Preprint of:

Nicholson, S. (2005). Understanding the foundation: The state of generalist search education in library schools as related to the needs of expert searchers in medical libraries. *Journal of the Medical Library Association*. 93(1). 58-65.

Understanding the Foundation: The State of Generalist Search Education in Library Schools as Related to the Needs of Expert Searchers in Medical Libraries

Scott Nicholson, Assistant Professor

srnichol@syr.edu

Syracuse University School of Information Studies

Abstract

Purpose: Explore the current state of generalist search education in library schools and consider that foundation in respect to the Medical Library Association's statement on expert searching.

Setting / Subjects: Syllabi from courses with significant searching components. Ten of the top library schools, as determined by the U.S. News and World Report rankings.

Methodology: Mixed methods, but primarily quantitative bibliometric methods.

Results: The educational focus in these searching components was on understanding the generalist searching resources and typical users, and performing a reflective search through application of search strategies, controlled vocabulary, and logic appropriate to the search tool. There is also a growing emphasis on Web-based search tools and a movement away from traditional set-based searching and toward free-text search strategies. While there is a core set of authors used in these courses, there is no core set of readings.

Discussion/conclusion: While library schools provide a strong foundation, there is still need for future medical librarians to take courses that introduce them to the resources, settings, and users associated with medical libraries. In addition, as more emphasis is placed on Web-based search tools and free-text searching, instructors of the specialist

medical informatics courses will need to focus on teaching traditional search methods appropriate for common tools in the medical domain.

Introduction

In order to become expert searchers as defined by the Medical Library Association (MLA) [1], library students need a strong theoretical and practical understanding of searching concepts. The primary way many students acquire this basis is through courses in library school with searching components. While a course specific to searching medical literature will help students hone their searching skills, the foundation for this advanced, subject-specific, searching course is laid through the core courses and more general searching courses. The goal of this article is to gain a better understanding of the topics covered in the non subject-specific searching components in library schools through a bibliometric study of literature used in the top library schools to teach generalist searching. By gaining this knowledge, instructors of specialty health science searching courses can focus on the unfulfilled needs of future expert searchers in medical libraries.

As digital information becomes pervasive in the scholarly environment, and therefore, the amount spent on digital resources in a library increases, the importance of strong searching skills also increases. Librarians of the 21st century are faced with a growing number of search interfaces for bibliographic records and full-text items as well as the variety of information through the public and deep Web. Knowledge of only Boolean logic, fielded searches, and cited reference searching no longer suffices for today's information professional. The problem continues to grow as new searching tools are introduced that do not come from the traditional information professions; even the basic Boolean terms such as AND do not carry their traditional meaning in Web search tools.

Library students interested in searching, therefore, need to develop a strongly grounded theoretical knowledge of searching while understanding the new demands of interactive searching and the heavily commercialized domain of Web search tools. These will lay the groundwork for courses in medical information resources and health care contexts that will create medical library professionals with strength in both the theoretical underpinnings of searching and also the applied knowledge needed. While this paper recognizes the need for continuing education, its focus will be on understanding the current state of search education in library schools and how it can provide some of the education needed to become an expert searcher in a health sciences library.

Prior Work

The topic of search education in library schools has been regularly examined in the library science education literature, although the questions considered have changed. Harter performed two studies in 1979 and 1982(with Fenichel) examining the number of library schools that offered some type of searching course, finding that about two-thirds of the programs offered a searching class and that most programs presented a basic understanding of searching somewhere in the curriculum[2,3]. Tenopir examined the issue in 1989 and found that over four-fifths of schools offered some type of searching class[4]. By 1997, most library schools surveyed by Hsieh-Yee reported some type of online searching component in the library science curriculum[5].

New types of searching became readily available in the 1990s. These new search tools, based on relevance ranking instead of precise Boolean searches, were growing in

popularity, both through public Web search tools as well as in traditional search tools such as DIALOG and OPACs. Xie and Cool took notice of these new tools and examined how well library schools were preparing students for these new types of search tools[6]. They presented how searching through these newer tools differs from searching more traditional Boolean-based tools and the problems encountered with using traditional search strategies in newer tools. They then selected the schools ranked in the top ten by *U.S News and World Report*[7] and reviewed their course descriptions to see which schools offered classes that covered proper use of these tools. They concluded by discussing the importance of considering these newer search tools in a library school curriculum.

Overview and Research Questions

The goal of this project was to gain an understanding of the content of search education in library schools and how it prepares library school students to become expert searchers. The evidence used is syllabi from these schools and, more specifically, the readings assigned during these courses. Undoubtedly, all of the courses examined cover some topics through lectures and assignments that are not covered in readings; however, the assumption is that the most important components of each course will be supplemented by reading material. These reading lists were examined with several research questions in mind:

Q1: How is search education integrated in library school curricula?

Q2: Is there a core group of readings across library schools used in search education?

Q3: How many readings cover different subtopics in these courses and how recent are these readings?

Q4: How does search education fit within the Medical Library Association's policy statement on Expert Searching?

These research questions are used as the organizational structure for most of the methodology and results in this work.

Methodologies

This study uses both qualitative and quantitative methods to explore the research questions, and is intended to be an exploratory study. Instead of using the population of all library schools, this study uses a method of sampling the best schools based upon prior research. Future researchers could take these methods and apply them to a larger sample of schools.

Creating the Sample: Selecting the Schools and Courses

The sampling method used is based on the method employed by Xie and Cool[6]. The top ten library schools according to *U.S. News and World Report* were selected; because of tied rankings, this list actually contains eleven schools:

- University of Illinois at Champaign-Urbana,
- University of North Carolina at Chapel Hill,
- Syracuse University,

- University of Michigan,
- University of Pittsburgh,
- Indiana University,
- Rutgers University,
- University of Wisconsin at Madison,
- Drexel University,
- University of California at Los Angeles, and
- University of Texas at Austin[7].

For each school, the course titles and descriptions were found on their Web sites and reviewed. Courses that contained searching components were identified, and then the site's navigation and search tools were used to locate the most recent syllabus for the course. If this syllabus could not be located for the course, the general Web search tools were used to seek a copy of the syllabus, and if that failed, the instructor of record and/or the last few instructors of the course were contacted directly in order to get syllabi for the courses. This strategy was successful for all but one of the schools, leading to a sample of 23 recent syllabi from courses containing searching components at ten of the top library schools in the country. Extracting the readings from the entire course in the cases of searching courses, or the searching components of courses primarily on other topics, resulted in 401 articles, books, Web sites, and other works.

Methodology for Q1. Integration of Search Education

In order to explore the first research question, the syllabi were examined for overall structure and content. This was done to gain an understanding at a macro level of the types of classes that had been selected and to better understand how different types of classes blend together to provide a generalist foundation of search education. This will set the stage for other studies which focus specifically on the specialist courses in medical information resources.

Methodology for Q2. Identification of a Core Group of Readings

Each of these syllabi was then opened in digital form, and the readings were extracted into an Excel file. Authority control was imposed on the readings; if two slightly different citations referred to the same work, they were combined into a single citation; however, citations to publications of similar works in different publication venues were left separate. The authors and years were manually extracted from the citations and external bibliographic resources were consulted as needed to fill in missing information.

The second research question required a traditional bibliometric study of the readings contained in the syllabi. To explore the question, first the authorship was counted across all courses, with co-authored papers counting as papers for each individual author. Zipf's law predicts that a small number of authors wrote a large number of the works. In addition, the specific readings assigned in these classes were explored in order to see if a similar pattern held; such a pattern would identify a core group of readings.

Methodology for Q3. Examination of subjects and years of articles

Each citation was assigned the most appropriate subject heading based upon the content of the paper from the following list, created inductively through a broad examination of the collected syllabi:

- DIALOG,
- Other Commercial Large-Scale Database (e.g. Lexis-Nexis, Dow Jones, Factiva,),
- Individual databases or OPACs,
- Web-based Search Tools,
- Multiple Databases (comparative works),
- Search Strategies, Logic, Thesauri and Controlled Vocabulary (practical works),
- Information Science / Information Retrieval (theoretical works),
- End-User Searching (focused on working with one individual), and
- Physical and Digital Library Services (e.g. Reference, End-User searching on a larger scale).

These topics were then mapped to the year of publication of the assigned works in order to explore patterns.

Methodology for Q4. Mapping courses into the MLA statement

As it stands, the MLA statement on Expert Searching (see Table 1) contains many points, some of which overlap. In order to aid a large-scale understanding of the issue, these points were examined and placed into four broader categories. These categories were portrayed in a model, and then the most commonly assigned works were placed into these categories. Any gaps between the model of expert searching and the commonly assigned works then point to educational needs for future medical librarians.

Identify and clarify an information need; Place an information need in the context of a discipline or practice; Perceive the implications of the information need in an institution; Use resources beyond the electronically available literature; Recognize limitations of searcher knowledge and institutional resources; Understand databases, including content, metadata, interfaces, matching and ranking algorithms; Develop a reflective, iterative, and heuristic search strategy; Combine deductive and inductive reasoning with domain knowledge to fulfill the information need; Evaluate results to match the user's recall and precision requirements and domain knowledge; Prepare results for the user by removing irrelevant material and identifying themes and gaps in retrieved information; Document the search process for user or legal purposes.

Table 1: Skills and Knowledge Areas for Expert Searching in Medical Libraries (paraphrased from [1])

Assumptions and Limitations

There are some assumptions and limitations made in creating this sample. First, this sample is not representative of all library schools, as the schools examined are only those from the *U.S. News and World Report* list of top schools. Second, this study does not examine subject-specific searching components; therefore, if a school focused its searching education through subject-specific courses, that school would be under-represented in the sample. Third, because only the required and optional readings were examined, topics covered solely in other forms (lectures or assignments) are not represented; it is assumed that readings are representative of the general topic areas covered in the class. Fourth, courses that have searching content that is not included in the official title or description may not have been included.

Results and Discussion

Results for Q1. Integration of Searching Skills in Library School Curriculum

Most of the schools examined followed the same pattern – a brief searching component in core courses and additional searching components in elective courses. The variants to the pattern came from schools that did not have a core-course framework; however, even these schools followed the pattern with a set of basic searching skills in one or more of the introductory courses. The elective courses were either courses that focused on searching or courses on a non-core topic that contained a generalist searching component.

The first category of courses were those core courses that contain searching components, commonly found in courses that covered reference work and information sources, but also found in core courses on cataloging, organization, or metadata. These components in the core courses typically covered search basics over a few weeks, with topics such as Boolean logic, database selection and use, Web searching, OPAC searching, and basic search strategies.

Students taking only these core classes have only these few weeks of search education accompanied by the searching they must do in order to perform research in their other courses. Given the significant number of digital information sources currently purchased by libraries in place of new print works[8], the amount of time dedicated to search education in the core courses should be reconsidered to ensure that library schools are preparing their students to survive as information professionals in the increasingly digital information future.

Another category of course is electives on a related topic with a searching component. The types of courses in this category varied considerably by program. Some courses were focused on information storage and retrieval systems through an examination of the back-end systems that power search tools; these courses typically integrated some sections on more advanced understanding of searching. Other classes focused on the human in the equation – either through a theory-driven information seeking/information behavior approach or centering on the interaction between a human and a computer. Another type of course with searching components was that focused on the intellectual organization of information, such as a course on cataloging or a course on indexing. Many schools offered courses on medical information resources that included a searching component focused on that topic; these courses are not considered here as these specialty courses are covered by another work in this symposium [9].

The final category of class with a searching component is comprised of those courses focused on searching. These courses, typically containing terms in the titles such as online, retrieval, search, or strategies, were offered in some form at all of the programs examined. They usually used classic DIALOG or similar tools for large portions of the class, but also contained information about end-user searching, Web search tools, other databases, search strategies, comparison of tools, and some information retrieval theory. In fact, the topic list presented in the methodology portion of this work is a good representation of the topics covered in this type of searching course.

The most significant concern about this type of class is inspired by Xie and Cool's work[6]; as more searching tools rely upon relevance ranking and interactive search features, strategies based on strict search control, manipulation of retrieval sets, and fielded searching (as compared to full-text searching) may no longer be as effective. In addition, as Web searching introduces a new level of competition between not only the search tool companies, but the information providers wishing to be indexed, courses should be altered to consider new challenges in discerning the trustworthiness and quality of the logic used, the resources indexed, and the results presented. This shift in generalist search education away from fielded searching suggests that those teaching specialist courses on medical searching must focus on the importance of fielded searching in medical databases.

Results for Q2. Core Authors and Readings Authors used in Search Education

As evidence for the rest of this research, the syllabi from these three categories of courses were collected and the readings were extracted. For the searching courses, all of the readings were extracted. For courses with search components, readings were extracted for only those components focused on searching. After authority control and cleaning, the final list of 401 readings was analyzed for patterns. The number of times each reading was used was also noted.

The first pattern explored was authorship. For this evaluation, an author was counted once for each time he/she was referenced in a syllabus. In addition, if a work had multiple authors, all authors were counted each time the work was referenced. This method resulted in a list of 404 different authors. As expected by bibliometric laws, a

small core of authors produced works used in many of the syllabi[10] . The most common authors were Carol Tenopir, Mary Ellen Bates, the DIALOG corporation, and Greg Notess. Table 2 contains the most frequently cited authors and how often a reading authored or co-authored appeared in the sample; the remaining 369 authors were only cited once or twice in the sample. Figure 1 is a graph of the frequency of author appearances with the authors ranked from high to low frequency. Zipf's law predicts that a few authors will write most of the works in a collection[10], and the expected Zipfian curve shape can be seen.

Author	**#**	**Author**	**#**	**Author**	**#**
Tenopir, C.	27	Basch, R.	5	Belkin, N.	3
Bates, M.E.	18	Feldman, S.	5	Berners-Lee, T.	3
DIALOG	15	Harter, S.	5	Bopp, R.	3
Notess, G.	11	Hock, R.	5	Borgman, C.	3
Bates, Marcia	9	Salton, G.	5	Dom, B.	3
Quint, B.	8	Somerville, A.	5	Ebbinghouse, C.	3
Janes, J.	7	Walker, G.	5	Fidel, R.	3
Ojala, M.	7	Berkman, R.	4	Kleinberg, J.	3
Kassel, A.	6	Saracevic, T.	4	McCain, K.	3
Kuhlthau, C.	6	Sherman, C.	4	Turnbull, D.	3
O'Leary, M.	6	Smith, L.	4	White, H.	3
Price, G.	6	Arnold, S.	3		

Table 2: Core List of Authors

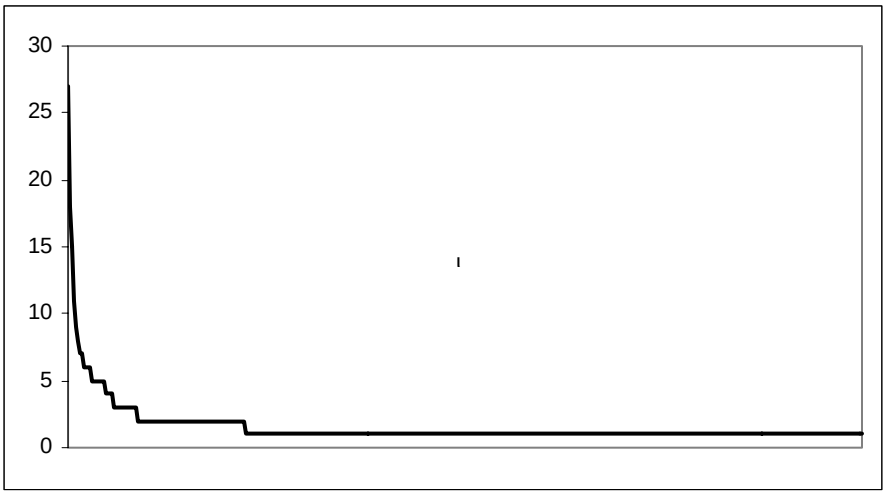

Figure 1: Curve of Frequency of Authors

As expected, there is a core set of authors used to teach general search skills in the top library schools. In order to explore a core set of readings, a similar technique was used. Readings were counted once for each reading list in which they were referenced. Readings with multiple authors were only counted once for each citation, however. This resulted in 401 separate readings used in the 23 courses.

Bibliometric laws suggest that a core group of readings should emerge across the courses. Surprisingly, a core set of readings did not emerge from this analysis. The three most common readings were seen in only 5 of the 23 courses, or 21% of these courses. Given that these courses (or the extracted sections of these courses) cover the same topic area, one would expect a much greater overlap. No readings were covered in 4 of the courses, and 10 readings were covered in three courses. About 90% of the readings extracted appeared in only a single syllabus. Table 3 contains the most common 13 readings and the number of different syllabi in which each appeared.

Reading	#
Bates, Marcia (1989). The design of browsing and berrypicking techniques for online search interface. *Online Review 13*, 407-424.	5
DIALOG Corporation (2002). *DIALOG Lab Workbook.* Cary, NC: Thomson-DIALOG.	5
Walker, G. and Janes, J. (1999). *Online Retrieval: A Dialogue of Theory and Practice.* Englewood, CO: Libraries Unlimited.	5
Basch, R. (1993). Secrets of the super searchers: Planning search strategies. *Online 17*(5), 52-58.	3
Bates, Marcia (1984). The fallacy of the perfect thirty-item online search. *RQ 24* (1), 43–50.	3
Bates, Marcia. (1988). How to use controlled vocabularies more effectively in online searching. *Online 12*(6), 45-56	3
Bopp, R.& Smith, L. (2001). *Reference and Information Services: An Introduction. 3rd ed.* Englewood, Co.: Libraries Unlimited.	3
DIALOG Corporation (2001). *DIALOG Pocket Guide*. Cary, NC: DIALOG.	3
Feldman, S. (2002). This is what I asked for? The searching quagmire. In Mintz. A. (Ed.) *Web of deception: Misinformation on the Internet*. Medford, N.J.: Information Today, 175-195.	3
Harter, S. (1986). *Online Information Retrieval: Concepts, Principles, and Techniques.* San Diego, CA: Academic Press.	3
Kuhlthau, C. (1991). Inside the search process: Information seeking from the user's perspective. *Journal of the American Society for Information Science 42*(5), 361-371.	3
Somerville, A. (1982) The pre-search reference interview -- a step by step guide. *Database 5*(1), 32-38.	3
Tenopir, C. (2001). Why I still teach DIALOG. *Library Journal 126*(8), 35-36.	3

Table 3: Most Common Readings in Searching Syllabi

While these thirteen works are the most common across all of the syllabi, one would be hard-pressed to consider them as a core set of readings. After all, there were 23 syllabi under consideration and all were focused on the same topic; it would be expected that core readings would appear more than 21% of the time. That said, both the author and reading lists could be considered as a starting point for those developing a general searching course for library schools.

What could cause this: there is no set of core readings across courses while there is a core list of authors across courses? Different instructors have chosen different works by the same author to support the same topic. Several hypotheses come to mind; for example, as syllabi are developed and re-developed during different years at different schools, faculty members seek out recent works by familiar authors. As these redevelopment cycles occur, the resulting set of readings across syllabi would be a patchwork of articles with common authors. Another hypothesis worth considering is that the field of electronic database searching is relatively young when compared to other portions of the library science curriculum such as cataloging, management, and reference. As searching is a young sub-discipline, classic core works have not yet emerged; over time, there will be a stabilization of works. The reality is most likely a combination of these two hypotheses.

Results for Q3. Topics Covered and Date of Publications

During the past ten years there has been a significant shift in the type of search tools available. Knowledge of only traditional Boolean-based search tools or cited reference search techniques as applied to gated resources such as databases and OPACs may not serve librarians effectively when dealing with newer tools focused on natural language searching and relevance ranking, as well as full-text searching tools. When revising a syllabus, faculty members are advised to ensure that they go beyond their traditional readings and authors and include readings that focus on these newer forms of search tools. Table 4 has a cross-tabulation of years and reading topics. This table demonstrates

that faculty members at the schools surveyed are appropriately changing focus and readings as times dictate.

Topic	Total Readings	Pre-1980	1981-1990	1991-1995	1996-2000	2001-2004	No date
Web-based Search Tools	101			1	51	48	1
Information Science and IR	85	1	16	20	35	13	
End-User Searching	46	1	11	3	16	15	
Multiple Databases	42		8	4	24	5	1
Library services	39	2	10	3	13	11	
Search Strategies/Logic/Thesauri	34	1	11	7	11	4	
DIALOG	25		2	6	10	7	
Other Commercial Tools	16			1	10	5	
Individual Databases and OPACs	13		3	3	6	1	
Grand Total	401	5	61	48	176	109	2

Table 4: Reading Topic by Year of Publication

By examining the reading topics, one can see that the highest number of Total Readings for each row support Web search tools and the more theoretical aspects of information science/retrieval. One consideration is that these frequency counts are on a per-reading basis; many of the DIALOG "readings" are entire books (such as the *DIALOG Lab Workbook*) and therefore the comparative number of pages read per topic may not correspond to the number of readings assigned. Faculty teaching in library schools should examine this list and their curricular offerings carefully to ensure that all of these topics are covered to some extent.

Results for Q4. General Search Education and Medical Librarianship

The statement released by the Medical Library Association titled "Role of Expert Searching in Health Sciences Libraries" defines a number of important skills and knowledge areas for a successful medical librarian searcher[1]. Table 1 summarizes

these skills from this report. Interestingly, these guidelines would apply in a general nature to all searchers.

Analysis of Skills Needed of Searchers

The skills enumerated in the MLA statement can be grouped into several broad categories as modeled in Figure 2:

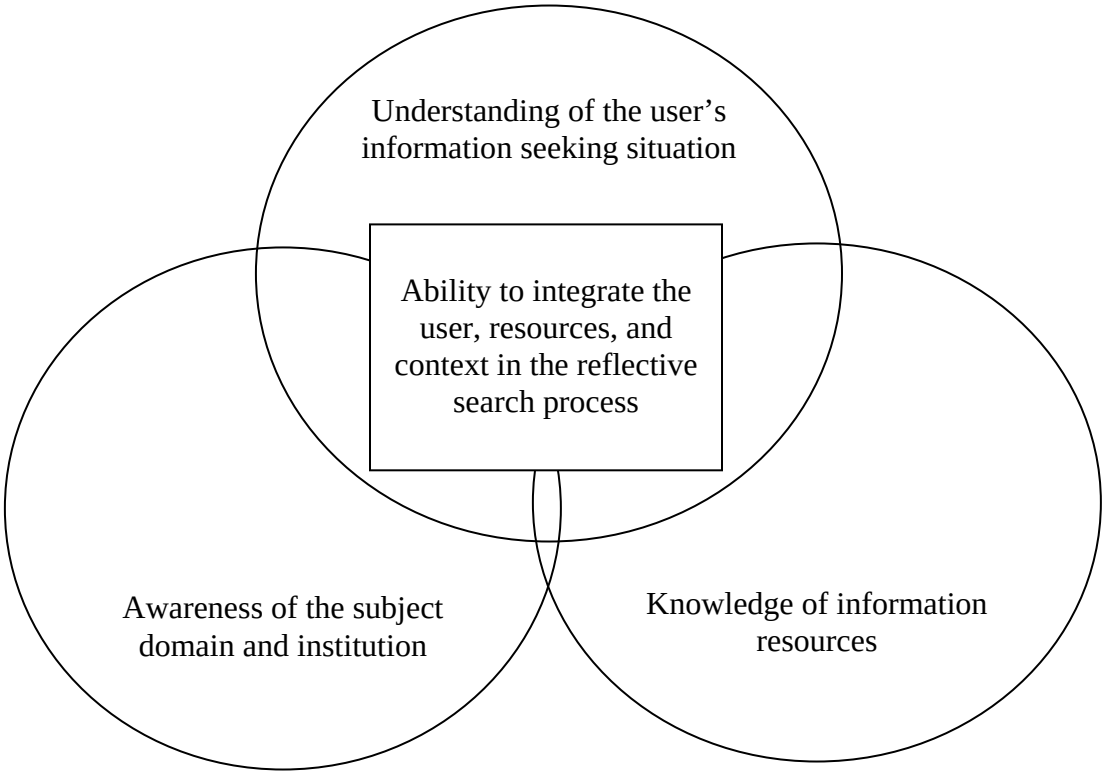

Figure 1: Model of the broad areas of expertise needed for searchers

Three of these categories are covered in the library school curriculum between the course covering the reference process and a basic searching course. The course covering reference typically focuses on understanding the user and learning about both print and electronic resources. Many times, these courses also contain a searching component,

which can give the student a strong base in search strategies. The searching course will continue this training, improving the student's ability to craft a reflective search while teaching them about the underlying structure of databases. The type of reflective searching is important to consider; as more focus is placed on the free-text search tools, less time is available for traditional set-based, fielded searching techniques.

The obvious need for students going into health science libraries will be to learn about the appropriate information sources, both traditional and digital Less obvious is an understanding of the contexts in which the students will be working, including the context for information seekers, the legal context of providing medical information, and the limitations on resources in different health library contexts. A similar contextual area of training centers on the understanding of the multiple employee levels, information needs, and information sophistication of users of library search services. Since these items are specific to the health science profession, they were not covered in the more generalist courses examined for this study. Finally, some of the items in this list that are challenging to deliver in a classroom format (such as balancing inductive and deductive reasoning), might be developed through experiences such as supervised internships.

Table 5 demonstrates how the commonly assigned articles map into these areas of the searching model. This supports the statement that while the standard library science searching courses cover some of the needs of these students, they are lacking in information about the subject domain and the users of medical library services. In order to properly prepare students for medical librarianship, students must have courses

focused on the specific applications of the conceptual lessons learned in generalist search classes. The generalist search classes provide a strong base upon which specialty classes can build.

	Understand users	Knowledge of resources	Awareness of domain	Reflective search
Bates, M. J. (1989)				X
DIALOG Corporation (2002)		X		
Walker and Janes (1999)		X		X
Basch, R. (1993)				X
Bates, Marcia (1984)	X			X
Bates, Marcia. (1988)		X		X
Bopp, R.& Smith, L. (2001)		X		
DIALOG Corporation (2001)		X		
Feldman, S. (2002)	X	X		
Harter, S. (1986)				X
Kuhlthau, C. (1991)	X			
Somerville, A. (1982)	X			
Tenopir, C. (2001)		X		

Table 5: Commonly assigned readings by topic area.

Conclusion

This study explores the foundation developed through a generalist search education and identifies the areas where specialist courses are needed to meet the skills and knowledge areas identified by MLA for expert searching in health sciences libraries. Courses that focus on the reference and organization procedures in libraries, searching, and the information sources and contexts in medical libraries should provide students with the core knowledge base, but, in fact, require supplemental work in subject-specific techniques and resources to create expert searchers as defined by the MLA.

Instructors of specialist courses should focus on the search techniques needed to locate medical information, such as fielded searching and thesauri, as well as the settings and specialized resources needed for health sciences librarianship. A supervised internship experience can help students fill in any gaps needed. All of these experiences should be available at the school selected by a student interested in becoming a health-science librarian. Library and medical informatics programs that are weak in one or more of these areas should seek partnerships to provide students with all of the experiences needed to help them become an expert searcher.

References

1. Medical Library Association. Medical Library Association Policy Statement: Role of Expert Searching in Health Sciences Libraries. [Web document]. Chicago, IL: Medical Library Association, 2003. [rev. 3 Sep 2003; cited 15 Mar 2004]. < http://www.mlanet.org/pdf/expert_search/policy_expert_search.pdf>.
2. Harter SP. An assessment of instruction provided by library schools in on-line searching. Information Processing and Management 1979;15(2):71-5.
3. Harter SP, Fenichel CH. Online searching in library education. Journal of Education for Librarianship 1982 Summer;23(1):3-22.
4. Tenopir C. Education for database intermediaries: How library schools have changed (and how they haven't). Online 1989 Nov;13(6):55-63.
5. Hsieh-Yee I. Teaching online and CD-ROM resources: LIS educators' views and practices. Journal of Education for Library and Information Science. 1997 Winter;38(1):14-34.
6. Xie H, Cool C. Online searching in transition: the importance of teaching ``interaction'' in library and information science education. Journal of Education for Library and Information Science 1998 Sep;39(4):323-30.
7. Library Science. usnews.com: Education: 2002 Graduate School Rankings [Web document]. 2002. [cited 24 Aug 2004]. < http://web.archive.org/web/20020126123211/www.usnews.com/usnews/edu/beyond/gradrank/gbinfos.htm>:

8. Kenney, B, Bryant, E. Reference budgets: A slow revolution. Library Journal 2003 Nov;128(19),8-9, 12-3.

9. Smith CA. An evolution of experts: MEDLINE in the library school. J Med Libr Assoc. 2005 Jan;93(1): xxxxxxx

10. Potter WG. Of making many books there is no end: Bibliometrics and libraries. Journal of Academic Librarianship 1988 Sep;14(4):238a-c.