DOCUMENT RESUME

ED 118 087 IR 002 988

AUTHOR Cooper, Michael D.; DeWath, Nancy A.
TITLE The Cost of On-Line Bibliographic Searching. Technical Report 003-75-01.
INSTITUTION Applied Communication Research, Stanford, Calif.
SPONS AGENCY National Science Foundation, Washington, D.C. Office of Science Information Services.
REPORT NO ACR-003-75-01
PUB DATE Dec 75
NOTE 47p.

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage
DESCRIPTORS Computers; *Cost Effectiveness; Data Bases; Experimental Programs; Information Retrieval; *Library Reference Services; *On Line Systems; *Public Libraries
IDENTIFIERS Bibliographic Searching; California (San Francisco); Computerized Literature Retrieval; Lockheed DIALOG; *Project DIALIB

ABSTRACT

A cost analysis was made of on-line bibliographic searching in four public libraries in the San Francisco Bay Area using the Lockheed DIALOG system. The study involved monitoring the time that 35 individuals in the libraries spent processing some 411 search requests. The study identified a set of seven tasks that are performed for each request and determined the average time and cost for each of the tasks. The average total search cost was $28.41, exclusive of telephone line charges. This figure does include the data base connect charges of $17.29. The average time required to process a request was 7.8 calendar days. A wide variation in the cost and time figures was found among the four libraries and among individual searchers. (Author/JY)

ACR-003-75-01

Technical Report

The Cost of On-Line Bibliographic Searching

Michael D. Cooper
Nancy A. DeWath

December 1975

Applied Communication Research
P.O. BOX 5849 STANFORD CALIFORNIA 94305

This paper reports a project commissioned by Applied Communication Research, Inc. as part of an on-going program studying potential user interfaces for on-line search services.

This research program is funded by the Office of Science Information Service, National Science Foundation, through grants to the Lockheed Missiles and Space Company (GN42299) and to Temple University (GN42271). Applied Communication Research, Inc. serves as an evaluation subcontractor to both projects.

Alice E. Ahlgren is program manager of the ACR evaluation of the Lockheed DIALIB project. Colin K. Mick manages the ACR-Temple effort.

As noted in this report, the study was conducted during the first year of a two-year project -- during the "free service" portion of the study. As a result, the various time estimates have been affected by a number of variables. The time estimates presented here may therefore be misleading for the following reasons.

1. The searchers were still in a learning mode -- the learning curve for searching (shown by time per search) started at 31 minutes per search and dropped to about 18 minutes at the end of the free period (mean for the free period was approximately 23 minutes).

2. Our data from the pay period (which began in June, 1975) indicate that the searchers are now becoming more sophisticated and tend to devote more time to off-line search preparation and less time to actual on-line searching.

ACR, Lockheed and the participating libraries are now discussing the possibility of replicating this study during the pay period.

ABSTRACT

A cost analysis of 411 on-line bibliographic search requests was conducted. The study involved monitoring the time that 35 individuals in four public libraries spent processing these requests. The study identified a set of seven tasks that are performed for each request and determined the average time and cost for each of the tasks. The average total search cost was $28.41 exclusive of telephone line charges. This figure included the data base connect charges of $17.29. The average time to process a request was 7.8 calendar days. A wide variation in the cost and time figures was found among the four libraries and among individual searchers.

INTRODUCTION

On-line bibliographic searching is becoming commonly used as an aid to the reference librarian and researcher. Commercially available systems, such as Lockheed's DIALOG and System Development Corporation's ORBIT, provide access to a multitude of machine searchable data bases for this purpose. Many issues remain unresolved with respect to the general process of on-line searching. These issues include questions such as the effectiveness of on-line searching, methods for training searchers, and optimal search strategies at the terminal. This paper examines the cost of on-line searching. On-line search costs include the charges that are incurred for connection to a commercial search service, the cost of printing bibliographic citations, and the cost of the reference librarians time.

There are a number of reasons for studying the cost of on-line searching. In the first place, it seems apparent that on-line searching is a close substitute for some forms of non-computerized bibliographic searching. If this is the case, and if the end product is the same, then it is important to know how the costs of the alternatives compare. Secondly, costs are an important tool to aid in resource allocation. The provision of any new service implies either additional funding or a diversion of funds from one type of service to another. With cost data this type of decision can be aided considerably. Finally, costs can provide a basis for making pricing decisions. It seems quite likely that users will have to pay for on-line bibliographic search services. The question is, how much? While there are numerous ways in which prices can be set (i.e., loss leader, marginal cost, cost recovery), a knowledge of costs can play an important part in their establishment.

The cost data reported in this paper were compiled as part of an ongoing project being conducted by the Lockheed Palo Alto Research Laboratory under the sponsorship of the National Science Foundation's Office of Science Information Service. As part of the study, Lockheed's on-line reference retrieval service (DIALOG) is being made available through four public libraries in the San Francisco Bay Area. On-line search services were provided at no cost to patrons during the first year of the project (August, 1974 through May, 1975). The search service is being provided at a reduced cost to patrons during the second year of the project (June, 1975 through May, 1976).

The four libraries participating in the project are all members of the Cooperative Information Network (CIN), a Bay Area consortium of public, special, and academic libraries. CIN, in cooperation with Lockheed, selected four public libraries in San Mateo and Santa Clara counties as sites for the placement of computer terminals. The sites included the Redwood City Public Library, the San Mateo County Library, the Santa Clara County Library, and the San Jose Public Library. All libraries in the CIN network were encouraged to participate in the experiment by either directly referring patrons or by forwarding patron requests to one of the above libraries.

The individuals who performed the on-line searching and search-related interviewing had (with rare exceptions) no previous experience with on-line searching (1). A core of eight librarians (two from each library) received the standard Lockheed introductory course on DIALOG (2 days) and were allowed time to practice searching. These librarians then provided the search training to other staff members of their respective libraries.

The experimental nature of this project differentiates it from the normal search situation. During the first year of the project, each of the four participating libraries received both 16 hours per month of free search time and 16 hours per month of free demonstration time. Since the search time was free, searchers were able to experiment with the system while conducting patron searches. They were not under any pressure to keep the searches short and efficient. On-line search times, and subsequently connect costs, reported in this paper may be somewhat higher than they might have been if this were not an experimental project in which free search time was available to the libraries.

PREVIOUS STUDIES

The development of adequate statistical (time, cost) monitoring mechanisms in on-line systems has been a slow process. Monitoring programs took a large leap forward with the advent of commercial systems which require the precise recording of user-system interaction time for accurate billing purposes.

Monitoring of the user-system interaction can take many forms. The simplest measurements deal with elapsed search time, time spent searching particular data bases, and frequency of command utilization. One of the earliest reports on user behavior as monitored by a retrieval system was presented by Summit (1969). Data reported included elapsed search time, number of index terms used in a search, number of Boolean expressions used in a search, and number of citations printed by the system. Subsequently, there have been a number of other analyses. Of particular note is a study by Benenfeld et al. (1974) which reported a comprehensive set of characteristics of on-line searches. Among the variables computed were the time required to discuss the search with the patron, the search time at the terminal, the number of citations printed out, and the total user cost of the search. Benenfeld's experience indicated user costs per search in the range of $28 to $56 depending on the data base used. Lawrence, Weil, and Graham (1974) also gathered cost data on bibliographic searching and Elman (1975) surveyed some of the previous studies and presented his own computations indicating that an average on-line search cost $47. Another aspect of on-line search costing that has received some attention in the literature is the cost of operating the computer equipment. Lancaster (1973) reviewed some of this literature. It is apparent that this area of cost analysis needs considerably more investigation before the internal economies of on-line searching can be settled.

(1) It is only fair to note that one of the authors of this paper was also one of the searchers being studied, an employee of San Mateo County Library. We hope that no bias resulted.

METHODOLOGY

The procedure for obtaining the data from which to compute the cost of on-line bibliographic searching required the cooperation of all individuals engaged in processing the on-line search requests. A set of seven tasks was defined including reference interview, originating library preparation, DIALOG library preparation, search, DIALOG library follow-up, originating library follow-up, and follow-up with patron. A form was then developed to collect data describing the amount of time spent on each task, the date on which the task was performed, and the individual who performed the task (Appendix A). This time sheet traveled with the search request through its processing. As completed time sheets were received, they were numbered and coded for computer processing. In order to protect the anonymity of the library employees performing the various tasks, employee names were replaced with code numbers for all computer tabulations. In the tables that refer to individuals performing search tasks, code numbers will be used.

Costs were developed both for each search performed and for each task within a search. Each task could be performed by a different individual. In order to arrive at the direct labor cost for a search, it was necessary to multiply the pay rate for each individual involved in the search by the number of minutes each spent at a particular task.

In addition to the labor cost of a search, two other costs must be considered. The first is the fee paid to the commercial search service (in this case, Lockheed) to access their data base. This fee is a function both of the length of time one is connected to the system via telephone line and of the particular data base (such as ERIC, NTIS, Psychological Abstracts, etc.) that is being searched. For each request, multiple data bases may be searched. In computing the on-line cost for a search, the cost per hour of accessing each data base was multiplied by the connect time to arrive at a total data base connect cost.

The second non-labor cost element is the number of bibliographic citations printed at the commercial vendor's computer center and sent by mail to the searcher's location. A charge is made for these off-line prints and this adds to the total cost of the search.

Several other factors could be included in the cost of a search. The major omission of this study was the cost of telephone line charges. Since all the libraries were relatively close to Lockkheed's computer center, they dialed directly to the center without the use of an intermediary service such as TYMSHARE. If an intermediary service had been necessary, these charges would have to be included in the total search cost.

Other cost elements that were excluded from the study were terminal rental costs, physical space charges, and indirect costs of overhead, administration, supplies, etc. Basically, the cost figures presented here summarize the direct costs of on-line searching.

The data for this study were collected during February, March, and part of April of 1975. Under the terms of the National Science Foundation grant, DIALOG service was provided free to the public for the first year of the experiment, and at a reduced rate for another year. The three month study period was one which occurred six months after the project had begun and during which the service was still free to all users.

It should be emphasized that the data comes from self-reports of the library personnel. Such a method is of course less accurate than an outside measure, but was the only practical way to collect the amount and variety of data needed.

At the end of the data collection period, 411 usable time sheets had been collected from the four DIALOG libraries (Table 1). Of these 411, 33.6% were from Redwood City Public Library, 25.1% from Santa Clara County Library, 22.6% from San Mateo County Library, and 18.7% from San Jose Public Library. Branch libraries and other CIN libraries took requests and relayed them to the DIALOG libraries; 16% of the requests during the reporting period showed some participation by these other libraries.

THE REQUEST PROCEDURE

A general outline of the DIALOG request procedure is flowcharted in Figure 1. For the data collection, the process was divided into seven tasks which formed the basic reporting units on the time sheet (Appendix A).

The tasks were defined by both where and at what stage of the search process the activity took place. A maximum of two possible types of libraries participated in each search: a library with a terminal (called DIALOG library) and a cooperating library, a branch library, or other CIN member (called originating library). The tasks were further divided according to whether the activity occurred before or after the actual on-line search. The DIALOG search itself was treated as a separate task. The tasks were defined as follows:

REFERENCE INTERVIEW: time that the user spent negotiating the request by submitting and discussing it with a library staff member. This could take place at either a DIALOG library or an originating library.

ORIGINATING LIBRARY PREPARATION: if the reference interview was at a cooperating library, the staff there might spend some time recording, researching, and transmitting the request.

DIALOG LIBRARY PREPARATION: time spent on the request at the DIALOG library with no patron present and before performing the on-line search. Typical activities included recording the receipt of the request, filling out forms, and researching search terms.

DIALOG CONNECT TIME or ON-LINE SEARCH TIME: the actual DIALOG search.

DIALOG LIBRARY FOLLOW-UP: time spent after the on-line search, with no patron present. Typically this would include filling out reporting forms, examining search results, and notifying the patron or originating library.

ORIGINATING LIBRARY FOLLOW-UP: analagous to DIALOG library follow-up, but at a cooperating library.

FOLLOW-UP WITH PATRON: at either type of library, the staff might spend time with the patron explaining the results of the search.

The only one of these activities that was essential was the actual on-line search. Any combination of the other activities might be performed, depending on the nature of the request and on the library or libraries involved. The individual libraries varied this general form in accordance with their own needs and policies.

The Santa Clara County Library is a large system with branches all over the county and a research library in Cupertino where the terminal was located. Users were free to submit requests through branch libraries or at the Cupertino library. Over the course of this study, virtually all requests originated at Cupertino. The requestor, on entering the Cupertino library, was directed to a DIALOG-trained staff member, if possible, who performed the reference interview. The librarian asked the patron if she/he wished to be present during the search, in which case an appointment was arranged for some time in the future. Scheduled searches were limited to three or four per day, due to the demands of other staff duties. If the user did not ask to be present, the search was performed at the searcher's convenience, often but not necessarily by the same person who had performed the reference interview. The user was called either when the on-line search was finished or when the off-line prints were received. Whenever possible a DIALOG staff member discussed the results with the patron.

The Redwood City Public Library consists of a main library and two small branches. Users with reference questions requiring the facilities of the main library were sent there directly, as were DIALOG requestors. When possible they were met by a DIALOG-trained reference librarian, who performed both the interview and search. The user was then called to come in to pick up the results and was met by a reference librarian if possible. Occasionally, the results were mailed to the patron.

The San Jose Public Library serves a metropolitan area. The users with questions on DIALOG were referred to the main library's science/business department whose personnel included two searchers. The other staff members in the department were also familiar with DIALOG. The user filled out the request form, and might have been briefly questioned by the staff. The primary searcher during the data collection period was a student volunteer who came in part-time to perform searches. No particular effort was made to have the patron talk with her, although she telephoned the requestor if she needed clarification of his/her request. The patron was notified, when the results were ready, by the searcher who explained the results, and the results were left in the science/business department to be picked up. A note was attached to the results encouraging the requestor to talk with the searcher if s/he did not understand the citation or was interested in obtaining specific articles.

The San Mateo County Library alone among the DIALOG libraries derived a substantial portion (about 60%) of its requests from other libraries. Like the Santa Clara County Library, users were allowed to place requests with the main library or with a branch or other CIN library; unlike Santa Clara, many users opted for the remote access. A librarian, although not necessarily one very familiar with DIALOG, generally performed a reference interview for each request. Searches received from other libraries or from walk-in business were placed with all other pending searches, to be performed at the librarians' convenience. The results were often sent to a branch library, even when the request originated at the main library.

Table 1

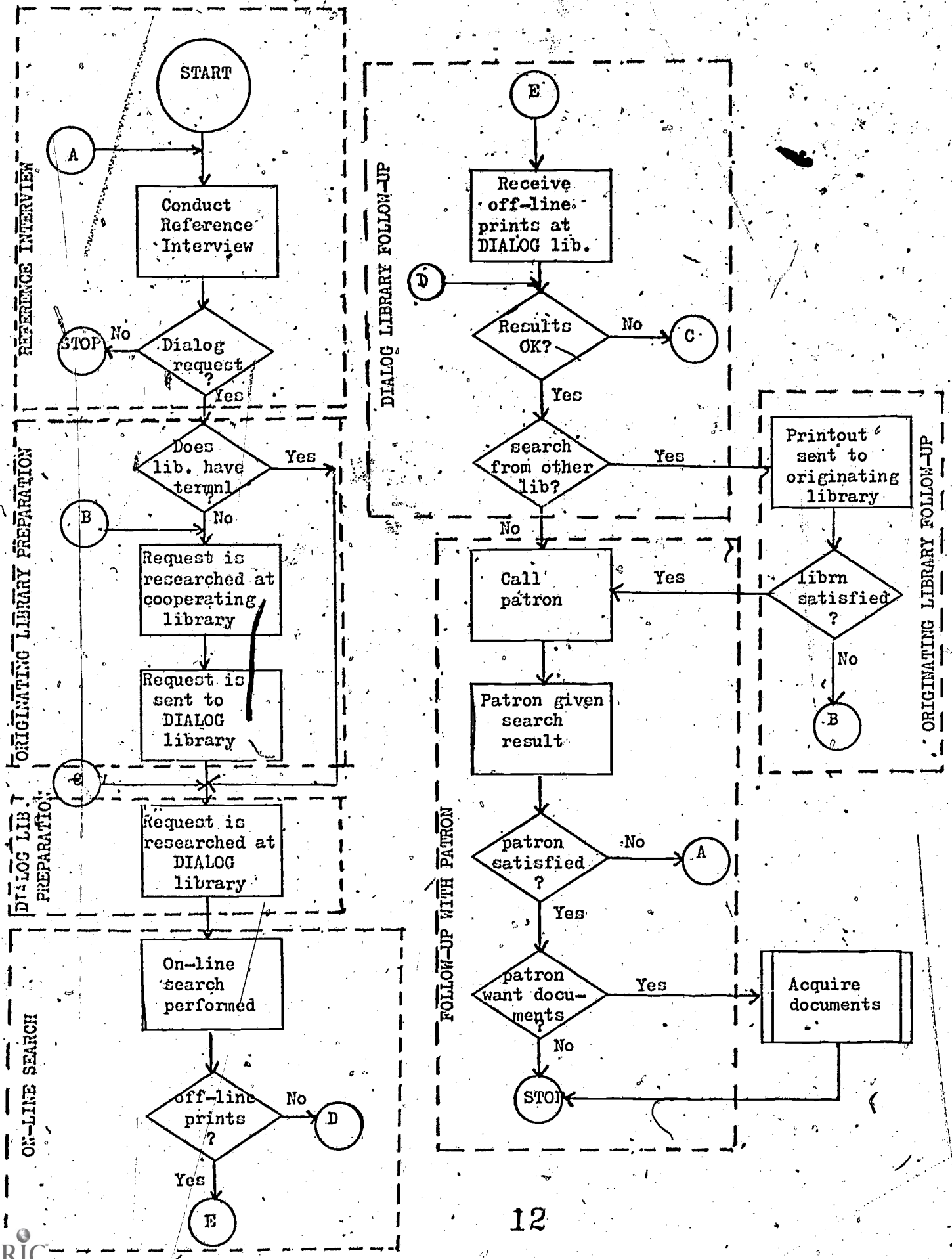
Frequency Distribution of Requests by Dialog Library

Library	Number of Requests	Percent Distribution of Requests
Redwood City (RC)	138	33.6
Redwood City	137	33.3
Menlo Park [1]	1	0.2
Santa Clara County (SC)	103	25.1
Cupertino Research Center	95	23.1
Gilroy [1]	3	0.7
Others [1]	5	1.2
San Mateo County (SM)	93	22.6
Central Library	37	9.0
Atherton	20	4.9
Belmont	6	1.5
Foster City	3	0.7
Millbrae	2	0.5
Woodside	1	0.2
Public Health and Welfare [1]	20	4.9
Other	4	1.0
San Jose Public Library(SJ)	77	18.7
Total	411	100.0

[1] An independent library that sends its requests to the indicated Dialog library.

FIGURE 1

FLOWCHART OF DIALOG REQUEST PROCESS



STAFF

One major decision for a library instituting this kind of service is that of how to staff the service. The libraries considered two major alternatives.

First, they could designate one or two people to do all the searches. The advantage to this would be that the searchers would quickly become highly skilled. Also, the responsibility for DIALOG and DIALOG-related activities would be clearly assigned. However, it would mean that at times the search staff would not be available to users with questions or problems.

The alternative was to have all of the reference personnel add DIALOG searching to their usual repertoire of reference skills. This would diffuse the searching, and therefore diffuse both the experience and the responsibility. However, it would increase the availability of trained personnel to users so that a request could be handled efficiently by virtually anyone at the reference desk.

Three of the libraries opted for the second alternative. At San Jose, however, due to special circumstances (1), most of the searching eventually devolved on one person.

All of Redwood City's staff who participated in DIALOG performed, at one time or another, almost every step of the procedure, including the on-line search. Generally, a librarian followed one request through from beginning to end. Santa Clara County showed roughly the same pattern, although they did have a few requests that originated at cooperating libraries so that the people who dealt with the patron did not perform searches.

At San Mateo County Library, staff members from other county system libraries frequently conducted some of the search tasks. Of the twenty people who were listed on time sheets, only five performed searches, and only four could be characterized as regular searchers.

Table 2 summarizes the salaries of the 37 employees involved in processing the on-line search requests. Participants' mean salaries for the job title ranged from a low of $658 per month for three library assistants to a high of $1325 for a supervisory librarian (2). For future reference, the table also shows the employee code numbers (instead of employee names) for each job title within each library.

(1) Near-by San Jose State University threatened to swamp San Jose Public Library with requests, so an arrangement was made whereby a library school student volunteer did a large proportion of the searching.

(2) Employees of each of the cities and counties in the study worked a different number of hours per week. Redwood City employees worked 38 hours per week while employees of Santa Clara, San Mateo and San Jose worked 40 hours per week. When labor costs were computed this factor was considered.

Table 2

Salary Schedule for Library Employees
Involved in Dialog Process

Library	Job Title	Mean Monthly Salary	Employee Code Numbers in this Classification	Number of Employees
Redwood City	Librarian I	$1,000.50	24, 25, 27	3
	Librarian II	1,083.00	26	1
	Librarian III	1,178.50	23	
Santa Clara	Libr.Assist. I	764.00	40	1
	Librarian I	974.50	35, 36, 37	3
	Librarian II	1,074.50	38, 41, 43	3
	Regional Librarian I	1,128.50	42	1
	Supervising Librarian	1,306.00	44	
San Mateo	Libr.Assist. I	658.00	2, 9, 11	3
	Libr.Assist. II	801.00	13	1
	Librarian I	992.00	1,10,12,14,16,17,18	7
	Librarian II	1,112.50	6, 7, 8	3
	Librarian III	1,232.00	4	1
	Supervising Librarian	1,325.00	15	1
San Jose	Librarian I	955.00	50*, 51, 53, 56	4
	Librarian II	1,052.50	52 54, 55	3

*Unpaid volunteer; salary calculated at Librarian I level for purpose of study.

Table 3

Task Frequency of Occurrence by Employee

Library	Employee Code Number	Task: Reference Interview	Originating Library Preparation	Dialog Library Preparation	Search	Dialog Library Follow-up	Originating Library Follow-up	Follow Up With Patron	Total Frequency
San Mateo County	3	15	20	16	19	8	17	9	104
	5	2		9	26	39		7	83
	6	2	3				4	3	12
	8	10		18	20	27		12	87
	9			33		3			36
	10	10		16	22	22		3	73
	14	8	12				10	2	32
	15	1		9	8	11		1	30
	Others[2]	16	8	4		1	7	6	42
Total SMCo.[1]		64	43	105	95	111	38	43	499
Redwood City	23	41		11	66	25		11	154
	24	16		13	24	22		17	92
	25	21		4	30	22		2	79
	26	11		15	22			4	52
Total RCPL[1]		89		43	142	69		34	377
Santa Clara County	35	43		13	41	41		25	163
	36	23		8	34	35		14	124
	37	16		26	30	29		18	119
	38	4							4
	40	1	1						2
	41	2	1					1	4
	42	3	1						4
Total SCCo.[1]		92	3	47	105	105		68	420
San Jose	50	12		27	36	34		6	115
	51	7		1	1				9
	52	7		5	15	7		2	36
	53	22		21	27	21		21	112
	54	2							2
	55	4		3	7			1	15
	56	2				6		1	9
	60				1				1
Total SJPL[2]		56		57	87	68		31	299
Total all Libraries[1]		301	46	252	429	353	38	176	1595

[1]Totals differ from other totals elsewhere because a) a task may be performed more than once per search, and b) tasks for which staff member could not be identified were eliminated from this tabulation.

In Table 3 data are presented on the frequency with which various employees performed each of the seven tasks connected with a search. For the entire experiment, 1595 tasks were performed for the 411 search requests averaging about 3.6 tasks per search. 429 searches were performed for the 411 requests indicating that very few duplicate searches were made for each request. Employee [redacted] at Santa Clara County performed the greatest number of tasks while employee 23 at Redwood City ranked second. Together these two employees accounted for 20% of all tasks performed.

DATA BASES

One possible source of variation among the libraries was the choice of data bases used for the search. Since the different bases cover different subjects, variations in base use could indicate that the libraries receive different types of requests, and/or that the librarians' search habits differ. Also, since the bases' hourly rates vary, differentiated use could affect the cost of searching (See Table 4).

The most commonly used bases (1), for all the libraries taken together, were NTIS (accounting for 17.84% of the bases used), Psychological Abstracts (16.57%), and ERIC (13.35%) (Table 5). NTIS is a highly varied base that covers a wide range of topics, and the documents abstracted were readily available to the project participants through purchase from NTIS or on loan from the California State Library. Pyschological Abstracts (PA) and ERIC cover psychology and education, respectively. All three of these could be expected to be of use to many public library patrons. The next most used bases were Social Science Citations Index (9.51%), COMPENDEX (COMPuterized ENgineering InDEX; 9.29%), and Chemical Abstracts Condensates (8.23%). The first ties in well with the first three bases; the latter two are equivalent to well-known hard-copy services. The least used bases were all Predicasts bases, including PATS Source (.11%), PATS Claims (.21%), and PATS Chemical and Electronic Market Abstracts Weekly (.85%). All are special-purpose bases.

Considering the libraries individually, their data base usage did differ from the group percentages, but not markedly. For example, the group proportion for NTIS is 17.84%; Redwood City's figure was 19% as were Santa Clara's and San Mateo's; while San Jose's was 13%. Psychological Abstracts' group percentage was 16.57%; the individual proportions were 18%, 14%, 22% and 13%.

(1) More than one base may be used on a search; these are the proportions of bases used, not of searches performed.

Table 4

Data Base Connect Charges Per Hour

Data base name	Charge per connect hour	Charge per off-line print
ERIC	$ 25	$ 0.10
CHEM ABSTRACTS	45	0.10
EXCEPT. CHILD ABS	25	0.10
NTIS	35	0.10
SSCI	70	0.10
COMPENDEX	65	0.10
AIM-ARM	25	0.10
NAL/CAIN	25	0.10
PSYCH ABSTRACTS	50	0.10
INSPEC-PHYSICS	45	0.10
INSPEC-ELECT. EN	45	0.10
INSPEC-COMPUTERS	45	0.10
ABI/INFORM	65	0.10
PATS CMA/EMA	90	0.20
PATS C/EMA WEEKLY	90	0.20
PATS F and S	90	0.20
PATS Source	90	0.20
IFI/Claims	$ 150	$ 0.10

These were the bases available during the data collection period.

Table 5.

Data Base Usage by Library

Data Base Name	Number of Data Base Uses by Library				Total Uses	Percent Distribution of Uses
	Redwood City	Santa Clara	San Mateo	San Jose		
ERIC	44	38	18	25	125	13.35
CHEM.ABSTRACTS	29	15	21	12	77	8.23
EXCEPT.CHILD,ABS.	2	3	5	6	16	1.71
NTIS	61	45	35	26	167	17.84
SSCI	32	27	5	25	89	9.51
COMPENDEX	24	27	5	25	89	9.29
AIM-ARM	1	1	3	5	10	1.07
CAIN	14	5	12	9	40	4.27
PSYCH. ABSTRACTS	56	33	40	26	155	16.57
INSPEC-PHYSICS	12	7	4	2	25	2.67
INSPEC-ELECT.ENG	4	12	2	2	20	2.14
INSPEC-COMPUTERS	2	4	3		9	0.96
ABI-INFORM	20	11	11	20	62	6.62
PATS CHEM ELECT	7	10		6	23	2.46
PATS CHEM-EL WEEK		3		5	8	0.85
PATS F-and S	10			10	20	2.14
PATS SOURCE	1				1	.11
CLAIMS/PAT PRES	1			1	2	.21
TOTAL	320	243	180	193	936	
Present Distribution	34.19	25.96	19.23	20.62		100.00

Although on the library level no significant differences appear, when broken down further to show the individual searches, differences become apparent (Tables 6 and 7). The most skewed figures are for searcher 3, who is a special librarian and whose search pattern reflects the specialized interests of her clients. The other searchers, however, have no such easy explanation for their preferences. Looking at the most frequently used base, NTIS, the frequency of use varies from 6.4% to 23.8% (Table 7). Even excluding searcher 3, Psychological Abstracts varies from a low of 4.9% all the way up to a high of 30.8%.

We can only guess at reasons for this variation. It may be that some searchers have certain subject expertise, and that the searches within a library are parcelled out accordingly. However, none of the libraries reported any such pattern to their searching. It is more likely that users fall into habits of searching and tend to use bases with which they are familiar. The most frequent searchers (23 and 50) show more of a scattering among the bases, so it may be that their added experience has encouraged them to branch out more, while other searchers stick with "safe" bases. Also, doing more searches will no doubt bring the searcher a greater variety of topics, and require a greater variety of bases. These are just conjectures, as the figures are not conclusive. If the searchers were either very conservative in their use of "new" bases or were strictly segregated by subject, they would center their searches on a handful of bases. However, almost all of them used each of the more popular bases at least once.

Table 8 breaks down the average time spent on each data base by the most active searchers. Again, the results show very little uniformity among the searchers. No one base tends to be a time-consuming base. The searchers varied widely in the average time that they spent on a single base. There seems to be some intra-library similarity, but it is very slight; the San Mateo County Library searchers all had fairly low averages, the Redwood City searchers a little higher, the San Jose people about the same as Redwood City or perhaps a little higher, and Santa Clara County highest of all. These figures are not directly comparable with the mean on-line time per search, since any number of bases can be used on a single search. However, they are something to keep in mind for the discussion of time per task that follows. If a search used the same number of bases, and the average time per base differed among the libraries, the average on-line time per search should vary accordingly.

OFF-LINE PRINTS

The number of off-line prints requested was a major variable in the cost of the search. San Mateo County tried not to print more than fifty citations per search, but that was not a hard and fast rule. There were two reasons for this guideline. First, the number of prints available under the terms of the grant was limited (although no library overran this limit during the first year of the study). Second, if the patron received too many prints with too much irrelevant information, the precision and therefore the quality of the search was lessened.

Table 6

Data Base Usage by Searcher

Library	Employee Code Number	Data Base Use Frequency											Total Uses by Searcher
		ERIC	CHEM ABS.	NTIS	SSCI	EI	CAIN	PA	INSPEC-P	ABI	PATS C/EMA	Other	
San Mateo County	3			7	3			15		4		1	30
	5	8	6	9	1	9	1	6	1	1		9	51
	8	7	3	8		7	1	8					34
	10	3	4	10	1	4	4	10	2	6		3	47
Redwood City	23	19	10	36	14	12	6	27	8	9	4	7	152
	24	13	2	5	7	1	2	16		2	1	3	52
	25	9	6	11	8	4	4	7	3	9	1	8	70
	26	2	10	8	5	7		5	1		1	3	42
Santa Clara County	35	17	2	14	13	9		14	3	3	2	9	86
	36	18	3	10	10	5	3	15	2	1	2	4	73
	37	3	10	20	4	15	2	4	2	7	4	9	80
San Jose	50	13	8	11	16	9	3	19	1	7	1	12	100
	52	2	3	7	2	2	4	2	1			5	30
	53	10	1	3	4	2	2	4		7	3	11	47
All Others		1	9	6	3	1	6	2	1	6	4	2	41
Total		125	77	165	91	87	38	154	25	62	23	86	93

Table 7

Proportionate Data Base Usage by Searcher[1]

Library	Employee Code Number	Percent of Searches Conducted by Searcher, by Data Base											Total Number of Uses	Percent of Total Uses, all Search
		ERIC	CHEM. ABS.	NTIS	SSCI	EI	CAIN	PA	INSPEC-P	ABI	PATS C/EMA	Other		
San Mateo County	3			23.3	10.0			50.0		13.3		3.3	30	3.2
	5	15.7	11.8	17.6	2.0	17.6	2.0	11.8	2.0	2.0		17.6	51	5.5
	8	20.6	8.8	23.5		20.6	2.9	23.5					34	3.6
	10	6.4	8.5	21.3	2.1	8.5	8.5	21.3	4.3	12.8		6.4	47	5.0
Redwood City	23	12.5	6.6	23.7	9.2	7.9	3.9	17.8	5.3	5.9	2.6	4.6	152	16.3
	24	25.0	3.8	9.6	13.5	1.9	3.8	30.8		3.8	1.9	5.8	52	5.6
	25	12.9	8.6	15.7	11.4	5.7	5.7	10.0	4.3	12.9	1.4	11.4	70	7.5
	26	4.8	23.8	19.0	11.9	16.7		11.9	2.4		2.4	7.1	42	4.5
Santa Clara County	35	19.8	2.3	16.3	15.1	10.5		16.3	3.5	3.5	2.3	10.5	86	9.2
	36	23.3	4.1	13.7	13.7	6.8	4.1	20.5	2.7	1.4	2.7	5.5	73	7.8
	37	3.8	12.5	25.0	5.0	18.8	2.5	5.0	2.5	8.8	5.0	11.3	80	8.6
San Jose	50	13.0	8.0	11.0	16.0	9.0	3.0	19.0	1.0	7.0	1.0	12.0	100	10.7
	52	6.7	10.0	23.3	6.7	6.7	13.3	6.7	3.3			16.7	30	3.2
	53	21.3	2.1	6.4	8.5	4.3	4.3	8.5		14.9	6.4	23.4	47	5.0
All Others		2.4	22.0	14.6	7.3	2.4	14.6	4.9	2.4	14.6	9.8	4.9	41	4.4
Total		125	77	165	91	87	38	154	25	62	23	86	933	
Percent of Total[1]		13.4	8.3	17.7	9.8	9.3	4.1	16.5	2.7	6.6	2.5	9.2		100.0

[1]Figures differ slightly from Table 5 due to rounding differences.

Table 8

Data Base Connect Time by Searcher

(In minutes)

Library	Employee Code Number	Mean Connect Time for Data Base										Mean search time, per base all bases
		ERIC	CHEM. ABS.	NTIS	SSCI	EI	CAIN	PA	INSPEC-P	ABI	PATS C/EMA	
San Mateo	3			8.88	6.21			8.46		11.13		8.70
	5	7.96	8.12	6.77	2.81	9.15	12.19	10.90	1.19	1.63		5.52
	8	4.23	7.67	3.08		2.54	4.57	4.82				4.06
	10	6.54	2.64	5.38	2.08	4.91	5.04	15.68	2.48	4.74		7.35
Redwood City	23	6.85	6.42	6.67	3.43	9.85	6.67	6.56	6.22	2.13	3.13	6.20
	24	11.66	6.79	7.86	7.27	3.34	9.20	13.29		1.19	10.20	10.29
	25	11.33	6.00	8.41	6.50	6.38	2.75	7.74	15.00	5.56	8.00	7.50
	26	26.00	16.50	10.25	8.00	14.07		7.80	27.00		14.00	13.27
Santa Clara	35	19.52	4.13	12.69	10.61	6.41		13.90	6.39	10.77	4.90	12.55
	36	21.92	14.24	18.74	10.18	11.90	12.33	15.08	16.66	14.73	6.83	16.09
	37	11.31	15.72	12.67	10.24	9.43	16.67	21.28	2.93	5.77	1.77	11.25
San Jose	50	9.23	10.98	6.33	6.26	7.67	14.49	14.49	8.95	9.06	3.21	6.43
	52	11.09	15.86	9.05	9.47	16.17	4.86	9.16	12.81			9.40
	53	33.93	12.75	17.80	19.59	4.10	10.17	19.67		10.05	25.70	20.53

Table 9

Frequency Distribution of Off-Line Prints by Searcher

Interval	Number of Off-Line Prints														Total
	San Mateo County				Redwood City				Santa Clara County			San Jose			
	Searcher No.				Searcher No.				Searcher No.			Searcher No.			
	3	5	8	10	23	24	25	26	35	36	37	50	52	53	
0	77	101	79	77	245	84	116	75	153	150	105	120	70	123	1575
1-19	10	15	5	13	31	15	13	18	11		21	22	1		175
20-39	5	9	7	10	18	5	5	10	10	2	14	12		2	109
40-59	3	3	5	6	7	6	2	2	12	5	2	11			64
60-79				2	10	1	5		4	3	3	3		4	35
80-99					5		5	1	3	4		6			24
100-119					2	2	4		2		1	3			14
120-139					2	2	1	1	4	1	1	2			14
140-159					1				1			1	2	1	6
160 & above					11		4		6	3	3		2	2	31
Total	95	128	96	108	332	115	155	107	206	168	150	180	75	132	2047
Mean Number of Off-Line Prints: All Entries	13.79	10.73	14.07	19.04	38.67	22.83	39.38	16.94	40.39	22.71	21.81	26.28	32.37	18.31	
Mean Number of Off-Line Prints Non-Zero Entries	20.40	13.48	17.73	18.56	37.81	29.63	38.21	20.61	44.13	39.46	21.17	26.00	51.57	42.44	

Whatever its effect on the quality of the search, San Mateo's practice of limited off-line prints had a marked effect on searches completed at that library. San Mateo County printed an average of 29 citations per search off-line, whereas San Jose printed 66, Santa Clara County 71, and Redwood City 75.

Table 9 breaks down into intervals the number of citations printed out off-line per data base (more than one base could be used on each search) by searcher and indicates the mean number of prints per base both for the searches for which prints were made (the non-zero entries), and for all searches, including those for which no prints were made. The figures differ strikingly.

The intervals show an expected decrease in the number of occurrences, moving from smaller numbers of printouts to larger. No off-line prints were requested from more than half of the data bases; this could be because the search results were not satisfactory, or because the relevant citations were printed on-line. No searcher made a practice of printing monumental bibliographies. The mean values are more informative, showing some fairly significant differences among searchers.

The mean number of off-line prints for all cases (including those searches for which no prints were made) was 14.26 for San Mateo County's searchers taken all together, ranging individually from 10.73 to 19.04 (Table 9). Redwood City's overall average was 32.98; its librarians ran from 16.94 to 39.38. Santa Clara's overall figure was 29.40, the individual figures from 21.81 to 40.39. San Jose's library-wide mean was 23.92, its individual low 18.31 and its high 32.37.

Generally, it can be said that searchers varied widely when deciding how many citations were necessary and/or acceptable to the user. The intra-library similarities are marked only for San Mateo County. Otherwise it was the individual and not the library who was the determining factor in the number of print-outs.

TIME PER TASK

Another major source of difference was the staff time devoted to each task. From the descriptions of the various libraries' request procedures, it is evident that some placed more emphasis on certain tasks than did others. Tables 10 and 11 quantify this difference in two ways. Table 10 is the mean time per task, by library, for all non-zero entries only; i.e. once it is decided that a task is to be performed, that is the average time for that task. Table 11 averages into Table 10's figures the entries with values of zero, i.e. those searches for which that task was not performed. In many cases zero entries are a significant proportion of the observations.

Table 10

Mean Time Per Task By Library
(in minutes)

Task	Mean Task Time By Library								Overall Mean time	
	Redwood City		Santa Clara County		San Mateo County		San Jose			
	Mean	# of observation	Mean	# of observation	Mean	# of observation	Mean	# of observation	Mean	# of observation
Reference interview	9.65	91	12.92	89	9.97	59	8.57	56	10.50	295
Originating library preparation	21.00	1	8.33	2	19.58	42	5.00	1	18.65	47
Dialog library preparation	12.07	45	11.44	45	9.72	76	9.47	57	10.48	223
Search	19.63	138	30.42	103	14.09	91	28.31	76	22.72	408
Dialog library follow-up	9.34	66	12.64	91	16.31	83	9.40	67	12.21	307
Originating library follow-up	18.00	2	15.00	2	10.60	38	-	-	11.48	42
Follow-up with patron	7.58	31	8.03	66	7.79	43	6.52	31	7.61	171

Table 11

Mean Time Per Task by Library, All Entries
(in minutes)

Task	Mean Task Time by Library				Overall Mean Time n = 411
	Redwood City n = 138	Santa Clara County n = 103	San Mateo County n = 93	San Jose n = 77	
Reference interview	6.36	11.61	6.33	6.23	7.54
Originating library preparation			8.84		
Dialog library preparation	3.94	5.00	7.94	7.01	5.69
Search	19.63	30.42	13.79	27.94	22.55
Dialog library follow-up	4.47	11.17	14.56	8.18	9.12
Originating library follow-up			4.33		
Follow-up with patron	1.70	5.15	3.60	2.62	3.17
Total	36.10	63.35	59.39	51.98	48.07

Table 10 shows differences among libraries, but none vary much from the group mean. The single exception is search time, which varies from a low of 14.09 minutes for San Mateo County to a high of 30.42 minutes for Santa Clara County, with an overall mean of 22.72 minutes. (More about this later.) However, Table 11 brings out more differences. For instance, reference interviews were around 6.3 minutes for all libraries except Santa Clara County, where the average was 11.61. The numbers of observations in the two tables explain the difference; 86% of Santa Clara County's requests were accompanied by a reference interview, while the next highest percentage was San Jose at 73%. Similarly, DIALOG library follow-up time as measured in Table 11 varies from a low of 4.47 minutes (Redwood City) to a high of 14.56 minutes (San Mateo County). Only 48% of Redwood City's requests had the DIALOG library follow-up performed, while 89% of San Mateo County's did.

The conclusion therefore is that the determining factor in time per task was whether or not the task was performed. Once the decision was made to perform the task, the time required was similar among libraries. The search time was the one important exception. This time is especially important since it determined not only salary expenditures, as did the other tasks, but also the charge for connect time, which, as will be seen, was a significant part of the overall cost.

In order to statistically determine if there were differences in the times taken to perform each of the seven tasks across the four libraries, a set of Analyses of Variances was conducted. These variances are reported in Table 12. The table shows that there were significant differences in the times taken to perform the reference interview, search, and DIALOG library follow-up.

While the Analysis of Variance indicated some differences in group means, by itself it did not indicate which of the four library's task times was significantly different from any other. To determine this, systematic comparisons of all possible combinations of task time mean values were performed and the results evaluated using Scheffe's test (1). Table 12 shows that for the Reference Interview there was a significant difference between Santa Clara's (SC) time and the times of Redwood City (RC), San Mateo (SM), and San Jose (SJ), but that there was no difference between the latter three means. Similarly for search time, Redwood City and San Mateo were significantly different from Santa Clara and San Jose.

PATRON'S PRESENCE

One decision that the participating libraries had to make was whether the requestor should be present during the search. On the one hand, the requestor was often much more familiar with the subject than the librarian was and could provide both an instant evaluation of how well the search was going and perhaps suggest alternative search strategies. The process of the search might also help draw out a non-communicative patron, and help him/her understand what the librarian needed to know.

(1)See (Kirk, 1968) for a discussion of this procedure. The test was conducted at the 0.05 level of significance.

Table 12

Analysis of Variance for Task Times

Variable Name	Source of Variance	Degrees of Freedom	Mean Squares	F Ratio	F Probability	Scheffe's Test
Reference Interview Time	Between Groups Within Groups	3 291	270.97 26.46	10.24	.000	(RC,SM,SJ) vs (SC)
Originating Library Preparation Time	Between Groups Within Groups	3 45	184.66 897.27	0.21	.892	
Dialog Library Preparation Time	Between Groups Within Groups	3 219	85.45 73.84	1.16	.327	
Total Search Time at Terminal	Between Groups Within Groups	3 406	5528.85 258.71	21.37	.000	(RC,SM) vs (SC,SJ)
Dialog Library Follow-Up Time	Between Groups Within Groups	3 304	830.86 41.96	19.80	.000	(RC,SJ) vs (SC) vs (SM)
Originating Library Follow-Up Time	Between Groups Within Groups	2 41	111.95 83.20	1.356	.272	
Follow-Up Time With Patron	Between Groups Within Groups	3 167	16.73 15.71	1.064	.366	
Total Time for Search	Between Groups Within Groups	3 407	14142.62 710.70	19.90	.000	(RC) vs (SC,SM,SJ)

On the other hand, a user unfamiliar with the search process could slow the search down, requiring lengthy explanations during the terminal sessions. The librarian might also prefer not to have someone watching over his/her shoulder during the search.

Ultimately, only Santa Clara County made a regular practice of allowing the patron to be present during the search. Thus figures are available on the relative lengths of searches with and without the patron, but only for Santa Clara County.

Of the 103 Santa Clara County searches reported, 60 were performed without the patron, and 43 with the patron. The average search time for those with no patron present was 25.0 minutes versus 37.9 for those with patron present. The chance of this large a difference resulting simply from random fluctuations is less than 1%; searches with the patron take significantly longer. However, the total time required for all tasks did not vary that much; without patron, the average was 56.3 minutes, with patron it was 61.0 minutes. There is a 7.5% chance that the sample means are not significantly different. Thus it is not conclusive that the patron's presence at the search has an adverse effect on total search time. It may be that the patron's presence at the terminal shortens the time required for either the reference interview or for the follow-up with the patron.

COSTS

All of these time figures can now be translated into costs, using the salary and fee schedules (Tables 2 and 4). Rather than calculate costs based on the average time figures arrived at earlier, the procedure in this study was to calculate directly the actual costs for each search based on the time required, the salary of the person performing the task, and the data base(s) used.

The mean salary costs for each task, by library, are presented in Table 13. The figures for the individual libraries cluster reasonably around the means for the group as a whole. The one exception is originating library figures, for which only San Mateo County had enough observations to be meaningful. The costs were quite low, averaging $1.03 for the reference interview, $1.95 for originating library preparation, $1.02 for DIALOG library preparation, $1.16 for follow-up at the DIALOG library, $1.07 for originating library follow-up, and finally $.74 for the post-search time spent with the user.

Table 14 presents another analysis of the costs incurred in the on-line search itself. The overall average was $17.29 for computer time, $9.16 for off-line printouts, and $2.24 for the labor at the terminal. This, combined with $5.02 for all other labor (detailed in Table 9), brings the cost for the entire search process, for all libraries, to an average of $28.41.

Table 13

Mean Salary Cost Per Task By Library
(in dollars)

Task	Mean Task Cost By Library				
	Redwood City	Santa Clara County	San Mateo County	San Jose	Overall
	Mean	Mean	Mean	Mean	Mean
Reference interview	$1.04	$1.21	$1.02	$ .80	$1.03
Orginating library preparation	1.14	.76	2.10	.45	1.95
Dialog library preparation	1.21	1.06	.97	.87	1.02
Search	2.03	2.83	1.47	2.70	2.24
Dialog library follow-up	.87	1.17	1.61	.87	1.16
Orginating library follow-up	.53	1.39	1.08	-	1.07
Follow-up with patron	.79	.75	.81	.60	.74

Table 14

Search Costs
(in dollars)

Cost Element	Mean Task Cost By Library				Overall Mean Cost
	Redwood City	Santa Clara County	San Mateo County	San Jose	
Data Base Charges	$14.51	$22.16	$10.55	$23.69	$17.29
Off-Line Print Charges	9.73	12.60	3.64	12.96	9.16
Search Labor Cost	2.03	2.83	1.47	2.70	2.24
Labor Cost for All Other Tasks	3.64	5.91	6.22	4.93	5.02
Total Cost of Search[1]	25.33	35.17	19.74	35.19	28.41

[1]The total cost is not additive due to differences in the number of observations for each cost element.

The individual libraries differed significantly from this mean, ranging from a low of $19.74 (San Mateo County) to a high of $35.19 (San Jose) and $35.17 (Santa Clara County). Redwood City was a little below the average at $25.33. No one item accounts for the differences.

San Mateo County, with the lowest overall average, had the highest labor cost for activity other than time at the terminal. They more than make up for this, however, by spending less on terminal time, both for computer time and labor, and also by printing far fewer citations. Referring back to the mean time per task in Table 10, it is evident that San Mateo County's searchers spent significantly less time at the terminal per search. The higher costs for other labor probably came from the added step incurred when requests originate at a library other than the one at which the terminal is located.

Santa Clara County and San Jose differed by only a few cents in the total search cost but their breakdowns were not identical. San Jose spent more on data base charges, though a little less on salaries for the other tasks. This indicates the use of more expensive bases and/or less expensive personnel (note that their most prolific searcher was the unpaid volunteer whose time was evaluated at the Librarian I rate). As indicated in Table 10, San Jose generally spent a little less time on each task than did Santa Clara County.

To look a little closer at the search process, Table 15 presents search costs for the high-frequency searchers. At this level there is more across-library variation. The three searchers who spent the least salary at the terminal belong to the low-ranking San Mateo County (searchers 8, 3, and 10). The next three lowest are from other libraries, two from Redwood City (23 and 25) and one from San Jose (52). One of Redwood City's searchers (23) is a relatively expensive Librarian III. The highest salary costs were incurred by Librarian I's from both San Jose (number 53) and Redwood City (26).

The data base charges, which are the largest single element in the search cost, follow a similar pattern. The six lowest searchers in terms of salary also account for five of the six lowest searchers in data base charges. The off-line print charges, as indicated earlier, differ strikingly, bringing all four of San Mateo County's searchers down to the four lowest total search costs. San Mateo County searchers are followed by searcher 23 from Redwood City and 32 from San Jose. The highest total search cost was attributed to searcher 37 of Santa Clara County, followed by searcher 36 from Santa Clara County.

Generally, however, differences within a library were not as striking as those between libraries. Thus it appears that although individual searchers do vary, the overriding consideration is the library at which they are searching. This could be a result of library policy, either written or implied (evidence the San Mateo County effort to keep down the number of prints). It could also be a result of interaction among the searchers at a library, arriving among themselves at a definition of what is an acceptable search length and number of prints. Also, since most of the searchers received at least part of their training from others in their own libraries, it could be that the intramural similarities reflect a common teacher.

Table 15

Mean Search Cost for Selected Searchers

(in dollars)

Library	Searcher Code	Search Labor Cost	Data Base Charges	Off-Line Print Charges	Total Cost
San Mateo	3	1.45	10.76	2.72	21.56
	5	2.24	12.74	3.15	19.82
	8	0.94	6.02	4.11	15.35
	10	1.48	12.17	4.25	21.59
Redwood City	23	1.81	11.80	12.29	23.86
	24	2.84	20.90	7.06	32.45
	25	1.61	13.02	12.34	24.86
	26	3.43	24.69	5.27	33.37
Santa Clara	35	2.56	19.19	13.04	33.40
	36	3.19	23.70	13.10	34.97
	37	2.80	24.70	10.97	37.57
San Jose	50	2.48	22.07	9.67	34.89
	52	1.89	15.03	32.95	24.35
	53	3.96	34.98	15.86	27.13

Two factors determine the computer charges for connect time: the time required for the search and the choice of data base. It is possible to separate out those two effects by considering the ratio of the cost of data base connect time to search time. This has the effect of normalizing the connect charge to reflect varying lengths of time spent searching. The actual ratios were:

Redwood City	$ .74/min.
Santa Clara Co.	$ .73/min.
San Mateo Co.	$ .77/min.
San Jose	$ .85/min.

From this it can be concluded that all libraries were spending roughly the same amount per connect minute for computer time. San Jose was slightly higher, denoting a slight tendency to make greater use of the more expensive bases.

Similarly, the salary cost for any part of the search depended both on the time spent and the salary of the person performing the task(s). Considering again the on-line search, the ratios of labor cost to time at the terminal for the libraries were:

Redwood City	$ .10/min.
Santa Clara Co.	$ .09/min.
San Mateo Co.	$ .11/min.
San Jose	$ .10/min.

The libraries did not vary greatly in the overall salary cost per minute at the terminal.

Broadening this analysis to the entire DIALOG process, the ratios of the total cost of the DIALOG process (labor, computer charges, printouts) to the total time, from reference interview through follow-up were:

Redwood City	$ .71/min.
Santa Clara Co.	$ .55/min.
San Mateo Co.	$ .33/min.
San Jose	$ .68/min.

A likely explanation for the low San Mateo County figure is that their searches tended to be more labor-intensive, while the others were more computer-intensive and computer time is much more expensive than people-time. Redwood City's searches also tended to use less computer time. However, their off-line printout charge was much higher than San Mateo County's and their chief searcher during the data collection was a relatively expensive Librarian III (the other libraries relied mainly on Librarian I's)(1).

(1)San Jose Public Library has recently indicated that with the advent of user charges for computer time, they are moving toward a more labor-intensive and less computer-intensive search pattern.

COMBINATIONS OF TASKS

In the description of the DIALOG process as it is performed in each library, it was pointed out that the libraries differed in the emphasis placed on different tasks. For instance, San Jose made no special effort to perform either a reference interview or to spend post-search time with the patron. Santa Clara County, on the other hand, tried to have the patron submit to a reference interview with a DIALOG staff member.

The greatest number of tasks that might possibly be performed would be for a search that originated in a library without a DIALOG terminal, for which all seven tasks listed on the time sheet could be performed. For a request originating at a DIALOG library, the maximum number of tasks would be five, eliminating the two tasks that are specific to non-DIALOG libraries.

Table 16 shows the twelve (out of 128 possible) combinations of tasks that were most frequently performed. These twelve accounted for 338 of the searches performed, or 82%. The most-used single pattern of tasks was number ten in Table 16. This pattern consisted of a reference interview, DIALOG preparation, search, DIALOG follow-up, and a follow-up with user; the "complete" DIALOG-originating pattern. The next most common was pattern seven, the same as ten but with the DIALOG preparation omitted. Pattern twelve, every possible task performed, occurred only ten times.

Given that different patterns of tasks mean that more or fewer steps are performed for each search, it follows that different patterns result in different costs.

Table 17 shows the mean total search cost, by library, for each of the frequently occurring combinations of tasks. Surprisingly, the lowest overall mean is the one with the most tasks (combination number 12). This figure is not conclusive, however, due to both the low number of cases (ten) and the fact that all cases were from San Mateo County, which had consistently low search costs.

The second least costly (again with figures only from one library, in this case Redwood City) was the sixth instance, which consisted of a reference interview, the search, and follow-up with patron.

Interestingly, the bare bones search - no task performed other than the actual on-line search - although inexpensive, at $15.67 (number one), was only the third least expensive combination.

The most expeive combination, considering all libraries together, was combination number eight, reference interview, search, DIALOG library follow-up and follow-up with user at $38.08. The next most expensive was pattern number ten, which includes all the tasks in combination eight plus DIALOG library preparation. The additional task lowered the cost to $34.23.

Table 16

Most Frequent Combinations of Tasks Performed for DIALOG Searches

Combination Number	Combination of Tasks							Frequency of occurrence
	Reference Interview	Originating Library Preparation	Dialog Library Preparation	Search	Dialog Library follow-up	Originating Library follow-up	Follow-up with User	
1	0	0	0	1	0	0	0	21
2	0	0	0	1	1	0	0	22
3	0	0	1	1	1	0	0	21
4	0	0	1	1	1	0	1	11
5	1	0	0	1	0	0	0	35
6	1	0	0	1	0	0	1	10
7	1	0	0	1	1	0	0	44
8	1	0	0	1	1	0	1	41
9	1	0	1	1	1	0	0	40
10	1	0	1	1	1	0	1	70
11	1	1	1	1	1	1	0	13
12	1	1	1	1	1	1	1	10
Total								338

Note: 0 indicates task not performed; 1 indicates task was performed.

Table 17

Total DIALOG Search Cost for Frequently Occuring Combinations of Tasks (in dollars)

Combination Number	Mean Total Search Cost				
	Redwood City	Santa Clara	San Mateo	San Jose	All Libraries
1	16,85	22.01	15.62	12.00	15.67
2	23.42	76.30	10.20	31.56	25.50
3	35.30	33.86	18.92		25.97
4	35.30	33.86	18.92		25.97
5	23.11	9.62		47.85	23.43
6	9.75				9.75
7	25.96	28.30		35.26	28.13
8	26.82	43.93	16.33	28.56	38.08
9	29.59	28.40	16.07	38.71	31.09
10	33.40	39.62	19.62	36.90	34.23
11			26.07		26.07
12			1.61		1.61

These data on task combinations imply that there is no direct correlation between the number of tasks performed and the cost of the search. San Mateo County had the lowest average cost, while generally performing more tasks per search than other libraries did.

ELAPSED TIME

Two measures of time are important to the user of any kind of search service. One is the time that the user or his/her surrogate actually spends on the search. The use of on-line searching reduces this measure from hours, even days, which must be spent leafing through abstracting and indexing publications, to only minutes spent conferring with the search personnel.

The second measure is waiting time - how long the user must wait from the instigation of the request until the results are received. The waiting time is an indication of how quickly the search service processes the request.

The mean time that requests spent in the system, (1) from the first patron contact until the last task (whatever that was) for all of the libraries together was 7.79 days (calender, not working days). For the individual libraries the figures were: Redwood City, 6.02 days; Santa Clara County, 4.87 days; San Mateo County, 14.67 days; and San Jose, 6.79 days. (2)

Table 18 analyzes this data further. This table shows the mean time between various pairs of tasks, by library. Not every task was performed for each search; the figures are only for those cases in which both elements of the pair in question were performed. It is also worth noting that the pairs examined were not necessarily successive tasks, but those for which a significant number of observations were available. There is overlap among the pairs examined, e.g. reference interview to search subsumes reference interview to originating library preparation, reference interview to DIALOG library preparation, and DIALOG library preparation to search.

Redwood City's data show a regular progression through the tasks. Reference interview to search required about two days of the six overall; search to follow-up with patron tended around three days. Considering that off-line printouts were sent first class mail, and required a minimum of two days to arrive, the figures were quite good.

(1)The minimum time possible is one day and that is for same day service.

(2)All of the libraries would expedite rush requests, if the patron had a good reason.

Table 18

Days Elapsed Between Tasks

Task Interval	Library: Redwood City Mean	Redwood City # of observation	Santa Clara Mean	Santa Clara # of observation	San Mateo Mean	San Mateo # of observation	San Jose Mean	San Jose # of observation	MEAN
Reference Interview to Originating Library Preparation					1.74	14			1.74
Reference Interview to Dialog Library Preparation	1.91	24	2.23	38	4.63	52	5.46	47	3.90
Originating Library Preparation to Dialog Library Preparation			4.50	2	4.94	37			4.9
Dialog Library Preparation to Search	.59	41	.88	45	3.77	76	0.75	58	1.79
Reference Interview to Search	1.93	92	3.05	89	8.77	58	6.12	56	4.9
Search to Dialog Library Follow-Up	1.60	66	0.16	92	1.85	83	1.08	67	1.19
Search to Originating Library Follow-Up	2.00	2	2.50	2	7.31	38			6.83
Search to Follow-Up With Patron	3.03	30	2.66	66	7.72	43	3.03	31	4.00
Dialog Library Follow-Up to Originating Library Follow-Up					6.00	38			6.00
Originating Library Follow-Up to Follow-Up With Patron					0.93	15			0.93
Dialog Library Follow-Up to Follow-Up With Patron	1.70	20	3.70	60	7.34	41	2.26	30	4.14
Average time in system, all cases	6.02	140	4.87	103	14.67	91	6.79	77	7.79

Santa Clara County shaved about a day off Redwood City's time, primarily in the post-search time. The Santa Clara County off-line print figures showed a greater tendency on the part of searchers to request no off-line prints. This tendency might have been a result of having the patron present frequently, (citations that might have been printed off-line in the patrons' absence might be printed on-line or not at all). This would eliminate the wait for the U.S. Postal Service and might account for Santa Clara County's lower post-search figure.

San Mateo County was the slowest of the group, at 14.67 days. Table 18 shows them with consistently higher figures than the other libraries, with the time divided almost equally between pre- and post-search (8.77 days vs. 7.72). San Mateo County was the only library with a significant number of tasks performed at the originating library, which added a step and a possible bottleneck to the search process. It may also be that the lack of direct contact with the patron made San Mateo County tend not to hurry in searching a request once it was received. Furthermore, DIALOG requests were competing with other subject requests received from the branches in much the same way as the DIALOG requests. In other words, DIALOG requests simply had to wait their turn.

San Jose's total of 6.79 days was a distant second to San Mateo County. The bulk of that time (6.12 days) was pre-search, echoing San Mateo County's pattern of having the request sit waiting. In this case, however, the wait was for San Jose's part-time searcher. Also, the San Jose staff had complained early in the project of being squeezed between DIALOG and their regular duties. This long pre-search wait and short post-search time suggest a general backlog. Items were handled speedily, once their turn came, but there were others competing for the staff's attention.

CONCLUSIONS

This study has reported the costs of on-line bibliographic searching in four public libraries in the San Francisco Bay Area using the Lockheed DIALOG system. This study was conducted during a period when search time was provided to the libraries without charge and the searchers were moderately experienced. It was found that there were seven different tasks involved in the search process. In the course of the study 35 individuals from the four libraries reported their own processing times for 411 search requests.

Two limitations should be considered in evaluating the results of this study. One is the self reporting nature of the data. Participants recorded time spent on search tasks rather than being observed and measured by others. Consequently, a possible bias is introduced depending on the accuracy with which the searchers recorded their times. A second limitation is the experimental nature of the project. On-line searching of the DIALOG system was provided at no cost to the public libraries through a grant from the National Science Foundation. Thus the searchers were under little economic pressure to perform effectively.

Considerable variation in the time required to perform search tasks was found between libraries and between searchers. For example, the average connect time to a data base for an individual search was 22.72 minutes. But this varied from 19.63 minutes at the Redwood City Public Library to 30.42 minutes at Santa Clara to 28.31 minutes at the San Jose Public Library. Similar variations were found in the overall cost of a bibliographic search: Redwood City's average was $25.38, Santa Clara's was $35.17, San Mateo's was $19.74, San Jose's was $35.19, and the overall average for the 411 searches was $28.41. The most expensive portion of the overall cost was the actual on-line search time, which averaged $17.29 for all searches. The next most expensive element of the total cost was that of off-line prints--this averaged $9.16 for all searches.

The only element that was omitted from the cost calculations was telephone line charges. This was due to the fact that the four libraries were in close physical proximity to Lockheed's computer center. If this situation did not hold for other searching locations it would be necessary to add the telephone costs. Assume a $10 per hour telephone connect charge through an organization such as TYMSHARE, and a mean search time of 22.72 minutes. This results in an average telephone charge of $3.79 which, when added to the $28.41 search total, totals $32.20 for the average search.

ACKNOWLEDGEMENTS

This project was supported 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 Communication Research, Inc., Palo Alto, California as a subcontractor for this evaluation of the project. We are indebted to Dr. Roger K. Summit and Mr. Oscar Firschein of Lockheed for their assistance. Dr. Colin K. Mick and Ms. Alice Ahlgren of Applied Communication Research, Inc. were especially helpful in all phases of the project. We acknowledge with considerable gratitude the assistance of Etiennette Manowski in the data coding and analysis phase of the project. But our major debt is to the 35 individuals who spent nearly three months keeping track of their time for this study.

REFERENCES

1. Summit, Roger K. REMOTE INFORMATION RETRIEVAL FACILITY. Lockheed Aircraft Corporation, Palo Alto, California. NASA Report CR-1318, April 1969.

2. Benenfeld, Alan R., Mary E. Pensyl, Richard S. Marcus, and J. Francis Reintjes. NASIC AT MIT PHASE 1 REPORT 16 JULY 1973 - 28 FEBRUARY 1974. Report ESL-R-543, Electronic Systems Laboratory, Department of Electrical Engineering Massachusetts Institute of Technology, Cambridge, Mass. March 29, 1974.

3. Lawrence, Barbara, Ben H. Weil, and Margaret H. Graham. MAKING ON-LINE SEARCH SERVICES AVAILABLE IN AN INDUSTRIAL - RESEARCH ENVIRONMENT. Paper presented at the American Society for Information Science, Third Mid-Year Meeting, May 1974.

4. Elman, Stanley A. "Cost Comparison of Manual and On-Line Computerized Literature Searching." SPECIAL LIBRARIES 66:1 (January 1975) 12-18.

5. Lancaster, F.W. and E.G. Fayen. "Information Retrieval On-Line." Melville Publishing Co., Los Angeles, California, 1973.

6. Kirk, Roger E. "Experimental Design: Procedures for the Behavioral Sciences." Brooks/Cole Publishing Co., Belmont, California, 1968.

DIALOG TIME SHEET

Please enter time in minutes to the nearest five minutes. Please date and initial each entry.

DIALOG LIBRARY: 1. RCPL____ 2. SCCO____ 3. SMCO____ 4._____

Patron name (optional): _____
Request number: _____

Originating Library _____

EVENT	TIME IN MINUTES	DATE	INITIALS
REFERENCE INTERVIEW initial patron contact			
ORIGINATING LIBRARY PREPARATION TIME without patron			
DIALOG LIBRARY PREPARATION TIME without patron			
DIALOG CONNECT TIME	don't enter time	(do date and initial)	
DIALOG LIBRARY FOLLOW-UP TIME without patron			
ORIGINATING LIBRARY FOLLOW-UP TIME without patron			
FOLLOW-UP TIME with patron			
MISCELLANEOUS describe:			

DIALOG CONNECT TIME	Patron present? Check if yes____	
DATA BASE	ELAPSED TIME	NO. OF OFF-LINE PRINTS

Use reverse to list additional bases searched on the same request as above.

Appendix B

Alternative Cost Calculations

In Tables 13 and 14 of the text the mean salary costs and mean search costs were summarized. These tables were calculated on the basis that each occurrence of a task resulted in a cost of performing the task and that the total cost was the sum of the task costs for all tasks which were performed.

Tables B-1 and B-2 are the analogies to Tables 13 and 14 with a different method of computation used. In Tables B-1 and B-2 the mean unit cost of a task is determined by dividing the total cost by the total number of valid observations, not just the number of observations for which the time was non-zero. While the mean search costs (Table B-2) remain relatively stable, there is a considerable reduction in the salary costs per task (Table B-1).

Table B-1

Mean Salary Cost Per Task by Library,
All Entries
(in dollars)

Task	Mean Task Time by Library				
	Redwood City	Santa Clara County	San Mateo County	San Jose	Overall
	Mean	Mean	Mean	Mean	Mean
Reference interview	.68	1.05	.65	.58	.75
Originating library preparation			.95		.22
Dialog library preparation	.39	.46	.81	.64	.55
Search	2.03	2.83	1.47	2.70	2.24
Dialog library follow-up	.42	1.04	1.44	.75	.87
Originating library follow-up			.44		.11
Follow-up with patron	.18	.48	.37	.24	.34

Table B-2

Mean Search Costs, All Cases
(in dollars)

Cost Element	Mean Task Cost by Library				Overall Mean Cost
	Redwood City	Santa Clara County	San Mateo County	San Jose	
Data base charges	14.51	22.16	10.55	23.69	17.29
Off-Line Print charges	7.48	7.09	2.90	6.56	6.17
Search Labor Cost	2.03	2.83	1.47	2.70	2.24
Labor Cost for All Other Tasks	1.67	3.03	4.66	2.21	2.81
Total Cost of Search	25.69	35.11	19.58	35.16	28.51