
End-user search behaviors and their relationship to search effectiveness*

By Barbara M. Wildemuth, Ph.D.
Associate Professor
School of Information and Library Science
Manning Hall, CB# 3360

Margaret E. Moore, M.P.H., A.M.L.S.
Head, Education Services
Health Sciences Library
CB# 7585

University of North Carolina at Chapel Hill
Chapel Hill, North Carolina 27599-3360

One hundred sixty-one MEDLINE searches conducted by third-year medical students were analyzed and evaluated to determine which search moves were used, whether those individual moves were effective, and whether there was a relationship between specific search behaviors and the effectiveness of the search strategy as a whole. The typical search included fourteen search statements, used seven terms or "limit" commands, and resulted in the display of eleven citations. The most common moves were selection of a database, entering single-word terms and free-text term phrases, and combining sets of terms. Syntactic errors were also common. Overall, librarians judged the searches to be adequate, and students were quite satisfied with their own searches. However, librarians also identified many missed opportunities in the search strategies, including underutilization of the controlled vocabulary, subheadings, and synonyms for search concepts. No strong relationships were found between specific search behaviors and search effectiveness (as measured by the librarians' or students' evaluations). 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 is known of end users' searches as mediated by information professionals, the search process involves several stages, beginning with 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 [1]. There is a need for increased understanding of how end users formulate and reformulate search strategies, particularly to as-

sist in the work of two groups: system specialists, who support the process through design of information retrieval systems, and librarians, who provide education and consultation in searching techniques. Therefore, a research study was undertaken that analyzed medical students' MEDLINE searches in detail, describing and evaluating the individual moves they made.

BACKGROUND

One approach to the study of search strategy formulation is to examine and categorize the individual moves made by a searcher. Bates identified twenty-nine search tactics, including those for monitoring search progress, optimizing use of system-file structure, formulating and reformulating the search, and selecting and revising specific terms [2-3]. These tactics provide a strong framework for examining search

* Funding for this research was provided by the Council on Library Resources Cooperative Research Program (Grant no. CLR 8040).

moves, although they have not been validated with empirical data from online bibliographic searching. The tactics have been found to be useful in categorizing moves made by medical students in searches of a factual database supporting their microbiology instruction [4-5].

A different set of categories was generated empirically by Fidel, based on observations of information professionals conducting bibliographic searches [6]. This set of thirty categories included moves to reduce the size of a retrieved set, increase the size of the set, and increase precision and recall simultaneously. Because they were generated from the activities 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 [7-8].

A knowledge-based approach to the categorization of search moves has been developed by Shute and Smith [9]. This approach views the search as a frame, in which concepts occupy slots and specific representations of those concepts (i.e., specific terms) fill the slots. Based on detailed observation of one search intermediary, this approach may or may not be applicable to end-user searching.

Analysis of errors made in search formulation is another way of examining end-user search moves. Sewell and Bevan analyzed errors made by pharmacists and pathologists searching TOXLINE and MEDLINE [10]. The most common errors were related to misspelled terms and misuse of the controlled vocabulary. In another study, users of BRS/After Dark had trouble "understanding the contents and structure of a database, understanding the use of appropriate search terms, and understanding Boolean logic" [11]. A recent examination of end users' "unproductive searches" of MEDLINE revealed problems associated with formulating the search and inappropriate use of the features in GRATEFUL MED [12]. Miller et al. studied both end-user searching errors and missed opportunities [13]. They found that 37% of the 500 search statements examined contained at least one error (resulting in no items retrieved) and that more than 75% of the search statements represented missed opportunities. Clearly, there is substantial room for improvement in end users' formulation of search strategies.

As Walker et al. point out, some of the problems experienced by end users in searching CD-ROM and online databases are associated with system design [14]. The complexity of representing an information need to a retrieval system often is 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 do

users who approach a system on a regular basis, yet interfaces meant for end users are usually identical to those intended for professional searchers. Collection of additional data describing end users' actual use of a database will help in improving the design of end-user interfaces for information retrieval systems.

Such information also would be useful to those who advise end users in search techniques. A series of studies at the University of North Carolina at Chapel Hill (UNC) indicated that students find MEDLINE searches to be useful in preparing case presentations [15-18]. Therefore, in spite of the long-standing debate on the value of user education [19], successively more advanced MEDLINE training has been integrated into the UNC School of Medicine curriculum each year. Collection of additional data on students' search strategy formulation and reformulation can help guide the development of future education and consultation services.

RESEARCH QUESTIONS

This study posed three questions: What are students doing during their searches? Are their searches effective? Are particular search behaviors related to search effectiveness?

The results provide a description of student search behaviors: which moves were used most frequently, the number of search statements in each search, the number of terms used in each search, and the number of citations displayed. The results also indicate the effectiveness of the students' searches. Finally, the results test the relationship between specific search behaviors and search effectiveness.

METHODS

Data collection

In 1992/93, during their third year of medical school, students in the internal medicine and pediatrics clerkships at UNC were required to search MEDLINE for patient care information. Participants attended brief MEDLINE orientation sessions given by the Health Sciences Library (HSL) staff. The objective of this program was to introduce students to the use of MEDLINE to find journal literature relevant to patient care.

One hundred sixty-one searches were completed between September 1992 and March 1993. Most of the searches (150) were conducted through the UNC Literature Exchange (UNCLE) service; UNCLE uses BRS searching software to make MEDLINE available through the campus network. The remaining eleven searches were conducted with MEDLINE on SilverPlatter compact discs. Some differences in search behaviors across the two systems were found and are

reported. Because each student had a unique information need, each search addressed a different topic.

As students performed their searches, search strategies and results were captured. For SilverPlatter searches, students printed out the strategy and results; for UNCLE searches, logs of the strategy were captured automatically, and the student printed out the results. Students then gave the searches to the clerkship coordinator, who in turn gave them to the library's Clinical Health Information Retrieval Program (CHIRP) coordinator to review. Prior to returning the search output to the student, the results were 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-type scale.

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 then were coded in two ways. First, using the moves or tactics identified by Bates [20–21] and Fidel [22], two members of the research team classified each search statement. The possible moves are defined briefly in Appendix A; for more details, see Wildemuth and Moore's "End User Searching of MEDLINE: Final Report" [23]. Second, one member of the research team coded the moves using Shute and Smith's method, based on changes in concept slots and fillers [24]. The possible codes are defined in Appendix B; for more details, see Wildemuth and Moore. [25]. Both coding schemes were pretested with a sample of six searches.

Three professional health sciences librarians, all experienced searchers, independently evaluated each search. Using five-point Likert scales, they 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, and the correct use of system syntax. Based on their expertise in searching MEDLINE, the three librarians also identified instances in which the student missed an opportunity to improve the search strategy [26].

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 statements per search, the average number of terms per search, the average number of citations displayed,

and the frequency of each type of move were calculated.

The second research question relates to the quality of the searches. Students' ratings of their satisfaction with each search yielded a self-evaluation of the quality of their searching behaviors. For this measure, averages of the responses were calculated (on a six-point scale) 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." The professional searchers' ratings of student searches provided an external evaluation. For each search, the ratings from the three professional searchers were averaged. In addition, the professional searchers' descriptions of a student's missed opportunities were analyzed qualitatively to identify both the most frequent errors and those with the most serious consequences.

Finally, the relationship between search behaviors and search effectiveness was examined. The quantitative descriptions of student search behaviors (number of search statements, etc.) were treated as independent variables, along with the frequency of each type of move (based on the Shute and Smith categorization [27]). In addition, the researchers took into account student characteristics, such as their training and prior experience with databases and their undergraduate majors. The effect of these variables on the measures of search effectiveness was tested with stepwise linear regression. It could be argued that the dependent variables were ordinal—rather than interval-level variables and that logistic regression would be more appropriate. However, because of the ease of data interpretation and the likelihood that the results would be essentially the same regardless of method, linear regression was used.

RESULTS

Student searching behaviors

Based on *t* test results, there were no statistically significant differences in student search behaviors between the students who returned questionnaires with their searches and students who did not, so descriptive statistics for all 161 searches were combined (Table 1). Students averaged fourteen statements in each search. They used six different terms in a typical search. The "limit" function was used relatively infrequently—only approximately once per search. Eleven citations were printed per search, on average.

There were some statistically significant ($P < 0.05$) differences between the 11 SilverPlatter searches and the 150 UNCLE searches. Students using UNCLE averaged more terms per search (6.5 versus 3.9 terms; $t = -4.4849$; $P < 0.01$) and used the limit function more often (1.4 times per search versus 0.2; $t = -5.8448$; $P < 0.01$) than did SilverPlatter users. Both of these differences probably are related to the way in which

the search logs were captured. For the SilverPlatter searches, only the printed search strategies handed in by the students were analyzed; for the UNCLE searches, multiple sessions relating to the search topic could be included in the analysis. Thus, a log printed for a SilverPlatter search was equivalent to the last session of an UNCLE search.

The moves used in all 161 searches, based on the tactics and moves defined by Bates [28–29] and Fidel [30], are reported in Table 2. The total number of moves was greater than the total number of search statements, because a student could make several changes in the search in one statement. All searches began with the Database move, because both systems require that a database be selected. The move used most frequently was Intersect 1, the combination of a set of terms or previously specified sets. Another common move was Weight 4, the use of term phrases and proximity operators. (The NEAR proximity operator is used by default on UNCLE searches when a term phrase is entered.) The third most frequent move was Select, the specification of a single-word term. Additional common moves included 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 common moves (except Database and Select) are tactics for reducing the size of the retrieved set. Syntax errors also were relatively common.

There were a few statistically significant differences in search moves between the students who returned questionnaires and those who didn't [31–33]. The students not completing questionnaires used the Vary move more often, substituting one term for another (0.7 times per search versus 0.2; $t = 3.4171$; $P < 0.01$). 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; $t = 2.4541$, $P = 0.02$). All sixteen uses of Limit 4, limiting terms to the title field, were by seven students who did not complete the questionnaire.

There were also a few statistically significant differences between the moves used with SilverPlatter and the moves used with UNCLE. The Select move, specifying a single-word term, was used more often in SilverPlatter searches (3.5 times per search versus 1.6; $t = 2.3908$; $P = 0.02$). Several moves were used more frequently in UNCLE searches: Limit 1, limiting by language (1.0 times per search versus 0.3; $t = -3.2835$; $P < 0.01$); Weight 3, limiting free-text terms to occur in a specified field (0.9 times per search versus 0.3; $t = -3.0486$; $P < 0.01$); Weight 5, limiting a search by publication form (0.7 times per search versus 0.2; $t = -2.5766$; $P = 0.02$); and typographical errors (0.5 times per search versus 0.1; $t = -3.4494$, $P < 0.01$).

Numerous other moves were used at least twice in

Table 1
Student searching behaviors

Variable	n	Mean	SD	Maximum	Median	Minimum
All searches						
Number of statements	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 statements	58	13.7	10.0	50	10	4
Number of terms	58	5.9	4.2	26	5	2
Use of limit (number of times)	58	1.3	2.1	12	1	0
Number of items printed	57	11.1	8.2	36	10	1

UNCLE but not at all in SilverPlatter searches: Add 2, Cancel, Exhaust, Expand 2, Fix, Limit 2, Limit 4, Limit 5, Mode, Narrow 2, Negate/Block, Neighbor, Respace, 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 whether additional searches would have included these moves. All these moves are syntactically possible on the SilverPlatter system.

Table 3 shows how the moves are categorized according to the Shute and Smith coding scheme [34]. 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 concept slot to the search. This move implies that a student included a new concept as part of a search statement that already contained one or more concepts. Deleting a concept slot, which involves repeating a search statement minus one of the concepts, was the next most common move. A fourth common move was to combine existing concept slots with the "AND" operator. This type of move is common in the "building-block" approach, in which individual concepts are specified, each in a separate step, then combined [35]. Moves involving the manipulation of concept slot fillers did not occur nearly as frequently as did moves manipulating concept slots. Students rarely used the "NOT" operator, "OR" to combine concept slots, the online thesaurus and index, or moves involving the manipulation of operators. On the other hand, both syntactical and typographical errors were common (the reported frequencies do not include "missed opportunities" identified by librarian evaluators).

There were no statistically significant differences between moves (adapted from Shute and Smith) used in searches accompanied by questionnaires and those not accompanied by questionnaires. However, there were several statistically significant differences between SilverPlatter and UNCLE searches. The UN-

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

Move	Students using move	Total uses	Mean frequency	SD	Maximum	Median	Minimum
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							
Narrow 1	1	1	0.0	0.1	1	0	0
Intersect 1	150	642	4.0	3.5	16	3	0
Narrow 2/Intersect 2	15	28	0.2	0.7	6	0	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 4	7	16	0.1	0.5	4	0	0
Limit 5	5	7	0.0	0.3	3	0	0
Weight 1	10	43	0.3	1.7	18	0	0
Weight 3	53	132	0.8	1.6	8	0	0
Weight 4	109	297	1.8	2.2	9	1	0
Weight 5	55	110	0.7	1.2	7	0	0
Negate/Block	5	6	0.0	0.2	2	0	0
Sub	2	2	0.0	0.1	1	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
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
Truncate	11	19	0.1	0.5	3	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							
Typo	48	75	0.5	0.9	4	0	0
Syntax	49	112	0.7	1.5	11	0	0
SilverPlatter flashback	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

CLE users added concept slots to their searches more often (4.2 times per search versus 1.3 for the SilverPlatter users; $t = -5.0229$; $P < 0.01$). The UNCLE users also replaced concept slot fillers with other slot fillers more often than did SilverPlatter users (0.9 times per search versus 0.3; $t = -3.1455$; $P < 0.01$). Only UNCLE users replaced a concept slot filler with a broader slot filler, replaced an operator with a broader operator, used the "OR" operator to combine concept slots, used the "NOT" operator, or checked the online thesaurus or index, all moves that are syntactically possible on the SilverPlatter system but for some reason were not used. These differences, as in the case of the number of terms and the number of limit commands reported earlier, may be due to the way in which the searching data were captured rather than system characteristics.

Search effectiveness

Search effectiveness was evaluated in three ways. Librarians evaluated four dimensions of the quality of the students' searches, the students evaluated themselves, and librarians noted missed opportunities in the students' search strategies. Fifty-eight searches (the number of searches accompanied by questionnaires) were included in the analysis reported here.

Librarian ratings. Before combining the sets of ratings from the three librarians, the inter-rater agreement was investigated. The Pearson correlations across librarians within each dimension ranged from 0.32 (between the first two raters, on the use of feedback to modify the search) to 0.67 (between the second and third raters, on the correct use of system syntax). In

Table 3
Frequency of moves (Shute and Smith), all searches (n = 161)

Move	Students using move	Total uses	Mean frequency	SD	Maximum	Median	Minimum
Beginning moves							
Database selection	161	267	1.7	1.3	8	1	1
New concept slot (initial set)	161	563	3.5	2.9	24	3	1
Moves to reduce the size of the set							
Combine concept slots with "AND"	87	204	1.3	2.0	15	1	0
Add concept slot(s)	140	650	4.0	3.3	16	3	0
Exclude ("NOT" operator)	5	7	0.0	0.3	2	0	0
Replace concept slot-filler with narrower slot-filler	55	99	0.6	1.2	8	0	0
Replace operator with narrower operator	10	12	0.1	0.3	2	0	0
Moves to increase the size of the set							
Delete concept slot(s)	98	277	1.7	2.0	10	1	0
Replace concept slot-filler with broader slot-filler	48	90	0.6	1.2	8	0	0
Replace operator with broader operator	9	9	0.1	0.2	1	0	0
Combine concept slots with "OR"	3	3	0.0	0.1	1	0	0
Moves to increase both precision and recall							
Replace concept slot-filler with other slot-filler	54	138	0.9	1.7	12	0	0
Errors and other moves							
Errors	87	231	1.4	2.5	19	1	0
Check index/thesaurus	7	10	0.1	0.3	3	0	0

addition, Cronbach's α was calculated, to estimate the internal consistency of the three librarians' ratings of each search. Cronbach's α ranged from 0.69 (on the use of feedback to modify the search) to 0.80 (on the correct use of system syntax) and was considered acceptable for the purposes of this research.

The results, as reported in Table 4, indicate that students' searches were adequate, receiving average ratings of approximately 3 (meaning "okay") on all four dimensions. The individual ratings covered the entire range of the five-point scale on all four dimensions.

Students' self-evaluations. Student estimates of their own performance were measured by 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 (Table 3) indicate that students generally

were satisfied with their searches. Because these two items were strongly related (Pearson's $R = 0.61$; Cronbach's $\alpha = 0.75$), they later were averaged for the purposes of the regression analysis.

Missed opportunities. The third measure of search performance was the librarians' identification of missed opportunities. These missed opportunities were cumulated across the librarians and then categorized by a member of the research team. Fifty-six (97%) of the fifty-eight searches evaluated contained missed opportunities of some kind (Table 5). The most common missed opportunity by far was failure to exploit the controlled vocabulary, such as by substituting Medical Subject Headings (MeSH) terms for free-text terms or by exploding a MeSH term. A related opportunity that was missed frequently was the use of subheadings. In addition, many of the searches included illogical Boolean combinations, such as intersecting a set with another set containing no hits.

Table 4
Measures of search effectiveness (n = 58)

Variable	Mean	SD	Maximum	Median	Minimum
Librarians' ratings (5-point scale)					
Initial selection of terms	2.8	1.0	4.7	2.7	1.0
Use of Boolean operators	3.2	0.8	4.7	3.3	1.7
Use of feedback to narrow or broaden search	3.0	0.8	4.7	3.0	1.3
Correct use of system syntax	3.1	0.9	4.7	3.2	1.3
Students' self-evaluations (6-point scale)					
I found what I was looking for in this search.	1.7	1.1	6	1	1
This search was an efficient use of my time.	1.9	1.3	6	1	1

Table 5
Missed opportunities (n = 58)

Missed opportunities	No. of students	Total frequency	Mean frequency	SD	Maximum	Median	Minimum
Should use the full database	2	2	0.0	0.2	1	0	0
Should use MeSH term	42	90	1.6	1.4	6	1	0
Should explode MeSH term	16	24	0.4	0.8	4	0	0
Should limit term to major descriptor	11	14	0.2	0.6	4	0	0
Should not use MeSH term (none available)	3	3	0.1	0.2	1	0	0
Should add synonyms with "OR"	10	12	0.2	0.5	2	0	0
Should truncate term/use truncation symbol	10	11	0.2	0.4	2	0	0
Should use broader term	5	7	0.1	0.5	3	0	0
Should use narrower term	6	6	0.1	0.3	1	0	0
Should use subheading	23	31	0.5	0.8	4	0	0
Should limit term to a specific field	10	10	0.2	0.4	1	0	0
Should limit to specific age groups	2	2	0.0	0.2	1	0	0
Should use a different proximity operator	18	26	0.4	0.8	3	0	0
Made an illogical Boolean combination	24	34	0.6	0.9	4	0	0
Other missed opportunities	2	2	0.0	0.2	1	0	0

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, or the measures of search effectiveness. The independent variables included the number of statements per search, the number of terms used (including limit functions), the number of citations displayed, and the frequency of each type of move [36].

In addition to the independent variables, several student characteristics were included in the regression equation to determine their effect. The first set of characteristics concerned a student's prior experience with three types of databases, rated on a four-point ordinal scale of frequency of use. These three types were personal computer database software (34% of the students had no experience), a locally developed factual database in bacteriology (20% of the students had no experience), and online bibliographic databases (all of the students had used them at least once). Because the second and third types were closely related ($X^2 = 13.608$, with 6 *df*; $P = 0.03$), experience with the local bacteriology database and experience with online bibliographic databases were combined into one variable for the regression analysis. The second background variable included in the regression analysis was a student's undergraduate major (science versus nonscience). Sixty-six percent of the students had an undergraduate degree in the natural or physical sciences.

Stepwise linear regression analysis was used to identify models that would predict each of the five dependent variables (initial selection of terms, use of Boolean operators, use of feedback to reformulate a

search strategy, 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 statements 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 the local bacteriology database and online bibliographic databases.

Some relationships were found, but, in each case, the prediction was weak and only marginally significant. In general, fewer statements in a search predicted greater success in initial selection of terms ($R^2 = 0.05$, $P > F = 0.09$). Similarly, students' past experience with the local bacteriology database and online bibliographic databases predicted their success in using Boolean operators ($R^2 = 0.13$; $P > F = 0.07$). An increased number of errors was associated with students' ability to use feedback to reformulate their searches ($R^2 = 0.06$; $P > F = 0.07$). The strongest relationship was found between the number of errors and the students' inability to correctly use system syntax ($R^2 = 0.17$; $P > F = 0.002$). None of the independent variables predicted students' evaluations of their own performance.

DISCUSSION

The results describing students' search behaviors provide a detailed view of the online searching process. A typical search involves fourteen statements, incorporates approximately seven different terms or concepts, and results in the retrieval of about eleven citations. The search is likely to incorporate the se-

lection of a database; selection of single-word terms, free-text term phrases, and phrases that appear in a particular field; combination of terms and phrases with the Boolean "AND" operator; and limitation of the output by language and publication form. Unfortunately, the search 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. The search also 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 search outcomes. Students' initial selection of terms was adequate, but it could be improved through increased use of an online thesaurus and expanded awareness of the importance of including synonyms in specifying each search concept. Syntactical and typographical errors lowered search performance even though they 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 improve outcomes in many cases. Students' self-evaluations indicated that most either were unaware of these problems in their search performance or were satisfied with the outcomes of their searches in spite of the problems.

This study did not find any strong links between particular search behaviors and the ratings of search performance. It seems that individual searches can be evaluated and recommendations made concerning how to improve them, but no generalizations can be made about the relationship between search performance and the number of statements executed, terms used, citations retrieved, or types of moves used. One avenue for further exploration would be to consider larger chunks of searching behavior; that is, to analyze the searches in terms of sequences of moves within a search, rather than the individual moves. Hsieh-Yee made a relevant point, noting the difficulty of analyzing complete search strategies and suggesting that sub-sequences of moves be the unit of analysis [37]. As independent variables in a regression equation, frequencies of individual moves were too weak to predict search performance.

IMPLICATIONS 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, student search performance could be improved if the number of syntactical errors were reduced. One way to effect this improvement would be to design systems that are more tolerant of variations

in syntax than are current systems. Some progress is being made in this area, as systems increasingly 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 able to focus on the substance of their searches rather than on 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 a 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; and students did not attempt to generate synonyms to specify fully a concept of interest. Each of these problems 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 through either system design or user education. If the online thesaurus were linked more closely to the search engine, the system could suggest synonyms from a controlled vocabulary when a term is entered. Common acronyms also can be added to the controlled vocabulary to ensure that users include both versions of a concept in their searches. Selection of terms from a list of possible synonyms is likely to be a more successful means of developing a coherent search strategy than is use of one representation of a concept selected from personal knowledge—particularly for students who are new to a domain. This problem also can 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. Almost all features of the information retrieval systems were used at least once. Each student may rely on only a few moves, but this group of students used more than thirty different kinds of moves, not including errors. For system designers, this finding implies that it is indeed worthwhile to make a wide range of features available, because at least some system users find 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 local users and can guide the development of customized training programs. At

UNC, these results will be used in these ways to identify needed UNCLE system enhancements and to help librarians develop advanced training, help screens, and user aids.

FUTURE RESEARCH

The results reported here are preliminary, in the sense that the data warrant further analysis. The results might become more meaningful if the search moves were re-analyzed using short sequences of moves as the unit of analysis. The starting point for such an analysis could be the search strategies outlined in Markey and Atherton [38]. The use of a slightly larger unit of analysis could prove fruitful in exploring the relationships between search behaviors and search outcomes.

ACKNOWLEDGMENTS

The authors wish to express their appreciation to five other members of the research team. Nancy Ryan, HSL graduate assistant and CHIRP coordinator, helped collect the data. Rikki Mangrum, research assistant, was instrumental in coordinating the coding of the search moves and preparing the data for analysis. HSL librarians Julia Shaw-Kokot, Cathryn Chiesa, and Bridget Loven were very generous with their time, evaluating the students' searches and identifying missed opportunities. Without the participation of each team member, this research project could not have been completed successfully.

REFERENCES

1. BELKIN NJ, VICKERY A. Interaction in information systems: a review of research from document retrieval to knowledge-based systems. London: British Library, 1985. (Library and information research report 35).
2. BATES MJ. Information search tactics. *J Am Soc Inf Sci* 1979;30(4):205-14.
3. BATES MJ. Search and idea tactics. In: White HD, Bates MJ, Wilson P, eds. For information specialists: interpretations of reference and bibliographic work. Norwood, NJ: Ablex, 1992:183-200.
4. WILDEMUTH BM, JACOB EK, FULLINGTON A, DE BLIEK R ET AL. A detailed analysis of end-user search behaviors. In: Proceedings of the Fifty-Fourth Annual Meeting of the American Society for Information Science. Medford, NJ: Learned Information, 1991:302-12.
5. WILDEMUTH BM, DE BLIEK R, HE S, FRIEDMAN CP. Search moves made by novice end users. In: Proceedings of the Fifty-Fifth Annual Meeting of the American Society for Information Science. Medford, NJ: Learned Information, 1992:154-61.
6. FIDEL R. Moves in online searching. *Online Rev* 1985; 9(1):61-74.
7. WILDEMUTH, A detailed analysis of end-user search behaviors.

8. WILDEMUTH, Search moves made by novice end users.
9. SHUTE SJ, SMITH PJ. Knowledge-based search tactics. *Inf Proc Manage* 1993;29(1):29-46.
10. SEWELL W, BEVAN A. Nonmediated use of MEDLINE and TOXLINE by pathologists and pharmacists. *Bull Med Libr Assoc* 1976;64(4):382-91.
11. SLINGLUFF D, LEV Y, EISAN A. An end user search service in an academic health sciences library. *Med Ref Serv Q* 1985;4(1):11-21.
12. WALKER CJ, MCKIBBON KA, HAYNES RB, RAMSDEN MF. Problems encountered by clinical end users of MEDLINE and GRATEFUL MED. *Bull Med Libr Assoc* 1991;79(1):67-9.
13. MILLER N, KIRBY M, TEMPLETON E. MEDLINE on CD-ROM: end user searching in a medical school library. *Med Ref Serv Q* 1988;7(3):1-13.
14. WALKER, op. cit.
15. MOORE ME. Clinical Health Information Retrieval Project: program evaluation report, 1989 [Unpublished paper]. Presented at the Librarians' Association of the University of North Carolina at Chapel Hill Research Forum, May 3, 1990.
16. RYAN N. Fostering skills in information retrieval: a study of the training and support needs of third year medical students searching the UNCLE MEDLINE database in support of their clinical clerkship experiences [Master's thesis]. Chapel Hill: School of Library and Information Science, University of North Carolina at Chapel Hill, 1993.
17. MOORE ME. UNCLE MEDLINE baseline survey report, 1992. [Unpublished].
18. MOORE ME. UNCLE MEDLINE two-year follow-up report, 1994. [Unpublished].
19. EADIE T. Immodest proposals: user education for students does not work. *Libr J* 1990;115(17):42-5.
20. BATES, Information search tactics.
21. BATES, Search and idea tactics.
22. FIDEL, op. cit.
23. WILDEMUTH BM, MOORE ME. End user searching of MEDLINE: final report. August 1993. (ERIC Document No. ED 363 357).
24. SHUTE, op. cit.
25. WILDEMUTH, End user searching of MEDLINE.
26. MILLER, op. cit.
27. SHUTE, op. cit.
28. BATES, Information search tactics.
29. BATES, Search and idea tactics.
30. FIDEL, op. cit.
31. IBID.
32. BATES, Information search tactics.
33. BATES, Search and idea tactics.
34. SHUTE, op. cit.
35. MARKEY K, ATHERTON P. ONTAP: online training and practice manual for ERIC data base searchers. Syracuse, NY: Syracuse University, ERIC Clearinghouse on Information Resources, June 1978. (ERIC Document No. ED 160 109.)
36. SHUTE, op. cit.
37. HSIEH-YEE IPY. The search tactics of novice and experienced searchers: a comparative study [Dissertation]. Madison: University of Wisconsin, 1990.
38. MARKEY, op. cit.

Received August 1994; accepted January 1995

APPENDIX A

Categories for coding moves based on Fidel (1985), Bates (1979, 1992), and Wildemuth (1991, 1992)

Beginning moves

Database	Select a specific database.
Rerun	To search a new set of records with a pre-existing search statement.
Resume	To store a search strategy and resume it at a later search session.
Select	"To break complex search queries down into sub-problems and work on one problem at a time" (Bates, 1979).
Exhaust	"To include most or all elements of the query in the . . . search formulation" (Bates, 1979).

Moves to reduce the size of the set

Narrow 1	Intersect a pre-existing set with a set created by more specific terms (adapted from Fidel, 1985).
Intersect 1	"Intersect a set with a set representing another query component" (Fidel, 1985).
Narrow 2/Intersect 2	"Qualify descriptors with role indicators [or] intersect sets with role indicators" (Fidel, 1985).
Limit 1	"Limit to documents written in a particular language" (Fidel, 1985).
Limit 2	"Limit to documents published, or indexed, in a particular period of time" (Fidel, 1985).
Limit 3	"Limit to documents retrieved from a specific portion of the database" (Fidel, 1985).
Limit 4	"Limit to sources that have, or do not have, a certain term in their titles" (Fidel, 1985).
Limit 5	Limit to studies on humans.
Weight 1	"Limit a descriptor to be a major descriptor" (Fidel, 1985).
Weight 3	"Limit free-text terms to occur in a predetermined field" (Fidel, 1985).
Weight 4	"Require that free-text terms occur closer to one another in the searched text" (Fidel, 1985).
Weight 5	"Limit to documents of a certain form" (Fidel, 1985).
Negate/Block	"Eliminate unwanted elements by using the NOT operator" (Bates, 1979; Fidel, 1985).
Sub	"To move downward hierarchically to a more specific (subordinate) term" (Bates, 1979).

Moves to increase the size of the set

Reduce	"To subtract one or more of the query elements from an already-pre-
--------	---

Cancel	"Eliminate restrictions previously imposed" (Fidel, 1985).
Include	"Group together a descriptor with all the descriptors that are its narrower terms" (Fidel, 1985).
Add 1/Parallel	"To make the search formulation broad (or broader) by including synonyms" (Bates, 1979; Fidel, 1985).
Add 2	"Add descriptors as free-text terms" (Fidel, 1985).
Expand 1/Super	"Enter [substitute] a broader descriptor" (Bates, 1979; Fidel, 1985).
Expand 2	"Group together search terms to broaden the meaning of a set" (Fidel, 1985).
Truncate	Truncate a term.

Moves to increase both precision and recall

Relate	"To move sideways hierarchically" (Bates, 1979).
Vary	"To alter or substitute one's search terms in any of several ways" (Bates, 1979).
Fix	"To try alternative affixes, whether prefixes, suffixes, or infixes" (Bates, 1979).
Respell	"To search under a different spelling" of a term (Bates, 1979).
Respace	"To try spacing variants" (Bates, 1979).

Errors and other moves

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.
Mode	To use a command from a different mode; e.g., to use a print command while in search mode.
Repeat	To use a search statement that was used in the previous move.
System	An inconsistency in system performance caused misexecution of a search statement.
Neighbor	"To seek additional search terms by looking at neighboring terms, whether proximate alphabetically, by subject similarity, or otherwise" (Bates, 1979).

APPENDIX B

Categories for coding moves based on Shute and Smith (1993)

Beginning moves

Database	Select a specific database
----------	----------------------------

New concept slot Enter term(s) for a concept that was not included in previous cycle.

Moves to reduce the size of the set

Combine with "AND" Combine two pre-existing slots using "AND."

Add concept 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 concept slot-filler "Replace a slot-filler with a narrower slot-filler in the same slot."

Narrow operator Replace an operator with a narrower operator.

Moves to increase the size of the set

Delete concept slot "Delete a slot (that was 'AND'-ed) from the [previous search cycle]."

Broaden concept slot-filler "Add a broader slot-filler to a slot already represented in the [previous search cycle] (using 'OR')."

Broaden operator Replace an operator with a broader operator.

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

Replace concept slot-filler "Replace a slot-filler with a sibling/cousin slot-filler (in the same slot)."

Errors and other moves

Error Typographical, syntactic, and other types of errors.

Neighbor Check the online thesaurus/index for (alphabetically or semantically) related terms.