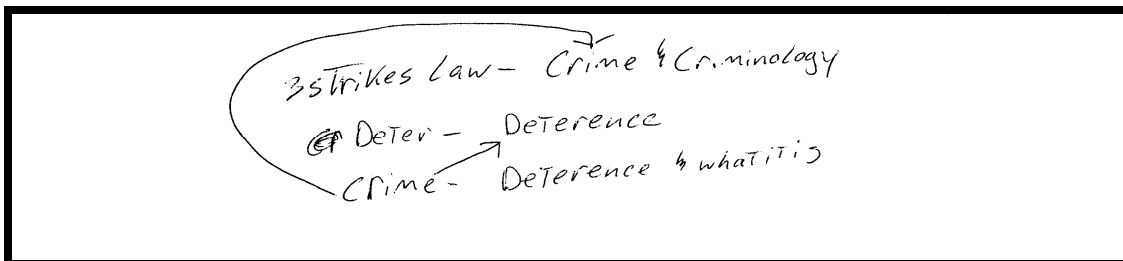
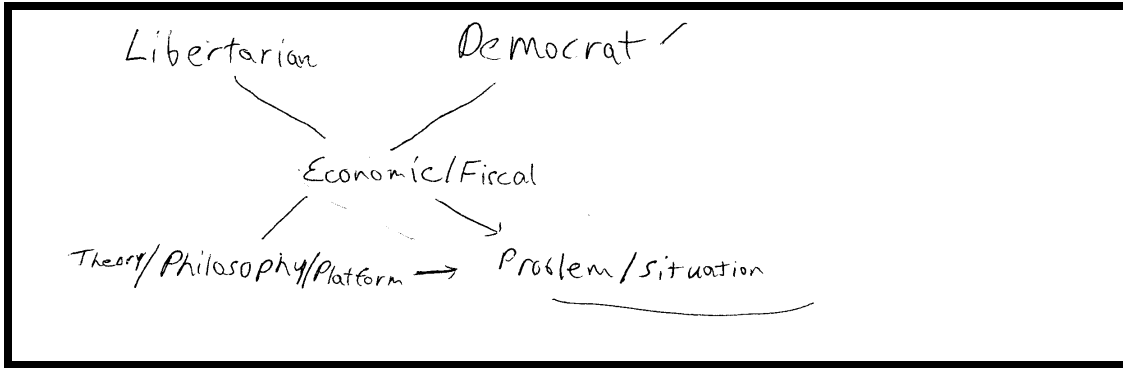Participant 7 – Network view
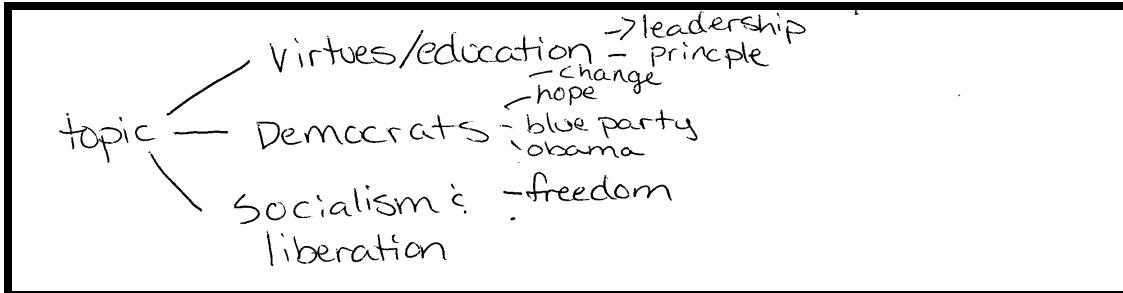
Participant 8 – Hierarchical view
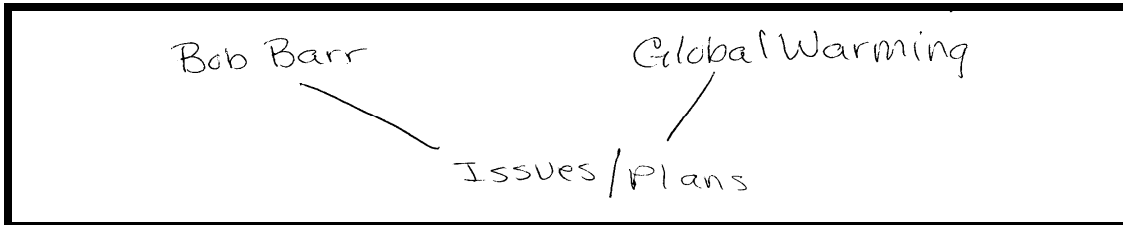
Participant 9 – Network view
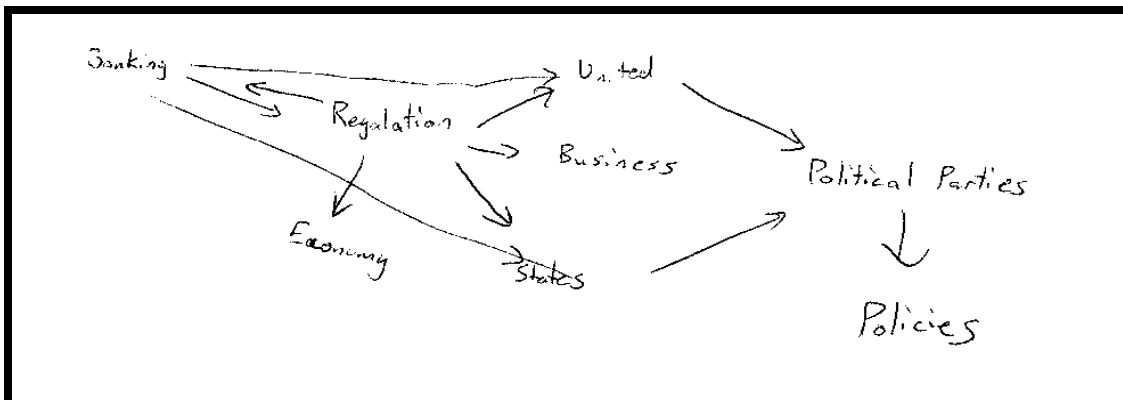
Participant 10 – Network view
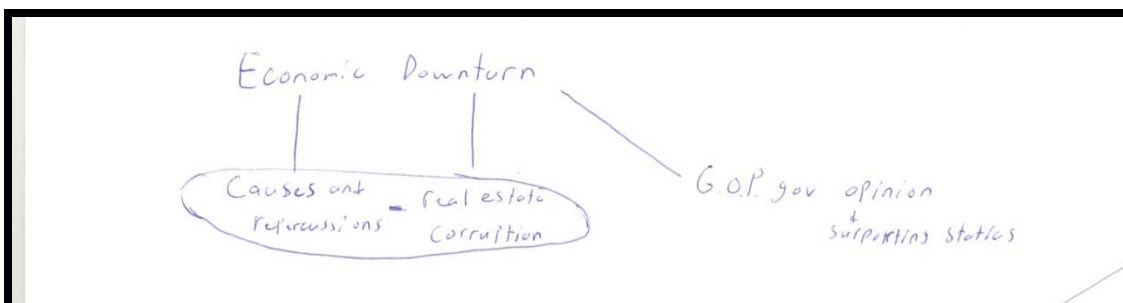
Participant 11 – Hierarchical view
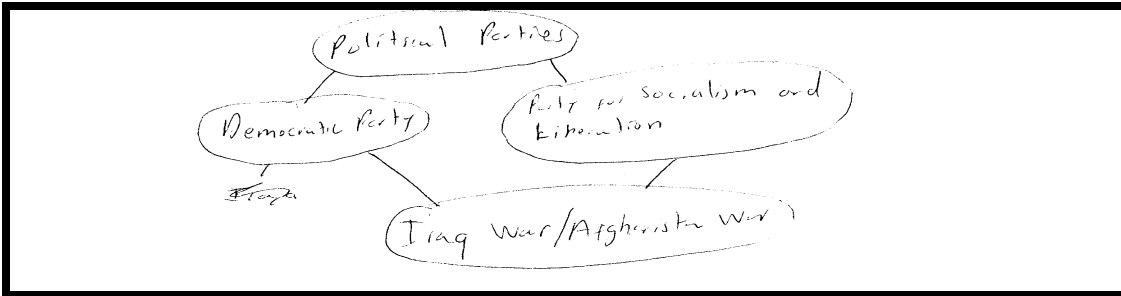
Participant 12 – Network view
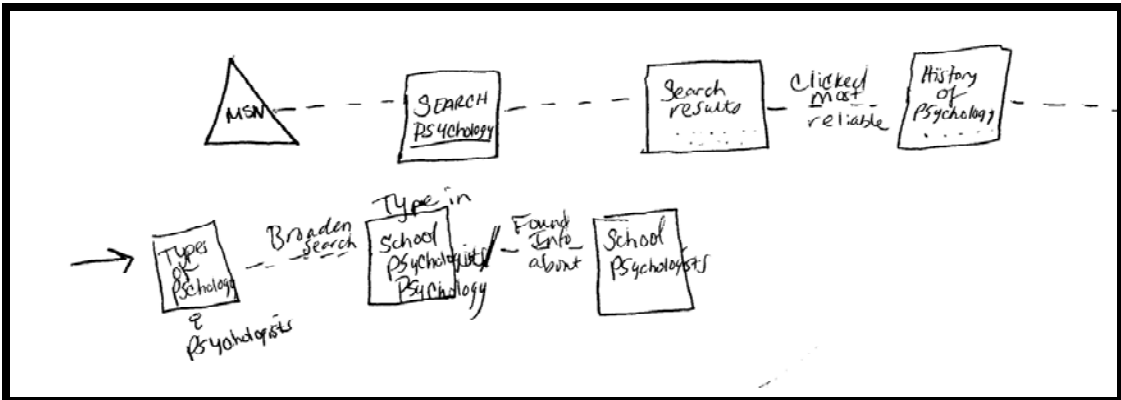
Participant 13 – Process view
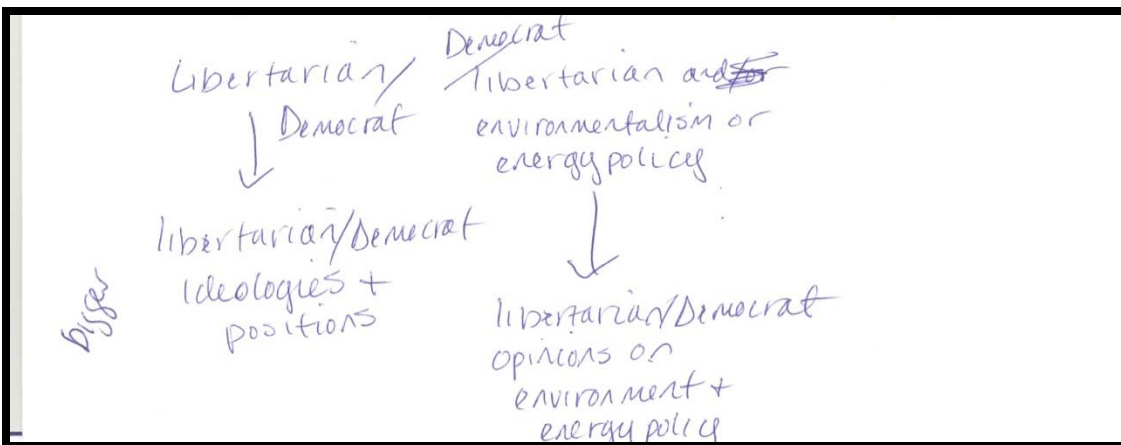
Participant 14 – Hierarchical view
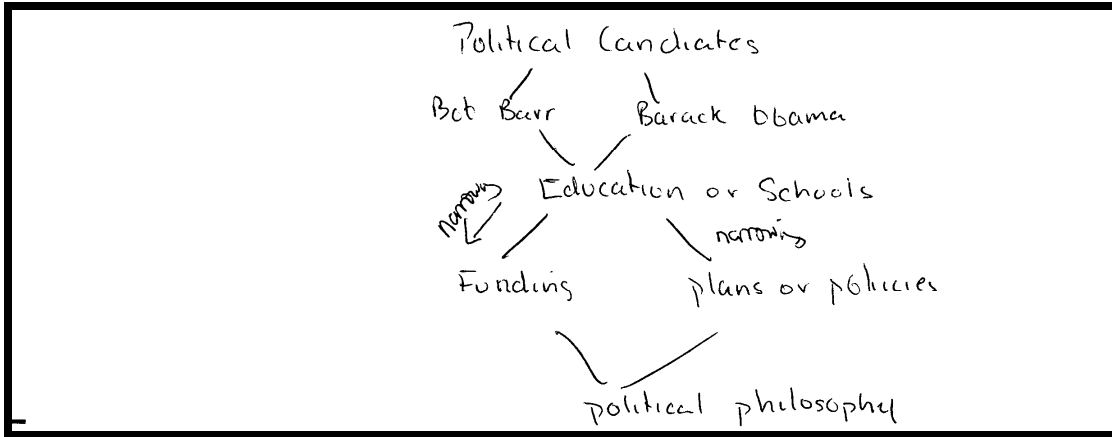
Participant 15 – Network view
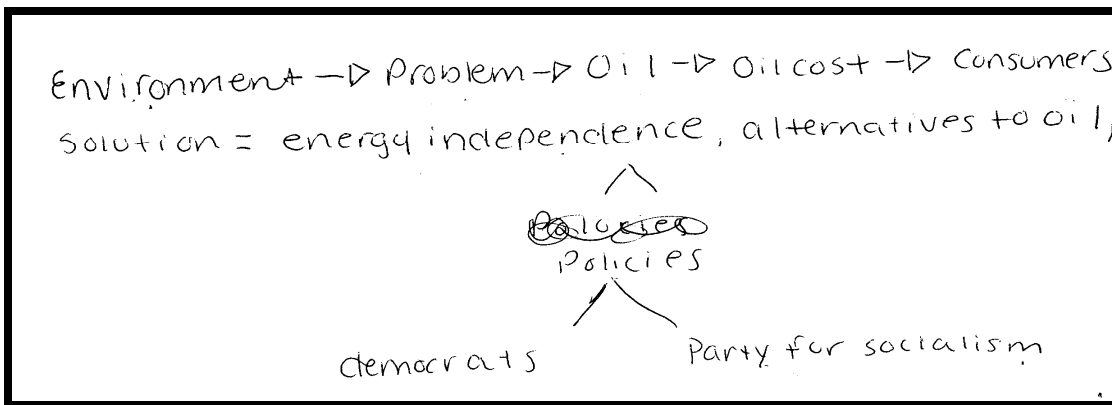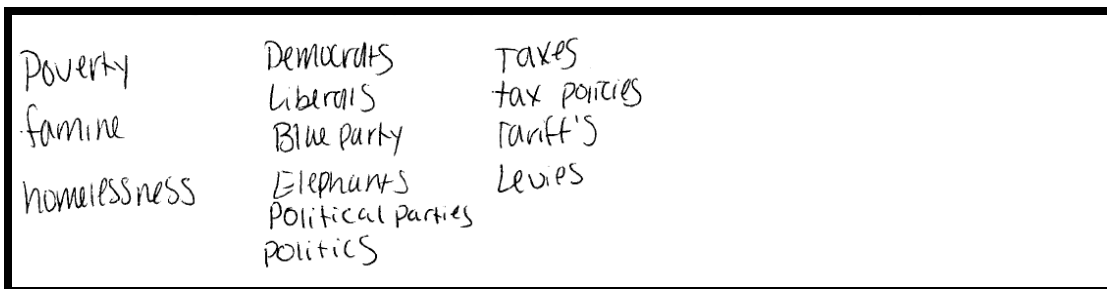

Participant 16 – Network view

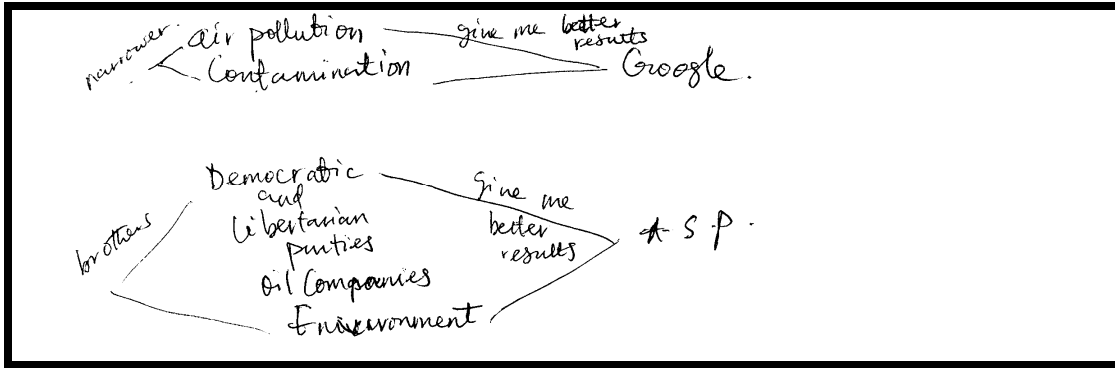Participant 17 – Process view

Participant 18 – Network view
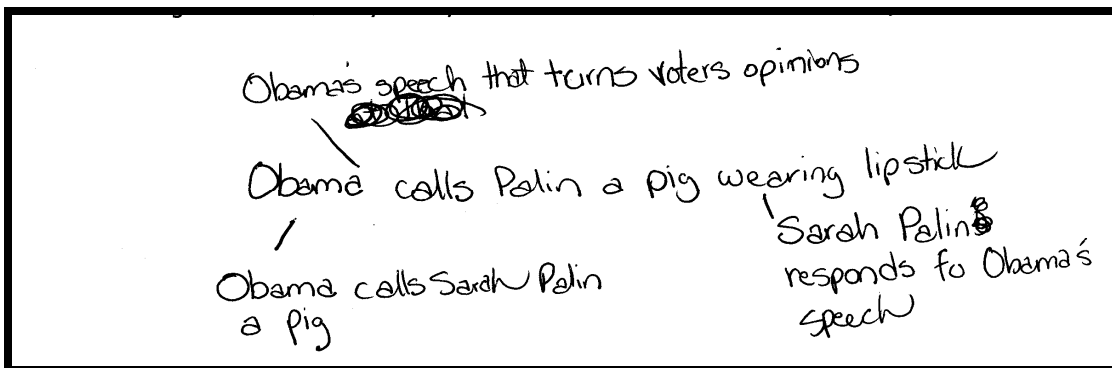
Participant 19 – Word list

Participant 20 – Word list

Participant 21 – Network view