ED 363 357 IR 054 711

AUTHOR Wildemuth, Barbara M.; Moore, Margaret E.
TITLE End User Searching of Medline. Final Report.
SPONS AGENCY Council on Library Resources, Inc., Washington, D.C.
PUB DATE 3 Aug 93
CONTRACT CLR-8040
NOTE 45p.
PUB TYPE Reports - Research/Technical (143) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Bibliographic Databases; College Libraries; Computer Software Development; Evaluation Methods; Higher Education; Library Instruction; Library Research; Medical Students; *Online Searching; Problems; *Search Strategies; Tables (Data); User Satisfaction (Information)
IDENTIFIERS *Medline

ABSTRACT

One hundred sixty-one MEDLINE searches conducted by third year medical students were analyzed and evaluated to determine which search behaviors were used, whether those individual moves are effective, and whether there is a relationship between specific search behaviors and the effectiveness of the search strategy as a whole. The typical search took 14 cycles, used 7 terms or concepts, and resulted in the display of 11 citations. The most common moves were selection of a database, entering single-word terms and free-text term phrases, and combining sets of terms. Syntactical errors were also common. Librarians judged the searches to be adequate, and students were quite satisfied with their own searches. Librarians identified many missed opportunities in the search strategies, including the underutilization of the controlled vocabulary, subheadings, and synonyms for search concepts. There were no strong relationships found between the librarians' and students' evaluations of the searches and the measures of searching behaviors. Implications of these findings for system design and user education are discussed. Eleven statistical tables presenting research results are included. The end user questionnaire, lists of categories for coding search strategy moves, a sample of coding of a search, and a search evaluation form are appended. (Contains 15 references.) (Author/KRN)

THE UNIVERSITY OF NORTH CAROLINA
AT
CHAPEL HILL

END USER SEARCHING

OF

MEDLINE

FINAL REPORT

CLR Grant No. 8040
August 3, 1993

END USER SEARCHING

OF

MEDLINE

FINAL REPORT

Submitted to the
Council on Library Resources
Cooperative Research Program
Grant No. CLR 8040

by

Barbara M. Wildemuth
School of Information and Library Science
University of North Carolina at Chapel Hill

and
Margaret E. Moore
Health Sciences Library
University of North Carolina at Chapel Hill

August 3, 1993

Table of Contents

List of Tables

Appendices

Acknowledgements

The authors would like to express their appreciation to three other members of the research team. Rikki Mangrum, our research assistant, was instrumental in coordinating the coding of the search moves and preparing the data for analysis. Julia Shaw-Kokot and Cathryn Chiesa, librarians at the UNC-CH Health Sciences Library, were very generous with their time, evaluating the students' searches and identifying missed opportunities. Without their participation, this research project could not have been completed successfully.

ABSTRACT

One hundred sixty-one MEDLINE searches conducted by third year medical students were analyzed and evaluated to determine which search behaviors were used, whether those individual moves are effective, and whether there is a relationship between specific search behaviors and the effectiveness of the search strategy as a whole. The typical search took 14 cycles, used seven terms or concepts and resulted in the display of 11 citations. The most common moves were selection of a database, entering single-word terms and free-text term phrases, and combining sets of terms. Syntactical errors were also common. Librarians judged the searches to be adequate, and students were quite satisfied with their own searches. Librarians identified many missed opportunities in the search strategies, including underutilization of the controlled vocabulary, subheadings, and synonyms for search concepts. There were no strong relationships found between the librarians' and students' evaluations of the searches and the measures of searching behaviors. Implications of these findings for system design and user education are discussed.

INTRODUCTION

End-user searching of databases is becoming more and more common, yet little is known about the ways in which end users formulate their search strategies. Based on what we know of end users' searches as mediated by information professionals, the search process involves ten stages--from the point at which the user has identified a problem, through presearch interaction with a human or computer intermediary, formulation and reformulation of a search strategy, and evaluation and use of the retrieved information (Belkin and Vickery, 1985). Increased understanding of end users' formulation and reformulation of search strategies is of particular interest to two audiences: systems designers who support the process through the design of information retrieval systems and librarians who provide instruction in searching techniques. Therefore, a research study was undertaken that analyzed medical students' MEDLINE searches in detail, describing and evaluating the individual moves they make.

BACKGROUND

One approach to the study of search strategy formulation is to examine and categorize the individual moves made by a searcher. Bates (1979, 1992) identified 29 search tactics, including tactics for monitoring the search progress, optimizing use of the system's file structure, formulating and reformulating the search, and selecting and revising specific terms. These tactics provide a strong framework for the examination of searcher moves, but have not been validated with empirical data from online bibliographic searching. They were found to be useful in categorizing moves made by medical students in searches of a factual database supporting their microbiology instruction (Wildemuth et al., 1991, 1992).

A different set of categories was empirically generated by Fidel (1985), based on observations of information professionals conducting bibliographic searches. This set of 30 categories included moves to reduce the size of a retrieved set, to increase the size of the set, and to increase precision and recall simultaneously. Because they were generated from observations of professional searchers, the applicability of these types of moves to end-user searching can be questioned. However, as with Bates' tactics, several of these moves were found to be applicable to medical student searching of a factual database in microbiology (Wildemuth et al., 1991, 1992).

The analysis of errors made in search formulation is another way of examining end-user search moves. Sewell and Bevan (1976) analyzed errors made by pharmacists and pathologists searching TOXLINE and MEDLINE. The most common errors were related to misspelled terms and misuse of the controlled vocabulary A study of the use of BRS/After Dark in a health sciences library found that users had trouble "understanding the contents and structure of a database, understanding the use of appropriate search terms, and understanding Boolean logic" (Slingluff, Lev and Eisan, 1985, p.18). More recently, Miller, Kirby and Templeton (1988) studied both end-user searching errors and missed opportunities. They found that 37% of the 500 search statements examined contained at least one error (resulting in 0 items retrieved), and over 75% of the search statements represented missed opportunities. A recent examination of end users' "unproductive searches" of MEDLINE revealed that 48% of the problems were associated with formulating the search and the remaining problems were related to inappropriate use of features in GRATEFUL MED (Walker et al., 1991). Both these studies indicate that there is significant room for improvement in end users' formulation of search strategies.

As Walker et al. (1991) point out, some of the problems that end users have in searching CD-ROM and online databases are associated with the system design. The complexity of representing an information need to a retrieval system is often exacerbated by arbitrary system syntax and overly-complex mechanisms for accomplishing common functions. It is a well-known maxim in systems design that novice and intermittent users require different interfaces than users who approach a system on a regular basis, yet interfaces meant for end users are most commonly identical to those intended for professional searchers. Additional data describing end users' actual use of a database will be helpful in improving the design of the end-user interfaces for information retrieval systems.

This information will also be useful to those who instruct end users in search techniques. A recent evaluation of end-user searching in a health sciences library (Moore, 1990) indicated that students found MEDLINE searches to be very useful in patient care and for preparing case presentations. However, one question that surfaced in the evaluation concerned the usefulness of end-user training. Students who attended the training found it helpful, but student attendance at educational sessions and comments from medical faculty serving as clinical clerkship directors indicated that few were convinced that instruction was needed. This is a long-standing debate in the field and deserves further investigation (Eadie, 1990). More data on students' search strategy formulation and reformulation can help to guide the development of future training programs.

RESEARCH QUESTIONS

The results of this study address three questions. First, they provide a description of student search behaviors: which moves are most frequently used, the number of search cycles students perform for each search, the number of terms used in each search, and the number of citations which students display for examination. Second, the results evaluate the effectiveness of the students' searches. Finally, the results test the relationship between specific student search behaviors and the effectiveness of the searches. These results have broad implications relating to interface design and user education.

METHOD

Data collection

During their third year of medical school, medical students at the University of North Carolina at Chapel Hill participate in the Clinical Health Information Retrieval Program (CHIRP). Medical students in the Internal Medicine and Pediatrics clerkships are required to search MEDLINE for patient care information. Participants attend brief MEDLINE orientation sessions given by the staff of the Health Sciences Library. The objectives of this program are to introduce students to using MEDLINE to find journal literature relevant to patient care. It is hoped that establishing this practice when students begin their clinical education will increase the likelihood of their continuing to read literature to support clinical decision-making.

One hundred sixty-one searches were completed from September 1992 to March 1993. MEDLINE on SilverPlatter compact disks was used for 11 of the searches; the remaining 150 searches were conducted through the new UNC Literature Exchange (UNCLE) service. UNCLE uses the BRS searching software to make MEDLINE available through the campus network. Some differences in search behaviors across the two systems used were found and are reported below. Since each student had a unique information need, each search addressed a different topic. As the students performed their searches, the search strategies and results were captured. For the SilverPlatter searches, the students printed the strategy and results; for the UNCLE searches, logs of the strategy were captured automatically and the student printed the results. Students then gave the searches to the clerkship coordinator who, in turn, gave them to the Library's CHIRP coordinator to review. Prior to returning the search output to the student, it was photocopied for later analysis.

In addition to turning in the searches, students were asked to fill out a questionnaire providing a brief description of the search topic, some demographic information and a rating of the student's satisfaction with the search using a six-point Likert scale. A copy of the questionnaire is attached as Appendix A. Questionnaires were completed for 61 of the searches (38%).

Coding of search moves

The student's description of the search topic was recorded at the top of each search strategy. The individual moves made by each student were then coded in two ways: one based on the moves/tactics identified and defined by Bates (1979, 1992) and Fidel (1985); the other based on changes in slots and fillers, as suggested by Shute & Smith (1993). Each coding method is described below.

Using the moves/tactics identified by Bates (1979, 1992) and Fidel (1985), two members of the research team classified each search cycle, i.e., search statement. Each search cycle consisted of one or more moves. A list of possible moves and their definitions is attached as

Appendix B. The classification method was pilot-tested on six SilverPlatter MEDLINE searches, two prepared by the researchers and four conducted by student end-users, to clarify the operationalization of the codes. Such fine-grained analysis will enable the comparison of these results from end-user searchers with Fidel's earlier study of professional searchers.

In addition, one member of the research team coded the moves using a method based on slots, representing concepts, and fillers, specific terms used to represent a concept (Shute and Smith, 1993). The possible codes and their definitions are attached as Appendix C. This method provided coding at a level of granularity appropriate for use in the later regression analyses. A sample of the two coding schemes for one search is attached as Appendix D.

Two professional health science librarians (both experienced searchers) independently evaluated each search, identifying and qualitatively describing missed opportunities (Miller, Kirby, and Templeton, 1988). They also rated the quality of the search in terms of the selection of initial terms (use of synonyms, truncation), the combination of terms (Boolean operators), the use of feedback to narrow or broaden the search, the correct use of system syntax, and the use of the online thesaurus. The Search Evaluation Form they used is attached as Appendix E. This rating form was also pretested with the six sample searches described above.

Data analysis

Analyses were conducted to address each of the three research questions. To provide a description of the students' searching behaviors, the search logs and output were examined. The average number of cycles students performed, the average number of terms per search, and the average number of citations which students displayed for examination were calculated. Frequency counts of the classifications of moves provided information about which moves were most often selected by these medical students.

The second research question relates to the quality of the searches conducted by the students. The students' ratings of their satisfaction with each search yielded a self-evaluation of the quality of their searching behaviors. For this measure, each student's responses to questionnaire items 5, "I found what I was looking for when I did this search," and 6, "This search was an efficient use of my time," were averaged. The professional searchers' ratings of student searches provided an external evaluation of the effectiveness of the search. Three of the ratings were reliable enough for inclusion in further analyses. For each search, each of these pairs of ratings from the professional searchers were averaged. The professional searchers' descriptions of a student's missed opportunities were analyzed qualitatively to identify those searching behaviors which are in need of improvement. This analysis focused on those errors which were committed most often and those errors which have the most serious consequences for the search results.

Finally, the relationship between the students' specific search behaviors and the measures of search effectiveness was tested. The quantitative descriptions of student search behaviors (number of search cycles, number of terms, number of citations printed) were treated as independent variables, along with the frequency of each type of move used (based on the Shute and Smith, 1993, categorization scheme). In addition, we took into account such student characteristics as their training and prior experience with databases and their undergraduate majors. The effect of these variables on the measures of search effectiveness were tested with stepwise linear regression[1].

RESULTS

Student searching behaviors

The average number of cycles per search, the average number of terms per search, the average number of times the "limit" function was used in each search, and the average number of items printed per search are reported in Table 1. There were no statistically-significant differences between the students who returned questionnaires with their searches and those students who did not, so descriptive statistics for all 161 searches conducted are reported in the top section of

[1] It could be argued that the dependent variables were ordinal, rather than interval, level variables, and that logistic regression would be more appropriate. Because of the ease of interpretation and the likelihood that the results would be essentially the same, linear regression was used in these analyses.

Table 1. The bottom section of the table includes only the 61 searches for which questionnaires were also returned (data to be used in later analyses).

Table 1. Student searching behaviors

Variable	n	Mean	Std Dev	Max	Median	Min
All searches						
Number of cycles	161	13.8	9.9	71	11	2
Number of terms	161	6.3	4.0	26	5	2
Use of limit (number of times)	161	1.3	2.1	12	0	0
Number of items printed	131	10.9	8.8	46	9	0
Searches with completed questionnaires						
Number of cycles	61	13.3	9.9	50	10	4
Number of terms	61	5.8	4.2	26	4	2
Use of limit (number of times)	61	1.3	2.0	12	1	0
Number of items printed	60	10.6	8.2	36	9.5	1

Students averaged 14 cycles, or search statements, in each search. They used six different terms in a typical search. The "limit" function was used relatively infrequently, only about once per search. Eleven citations were printed per search, on average.

There were some statistically-significant ($p<.05$) differences between the SilverPlatter searches and the UNCLE searches. Students using UNCLE averaged more terms per search (6.5 versus 3.9) and used the limit function more often (1.4 times per search versus 0.2). It is likely that both of these differences relate to the way in which the search logs were captured, rather than real differences in the searches performed. For the SilverPlatter searches, only the printed search strategies handed in by the students were analyzed; for the UNCLE searches, any sessions relating to the search topic were included in the analysis. Since many of the students' searches involved multiple sessions over several days, the UNCLE searches probably included terms that were later dropped and additional uses of the limit function. A log printed for a SilverPlatter search was equivalent to the last session of an UNCLE search. This difference in data capture method could account for the higher means calculated for the UNCLE searches.

The moves used in all 161 searches, based on the tactics/moves earlier defined by Bates (1979, 1992) and Fidel (1985), are reported in Table 2. The number of students using each move and the number of times the move was used are reported, as well as the average, maximum, median, and minimum number of uses of that move per search. Each search cycle consisted of one or more moves, since a student could make several changes in the search in one cycle. Therefore, the total number of moves was greater than the total number of cycles. The moves are grouped roughly following Fidel's (1985) scheme.

All searches began with the Database move, since the system required that a database be selected. The most frequently-used move was Intersect 1, intersecting a set with another query component. This category included the combination of a set of terms, the addition of terms to previously-specified sets, and the combination of previously-specified sets. One hundred fifty of the 161 students used this move at least once, and it was used, on average, four times per search. Another common move was Weight 4, the use of term phrases and proximity operators.[2] One hundred nine of the students used this move, and it was used, on average, twice per search. Additional common moves included Select, the specification of a single-word term; Limit 1, limiting a search by language; Weight 3, limiting free-text terms to occur in a specific field; and Weight 5, limiting a search to documents of a certain form. It should be noted that all these frequently-used moves (except Database and Select) are tactics for reducing the size of the retrieved set. Syntax errors were also relatively common, made in 49 of the searches, and occurring, on average, 0.7 times per search.

[2] The NEAR proximity operator is used by default on UNCLE searches, when a term phrase is entered.

Table 2. Frequency of moves (based on Bates and Fidel), all searches (n=161)

Move	Students using move	Total uses	Mean frequency	Std Dev	Max	Median	Min
Beginning moves							
Database	161	267	1.7	1.3	8	1	1
Rerun	8	20	0.1	0.7	6	0	0
Resume	1	1	0.0	0.1	1	0	0
Select	95	283	1.8	2.5	20	1	0
Exhaust	3	5	0.0	0.3	3	0	0
Moves to reduce the size of the set							
Intersect 1	150	642	4.0	3.5	16	3	0
Limit 1	80	154	1.0	1.3	7	0	0
Limit 2	32	71	0.4	1.1	6	0	0
Limit 3	1	1	0.0	0.1	1	0	0
Limit 5	5	7	0.0	0.3	3	0	0
Weight 5	55	110	0.7	1.2	7	0	0
Narrow 1	1	1	0.0	0.1	1	0	0
Sub	2	2	0.0	0.1	1	0	0
Weight 1	10	43	0.3	1.7	18	0	0
Weight 3	53	132	0.8	1.6	8	0	0
Limit 4	7	16	0.1	0.5	4	0	0
Weight 4	109	297	1.8	2.2	9	1	0
Narrow 2/Intersect 2	15	28	0.2	0.7	6	0	0
Negate/Block	5	6	0.0	0.2	2	0	0
Moves to increase the size of the set							
Reduce	35	47	0.3	0.6	3	0	0
Cancel	11	13	0.1	0.3	2	0	0
Truncate	11	19	0.1	0.5	3	0	0
Include	1	1	0.0	0.1	1	0	0
Add 1/Parallel	11	18	0.1	0.4	3	0	0
Add 2	7	8	0.0	0.2	2	0	0
Expand 1/Super	5	5	0.0	0.2	1	0	0
Expand 2	14	33	0.2	0.9	7	0	0
Moves to increase both precision and recall							
Relate	1	1	0.0	0.1	1	0	0
Vary	38	80	0.5	1.1	5	0	0
Fix	9	10	0.1	0.3	2	0	0
Respell	33	42	0.3	0.6	3	0	0
Respace	10	10	0.1	0.2	1	0	0
Errors and other moves							
Syntax	49	112	0.7	1.5	11	0	0
Typo	48	75	0.5	0.9	4	0	0
SPflash	6	11	0.1	0.4	4	0	0
Mode	5	20	0.1	1.0	9	0	0
Repeat	9	10	0.1	0.3	2	0	0
System	2	2	0.0	0.1	1	0	0
Neighbor	7	10	0.1	0.3	3	0	0

There were a few statistically-significant differences in the moves made between the students who returned questionnaires and those who didn't. All 16 uses of Limit 4, limiting terms to the title field, were by seven students who did not complete the questionnaire. The students not completing questionnaires used the Vary move more often, substituting one term for another (0.7 times per search versus 0.2). The use of term phrases and proximity operators, Weight 4, was more common among those who did not fill out the questionnaire (2.1 times per search versus 1.4). Because only the data from the searches accompanied by completed surveys is to be used in the later regression analysis, Table 3 includes only the data from those 61 searches. Other than the differences described above, the results for the 61 searches reported in Table 3 are very similar to those for the entire 161 searches.

There were a few statistically-significant differences between the moves used on SilverPlatter and the moves used on UNCLE. The Select move, specifying a single-word term, was used more often in SilverPlatter searches (3.5 times per search versus 1.6). Several moves were used more commonly in UNCLE searches: Limit 1, limiting by language (1.0 times per search versus 0.3); Weight 3, limiting free-text terms to occur in a specified field (0.9 times per search versus 0.3); Weight 5, limiting a search by publication form (0.7 times per search versus 0.2); and typographical errors (0.5 times per search versus 0.1). In addition, there were several moves that did not occur in SilverPlatter searches (no statistical significance test could be performed for these differences): Add 2, Cancel, Exhaust, Expand 2, Fix, Include, Limit 2, Limit 3, Limit 4, Limit 5, Mode, Narrow 1, Narrow 2, Negate/Block, Neighbor, Respace, Resume, SilverPlatter flashbacks, Sub, Super, System, and Weight 1. Because so few SilverPlatter searches were conducted, it is impossible to determine whether system characteristics affected students' choices of moves in these cases, or additional searches conducted would have included these moves. All these moves are syntactically possible on the SilverPlatter system.

Table 3. Frequency of moves (Bates and Fidel), searches with questionnaires (n=61)

Move	Students using move	Total uses	Mean frequency	Std Dev	Max	Median	Min
Beginning moves							
Database	61	98	1.6	1.1	6	1	1
Rerun	6	11	0.2	0.6	4	0	0
Resume	0						
Select	35	117	1.9	3.1	20	1	0
Exhaust	2	4	0.1	0.4	3	0	0
Moves to reduce the size of the set							
Intersect 1	56	223	3.7	3.5	15	2	0
Limit 1	34	65	1.1	1.3	5	1	0
Limit 2	10	20	0.3	0.9	5	0	0
Limit 3	0						
Limit 5	3	5	0.1	0.4	3	0	0
Weight 5	21	46	0.8	1.4	7	0	0
Narrow 1	1	1	0.0	0.1	1	0	0
Sub	1	1	0.0	0.1	1	0	0
Weight 1	4	14	0.2	1.3	10	0	0
Weight 3	18	52	0.9	1.8	8	0	0
Limit 4	0						
Weight 4	38	83	1.4	1.5	5	1	0
Narrow 2/Intersect 2	5	9	0.1	0.6	3	0	0
Negate/Block	1	1	0.0	0.1	1	0	0
Moves to increase the size of the set							
Reduce	12	14	0.2	0.5	2	0	0
Cancel	6	7	0.1	0.4	2	0	0
Truncate	6	9	0.1	0.5	3	0	0
Include	0						
Add 1/Parallel	3	4	0.1	0.3	2	0	0
Add 2	4	4	0.1	0.2	1	0	0
Expand 1/Super	2	2	0.0	0.2	1	0	0
Expand 2	5	11	0.2	0.9	7	0	0
Moves to increase both precision and recall							
Relate	1	1	0.0	0.1	1	0	0
Vary	8	11	0.2	0.6	4	0	0
Fix	3	3	0.0	0.2	1	0	0
Respell	11	14	0.2	0.5	2	0	0
Respace	2	2	0.0	0.2	1	0	0
Errors and other moves							
Syntax	19	40	0.7	1.2	5	0	0
Typo	15	20	0.3	0.7	4	0	0
SPflash	5	10	0.2	0.6	4	0	0
Mode	4	19	0.3	1.5	9	0	0
Repeat	3	4	0.1	0.3	2	0	0
System	1	1	0.0	0.1	1	0	0
Neighbor	3	4	0.1	0.3	2	0	0

The moves used in all 161 searches, based on the Shute and Smith (1993) coding scheme, are reported in Table 4. There were no statistically-significant differences in the moves made between the searches that were accompanied by questionnaires and those that weren't. The number of students using each move and the number of times the move was used are reported, as well as the average, maximum, median, and minimum number of uses of that move per search.

Table 4. Frequency of moves (Shute & Smith), all searches (n=161)

Move	Students using move	Total uses	Mean frequency	Std Dev	Max	Median	Min
Database selection	161	267	1.7	1.3	8	1	1
New slot (initial set)	161	563	3.5	2.9	24	3	1
Combine existing slots	87	204	1.3	2.0	15	1	0
Combine slots with OR	3	3	0.0	0.1	1	0	0
Add slot(s)	140	650	4.0	3.3	16	3	0
Delete slot(s)	98	277	1.7	2.0	10	3	0
Exclude (NOT operator)	5	7	0.0	0.3	2	0	0
Replace slot-filler with broader slot-filler	48	90	0.6	1.2	8	0	0
Replace slot-filler with other slot-filler	54	138	0.9	1.7	12	0	0
Replace slot-filler with narrower slot-filler	55	99	0.6	1.2	8	0	0
Replace operator with broader operator	9	9	0.1	0.2	1	0	0
Replace operator with narrower operator	10	12	0.1	0.3	2	0	0
Check index/thesaurus	7	9	0.1	0.3	2	0	0
Errors	87	231	1.4	2.5	19	1	0

All the students, of course, selected a database and included at least one New slot (the first concept) in their searches. Another common move was to Add a slot to the search. This category implies that a student included a new concept as part of a search statement that also contained an existing concept. One hundred forty of the 161 searches included this move, and it occurred an average of four times per search. Deleting a slot, i.e., repeating a search statement minus one of the concepts, was the next most common move. It was used in 98 of the searches, and occurred an average of twice in each search. A third common move was to combine existing slots. This type of move is common in the "building-block" approach (Markey and Atherton, 1978), in which individual concepts are specified, each in a separate step, then combined. This move was included in 87 of the searches, occurring once, on average, in each search. Unfortunately, the next most common type of move was an error, occuring in 87 of the searches--just over half. These errors included both syntactical and typographical errors, but did not include the "missed opportunities" identified by the librarian evaluators. Moves including the manipulation of slot-fillers did not occur nearly as frequently as moves manipulating slots. Only about one-third of the searches included any changes in slot-fillers, averaging less than one occurrence per search. The use of the NOT operator, the use of OR to combine slots, the use of the online thesaurus/index, and the manipulation of operators were used very infrequently.

The frequencies of the moves used in the 61 searches accompanied by questionnaires are reported in Table 5 (data to be used in the later regression analysis). The number of students using each move and the number of times the move was used are reported, as well as the average, maximum, median, and minimum number of uses of that move per search. There were no statistically-significant differences between the moves used in searches accompanied by questionnaires and those not accompanied by questionnaires.

There were several statistically-significant differences between the SilverPlatter searches and the UNCLE searches. The UNCLE users added slots to their searches more often (4.2 times per search versus 1.3 for the SilverPlatter users). The UNCLE users replaced slot-fillers with other slot-fillers more often than SilverPlatter users (0.9 times per search versus 0.3). Only UNCLE users replaced a slot-filler with a broader slot-filler, replaced an operator with a broader operator, used the OR operator to combine slots, used the NOT operator, and checked the online thesaurus/index, though all these moves are syntactically possible on the SilverPlatter system. As in the case of the number of terms and the number of limit commands, reported earlier, these differences may be due to the way in which the searching data were captured, rather than real differences in use of these two systems.

Table 5. Frequency of moves (Shute & Smith), searches with questionnaires (n=61)

Move	Students using move	Total uses	Mean frequency	Std Dev	Max	Median	Min
Database selection	61	99	1.6	1.1	6	1	1
New slot (initial set)	61	204	3.3	2.4	12	2	1
Combine existing slots	36	81	1.3	1.7	7	1	0
Combine slots with OR	1	1	0.0	0.1	1	0	0
Add slot(s)	55	217	3.6	3.0	14	3	0
Delete slot(s)	33	85	1.4	1.8	6	1	0
Exclude (NOT operator)	1	2	0.0	0.3	2	0	0
Replace slot-filler with broader slot-filler	13	30	0.5	1.3	8	0	0
Replace slot-filler with other slot-filler	16	58	1.0	2.2	12	0	0
Replace slot-filler with narrower slot-filler	18	33	0.5	1.3	8	0	0
Replace operator with broader operator	4	4	0.1	0.2	1	0	0
Replace operator with narrower operator	6	8	0.1	0.4	2	0	0
Check index/thesaurus	3	4	0.1	0.3	2	0	0
Errors	30	94	1.5	3.1	19	0	0

Search effectiveness

Search effectiveness was evaluated in three ways: librarians evaluated the quality of the students' searches on a rating scale; the students evaluated themselves; and librarians noted missed opportunities in the students' search strategies. Each of these measures of search effectiveness is reported below. Only the searches for which the student completed a questionnaire could be included in this analysis (n=61).

Two librarians, both very experienced in searching MEDLINE, independently rated the quality of each search on five dimensions: initial selection of terms, use of Boolean operators to combine terms and sets of terms, the use of system feedback to narrow or broaden the search, the correct use of system syntax and commands, and use of the online thesaurus. Each of these dimensions was evaluated on a five-point scale (1=poor, 3=OK, 5=excellent), with the option of any dimension being noted as not applicable to this search. One evaluator marked the use of feedback as not applicable to one search; this case was analyzed as missing data. Both evaluators marked the use of the online thesaurus as not applicable to all the searches except one, so this dimension was dropped from further analysis.

The average ratings for the students' searches on the four dimensions are reported in Table 6, separately for each librarian/evaluator. The evaluators used the entire five-point range in their evaluations, averaging near 3 (=OK) on each of the dimensions. Evaluator 2 seemed to rate the searches slightly higher, on average, than did Evaluator 1, but the difference was not statistically significant for any of the four dimensions or for a composite of the ratings.

Table 6. Search evaluations by librarians, data with surveys (n=61)

Variable	Mean	Std Dev	Max	Median	Min
By first evaluator					
Initial selection of terms	2.7	1.1	5	3	1
Use of Boolean operators	3.1	0.6	4	3	2
Use of feedback to narrow or broaden search	3.0	0.8	5	3	2
Correct use of system syntax	3.2	0.9	5	3	1
By second evaluator					
Initial selection of terms*	2.8	1.3	5	3	1
Use of Boolean operators*	3.3	1.2	5	3	1
Use of feedback to narrow or broaden search**	3.3	1.3	5	3	1
Correct use of system syntax*	3.4	1.2	5	3	1

* Note: Only 60 responses because this item was coded as "not applicable" to one search by the second evaluator.
**Note: Only 59 responses because this item was coded as "not applicable" to two searches by the second evaluator.

Before combining the two sets of ratings, as originally planned, the interrater agreement was investigated further. Several measures of interrater agreement, differing on their assumptions about the level of the data (ordinal versus interval), were calculated for each dimension and are reported in Table 7. In general, they indicate that the two evaluators did not have a high level of agreement. For the purposes of the analysis reported here, the third dimension--use of system feedback to narrow or broaden the search--was dropped because of its low reliability. The other three dimensions were retained and the scores were averaged (see Table 8). Prior to more formal publication of these results, a third evaluator will independently rate the searches and interrater agreement will be evaluated again.

Table 7. Interrater agreement for librarians' ratings of student searches

Measure of agreement	Initial selection of terms	Use of Boolean operators	Use of feedback to narrow or broaden search	Correct use of system syntax
Pearson's r	0.47	0.46	0.32	0.56
Coefficient alpha (for raw variables)	0.63	0.56	0.44	0.70
Spearman rank correlation	0.46	0.53	0.30	0.56
Kendall's tau-b	0.39	0.46	0.25	0.48
Cohen's kappa	0.13	0.28	0.00	0.13
Cohen's weighted kappa	0.31	0.33	0.13	0.32

Table 8. Average ratings across the two librarians/evaluators

Variable	Mean	Std Dev	Max	Median	Min
Initial selection of terms	2.8	1.0	5	2.5	1
Use of Boolean operators	3.2	0.8	4.5	3	2
Correct use of system syntax	3.3	0.9	5	3.5	1

The results, as reported in Table 8, indicate that students' searches are adequate, receiving a rating of approximately 3 (=OK) on all three dimensions. Students' initial selection of terms and the correctness of their system syntax covered the entire range of ratings; their use of Boolean operators was rated between 2 and 4.5 on a five-point scale.

The second measure of search quality was a student's estimate of his or her performance, as measured with two items on the questionnaire: Item 5, "I found what I was looking for in this search," and Item 6, "This search was an efficient use of my time." Each item was rated on a scale from 1 (strongly agree) to 6 (strongly disagree). The results from these two items are reported in Table 9. The mean for each question was less than 2 and the median for each was 1, indicating that students generally were satisfied with their searches.

Table 9. Students' self-evaluations (n=61)
(Strongly agree = 1; strongly disagree = 6)

Item	Mean	Std Dev	Max	Median	Min
5. I found what I was looking for in this search.	1.7	1.0	6	1	1
6. This search was an efficient use of my time.	1.9	1.3	6	1	1

For the purpose of defining variables for the regression analysis, the relationship between these two questions was explored. If they are highly related, they should be combined as one variable in the regression equation; if they are not highly related, they should be treated as two separate variables. The correlation (Pearson's *r*) between the two questions was 0.61, and Cronbach's alpha for the combined scale of two items was 0.75. Therefore, these two items were combined into one variable for the regression analysis and considered to be a measure of the students' overall satisfaction with their search performance.

The third measure of search performance was the identification of missed opportunities by the two librarians. Each librarian independently reviewed the search strategies used by the students. Based on their expertise in searching MEDLINE, they noted instances in which the student missed an opportunity to improve the search strategy. These notes were categorized by a member of the research team. The types of missed opportunities identified and the frequency of each are reported in Table 10. The errors identified in the analysis of moves are included also.

Table 10. Missed opportunities and errors (n=61)

Missed opportunity or error	Number of students	Total frequency	Mean frequency	Std Dev	Max	Median	Min
Missed opportunities							
Should use MeSH term	38	73	1.3	1.3	5	1	0
Should not use MeSH term (none available)	2	3	0.1	0.3	2	0	0
Should limit term to major descriptor	2	2	0.0	0.2	1	0	0
Should explode MeSH term	5	6	0.1	0.4	2	0	0
Should add synonyms with OR	13	19	0.3	0.7	3	0	0
Should truncate term/use truncation symbol	9	14	0.2	0.7	3	0	0
Should use broader term	2	3	0.1	0.3	2	0	0
Should use narrower term	1	1	0.0	0.1	1	0	0
Should use subheading	15	25	0.4	1.0	5	0	0
Should limit to specific age groups	8	9	0.2	0.4	2	0	0
Should use a different proximity operator	5	12	0.2	0.8	4	0	0
Made an illogical Boolean combination	7	14	0.2	0.8	5	0	0
Other missed opportunities	15	16	0.3	0.5	2	0	0
Errors							
Syntactical errors	19	40	0.7	1.2	5	0	0
Typographical errors	15	20	0.3	0.7	4	0	0
SilverPlatter flashbacks	5	10	0.2	0.6	4	0	0
Mode errors	4	19	0.3	1.5	9	0	0
Repeated statements	3	4	0.1	0.3	2	0	0
System error	1	1	0.0	0.1	1	0	0

As can be seen from the data in Table 10, these students missed many opportunities to improve their searches. Fifty-two of the 61 searches evaluated contained missed opportunities of some kind and 30 contained errors. All together, 56 (92%) of the searches contained either a missed opportunity or an error or both. By far, the most common missed opportunity was exploitation of the controlled vocabulary, MeSH. Thirty-seven searches could have been improved with the inclusion of MeSH terms in place of free-text terms. A similar vocabulary-related problem was the lack of inclusion of appropriate synonyms for a search concept. Thirteen of the searches would have been improved by the addition of synonyms. The other opportunity that was commonly missed was the use of subheadings, which would have improved 15 of the searches. As noted earlier, the most common errors were syntactical and typographical, occurring in 19 and 15 searches, respectively.

Relationship between searching behaviors and effectiveness

The third research question concerns the relationship between the process of searching and the effectiveness of a search. In this study, the librarians' and students' ratings of a search's quality were used as dependent variables: the measures of search effectiveness. The independent variables included the number of search cycles per search, the number of terms used (including limit functions), the number of citations printed, and the frequency of each type of move (based on the Shute and Smith, 1993, categorization).

In addition to the independent variables, several characteristics of the students were included in the regression equation to determine their effect. One background variable of interest was the student's experience with computerized databases. Descriptive statistics for the questionnaire items measuring the students' searching background are reported in Table 11. There was a statistically-significant relationship between item 9, "Have you ever searched INQUIRER for microbiology information?," and item 10, "Before this search, had you ever used computers to search bibliographic databases to find journal articles?," so those two items were combined into one group of dummy variables for the regression equation. Item 8, "Have you ever used database management software like dBase or Microsoft Works?," was included as a separate set of dummy variables in the regression equation. A second background variable of interest was the student's undergraduate major (science versus non-science). Fifty-eight students provided information about their undergraduate background. Of those, 66% had an undergraduate degree in a natural or physical science.

Table 11. Experience with computerized databases (n=59; 2 students did not respond)

Item	No, never	Yes, once or twice	Yes, 3-4 times	Yes, 5+ times
8. Have you ever used database management software like dBase or Microsoft Works?	20	14	4	21
9. Have you ever searched INQUIRER for microbiology information?	12	12	8	27
10. Before this search, had you ever used computers to search bibliographic databases to find journal articles?	0	4	7	48

Stepwise linear regression analysis was used to identify models that would predict each of the four dependent variables: initial selection of terms, use of Boolean operators, correct use of system syntax, and the students' evaluations of their performance. The independent variables were entered into the model individually or in groups. Individual variables included the number of cycles per search, the number of citations printed per search, the number of moves coded as errors, student experience with microcomputer database management software, and student undergraduate major. Groups included the frequencies of the moves that were not errors and the students' experience with INQUIRER and online bibliographic databases.

No variables or groups of variables predicted students' performance in the initial selection of terms. Students' past experience with INQUIRER and online bibliographic databases predicted their success in using Boolean operators, but the prediction was very weak and only marginally significant ($R^2 = 0.11$, prob>F = 0.10). The number of errors predicted librarians' evaluations of students' correct use of system syntax, but only weakly ($R^2 = 0.11$, prob>F = 0.01). The number of errors also predicted students' evaluations of their own performance, but again, only very weakly ($R^2 = 0.05$, prob>F = 0.11).

DISCUSSION

This study of end-user searching behavior addressed three specific research questions: what happens when students search a large bibliographic database, are they effective in their searches, and does any individual aspect of the search process predict successful performance?

Using a large sample of naturalistic searches performed by third-year medical students, each of these questions was answered.

The results describing students' search behaviors provide a detailed view of the online searching process. A typical search takes 14 cycles, incorporates about seven different terms or concepts, and results in the retrieval of about 11 citations. It is likely to incorporate the selection of a database; selection of single-word terms, free-text term phrases, phrases that appear in a particular field, combinations of terms and phrases with the Boolean AND operator, and limitation of the output by language and publication form. Unfortunately, it is also likely to include syntactical or typographical errors and is not likely to draw on a controlled vocabulary as often as would be beneficial. It is unlikely to include extensive manipulation of synonyms, reliance on an online thesaurus, or the use of the NOT operator.

Several of these search behaviors have a direct impact on the effectiveness of the searches. Students' initial selection of terms was adequate, but could be improved. Increased use of an online thesaurus and more awareness of the importance of including synonyms in the specification of each search concept are possibilities for improved performance. Syntactical and typographical errors affected search performance negatively, though usually were noticed and corrected quickly. The students' use of Boolean logic was adequate, but there were some errors and the increased use of OR to combine synonyms would result in improved outcomes in many cases. Unfortunately, students' self-evaluations indicate that most are unaware of these problems in their search performance or are satisfied with the outcomes of their searches, in spite of the problems.

This study was not successful in finding any links between particular search behaviors and search performance. It seems that individual searches can be evaluated and recommendations made for their improvement, but no general statements can be made about the relationship between search performance and the number of cycles executed, terms used, or citations retrieved, or the types of moves used. One avenue for further exploration is to consider larger chunks of searching behavior, i.e., to analyze the searches in terms of sequences of moves within a search, rather than the individual moves. Hsieh-Yee (1990) made a similar point, noting the difficulty of analyzing complete search strategies and suggesting that sub-sequences of moves be the unit of analysis. As independent variables in a regression equation, frequencies of individual moves are too weak to predict search performance.

IMPLICATIONS OF THE RESULTS FOR LIBRARIES

In spite of the lack of results from the regression analysis, the analysis of moves and the identification of missed opportunities can provide some guidance for both designers of information retrieval systems and librarians who offer user education in searching.

First, students' search performance could be improved if the number of syntactical errors were reduced. One way to make this improvement would be to design systems that are more tolerant of variations in syntax. Some progress is being made in this area, as more systems are designed for intermittent users, rather than professionals who have a responsibility to develop syntactical expertise. As information retrieval systems become "smarter," end users will be allowed to focus on the substance of their searches, rather than the syntax. Until then, user education must fill the gap. Common syntactical errors can be identified through examination of search logs, and training sessions and user aids can highlight the errors that are most problematic in the execution of the search.

Second, students' search performance could be improved with improved vocabulary support. Students made typographical errors, selecting the correct term but entering it in a form unrecognizable to the system; students did not use the online thesaurus available to them; and students did not attempt to generate synonyms to fully specify a concept of interest. Each of these mistakes had a negative effect on search outcomes. Typographical errors can best be addressed through system design, automatically referring the user to a list of possible terms when an entered term retrieves no citations. Generation of synonyms and selection of descriptors when appropriate can be addressed either through system design or user education. If the online thesaurus is more closely linked to the search engine, the system can suggest synonyms from a controlled vocabulary when a term is entered. Common acronyms can also be added to the controlled vocabulary to ensure that users include both versions of the concept in their searches. These students were trying to find only a small number of articles relevant to specific clinical cases, but selecting terms from a list of possible synonyms is likely to be a more successful means of developing a coherent search strategy than using one representation of a concept selected from personal knowledge -- particularly for students who are new to a domain. This problem can also be addressed in user education that emphasizes the ambiguity of natural language and the usefulness of a controlled vocabulary in guiding a search through a large database.

One other finding of interest to system designers and librarians is the wide range of moves used by these students. There are few features of the information retrieval systems available that were not used, at least once. Each student may rely on only a few moves, but this group of students used over 30 different kinds of moves, not including errors. For system designers, this finding implies that it is indeed worthwhile to make these features available. At least some system users are finding them helpful. For librarians, this finding implies that advanced training sessions and user aids focused on particular features may be useful to their clients. Examination of search logs at a particular institution may reveal which features are important to the users at that institution and can guide the development of customized training programs. Locally, these results will be used in such ways: to identify needed UNCLE system enhancements and to help librarians in developing advanced training, help screens, and user aids.

FUTURE RESEARCH

The results reported here are preliminary, in the sense that the data collected in this study warrant further analysis. As mentioned above, a third librarian will rate the quality of the student searches to improve the reliability of those evaluations. To expand the meaningfulness of the results, the search moves themselves will be re-analyzed using short sequences of moves as the unit of analysis. The starting point for this analysis will be the search strategies outlined in Markey and Atherton (1978): the building block approach, the citation pearl growing approach, the successive fractions approach, the most specific facet first approach, and the lowest postings facet first approach. Graphical representation of the search strategies will also be explored to provide new perspectives on the searching process. It is hoped that using a slightly larger unit of analysis will prove fruitful in exploring the relationships between search behaviors and search outcomes.

REFERENCES

Bates, Marcia J. Information search tactics. Journal of the American Society for Information Science 30 (4): 205-14, July 1979.

Bates, Marcia J. Search and idea tactics. In For Information Specialists: Interpretations of Reference and Bibliographic Work, by Howard D. White, Marcia J. Bates, and Patrick Wilson. Norwood, NJ: Ablex, 1992. Pages 183-200.

Belkin, Nicholas J.; Vickery, Alina. Interaction in Information Systems: A Review of Research from Document Retrieval to Knowledge-Based Systems. Library and Information Research Report 35. British Library, 1985.

Eadie, Tom. Immodest proposals: User education for students does not work. Library Journal 115 (17): 42-45, October 15, 1990.

Fidel, Raya. Moves in online searching. Online Review 9 (1): 61-74, 1985.

Hsieh-Yee, Ingrid P. Y. The Search Tactics of Novice and Experienced Searchers: A Comparative Study. Ph.D. dissertation, University of Wisconsin - Madison, 1990.

Markey, Karen; Atherton, Pauline (based on the earlier unpublished work of Charles P. Bourne, Barbara Anderson, and Jo Robinson). ONTAP: Online Training and Practice Manual for ERIC Data Base Searchers. Syracuse, NY: Syracuse University, ERIC Clearinghouse on Information Resources, June 1978. ED 160 109.

Miller, Naomi; Kirby, Martha; Templeton, Etheldra. MEDLINE on CD-ROM: End user searching in a medical school library. Medical Reference Services Quarterly 7 (3): 1988.

Moore, Margaret E. Clinical Health Information Retrieval Project: Program evaluation report, 1989. Paper presented at the Librarians' Association of the University of North Carolina at Chapel Hill Research Forum, May 3, 1990.

Sewell, Winifred; Bevan, Alice. Nonmediated use of MEDLINE and TOXLINE by pathologists and pharmacists. Bulletin of the Medical Library Association 64 (4): 382-91, October 1976.

Shute, Steven J.; Smith, Philip J. Knowledge-based search tactics. Information Processing & Management 29 (1): 29-46, 1993.

Slingluff, Deborah; Lev, Yvonne; Eisan, Andrew. An end user search service in an academic health sciences library. Medical Reference Services Quarterly 4 (1): 11-21, Spring 1985.

Walker, Cynthia J.; McKibbon, K. Ann; Haynes, R. Brian; Ramsden, Michael F. Problems encountered by clinical end users of MEDLINE and GRATEFUL MED. Bulletin of the Medical Library Association 79 (1): 67-69, January 1991.

Wildemuth, Barbara M.; de Bliek, Ruth; He, Shaoyi; Friedman, Charles P. Search moves made by novice end users. ASIS '92: Proceedings of the 55th ASIS Annual Meeting 29: 154-61, 1992.

Wildemuth, Barbara M.; Jacob, Elin K.; Fullington, Angela; de Bliek, Ruth; Friedman, Charles P. A detailed analysis of end-user search behaviors. ASIS '91: Proceedings of the 54th ASIS Annual Meeting 28: 302-12, 1991.

APPENDICES

Appendix A. Questionnaire

7/11/91

END USER SEARCHING OF MEDLINE

We are currently studying end-user MEDLINE searching. The results will be used to guide development of educational services and future search systems. One MEDLINE search is required for this clerkship, but your participation in this study is voluntary. All we ask is that you give us permission to use your search, complete this brief questionnaire, and turn in the questionnaire with the search printout. We will return your search printout with feedback and provide educational services and search assistance.

Are you willing to let us use your search for research purposes? YES NO

If you would like further information about the study, please contact either of the two principal investigators – Barbara Wildemuth, UNC-CH School of Information and Library Science (962-8072) or Margaret Moore, UNC-CH Health Sciences Library (962-0700). Fur further information about your rights as a participant, please contact the Academic Affairs Institutional Review Board at 966-5625.

IMPORTANT! If you want this search to count towards your clerkship requirement and/or feedback, please print your name on the search printout. Your name will be blacked out for study purposes. Your questionnaire responses and search results will remain anonymous and confidential.

1. Please describe your search topic.

2. What is the purpose of this search? (Please circle all that apply.)

   (a) Working up patient(s) on this rotation

   (b) Preparing for a CPC, rounds, or case presentation

   (c) Other (Please describe.) ______________

After searching, please circle appropriate responses to the following statements.

	Strongly agree					Strongly disagree
3. This system was easy to use.	1	2	3	4	5	6
4. Computerized bibliographic search technology should be available to all clerkship students.	1	2	3	4	5	6
5. I found what I was looking for in this search.	1	2	3	4	5	6
6. This search was an efficient use of my time.	1	2	3	4	5	6

7. Did you take advantage of any of the following search aids? (Please circle all that apply.)

   (a) SilverPlatter help screens

   (b) introduction or guide on computer

   (c) printed user guide next to computer

   (d) asked other students for help in searching

   (e) asked library staff for help in searching

   (f) attended open help sessions at Library

   (g) SilverPlatter training in Clinical Epidemiology

   (h) SilverPlatter workshop at Library

OVER

8. Have you ever used database management software like dBASE or Microsoft Works?

(a) No, never

(b) Yes, once or twice.

(c) Yes, three or four times.

(d) Yes, five or more times.

9. Have you ever searched INQUIRER for microbiology information?

(a) No, never.

(b) Yes, once or twice.

(c) Yes, three or four times.

(d) Yes, five or more times.

10. Before this search, had you ever used computers to search bibliographic databases to find journal articles?

(a) No, never

(b) Yes, once or twice.

(c) Yes, three or four times.

(d) Yes, five or more times.

If yes, what system(s) did you use? (Please circle all that apply.)

(a)	SilverPlatter MEDLINE	(e)	InfoTrac
(b)	BRS Colleague	(f)	Other (Please specify)
(c)	Grateful Med		________________
(d)	PaperChase	(g)	Don't remember

11. What was your undergraduate major? ________________

12. What field of medicine do you plan to enter? ________________

Additional comments or suggestions:

Please attach this questionnaire to your search printout. Return to clerkship office or in box next to the computer.

IMPORTANT! You are required to turn in one MEDLINE search for this clerkship. If you want the search to count towards that requirement, please print your name on the search printout. It will be blacked out for study purposes. Your questionnaire responses and search results will remain anonymous and confidential.

Thank you!!

clrquest
7/11/91

Appendix B. Categories for coding moves based on Fidel (1985) and Bates (1979, 1992)

Individual move definitions were adapted from moves and tactics proposed by Fidel (1985), Bates (1979, 1992), and Wildemuth (1991, 1992). Quoted definitions are from the original source for each move definition.

BEGINNING MOVES

Move	Definition	Notes
Database	Select a specific database	Operationalized as the first move of each day/session and at change of database.
Rerun	To search a new set of records with a pre-existing search statement.	
Select	"To break complex search queries down into sub-problems and work on one problem at a time."	Originally defined by Bates (1979); operationalized as one single-word descriptor.
Exhaust	"To include most or all elements of the query in the.. search formulation."	Originally defined by Bates (1979); operationalized as four or more terms combined with ANDs.
Weight 3	"Limit free-text terms to occur in a predetermined field." (This category includes terms limited to any of the descriptor fields. Those limited to Language, Update, Subset, or Title fields are covered by Limit 1, 2, 3, and 4, respectively. Those limited by publication type are covered by Weight 5.)	Originally defined by Fidel (1985); also included as a move to reduce the size of the set.
Weight 4	"Require that free-text terms occur closer to one another in the searched text."	Originally defined by Fidel (1985); operationalized as term phrases; also included as a move to reduce the size of the set.
Expand 2	"Group together search terms to broaden the meaning of a set."	Originally defined by Fidel (1985); operationalized as multiple terms combined with ORs; also included as a move to increase the size of the set.
Truncate	Truncated term.	Also included as a move to increase the size of the set.

MOVES TO REDUCE THE SIZE OF THE SET

Move	Definition	Notes
Intersect 1	"Intersect a set with a set representing another query component."	Originally defined by Fidel (1985).
Limit 1	"Limit to documents written in a particular language."	Originally defined by Fidel (1985).
Limit 2	"Limit to documents published, or indexed, in a particular period of time."	Originally defined by Fidel (1985).
Limit 3	"Limit to documents retrieved from a specific portion of the database."	Originally defined by Fidel (1985).
Limit 5	Limit to studies on humans.	Operationalized as any subset identified with a checktag.
Weight 5	"Limit to documents of a certain form."	Originally defined by Fidel (1985).
Negate/Block	"Eliminate unwanted elements by using the.. NOT operator."	Defined earlier by both Bates (1979) and Fidel (1985).
Narrow 1	Intersect a pre-existing set with a set created by more specific terms. (Adapted from original definition.)	Originally defined by Fidel (1985); "more specific terms" operationalized as narrower terms from the MeSH tree.
Weight 2	Intersect pre-existing set with a broader term. (Adapted from original definition.)	Originally defined by Fidel (1985); "broader term" operationalized as broader terms from the MeSH tree.
Sub	"To move downward hierarchically to a more specific (subordinate) term."	Originally defined by Bates (1979).
Weight 1	"Limit a descriptor to be a major descriptor."	Originally defined by Fidel (1985).
Weight 3	"Limit free-text terms to occur in a predetermined field." (See listing in *Beginning moves* for details.)	Originally defined by Fidel (1985).
Limit 4	"Limit to sources that have, or do not have, a certain term in their titles."	Originally defined by Fidel (1985).
Weight 4	"Require that free-text terms occur closer to one another in the searched text."	Originally defined by Fidel (1985); operationalized as term phrases.
Narrow 2/ Intersect 2	"Qualify descriptors with role indicators [or] intersect sets with role indicators."	Originally defined by Fidel (1985); operationalized as the inclusion of subheadings.

MOVES TO INCREASE THE SIZE OF THE SET

Move	Definition	Notes
Reduce	"To subtract one or more of the query elements from an already-prepared search formulation."	Originally defined by Bates (1979); operationalized as the repetition of a set minus at least one term.
Cancel	"Eliminate restrictions previously imposed," such as restricting the search to particular fields, use of proximity operators, or limitations imposed with the "limit" function.	Originally defined by Fidel (1985).
Include	"Group together a descriptor with all the descriptors that are its narrower terms."	Originally defined by Fidel (1985).
Add 1/Parallel	"To make the search formulation broad (or broader) by including synonyms."	Defined earlier by Bates (1979) and Fidel (1985).
Add 2	"Add descriptors as free-text terms."	Originally defined by Fidel (1985).
Expand 1/Super	"Enter [substitute] a broader descriptor."	Earlier defined by both Bates (1979) and Fidel (1985); "broader descriptor" operationalized as broader term from the MeSH tree.
Expand 2	"Group together search terms to broaden the meaning of a set."	Originally defined by Fidel (1985).
Truncate	Truncate a term.	

MOVES TO INCREASE BOTH PRECISION AND RECALL

Move	Definition	Notes
Relate	"To move sideways hierarchically," i.e., to substitute a related term.	Originally defined by Bates (1979); "related term" operationalized based on the top two levels of the MeSH tree.
Vary	To substitute one term for another, with no change in the number of terms; the new term may be unrelated to the original term.	Originally defined by Wildemuth (1992).
Fix	"To try alternative affixes, whether prefixes, suffixes, or infixes."	Originally defined by Bates (1979); truncation coded as a move to increase the size of the set.
Respell	"To search under a different spelling" of a term.	Originally defined by Bates (1979); also includes her monitoring tactic, Correct, i.e., to correct spelling errors.
Respace	"To try spacing variants."	Originally defined by Bates (1979).

ERRORS AND OTHER MOVES

Move	Definition	Notes
SPFlash	SilverPlatter flashback: To use SilverPlatter syntax that does not work in UNCLE.	
Typo	To mistype a search term.	
Syntax	To use the wrong syntax in a search statement.	
Repeat	To use a search statement that was used in the previous move.	This move was considered an error.
System	An inconsistency in system performance caused mis-execution of a search statement.	
Neighbor	"To seek additional search terms by looking at neighboring terms, whether proximate alphabetically, by subject similarity, or otherwise."	Originally defined by Bates (1979).

Appendix C. Categories for coding moves based on Shute & Smith (1993)

The categories are based on the knowledge-based search tactics defined by Shute and Smith (1993). In this coding scheme, the idea of frames, made up of slots populated with fillers, is used to represent the concepts of a search strategy represented by particular terms. A slot is a particular search concept; a slot-filler is a term representing that concept.

BEGINNING MOVES

Move	Definition	Notes
Database	Select a specific database	Operationalized as the first move of each day/session and at change of database. Also included rerunning the same search in another database.
New slot	Enter term(s) for a concept that was not included in previous cycle.	

MOVES TO REDUCE THE SIZE OF THE SET

Move	Definition	Notes
Combine	Combine two pre-existing slots using AND.	The slots were referred to by set number and did not include a reference to the previous search cycle.
Add slot	"Add a slot-filler for a slot that is not represented in the [previous search cycle] (using AND)."	
Exclude	"Exclude a slot-filler (using NOT)."	
Narrow slot-filler	"Replace a slot-filler with a narrower slot-filler in the same slot."	
Narrow operator	Replace an operator with a narrower operator.	For example, AND might be replaced with NEAR.

MOVES TO INCREASE THE SIZE OF THE SET

Move	Definition	Notes
Delete slot	"Delete a slot (that was ANDed) from the [previous search cycle]."	
Broaden slot-filler	"Add a broader slot-filler to a slot already represented in the [previous search cycle] (using OR)."	This move might also involve replacing a slot-filler with a broader slot-filler.
Broaden operator	Replace an operator with a broader operator.	For example, NEAR might be replaced with AND.
Combine with OR	"Add a slot-filler to a slot that is not filled in the [previous search cycle] (using OR)."	

MOVES TO INCREASE BOTH PRECISION AND RECALL

Move	Definition	Notes
Replace slot-filler	"Replace a slot-filler with a sibling/cousin slot-filler (in the same slot)."	The new slot-filler is not in a hierarchical relationship (broader/narrower) to the slot-filler being replaced.

ERRORS AND OTHER MOVES

Move	Definition	Notes
Error	Typographical, syntactic and other types of errors.	Includes all the types of errors delineated in Appendix B.
Neighbor	Check the online thesaurus/index for (alphabetically or semantically) related terms.	The same move as "Neighbor", defined by Bates (1979).

Appendix D. Sample coding of a search

Search log	Number retrieved	Codes, based on Bates and Fidel	Codes, based on Shute & Smith
/dev/ttyp5		Database	Database
001, mitral-regurgitation	0	Weight 3	New slot
002, murmurs	1175	Select	New slot
003, mitral	9669	Select	Broaden slot-filler
004, 2 and3	0	Intersect 1; Typo	Error
005, 2 and 3	375	Respace	Combine
006, clinical	317879	Select	New slot
007, 5 and 5	375	Intersect 1; Typo	Error
008, 5 and 6	133	Intersect 1	Add slot
009, diagnosis	323061	Select	New slot
010, 8 and 9	92	Intersect 1	Add slot
011, clinical and diagnosis and aortic and murmurs	76	Exhaust	Add slot; Delete slot
012, aortic and murmurs and diagnosis	179	Reduce	Delete slot

Appendix E. Evaluation form used by librarian raters

SEARCH EVALUATION FORM

1. MISSED OPPORTUNITIES

The attached search has been segmented into individual search moves. Examine each move and identify those which are "missed opportunities."

A move is considered a *missed opportunity* if it could have been improved in some significant way. For example,

- Truncation was not used,
- The searcher failed to use appropriate MeSH headings,
- A MeSH heading was used, but without the appropriate punctuation to search the MJ and MN fields,
- The searcher specified only a single field, such as Title, when other fields would also have been appropriate, or
- The searcher failed to explode a term when appropriate.

Please mark the missed opportunities with an asterisk. Briefly explain each missed opportunity, identifying the move which you believe would have been more appropriate.

2. OVERALL EVALUATION

Please provide your overall assessment of the quality of the attached search on each of the following criteria:

	Poor		OK		Excellent	
Initial selection of term(s)	1	2	3	4	5	N/A
Use of Boolean operators to combine terms and sets of terms	1	2	3	4	5	N/A
The use of system feedback to narrow or broaden the search	1	2	3	4	5	N/A
The correct use of system syntax and commands	1	2	3	4	5	N/A
Use of the online thesaurus	1	2	3	4	5	N/A