

The Effect of User Fees on the Cost of On-Line Searching in Libraries

Michael D. COOPER and Nancy A. DeWATH: School of Library and Information Studies, University of California, Berkeley, California.

A comparative analysis is presented of the time and cost to a library of providing on-line searching services without charge or for a fee. When the service was free to the user the cost to the library averaged \$26.73, and when a fee was charged the cost was \$28.78. Shifts in library resources were found to have taken place, with the librarian spending relatively more time preparatory to and subsequent to the search and less time at the terminal during the pay period than during the free period. Data base connect charges were found to have decreased by \$5.75 and off-line print costs to have increased by \$2.00 from the free to the pay period. A sharp decrease of 3.4 days (44 percent) occurred between periods in the time required to process a request.

INTRODUCTION

Libraries and information dissemination agencies that have implemented or are contemplating on-line bibliographic searching services are faced with the question of the effect of offering such a service. This paper compares the costs that libraries incur in providing the service under two conditions: when the service is free to the user and when the user has to pay a fee for the service.

Traditionally, library service is provided to patrons without charge. Presumably this is done because library service has benefits to the community in excess of the benefits received by the individual patron, and also perhaps because a unit of library service is not easily identifiable or measurable in order to apportion charges. The introduction of on-line bibliographic searching, however, makes the issue of whether to charge the user for library service more complex. The library, in providing on-line bibliographic search facilities, is acting as an intermediary between the user and a commercial vendor. Each search incurs a charge from the search service vendor. Second, search requests are unique in that their results cannot usually be used by other library patrons, in contrast to library services such as the provision of books, serials, etc. Third, the unit of service is easily identifiable.

In order to make a rational decision about charging or not charging for on-line searching services, a number of issues need to be considered:

1. Should fees be charged for all or part of these library services, and, if so, which parts?
2. Will individuals who need on-line searching services be deterred by imposition of fees?
3. Does charging a fee for on-line searching result in more efficient and/or better quality searching on the part of the library staff?
4. Should library administrators be influenced in their resource allocation decisions by the fact that users are willing to pay for on-line searching? Would the long-range consequences of this be to shift library service into paying activities or away from paying activities?
5. How should prices be set for the services if it is decided to charge the user?

These are but a few of the many issues that a library must resolve in deciding to implement user charges. In this paper we concentrate on the question of how the introduction of fees for library services impacts the library's resources, in terms of staff time and costs to the library of providing the services. The paper draws on experiences from the Lockheed/Cooperative Information Network DIALIB project.

The DIALIB project has been reported elsewhere in the literature.¹⁻³ The project has been offering Lockheed DIALOG searches to public library patrons in the San Francisco Bay area since 1974. During the first two years of the project (June 1974 to May 1976), four public libraries (Redwood City Public Library, San Jose Public Library, Santa Clara County Library, and San Mateo County Library) performed searches using regular reference staff who had been trained in DIALOG searching. During the first year, DIALOG charges were paid in full by a National Science Foundation (NSF) grant and the libraries contributed the necessary staffing. During the second year, the grant covered half the connect charges from the search service vendor and the libraries passed on the other half to the users. The libraries were compensated by NSF for their staff time at the rate of \$10 for each hour that the staff spent on the terminal.

During the pay period, users of the on-line searching services could have a "standard" or a "custom" search performed. The standard search cost \$5 and was limited to one data base, a maximum of ten search sets (combinations of logical operations and index terms), and twenty off-line prints. For a custom search the user paid half the actual data base charges incurred as well as half the off-line print charges. Only 14 percent of the searches performed during the entire second year of the project were standard searches.*

* In the third and final year of the project (June 1976-May 1977), two libraries have continued to offer the service, charging users with the full search cost. The terminals are subsidized by NSF, and the staff are provided by the participating libraries.

METHODOLOGY

In a previous paper, Cooper and DeWath analyzed the cost of providing on-line searching when the service was free to all users.^{4, 5} The current study extends the analysis by investigating the costs during the pay period and statistically comparing the pay and free periods.

The two hypotheses tested in the present paper were:

1. That within each library, the differences in the time and cost of a search between the free and pay periods are not significant.
2. That within the pay period and within the free period, the differences in the time and cost variables between libraries are not significant.

The statistical methodology used compared the mean values of the time and cost variables using contrasts. Each contrast compares the mean values of a given variable for two groups, and tests to see whether the difference is statistically significant. For a detailed discussion of the statistical tests, see appendix A.

In the presentation of the experimental results, the paper will indicate whether a comparison between free and pay period variables or between two libraries' variables was significant. This means that the statistical test described above was employed and that the contrasts between the means were significant. Any exceptions to this procedure will be noted.

Differences between Free and Pay Periods

The methodology used in the present study was kept as similar as possible to that of the free period cost analysis to make comparisons possible. A time sheet was filled out by library staff members as they performed the various tasks associated with each search.⁶ Seven possible tasks were defined, but not all tasks were necessarily performed for each search.* In addition, two kinds of libraries were identified: a DIALOG library, which had a terminal where the search was performed, and an originating library without a terminal (often a branch of a DIALOG library) that might take a request from a patron and relay the request to the DIALOG library

* The seven tasks are defined in detail in reference 4. They are:

1. Reference interview: the time spent with the patron defining the question.
2. Originating library preparation: activities at the originating library, performed without the patron, involving preparing the question prior to relaying it to a DIALOG library.
3. DIALOG library preparation: activities at the DIALOG library without the user, preparing for the terminal session.
4. Search: the actual on-line search.
5. DIALOG library follow-up: post-search activities at the DIALOG library with no patron present.
6. Originating library follow-up: post-search activities at the originating library without the patron.
7. Follow-up with patron: the time spent with the patron explaining the results of the search.

for searching. Each person who processed the request entered his or her initials, the time spent on the request, and the date on the time sheet. The time sheets were collected and analyzed to compute the time and costs per search. The actual salary schedules for the participants' job classifications, and the DIALOG system's data base fee schedule, were used to compute the costs for each search. The data analysis resulted in both time and cost figures for each of the seven tasks.

The first (free) and second (pay) years of the project differ notably in some areas:

1. During the free period, all the library personnel were relative novices at on-line searching. Those who remained with the project through the second year can be described as relatively experienced searchers, while others left and were replaced with novices. Thirty-one percent of the personnel who participated in these analyses were active during both periods. Thus the population of searchers whose activities were analyzed was only partially the same for both periods.

2. Seventeen new data bases were added to the DIALOG system between the end of the first data collection period and the end of the second, so the available information sources were not completely the same between periods.

3. All participants' salaries were increased by cost-of-living adjustments from one year to the next. In addition, some of the staff who participated in both periods were promoted, and received commensurate salary increases. Adjustments to allow comparisons in costs between the free and pay period were made for these changes in salary levels.

4. The nature of the requests may have changed with the institution of fees; the nature of the requestors did change, with relatively more graduate students and fewer undergraduates and professionals requesting searches.⁷

In addition, the number of search requests for which cost data were recorded was different from the free period to the pay period. The free period analysis was based on a sample of 411 (21 percent) of the 1,929 total searches performed during the first year of the project. The sample consisted of almost all searches from roughly the middle of the first year (January to March 1975). The second year's sample included 359 (62 percent) of the total 581 searches performed during the second year. The sample consisted of almost all searches performed during approximately the last seven months of the second year (November 1975 to May 1976). Aside from sample size variations, the distribution of the sample among the four libraries varied between the two data collection periods (table 1). The number of searches performed varied widely from one library to another.

COMPARATIVE ANALYSIS OF FREE AND PAY PERIODS

The previous section outlined some of the conceptual difficulties in comparing the pay period time and cost variables with those of the free

period. This section analyzes data base usage, citations printed, time spent on the search, and cost of the search, to ascertain whether search characteristics changed between periods.

Data Base Usage and Charges

The same concentration of data base use that was found in the free period was also found in the pay period. In the former period, the ERIC, NTIS, and Psychological Abstracts bases accounted for 48 percent of all uses, while in the pay period the same bases accounted for 55 percent of all uses. Changes in use from one period to the other occurred most heavily with the ERIC base where there was a more than 9 percent increase in use (table 2). Although seventeen new bases were added by the search service vendor since the previous study, their availability made little difference in the pattern of data base use. All of the eighteen available bases were used by searchers during the free period, but only twenty-eight of the thirty-five available bases were used during the pay period. Other

Table 1. Sample Size

| <i>Library</i> | <i>Free Period</i> | | <i>Pay Period</i> | | <i>Percent Change in Number of Requests from Free to Pay Period</i> |
|--------------------------------|---------------------------|----------------|---------------------------|----------------|---|
| | <i>Number of Requests</i> | <i>Percent</i> | <i>Number of Requests</i> | <i>Percent</i> | |
| Redwood City Public Library | 138 | 33.6 | 84 | 23.4 | -39.1 |
| Santa Clara County Library | 103 | 25.1 | 118 | 32.9 | +14.6 |
| San Mateo County Library | 93 | 22.6 | 27 | 7.5 | -71.0 |
| San Jose Public Library | 77 | 18.7 | 130 | 36.2 | +68.8 |
| Total Sample Size | 411 | 100.0 | 359 | 100.0 | +12.7 |

Note: There is a significant difference between the percent distribution of requests in the free versus the pay period using the chi square test at $\alpha = .05$.

Table 2. Data Base Connect Charges and Usage

| <i>Data Base Name</i> | <i>Charge per Connect Hour</i> | <i>Pay Period</i> | | <i>Percent of Data Base Uses</i> | <i>Change in Percent of Uses from Free to Pay Period</i> |
|------------------------|--------------------------------|----------------------------------|-------------------|----------------------------------|--|
| | | <i>Charge per Off-Line Print</i> | <i>Total Uses</i> | | |
| ERIC | \$25 | \$.10 | 159 | 22.78 | +9.43 |
| NTIS | 35 | .10 | 119 | 17.05 | -0.79 |
| Psych. Abstracts | 50 | .10 | 107 | 15.33 | -1.24 |
| COMPENDEX | 65 | .10 | 69 | 9.89 | +0.60 |
| SSCI | 70 | .10 | 42 | 6.02 | -3.49 |
| Chem. Abstracts | 45 | .08 | 35 | 5.01 | -3.22 |
| ABI/INFORM | 65 | .10 | 23 | 3.30 | -3.32 |
| INSPEC. ELECT. ENG. | 45 | .10 | 19 | 2.72 | +0.58 |
| All Other Bases | — | — | 125 | 17.90 | +1.45 |
| Total | — | — | 698 | 100.00 | — |

factors that might have influenced the choice of data bases include the types of requests received and the searchers' preferences among data bases. The searchers indicated that they experienced difficulty in maintaining their competence with the large number of bases available.⁸ The observed pattern of limited data base use might be at least partly due to the searchers' choice of the data bases that are (1) easiest to keep current with, (2) most worth the investment for keeping up to date in terms of data base demand, or (3) easiest to use without studying changes and data base idiosyncrasies. The relatively low use of some of the more expensive data bases, e.g., the Predicasts bases, may also indicate some discrimination in favor of bases that (in the judgment of the searchers) offer greater value per dollar spent.*

There were no differences in the cost per connect hour to use a particular data base between the free and the pay period. This was advantageous in that variations in costs of a given data base could not directly influence the use of the data base.

Off-line Prints

The average number of off-line prints per search rose from sixty-one during the free period to eighty-eight in the pay period (table 3). This increase is primarily a result of the much higher averages for two libraries (Santa Clara and San Mateo counties). Only 18 percent of the searches during the free period resulted in no prints at all, compared to 32 percent in the pay period. Many explanations can be proposed for this increase in the mean value: (1) the librarians were more careful about accepting searches for which they expected to find a fair amount of information, once user fees were instituted; (2) the \$0.05 per citation that users actually were charged during the pay period was too low to discourage printing, while the high cost of the search encouraged the searcher to print whatever was found, even if it was not highly relevant;[†] (3) most of the searchers were more experienced than during the earlier study and had more success at finding relevant citations.

Data Bases

Multiple data bases were sometimes searched for a given request. During the free period an average of 2.3 data bases per search were used, while in the pay period the average was 1.9. Forty-eight percent of all searches in the pay period and 31 percent of the searches in the free pe-

* The comments of the staff members indicated a continual awareness of the cost of the search to the patron during the pay period. In fact, the participants agitated throughout the study for a DIALOG feature that would give the accumulated cost of a search at any time during the search session when issued a special DIALOG command.

† Recall that the user paid only the actual data base charges and half the off-line print charges. Thus the figure \$0.05, rather than \$0.10, per citation printed.

riod used only one data base. The standard searches accounted for 14 percent of the total. These searches were by definition limited to one data base and explain most of the variation in the two figures.

It is interesting to note that during the pay period the first data base used for a search accounted for only 19 percent of the off-line prints, while the second data base produced 55 percent of the citations printed.

Staff Time

The library staff involvement with a search request neither begins nor ends at the terminal. Considerable time is required for other search-related activities.

The total time required to process a request during the pay period averaged 54.9 minutes, which did not differ much from the free period's 48.7 minutes (table 4). Two of the libraries did differ significantly from their own previous year's averages: Redwood City almost doubled the total time spent on a search from 35.1 minutes during the free period to 67.6 minutes in the pay period, and Santa Clara County decreased its time on a given request from 58.3 minutes to 46.9. San Mateo County's appar-

Table 3. Mean Number of Off-line Prints per Search (All Cases)

| <i>Library</i> | <i>Free Period</i> | <i>Pay Period</i> | <i>Difference</i> |
|--|--------------------|-------------------|-------------------|
| Redwood City Public Library | 79 | 71 | -8 |
| Santa Clara County Library | 69 | 146 | +77 |
| San Mateo County Library | 29 | 72 | +43 |
| San Jose Public Library | 60 | 51 | -9 |
| Mean Prints per Search (All Libraries) | 62 | 88 | +26 |

Table 4. Mean Time per Task by Library (in Minutes) (Non Zero Entries)

| <i>Task</i> | <i>Mean Task Time by Library—Pay Period</i> | | | | | <i>Overall Mean</i> | <i>Free Period Mean Task Time</i> |
|---------------------|---|--------------------|------------------|-----------------|-------|---------------------|-----------------------------------|
| | <i>Redwood City</i> | <i>Santa Clara</i> | <i>San Mateo</i> | <i>San Jose</i> | | | |
| Reference Interview | 24.04* | 13.03 | 16.17 | 13.87* | 15.87 | 10.50 | |
| Originating Library | | | | | | | |
| Preparation | — | 10.00 | 17.50 | 10.00 | 15.45 | 18.65 | |
| DIALOG Library | | | | | | | |
| Preparation | 23.50* | 12.80 | 27.55 | 15.38* | 17.83 | 10.48 | |
| Search | 16.53 | 17.16* | 22.27 | 13.21* | 15.95 | 22.72 | |
| DIALOG Library | | | | | | | |
| Follow-up | 18.58* | 14.38* | 22.63* | 10.17 | 14.17 | 12.21 | |
| Originating Library | | | | | | | |
| Follow-up | — | — | 10.83 | — | 10.83 | 11.48 | |
| Follow-up with | | | | | | | |
| Patron | 15.00* | 9.46* | 10.67 | 9.00 | 10.25 | 7.61 | |
| Total Time Spent on | | | | | | | |
| Request—Pay | | | | | | | |
| Period | 67.60* | 46.92* | 79.15 | 48.79 | 54.86 | — | |
| Total Time Spent on | | | | | | | |
| Request—Free | | | | | | | |
| Period | 35.07 | 58.28 | 57.23 | 50.76 | 48.73 | — | |

* Significant difference between free and pay period at family $\alpha = .05$ (see appendix A).

ent increase of 22 minutes (from 56.2 minutes in the free period to 79.2 minutes in the pay period) is dramatic. But due to the small number of observations and a large variability among them, it is not statistically significant. The San Jose Public Library had a very consistent total search time, recording 48.8 minutes during the pay period and 50.8 minutes during the free period. Since no additional staffing was available for this service, the one library's dramatic increase represents a shifting of resources from other (nonfee) library services to the search service.

For each of the five DIALOG library tasks that could be compared, the time spent increased from the free to the pay period (see table 4). Reference interview time increased from 10.5 to 15.8 minutes, DIALOG preparation from 10.5 to 17.8 minutes, DIALOG follow-up from 12.2 to 14.2 minutes, and follow-up with patron from 7.6 to 10.3 minutes. Contrasts were performed on the individual libraries' values using means of time and cost variables, and two libraries showed several significant increases in task times (table 4).

The off-line tasks also were performed more frequently during the pay period. Most notably the percent of searches having reference interviews increased from 72 percent to 91 percent from one period to the next.*

With regard to staff time, it can be concluded that with the institution of user fees for on-line searching, the librarians are apparently substituting off-line time for on-line time. They are performing the off-line tasks more frequently and spending longer at them when they are performed. It is possible on DIALOG to reduce on-line time (and search service charges) to a certain extent by spending more time off-line structuring the request, using hard copy thesauri, and otherwise preparing for the search so that less time is required to search for synonyms and try various possible search formulations on-line. It appears that the librarians did just that.

Search Costs

The total cost of an on-line search includes the payment to the search service vendor for data base connect charges and off-line citation printing. It also includes the direct salary costs for the individuals who process the search request. Other possible costs that can be considered, but were not included in the calculations below, include telephone line charges, terminal rental, and overhead. The search service costs reflect those costs actually incurred and not those costs charged to the patron.†

* During the pay period 67 percent of the searches involved DIALOG library preparation as opposed to 54 percent during the free period. DIALOG library follow-up was performed in 84 percent of the pay period searches and 75 percent of the free period searches. There was only 1 percent difference in the number of follow-ups with patron between periods (43 percent during the pay and 42 percent during the free period).

† Recall that under the terms of the NSF grant, the user, even during the pay period, only paid half of the charges for a custom search.

Twenty-eight library staff members, ranging from clerks through supervising librarians, participated in the search process during the sampled pay period. This compares with forty people during the free period. Their average salaries during the pay period are given in table 5 together with their average salaries during the free period. For all staff involved in the searching process, salaries increased by 9.5 percent between the two periods. Aside from the general salary increase between periods, there were shifts in the number of people in a particular job classification who were involved in the searching. For example, the number of Librarian I persons declined from nineteen in the free period (47.5 percent of the total staff) to eleven in the pay period (39.3 percent). Similarly the number of Librarian II persons declined from eleven in the free period to eight in the pay period, although the overall percentage of library employees in that classification remained steady at about 28 percent.

The average costs reported in tables 6 and 7 are calculated from the actual costs of each reported search, based on the times required, the individuals performing the tasks, the data bases used, and the number of citations printed off-line. In order to provide valid cost comparisons between the free period and the pay period, adjustments were made to the original free period data. All of the costs for free period searches were re-

Table 5. *Salary Schedule*

| Job Title | Mean Monthly Salary | | Number of Participants in This Classification | |
|-------------------------|---------------------|------------|--|------------|
| | Free Period | Pay Period | Free Period | Pay Period |
| Typist Clerk | — | \$ 821.00 | — | 2 |
| Library Assistant I | \$ 658.00 | 706.00 | 3 | 2 |
| Library Assistant II | 782.50 | 976.50 | 2 | 2 |
| Librarian I | 980.25 | 1,082.25 | 19 | 11 |
| Librarian II | 1,080.25 | 1,171.25 | 11 | 8 |
| Librarian III and above | 1,233.80 | 1,340.80 | 5 | 3 |

Table 6. *Mean Salary Cost for Task by Library (in Dollars)*

| Task | Mean Task Cost by Library—Pay Period | | | | | Overall Mean | Mean Task Cost— Free Period— in Constant Dollars |
|--|--------------------------------------|-----------------------|---------------------|--------------------|------|-----------------|--|
| | Redwood City | Santa Clara County | San Mateo County | San Jose Public | | | |
| Reference Interview Originating Library | 2.63 | 1.35 | 1.79 | 1.56 | 1.72 | 1.12 | |
| Preparation | — | 1.03 | 1.90 | 1.06 | 1.66 | 1.96 | |
| DIALOG Library Preparation | 2.54 | 1.32 | 3.04 | 1.69 | 1.93 | 1.06 | |
| Search | 1.88 | 1.79 | 2.17 | 1.50 | 1.74 | 2.43 | |
| DIALOG Library Follow-Up | 2.03 | 1.49 | 2.66 | 1.10 | 1.53 | 1.26 | |
| Originating Library Follow-Up | — | — | 1.21 | — | 1.21 | 1.08 | |
| Follow-Up with Patron | 1.65 | .97 | 1.23 | 1.05 | 1.15 | .81 | |

Table 7. Search Cost

| Cost Element | Mean Task Cost by Library—Pay Period | | | | | Overall Mean | Mean Task Cost— Free Period— in Constant Dollars |
|--------------------------------------|--------------------------------------|-----------------------|---------------------|--------------------|-------|-----------------|--|
| | Redwood City | Santa Clara County | San Mateo County | San Jose Public | | | |
| Data Base Charges | 11.83 | 13.30* | 18.22 | 8.60* | 11.60 | 17.35 | |
| Off-Line Print Charges | 8.14 | 17.78 | 8.16 | 6.54* | 10.87 | 8.83 | |
| Search Labor Cost | 1.88 | 1.79* | 2.17 | 1.50* | 1.74 | 2.43 | |
| Labor Cost for All Other Tasks | 5.98* | 3.30 | 7.17 | 4.58* | 4.68 | 2.98 | |
| Total Cost of Search —Pay Period† | 26.46 | 33.37 | 33.15 | 19.55 | 26.73 | 28.78 | |
| Total Cost of Search —Free Period | 25.40 | 35.84 | 19.77 | 36.15 | 28.78 | | |

* Significant difference between free and pay period at family $\alpha = .05$ (see appendix A).

† Totals are not additive due to differences in the number of observations in each cell.

calculated using data base charges, off-line print charges, and salaries that were in effect during the pay period rather than the free period.

For example, for a particular free period search, the salary of the searcher might have increased from \$1,000 to \$1,100 per month between the two periods. In computing the adjusted free period cost of the search, the \$1,100 salary would be used. In general, salaries of individuals involved in a free period search were adjusted by using equivalent pay period salaries corresponding to the individual's job title. The effect of the process is to change free period search costs into constant dollars that can then be compared to the pay period costs in a consistent manner. After adjustment, the differences between free and pay period costs are due to the different times required by the various tasks, plus any differences in the job classifications of the people involved, choice of data bases, and number of citations printed.

Table 7 summarizes the major cost elements of a search for both the free period and the pay period in "pay period dollars," and table 6 breaks down the labor costs according to the various tasks. Table 7 shows that the labor costs for all the search tasks except the actual search have increased. (Comparisons for originating library preparation and originating library follow-up should be ignored since the number of observations is not adequate to make valid comparisons.) For example, reference interview labor cost increased from \$1.12 during the free period to \$1.72 during the pay period. Similarly, the labor cost for the follow-up with the user increased from \$.81 in the free period to \$1.15 in the pay period. Search labor cost, however, decreased from \$2.43 in the free period to \$1.74 in the pay period. It is interesting to note that the adjustment of the free period costs into constant dollars results in a very small change in the actual search labor costs. For example, the greatest change in any search labor cost figure in table 7 was \$0.14 for one task. Although adjusting the costs into constant dollars in this experiment made little difference in the values, if the

time difference between observations had been greater, the effect would obviously have been more noticeable.

Table 7 shows some major cost shifts between the free and the pay periods. Data base charges declined from \$17.35 in the free period to \$11.60 in the pay period, and off-line print charges increased from \$8.83 to \$10.87. Labor costs for all search activities increased between the two periods. Total search cost declined from \$28.78 in the free period (constant dollars) to \$26.73, reflecting mainly the decline in data-base charges in the pay period. Among the libraries there were small shifts in costs for two of the libraries, and large changes for two others.* Only San Jose Public Library showed a significant decrease in total search cost.

The cost figures reinforce the findings of the time figures; the cost of the on-line search itself has been reduced, but the cost of related activities has risen as the librarians spend more time at them so that the overall cost of the search is only slightly lower, if at all.

Under the conditions reported in this study (namely, that the user during the pay period was charged 50 percent of the connect charges only), a proportion of the total costs of an on-line search was shifted from the user to the library on the introduction of user fees.

Patron Presence during Search

While a search was conducted, the user was invited to be present much more frequently during the pay period (50 percent of the searches) than the free period (15 percent).† The librarians in the DIALIB project who preferred to have the user present noted that the user can often provide useful information during a search, evaluate the results as they appear, and help to alter the course of the search, if required. In addition, a user who has seen what is and is not available on-line is more likely to be satisfied with the results. Those searchers who preferred not to have the user present felt that the user tended to slow the search down because of unfamiliarity with the system.

The actual effect of the user's presence on the search during the pay period was slight, while during the free period it was much greater. During the free period the average time at the terminal with the patron present was 33.85 minutes; it was 20.93 minutes with the patron not present. During the pay period similar figures were 16.51 minutes and 15.38 minutes.

Some care is needed in interpreting time differences due to patron presence at the search. Mitigating factors that could confound the results include the fact that searches with the patron present could have been more

* Note again that the number of observations for San Mateo County is relatively small, and variability is quite large. The large change in cost should be treated with caution.
 † In a study conducted by the System Development Corporation, it was found that more experienced searchers tend to be more inclined to allow the user to participate in the search than less experienced searchers.⁹

complex than others, or that the user stayed with the searcher because of anticipated problems. It is also possible that those users present during a search were in some way more demanding, requiring more of the librarian's time. The only conclusion possible is that fears that the patron will generally slow down the search are not justified, since under the circumstances tested that did not hold true.

Elapsed Time

The time required to process a search was compared between the free and the pay period to see if the process was more efficiently performed when the user was paying for the search. There is strong evidence that a considerable reduction in total processing time did take place. During the free period, the elapsed time from performance of the reference interview to completion of the follow-up with the patron was 7.8 days. The average during the sampled searches in the pay period was 4.4 days, for a 44 percent reduction. This reduction may be due to the integration of the DIALOG procedures into the other library activities, a more experienced staff during the pay period, as well as perhaps some pressure to provide prompt service to paying users. The volume of requests was also much lower during the pay period, which no doubt helped reduce backlog problems considerably.

Table 8 summarizes the elapsed time information for the free and the pay period by library. Between the free and pay period, a large reduction in the elapsed time took place between the time a search request was made in a reference interview and the actual search was performed. During the free period this process required 4.9 days while in the pay period it took 2.1 days. The large number of days required for San Mateo County to process requests reflects the geographic dispersion of their branch structure and also the relatively small number of searches performed by them.

COMPARISONS BETWEEN LIBRARIES

The discussion to this point has dealt with the first hypothesis—that within a given library the variables differ between the free and the pay periods. The second hypothesis tested was whether within a given period the libraries differed significantly from one another.

To test this second hypothesis, a series of paired contrasts were performed on the data for each period, using the mean values for each library for the time and cost variables. Each pair of libraries' values for each of seven major variables within each of two periods was compared.

The results, reported in tables 9 and 10, indicate a greater conformity among the libraries during the pay period than during the free period. Table 9 reports the contrasts between all possible pairs of libraries during the free period. The seven major variables contrasted are represented by abbreviations; within the table, each variable that tested significantly different between a pair of libraries is entered in the intersection of that

Table 8. *Elapsed Time to Process a Search Request (in Calendar Days)*

| Period | Redwood City | Mean Elapsed Time—Pay Period | | | Libraries | Mean Elapsed Time—Free Period—All Libraries |
|---|--------------|------------------------------|------------------|-----------------|-----------|---|
| | | Santa Clara County | San Mateo County | San Jose Public | | |
| Reference Interview to On-Line Search | 1.1 | 2.6 | 6.5 | 1.3 | 2.1 | 4.9 |
| On-Line Search to Follow-up with Patron | 3.0 | 3.0 | 5.6 | 3.3 | 3.4 | 4.0 |
| Total Time in System—Pay Period | 3.2 | 3.3 | 10.8 | 4.8 | 4.4 | 7.8 |
| Mean Total Time in System—Free Period | 6.0 | 4.9 | 14.7 | 6.8 | 7.8 | |

Table 9. *Significant Differences between Libraries for Selected Search Variables—Free Period*

| Library | Santa Clara | Library San Mateo | San Jose |
|--------------|--------------------|-------------------|---------------|
| Redwood City | RI, ST, DF, TT, TC | ST, DF, TT, TC | ST, TT |
| Santa Clara | — | RI, ST, TC | RI, DF, F, TT |
| San Mateo | — | — | ST, DF, TC |

Note: The symbols for the variables are defined as follows: RI—reference interview time; DP—DIALOG library preparation time; ST—search time at terminal; DF—DIALOG library follow-up time; F—follow-up time with patrons; TT—total time for all search and search-related tasks; TC—total cost of search.

Table 10. *Significant Differences between Libraries for Selected Search Variables—Pay Period*

| Library | Santa Clara | Library San Mateo | San Jose |
|--------------|-------------|-------------------|--------------------|
| Redwood City | RI, DP, TT | — | RI, DP, DF, TT, TC |
| Santa Clara | — | — | ST, DF, TC |
| San Mateo | — | — | DF |

Note: The symbols for the variables are defined as follows: RI—reference interview time; DP—DIALOG library preparation time; ST—search time at terminal; DF—DIALOG library follow-up time; F—follow-up time with patron; TT—total time for all search and search-related tasks; TC—total cost of search.

pair. For example, Redwood City and Santa Clara County proved to be significantly different in their free period values for reference interview time, time at terminal, DIALOG library follow-up time, total time for all search-related activities, and total cost of the search. Overall, twenty-one of forty-two tested comparisons proved significant.

Table 10 repeats the analysis for the pay period data. Only twelve of the differences proved significant this time, a reduction of almost one half.

Since the major single difference between the two periods was the institution of user fees, it can be concluded that the great change in the number of significant contrasts from the free period to the pay period is probably due at least in part to those fees. Since the movement from one year to another is in the direction of greater conformity among the libraries, it is possible that this conformity represents a movement toward some optimal

state. It has been shown that with the institution of user fees the searchers apparently tried to eliminate unnecessary cost to the patron by reducing on-line time and increasing off-line time. It is possible that this increasing conformity represents the same trend toward eliminating unnecessary identifiable costs. When the service was free, the searcher was at liberty to experiment with the service, to try different approaches to the same question, and generally vary the search procedure. With the introduction of search fees and the accompanying pressure on the searcher to perform effectively at the terminal, this variation and experimentation was no longer possible.

The data are not persuasive enough to allow any inferences about what this optimal level might be. But the greater similarity among the libraries tested argues for a greater generalizability of the result from these sample libraries to other libraries. Another library considering instituting such a service can be encouraged by the similarity of the values among these libraries, despite their different populations, organizational structures, and means of advertising the search service to potential users. This leads one to believe that the results are of some value in predicting the influence of instituting on-line searching in another library.

SUMMARY AND CONCLUSIONS

This paper has analyzed the time required and cost incurred to conduct on-line searching in four public libraries when the user received the service free of charge and when a fee is paid for the search. The primary effect of the institution of search fees is a shifting of a proportion of the costs from the user, in terms of search service charges, to the library, in the form of staff time devoted to search-related tasks.

The time and cost required to perform a task in the free period was always higher than in the pay period, except in the case of the labor cost of searching. For example, it took an average of about six minutes (12.7 percent) longer to do all the search-related tasks in the pay period than in the free period. The actual searcher time at the terminal, however, declined by close to seven minutes, and the cost of the connect time to the data base declined by \$5.75 from the free period to the pay period.

The cost of the off-line prints, however, increased by about \$2.00 from the free to the pay period. This could be a reaction to the institution of user fees whereby the searcher felt the necessity to produce more output since the user was paying for the service. Overall, there was a modest decrease in the average total cost of a search from \$28.78 in the free period to \$26.73 in the pay period, and a sharp decrease (44 percent) of 3.4 days in the number of days required to process a request.

The data presented in this paper do not strongly support the conclusion that libraries are more efficient in providing on-line searching services when the user pays a fee. But there is some statistical evidence to that effect. Total costs have come down since the institution of fees, and for

the task that requires reimbursement from the user there was a decline in costs. There is another aspect to the problem. The data support a view that there are fewer significant differences between the libraries in search costs during the pay period than the free period. This suggests that institution of a fee may have the effect of reducing the heterogeneity of the libraries' services, at least with respect to costs.

The efficiency with which a search is performed is, of course, not solely a function of cost. This study cannot say whether the searchers themselves were more or less efficient in one period or another, or in one library or another. The analysis is confined to the cost of the search. The larger question of cost-effectiveness will have to await further research.

ACKNOWLEDGMENTS

Funds for this project were provided in part by a grant from the National Science Foundation Office of Science Information Service to the Lockheed Palo Alto Research Laboratories and to Applied Communications Research, Inc., Palo Alto, California. Our thanks go to the librarians who participated in the study, to Colin Mick and Alice Ahlgren at Applied Communications Research, and especially to Gale G. Hannigan for her research assistance.

REFERENCES

1. Roger K. Summit and Oscar Firschein, *Two Year Interim Report: Investigation of the Public Library as a Linking Agent to Major Scientific, Educational, Social, and Environmental Data Bases*. (Palo Alto, Calif.: Information Systems Program, Lockheed Palo Alto Research Laboratory, Lockheed Missiles and Space Company, Inc., September 1976.)
2. Roger K. Summit and Oscar Firschein, *Two Year Interim Report: Investigation of the Public Library as a Linking Agent to Major Scientific, Educational, Social, and Environmental Data Bases. Annex I: Evaluation Results*, by Alice E. Ahlgren. (Palo Alto, Calif.: Information Systems Program, Lockheed Palo Alto Research Laboratory, Lockheed Missiles and Space Company, Inc., October 1976.)
3. Roger K. Summit and Oscar Firschein, *Two Year Interim Report: Investigation of the Public Library as a Linking Agent to Major Scientific, Educational, Social, and Environmental Data Bases. Annex II: Publicity Aspects*. October 1976. (Palo Alto, Calif.: Information Systems Program, Lockheed Palo Alto Research Laboratory, Lockheed Missiles and Space Company, Inc., October 1976.)
4. Michael D. Cooper and Nancy A. DeWath, "The Cost of On-Line Bibliographic Searching," *Journal of Library Automation*, 9:195-209 (September 1976).
5. Michael D. Cooper and Nancy A. DeWath, *The Cost of On-Line Bibliographic Searching*, Technical Report ACR-003-75-01 (Palo Alto, Calif.: Applied Communication Research, December 1975.)
6. *Ibid.* See appendix A for a copy of the data collection form.
7. Roger K. Summit and Oscar Firschein, *Two Year Interim Report . . . Annex I: Evaluation Results*.
8. *Ibid.*
9. Judith Wanger, Carlos A. Cuadra, and Mary Fishburn, *Impact of On-Line Retrieval Services: A Survey of Users, 1974-75*. (Santa Monica, Calif.: System Development Corporation, 1976.)

APPENDIX A

Statistical Tests

Comparisons of the means for the various hypotheses were performed using contrasts. To form a contrast, the initial observations of one search variable are recorded. The variable may be the time or cost of a reference interview, search preparation, or follow-up; or it may be the total time for a search, the total number of off-line prints, or the data-base connect charges. Define X_{LPSO} as the observed value for a variable for library L, time period P (where the period is either the free period or the pay period), staff member S, and observation O.

There are usually multiple observations of a variable for a particular staff member, and the mean of these n_{LPS} observations is given by

$$\bar{X}_{LPS} = \sum_0 X_{LPSO} / n_{LPS}$$

and the standard deviation by

$$S_{LPS} = \sqrt{\frac{\sum_0 (X_{LPSO} - \bar{X}_{LPS})^2}{n_{LPS}}}$$

The mean of a variable X for library L during time period P is given by

$$\bar{X}_{LP..} = \frac{\sum_s n_{LPS} \bar{X}_{LPS}}{\sum_s n_{LPS}}$$

This is simply the weighted average of each staff member's mean value for the variable. Similarly, the standard error of the mean is given by

$$SE_{\bar{X}_{LP..}} = \frac{\sum_s n_{LPS} S_{LPS}^2}{\sum_s (n_{LPS})^2}$$

and the number of observations is

$$n_{LP..} = \sum_s n_{LPS}$$

To compute the confidence interval for a contrast (say, the difference between the pay and free period values for a variable for library 1), the following equation is used:

$$\bar{X}_{11..} - \bar{X}_{12..} \pm t_{DUNN} SE_{(\bar{X}_{11..} - \bar{X}_{12..})}$$

where the standard error (SE) is defined as

$$SE_{(\bar{X}_{11..} - \bar{X}_{12..})} = \sqrt{SE^2_{\bar{X}_{11..}} - SE^2_{\bar{X}_{12..}}}$$

and the value of t_{DUNN} is given in a standard table* (4, p.551).

The appropriate value of the Dunn coefficient depends on the number of error degrees of freedom, the number of comparisons made, and the alpha level used. All tests were made at alpha = 0.05 per family of contrasts. The number of error degrees of freedom was assumed to be infinite. For comparison of the free versus the pay period, four contrasts were calculated (one for each library) and the Dunn value used was 2.50. Comparisons of the libraries for the combined period involved six contrasts (four libraries, compared two at a time) and the Dunn coefficient was 2.64. Comparisons of pairs of libraries for the pay period (and then the free period) involved twelve contrasts. The Dunn value used was 2.86 (interpolated).

* Roger E. Kirk, *Experimental Design: Procedures for the Behavioral Sciences* (Belmont, Calif.: Brooks/Cole Publishing Co., 1968.)