

The Effect of CD-ROM Instruction on Search Operator Use

Trudi E. Jacobson and Janice G. Newkirk

This study reports the results of a survey of 675 CD-ROM users, obtaining data from both questionnaires and actual search strategies. One of the authors' primary concerns involved investigating the effect of prior instruction or assistance on search strategy. The authors found a statistically significant but weak relationship between prior instruction and level of search skill. Other measures such as user status, department affiliation/major, age, sex, and number of previous CD-ROM uses were compared to search skill. Only user status and department affiliation/major were significantly correlated to search skill. Additional research is needed on the effectiveness of various types of CD-ROM instruction.



D-ROM searching has substantially increased the service demands in libraries since its advent almost a decade ago. Public service academic librarians spend a considerable amount of time assisting and instructing patrons in the use of these end-user databases. Yet, for all the time and money devoted to these activities, little research has been done on their effectiveness. The authors based their instruction and assistance approaches on what they discovered helping people with paper resources, adding, of course, information about the technology.

Research indicates that many library patrons do not search very well.¹ Con-

versely, research shows that patrons enjoy using CD-ROM indexes, whether they use them successfully or not. Research also indicates that librarians are assisting and instructing patrons in increasing numbers. (At the University at Albany, State University of New York, CD-ROM questions as a percentage of total reference questions have steadily increased at a rate of two percent per year since 1992.) But does such assistance and instruction affect patrons' search skills? Does the immense investment in patron assistance with CD-ROMs matter?

These may be two of the most important questions in academic library public services today, yet remarkably little re-

Trudi E. Jacobson is Coordinator of User Education Programs for the University Libraries at the University at Albany, State University of New York, and the late Janice G. Newkirk was CD-ROM Coordinator for the University Libraries at the University at Albany. The research reported here was assisted by a grant from the Spencer Foundation. The data presented, statements made, and views expressed are solely the responsibility of the authors. The authors would like to thank Rhonda Allen, a doctoral student at the University at Albany, and John Wuthrich for their assistance and valuable insights on the data analysis and on this article. The authors acknowledge the assistance of Thomas J. Galvin, Director of the Information Science Doctoral Program in the University at Albany's Nelson A. Rockefeller College of Public Affairs and Policy, in connection with the Spencer Foundation grant that supported this study.

search has directly addressed them. At the University at Albany, the authors thought it was time to examine these issues. With a grant from the Spencer Foundation, the authors set out to study what impact librarian instruction and assis-

Although the CD-ROM and end-user literature include evaluations of end-user searching . . . there are few studies that assess the impact of instruction on end-user searching.

tance were having on CD-ROM searching. We hoped to find out whether classes and one-on-one assistance from librarians had a considerable impact on end-user search proficiency. We hypothesized that patrons who had attended searching classes or who had been directly assisted by librarians (or trained library assistants) would be far more sophisticated searchers. Specifically, we wanted to test whether end-users having training or assistance would use more Boolean operators, field searches, and referrals to previous search terms than end-users who had no assistance or instruction.

Literature Review

Although the CD-ROM and end-user literature includes evaluations of end-user searching, accomplished by either surveying patrons on their satisfaction with their searches or actually analyzing the results of their searches, there are few studies that assess the impact of instruction on end-user searching. Domenica M. Barbuto and Elena E. Cevallos found that students generally do not use advanced searching techniques emphasized in the training sessions.² Rebecca Bostian and Anne Robbins examined PsycLIT search results from students who had a range of instructional experiences. They concluded that the only instruction that caused a significant difference in search capability was a live search demonstration.³ Dorothy F. Davis, who also studied the effectiveness of different types of

PsycLIT CD-ROM instruction, suggests that video instruction is the most effective method.⁴ Linda Stewart and Jan Olsen found that students given formal instruction on the CD-ROM version of ERIC did search more effectively.⁵ Their instruction concentrated on the use of Boolean logic and vocabulary control, and included a search demonstration.

Methodology

Project Design

The current project involved a two-part collection of data. One part involved administering a survey to all library patrons using SilverPlatter CD-ROMs during selected periods. This survey included a questionnaire (available from the authors) to collect basic demographic data, information on the respondents' experience with computers, and information on what type of prior instruction and/or assistance respondents had in CD-ROM searching. The second part of the data collection involved saving each respondent's search strategy and attaching it to the questionnaire. We employed three students to administer the survey during morning, afternoon, and evening hours, weekends as well as weekdays, from February through May of 1994. These students were available to answer questions that arose as respondents filled out the questionnaire. In all, the students approached more than a thousand patrons for the survey. We discarded several hundred surveys because the patrons indicated that they had received the immediate and direct assistance of a librarian with the search in question. We concluded that these search strategies would represent the librarians' rather than the patrons' efforts. We collected a total of 675 usable surveys.

Survey Instrument

The questionnaire was divided into four sections: section 1 gathered basic demographic data, and section 2 evaluated the respondents' computer experience. We

hypothesized that CD-ROM searchers with extensive experience using other computer applications are probably more effective searchers than those who are unfamiliar with computers. Section 3 asked about the purpose of, and satisfaction with, the current search. It also asked respondents to indicate how often they had searched using CD-ROMs in the past. We used this section minimally in our analysis. Section 4, the most important section of the questionnaire, identified the types of CD-ROM training and assistance that the respondents had previously received. Because we were trying to identify the most effective types of assistance and instruction, this last section of the survey contained thirty categories, which included various types of classroom instruction, as well as other forms of instructional assistance, classified by time and location of instruction.

The second part of the data collection involved the respondents' search strategy. Though the survey data were subject to the usual problems of self-reported data, we determined the search skill level based on students' actual search strategies. One of the authors coded all the strategies for the number of different Boolean operators, field searches, and referrals to previous search terms used. In the following discussion, references to "operators" or "Boolean operators" should be understood to include the concepts of field searches and referrals to previous search statements. Although we recognize that the use of a variety of search operators is not always necessary for a good or useful search, correct use of a variety of operators is a clear indication of knowledge of, or skill at, searching and should directly relate to good searching. In addition, these elements are easy to measure objectively.

The Survey Population

Although we had hoped to collect surveys from a representative cross section of the population that uses the Univer-

sity at Albany's CD-ROMs, the persons participating in the survey were predominantly female (almost 66%) and were overwhelmingly involved in research in the social sciences (over 28% in psychology, over 21% in education, and over 14% in other social science disciplines) (see table 1).

In other words, over 60 percent of those included in the survey were working in one of these three areas. We cannot say for certain how representative our survey

TABLE 1
Summary of Characteristics for Survey Respondents

Age	% of respondents
0-17	0.6
18-25	59.3
26-34	22.1
35 or older	17.9
Missing	0.1
Sex	% of respondents
Female	66.0
Male	34.0
Major/dept.	% of respondents
Agriculture	0.1
Area Studies	1.9
Biological Scis.	4.3
Business & Mgt.	2.4
Communications	1.0
Comp. & Info. Sci.	0.9
Education	21.5
Fine & Applied Arts	0.4
Foreign Language	0.7
Health Professions	1.5
Humanities	8.3
Library Sciences	0.1
Mathematics	1.0
Physical Sciences	1.0
Psychology	28.6
Public Affairs	4.9
Social Sciences	14.5
Interdisciplinary	0.3
University Admin.	0.9
Not Applicable	1.2
Missing Data	4.3

TABLE 2
Cross-tabulation of Training by Skill Level

	None or Incorrect Use	Low Skill Level	Moderate Skill Level	High Skill Level
Ever-trained	56 19.5%	60 20.9%	98 34.1%	73 25.4%
Never-trained	116 29.9%	106 27.3%	111 28.6%	55 14.2%

is because no previous survey has been done to explore the demography of our CD-ROM users. However, with the exception of the gender breakdown, these figures closely reflect our overall user population and use of our SilverPlatter databases. A majority of our graduate students (73.7%) are in the social sciences, as are 30.8 percent of undergraduates who have declared majors (almost half of our undergraduates are undecided or not in a program). Overall, 44 percent of the students are in the social sciences. This is

Although a low skill level was exhibited by 20.9 percent of those who had received instruction and 27.3 percent of those who had not, a moderate level of skill was used by 34.1 percent of those having received instruction and 28.6 percent of those who had not.

less than in our study population, but a large percentage of those who are undecided are taking courses in the social sciences. The survey population includes a higher percentage of females than does our campus population, where 47.9 percent of undergraduates and 58.5 percent of graduate students are female. This representation, interestingly, is at odds with research reviewed by Elizabeth Cardman in her article on computer use and gender.⁶ Statistics on overall database use for the period of the study (February through May 1994) show that use of ERIC and PsycLIT closely reflects the survey popu-

lation: 26.1 percent of all searches done on SilverPlatter databases were conducted on PsycLIT, and 22 percent were conducted on ERIC.

As for age distribution, over 59 percent of the respondents were be-

tween eighteen and twenty-five years old, over 22 percent were between twenty-six and thirty-four, and almost 18 percent were over thirty-five. Most of those participating in the survey were graduate students (53.6%) or undergraduate students (40.9%). A huge majority (over 96%) indicated that they had some computer experience outside the library. In addition, in keeping with previous surveys, over 78 percent indicated that they thought the results of their search were valuable (with another 18 percent indicating that they did not yet know). Although we did not keep an exact count, very few of the people approached refused to complete the survey.

Results

The authors coded and analyzed responses to the survey through the use of the statistical package SPSS. We created a measure of the level of success of searching (in this case determined by the successful use of Boolean operators) by collapsing all the possible combinations of use of Boolean operators into three distinct categories. The first category—high skill level—included those persons who used three or more Boolean operators in their search; the second category—moderate skill level—included all those who used two Boolean operators; and the third category—low skill level—included all those who used one operator. We treated persons who used no operators or who only used incorrect operators separately in the analysis.

The authors then performed cross-tabulations on all the important measures. We compared "ever-trained" with "never-trained" searchers, and combined all the measures pertaining to any sort of end-user training into one category. The authors treated respondents who indicated they had never had any kind of training or class instruction on CD-ROM searching as another category. Although a low skill level was exhibited by 20.9 percent of those who had received instruction and 27.3 percent of those who had not, a moderate level of skill was used by 34.1 percent of those having received instruction and 28.6 percent of those who had not. Furthermore, almost twice as many of those who had prior instruction used the high level of search skill as defined in our study (25.4%), compared to those who received no instruction (14.2%) (see table 2 and figure 1).

To study the relationship between the two variables of class instruction and skill level indicated by the search strat-

egy, we first examined the various row percentages in the cross-tabulations (the percent at each skill level divided between those who had had instruction and those who had not). We then conducted a chi-square test to see if there was a statistically significant relationship between the two variables. The chi-square for the measure of class instruction by the measure of search skill-level cross-tabulation resulted in a value of 22.4 ($df=3$, $p<.05$), indicating a significant relationship. Cramer's V was then used to determine the strength of the relationship. The value from this test "ranges from 0 to +1; the larger the value of V , the greater is the association in the variables."⁷ The measure of the association was 0.182, indicating a weak association. We next considered each type of instruction as an individual measure and performed cross-tabulations. These cross-tabulations showed no statistically significant relationships between type of training and skill level.

FIGURE 1
Relationship between Training and Skill Level

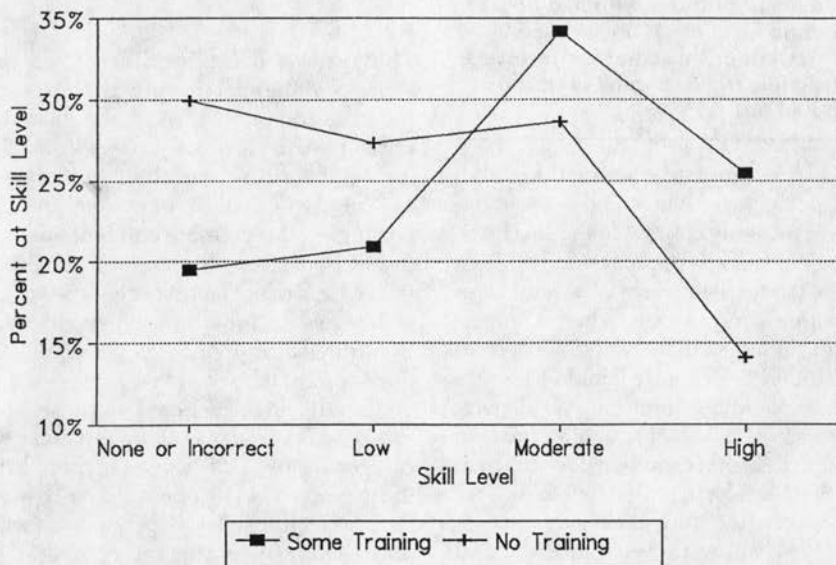


TABLE 3
User Groups by Skill Level

	None or Incorrect Use	Low Skill Level	Moderate Skill Level	High Skill Level
Faculty	10 45.5%	2 9.0%	6 27.3%	4 18.2%
Graduate	77 21.6%	77 21.6%	121 34.0%	81 22.8%
Undergrad.	77 28.3%	84 30.9%	76 27.9%	35 12.9%
Staff	4 26.7%	2 13.3%	5 33.3%	4 26.7%

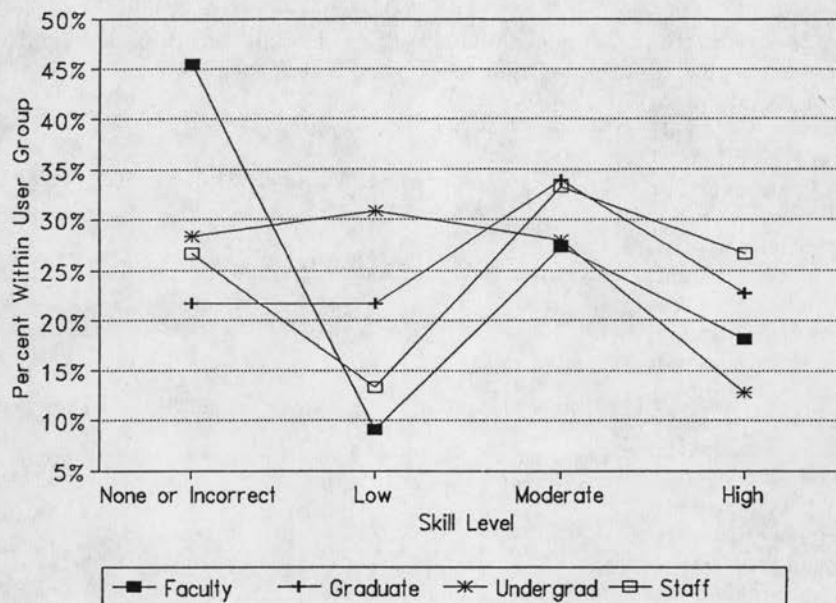
Chi-square 25.5 (df=9, p<.05)

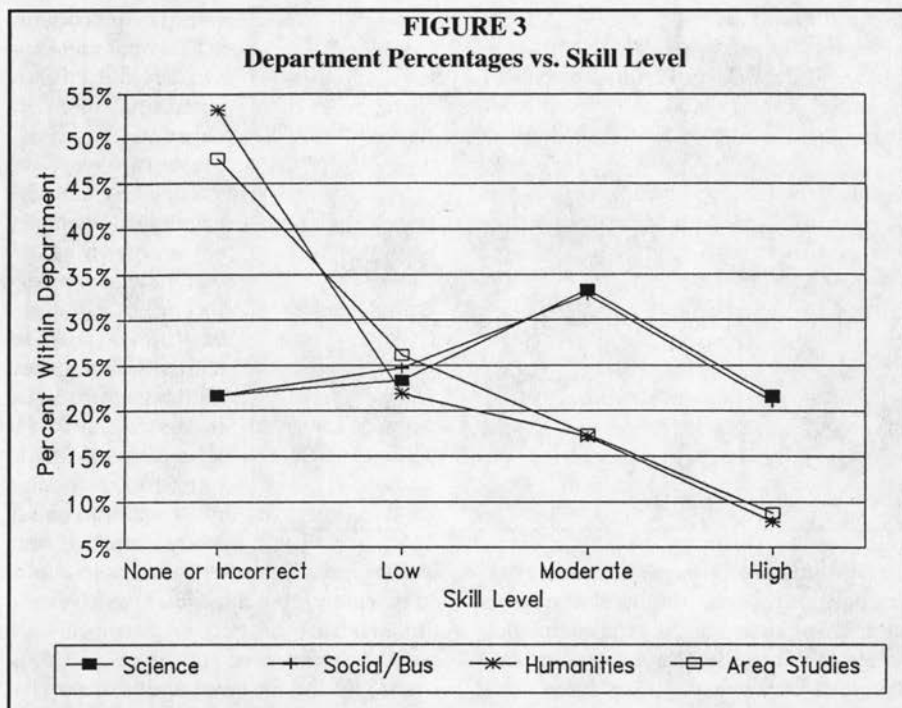
cells did not contain at least five items, a requirement for a statistically valid chi-square analysis, the results were revealing (see table 3). Faculty do the highest percentage of searches using incorrect or no operators. Staff and graduate students do the most searches indicating a moderate skill level, whereas staff members perform the most searches using a high skill level.

The authors also performed cross-tabulations that did not involve instruction. We compared user group with skill level, combining the University at Albany and non-University at Albany respondents in each category. Even though five

Undergraduate students hover at about 30 percent for the three lower skill levels, but drop to 12.9 percent for high-skill level searches (see figure 2). The chi-square for the measure of user group by the measure of skill-level cross-tabulation

FIGURE 2
User Group Percentages vs. Skill Level





resulted in a value of 25.5 ($df=9$, $p<.05$), indicating a significant relationship. Cramer's V was 0.113, signifying a weak association.

We also compared department affiliation/major with skill level. In order to obtain an adequate count in each category, we collapsed departments into four broad areas: sciences, social sciences/business, humanities, and area studies. This combination of departments still resulted in two cells with fewer than five items, so the conclusions should be viewed with caution (see table 4). Searchers affiliated with departments in the humanities and area studies did the most searches

using no operators or using operators incorrectly, and had the fewest high skill-level searches (see figure 3). Although this might suggest that these users would benefit most from instruction, it should be noted that, for the humanities, our only

TABLE 4
Department/Major by Skill Level

	None or Incorrect Use	Low Skill Level	Moderate Skill Level	High Skill Level
Science	13 21.7%	14 23.3%	20 33.3%	13 21.7%
Social Sci./ Business	108 21.6%	123 24.6%	163 32.7%	105 21.0%
Humanities	34 53.1%	14 21.9%	11 17.2%	5 7.8%
Area Studies	11 47.8%	6 26.1%	4 17.4%	2 8.7%
Chi-square	39.5 ($df=9$, $p<.05$)			

SilverPlatter database is the MLA Bibliography. Although not all humanities-affiliated searchers would necessarily have used MLA, those that did were often searching the title of a literary work or an author's name—searches that do not always require operators to be good searches. The chi-square for the measure of department affiliation by the measure of skill-level cross-tabulation resulted in a value of 39.5 ($df=9, p<.05$), indicating a significant relationship. Cramer's V was 0.247 in the range from 0 to 1.

Neither age nor sex showed a relationship with search skill. However, when the age categories zero to seventeen and eighteen to twenty-five were combined to provide sufficient frequencies in each cell, we did note that there was a larger differentiation in the skill levels of trained searchers in the twenty-six to thirty-four age range. Of those searchers who used no operators or who used them incorrectly, 17.4 percent fell into this age range. Of those who had a low skill level, 23.5 percent were between twenty-six and thirty-four. Of those with a moderate skill level, 21 percent were in this age range, and 28.4 percent of those with a high skill level were between twenty-six and thirty-four. The percentages in the zero to twenty-five and the thirty-five-plus age ranges stayed more constant across skill levels.

In asking survey respondents how many times they had searched CD-ROMs in the past, we hypothesized that those who were frequent searchers would have the best search strategies. Although 31.3 percent of those who had never searched before did not use any operators or used them incorrectly, there was not a clear trend in increased operator use among those who had never searched

through the most experienced searchers. In addition, whereas 19.7 percent of the most experienced searchers did the best searches, so did 18.8 percent of the new searchers (see table 5). Similarly, experience in online catalog searching, use of interactive learning software, online database searching, and e-mail experience did not relate to search competency.

Conclusions

Our hypothesis—that patrons who had attended searching classes or who had been directly assisted by librarians (or trained library assistants) would be far more sophisticated searchers—seems to be correct based on one test, but additional research should be undertaken before we can state with resounding affirmation that classes help students. Several of the analyses of the results of this study showed no statistically significant relationships between the skill level and other variables. The chi-square test did demonstrate that there is a positive association between class instruction and more skillful searching; however, based on our Cramer's V test result, class instruction makes a less pronounced difference than we expected. None of the other individual instruction analyses detected statistically significant associations, in part because insufficient data existed. In retrospect, we realize that the last section of the ques-

TABLE 5
Prior Use of CD-ROMs by Skill Level

	None or Incorrect Use	Low Skill Level	Moderate Skill Level	High Skill Level
Never	15 31.3%	14 29.2%	10 20.8%	9 18.8%
1 time	10 19.2%	14 26.9%	15 28.8%	13 25.0%
2-5 times	40 25.2%	43 27.0%	52 32.7%	24 15.1%
6 or more times	107 25.7%	95 22.8%	132 31.7%	82 19.7%

tionnaire was too complex to net sufficient responses in any specific category to allow for a useful analysis. If we were to repeat this study, we would only differentiate hands-on versus demonstration classes during the current semester or an earlier semester.

The first question librarians need to ask after reviewing this research is: Are public service librarians spending their time wisely teaching and assisting end-users? The answer, we believe, is yes. This study involved only the use of Boolean operators, field searching, and referrals to previous search statements, and not other search skills. It is possible that Boolean operators are the most "foreign" concepts; the least easily grasped of any that librarians try to impart.⁸ Possibly the amount of information presented in instruction sessions overwhelms patrons. When teaching the local setup, the concept of a database, the variety of databases available, and specific searching techniques,

much is bound to be lost. In addition, if students learn little but become comfortable with CD-ROM technology or with the library in general, our work is clearly purposeful.

Thus, although we do not believe that the results of our research refute the importance of CD-ROM instruction and assistance, we do believe the results make a persuasive case for more research into the issues we've explored. If there isn't a strong association between what we are teaching and what end users are learning, we need to discover why. We need to know how much end users are motivated to learn search skills (perhaps their high satisfaction level with any search, however poor, prevents them from appreciating the value of instruction) or if the software itself somehow leads end users away from employing more sophisticated search skills. Most important, we need to discover if our teaching and assistance techniques are inappropriate for the technology.

Notes

1. Michael Culbertson, "Analysis of Searches by End-Users of Science and Engineering CD-ROM Databases in an Academic Library," *CD-ROM Professional* 5 (Mar. 1992): 76-79; Susan K. Charles and Katherine E. Clark, "Enhancing CD-ROM Searches with Online Updates: An Examination of End-User Needs, Strategies, and Problems," *College & Research Libraries* 51 (July 1990): 321-28; Janet L. Nelson, "An Analysis of Transaction Logs to Evaluate the Educational Needs of End Users," *Medical Reference Services Quarterly* 11 (winter 1992): 11-21.

2. Domenica M. Barbuto and Elena E. Cevallos, "End-User Searching: Program Review and Future Prospects," *RQ* 31 (winter 1991): 214-27.

3. Rebecca Bostian and Anne Robbins, "Effective Instruction for Searching CD-ROM Indexes," *Laserdisk Professional* 3 (Jan. 1990): 14-17.

4. Dorothy F. Davis, "A Comparison of Bibliographic Instruction Methods on CD-ROM Databases," *Research Strategies* 11 (summer 1993): 156-63.

5. Linda Stewart and Jan Olsen, "Compact Disk Databases: Are They Good for Users?" *Online* 12 (May 1988): 48-52.

6. Elizabeth R. Cardman, "The Gender Gap in Computer Use: Implications for Bibliographic Instruction," *Research Strategies* 8 (summer 1990): 116-28.

7. Peter Hernon, *Statistics for Library Decision Making: A Handbook* (Norwood, NJ: Ablex, 1989), 104.

8. Rasmussen lists "user-unfriendliness" as one of the problems identified with online systems using Boolean retrieval. Edie M. Rasmussen, "Current Research in Information Science and its Impact Upon Librarianship," *Organizing a Research Agenda: Information Studies for the 1990s* (Halifax, Nova Scotia: Dalhousie Univ. School of Library and Information Studies, 1990), 34.