

AN ANALYSIS OF COST FACTORS IN MAINTAINING  
AND UPDATING CARD CATALOGS

J. L. DOLBY and V. J. FORSYTH:  
R&D Consultants Company, Los Altos, California

*This study enumerates and compares costs of manual and computerized catalogs. The difficulties of making comparative cost studies are examined. The report concentrates on the problems of cost element definition and on the reporting of as many comparable sources as possible. Results of cost studies are presented in the form of tables that show comparative costs of cataloging, card processing, conversion, and manual and computerized processing. There are also tables on card catalog costs. Conclusions are that the costs of manual and automated methods are essentially the same for short entries, and that there is a substantial economic advantage for automated methods in full entries.*

A side benefit of the present interest in library automation is the amount of attention now being given to study of the traditional methods of librarianship. This phenomenon is hardly unique to librarianship; in almost every area of human endeavor where attempts have been made to introduce the use of computers, workers in the field have suddenly discovered that they did not understand some of their long-standing methods quite as fully as they had believed. The source of this seeming anomaly is easy to find: to program a computer, it is necessary to specify the work to be done in much greater detail than is necessary to explain the same problem to a human being, that curious human phenomenon known variously as "common sense" or "experience" making up the difference. It has not been uncommon over the past decade to hear many survivors of the "automation experience" admit that a main benefit of use of the machine was

acquisition of better procedures through a more detailed understanding of the process involved.

Improved knowledge of "processes about to be automated" extends to the cost of the process as well, and with added force. In recommending the substitution of one procedure for another in a cost-conscious atmosphere, it behooves one to proffer sound financial reasons for doing so. Computers are expensive devices. They also represent expenditure of a different kind of money: capital or lease funds in place of labor expense. Thus, although one can still hear the occasional cry that it is difficult to obtain reasonable cost data on various parts of library operations, it is becoming increasingly difficult to pick up an issue of almost any library journal that does not include at least one piece of cost information.

This paper is concerned with the cost of maintaining and updating card catalogs. As the authors have observed elsewhere (1), the cost of computing is going down at a rather spectacular rate, while the cost of labor is increasing. If this trend continues, almost every library will be forced to automate certain aspects of the catalog operation at some point in time. The cited report provided some information about the cost of computerized library catalogs. By adding a summary of the cost factors in the use of card catalogs, this article should place in slightly better perspective the more difficult problem of deciding (in the context of a particular library) when the crossover point between manual and automated methods is to be reached.

The plan of attack remains essentially the same as in the previous report: selection from among the growing number of papers on the subject those that provide comparable sets of cost information pertinent to the various cost elements of the card catalog operation. It is appropriate, therefore, to begin this study with a brief description of the difficulties in comparing cost statistics in such a way.

#### PROBLEMS OF COMPARATIVE COST STUDIES

Although comparative cost studies have much to recommend them, they are fraught with certain difficulties (2). In the first place, few librarians would group elementary cost operations in precisely the same way. One library may consider a particular element of cost as part of the acquisitions operation and a second as a part of the cataloging operation; a third may ignore it altogether, or include it in the burden or overhead cost. Nor is this mere capriciousness on the part of members of the library community. Library operations not only differ from one another, but they also change with time.

Consider, for example, the problem of obtaining a set of catalog cards for a particular monograph. Any or all of the following alternatives might be in use at a given library: the cards may be 1) supplied with the book as a service provided by the bookseller at some extra cost; 2) ordered from the Library of Congress; 3) provided by a centralized cataloging

operation serving several libraries (as in a county or state library system); 4) prepared by catalogers working in the library; or 5) generated by computer program from standard listings (e.g., from MARC tapes).

Comparing any two of these procedures within a given library does not present any overwhelming problems, although minor questions of definition do occur (for example, how much of the cost of ordering should be allocated to the acquisitions department and how much to the cataloging department when both the book and the catalog cards are obtained simultaneously from the same source?). However, to compare costs from two different libraries, it is essential to know what proportion of each card source was used by each library. Fortunately for the purposes of this study most libraries are presently using a mix of method 2 (LC) and method 4 (own catalogers), and at least some provide sufficient information to enable determination of the appropriate mix for each. However, the problem is indicative of one essential difficulty in comparative cost analyses; and one that, although eased, would not be eliminated by having all libraries band together for adoption of a standard costing procedure.

A second difficulty arises from temporal and geographic differences in the cost of manpower. On the surface, this problem can be eliminated, or substantially reduced, by having all studies based on man-hours spent, rather than on dollars required per item, and a number of writers have suggested such a change in reporting procedure. However, the problem is not quite so simple. For example, determining the number of man-hours spent on cataloging adds cost to the study that tends to reduce the number of libraries willing to report; those that do report may or may not be a representative sample of the total.

However, there is a more basic problem. In almost all libraries the real restraint on activities is financial: there are just so many funds available for cataloging and these must be used to at least keep the backlog of uncataloged material down to the amount of space available to store it. Suppose, for instance, that the amount of material to be cataloged increases by ten percent from one year to the next and that the catalogers are fortunate enough to obtain ten percent salary increases over the same period. It is not impossible to consider that in some libraries the catalogers may be forced to "earn" this raise by absorbing at least a part of the increased load without extra help. Balancing salary increases by productivity increases is, of course familiar in industry and may well exist in libraries. As evidence that such an effect is present, it is noted later in this report (see Table 4) that three studies made in three rather different libraries over a period of six years showed costs of from \$0.228 to \$0.235 per card for preparation, production, and filing.

The total range (\$0.007) is only three percent of the average cost per card. (\$0.230). Such close agreement would be startling if it were found in three simultaneous studies of three nearly identical library operations. To set this agreement aside as pure coincidence seems unwarranted. It is

more reasonable to assume that librarians are forced to operate under strong financial constraints and that they adjust their performance to those constraints through hiring of less well-trained personnel, increased time pressures on all personnel, etc. If this is the case, "standardized" reporting through time figures might be quite misleading unless cost figures were reported as well.

Finally, there is the question of allocating burden or overhead. Potentially, burden could present a severe problem, and occasionally it may. However, in most of the reports cited here, burden is either ignored or separately stated and there is no reason to suspect that the results given in the summaries are noticeably biased by unseen burden differences. Nevertheless, it would be of interest to determine proper overhead figures for library operations, as the switch to automation (which seems inevitable), will entail the use of more machines and fewer people, which in turn may drastically alter the overhead structure.

#### THE USE OF COST INFORMATION

Having noted some of the difficulties that tend to cloud cost comparisons, it is perhaps useful to investigate how cost information is likely to be used. The nature of the problem can be illustrated by two rather different situations. One is exemplified by Library "A", a large public library of some years' standing. It is considering the possibility of changing from its present manual procedures to some form of automation, and wishes to determine a reasonable strategy for implementing such a change over the next five years. Library "B", otherwise comparable to "A", has been keyboarding the catalog records of its current acquisitions for the last three years. It has now decided to convert its retrospective catalog and wishes to choose the most economic procedure for this step.

The differences in the problems facing two such libraries are basically the classic differences between strategy and tactics. Library "A", must lay out a long-term plan, taking into account the growth in its collection over the five-year period, likely changes in equipment and personnel available to it, increases in labor costs, decreases in equipment usage costs, etc. Library "B", on the other hand, is in the position of making a specific set of decisions as to whether the work should be done in-house or subcontracted; whether the Library should use punched cards, punched paper tape, or optical character-recognition devices; and so forth.

In terms of cost, Library "B", has to prepare a specific budget request for its funding agency, and it is reasonable to assume that that funding agency will require assurance that the task is to be accomplished at the minimum cost consistent with the designated quality level. Cost differences of as little as five percent may be quite important to Library "B". General cost summaries can be of use only in enumeration of the possible alternatives. Even the accounting procedures in effect in the local system will have a bearing on the final decision.

Thus, the primary utility of a general cost summary to the library about to commit itself in a tactical situation is the information it can provide about the problem statement: which cost factors other libraries have been able to identify in similar situations; which of the various alternatives may be safely eliminated from consideration on the grounds that their present costs are considerably higher than other existing methods; and so forth. The likelihood seems remote that any general study, or, for that matter, any particular study, will be sufficiently applicable to the library now undertaking the problem to enable it to take over cost structures unchanged.

Library "A", faced with establishing a long-range plan, has much more flexibility available to it. Its interest in specific costs will be established by some gross notion as to what quantity of funds are likely to be available over the period under plan. Some procedures may be seriously considered because they are relatively new and untried and hence of potential interest to national funding agencies who would not consider funding further experiments with procedures that have been thoroughly tested. Access to good cost information of such well-tested procedures will help in establishing the likely costs for important aspects of the overall plan. Of even greater interest is the possibility that certain costs are likely to undergo substantial change over the planning period. For instance, in Reference 1 it was noted that optical character recognition may be a very attractive long-run option for catalog conversion problems precisely because it is so new, and hence has not had time to allow a sufficient number of service centers to spring up to provide truly competitive service capabilities. Computer typesetting with the new generation of hardware is in much the same category.

In both situations it is clear that what is most needed is the enumeration of cost elements on the one hand and operating cost experience on the other. Precise estimates of any one cost element are of relatively little importance, either because they are so likely to change over the long run, or because they are likely to be not appropriate to a specific application even in the short run.

Comparative cost information would therefore seem to provide a good basis for either application. The comparison forces an enumeration of cost elements precisely because one must evaluate the cost structure of each source to be sure that a reasonable comparison is possible. Reporting of the actual experience of several libraries provides a range of experience, not only over several libraries but also over time, so that the extremes reported give an indication of the variability that must be allowed for. In what follows, therefore, concentration is on the problems of cost element definition and on the reporting of as many sources as are comparable in the broad sense. Because precise estimates are not only difficult to obtain, but also unlikely to be relevant to most users, no attempt has been made to provide formal estimates either of the average cost figures or of their underlying variability.

## THE COST OF CATALOGING

The preparation of catalog information for a given monograph is perhaps the most sophisticated operation in the entire catalog operation. As such it is probably the last to be considered a candidate for automation, although it is not unreasonable, even now, to consider the use of computers as aids to the cataloger. Consequently in many operations the cost of cataloging will continue to be an invariant regardless of whether automation is introduced into other aspects of the catalog operation or not. Nevertheless, it is useful to study the cost of catalogs, both to establish the relative cost of cataloging and the subsequent processing steps, and to establish the line of demarcation between the catalog step and the subsequent steps.

Any enumeration of the detailed steps involved in a complex process must be tentative. This is nowhere more true than in the cataloging operation. Fortunately the number of descriptions in detail is growing. For the cataloging operation, three sources of information were used: 1) a detailed analysis made as part of an overall time and motion study of operations in the Lockheed Research Library (3); a detailed study of the cataloging and processing activities of the New York Public Library as a preliminary to possible automation of some of these operations (4); and a detailed study of the acquisitions, cataloging, and other processing operations for the Columbia University science libraries (5). A summary of these studies is given in Table 1.

In addition to the eight items in Table 1, the Lockheed Library study included five other items that we have chosen to include in subsequent operations.

It is generally true that professionals do not like to have their jobs subjected to the minutiae of time and motion study. There is always the ugly feeling that the creative (and most important) aspects of the job cannot be subjected to simple measurement. Nevertheless, cataloging is a continuing effort in most libraries and it is possible to establish some average production rates in terms of number of books cataloged per month or the number of minutes needed per book. The problem, as with most statistical studies, is not with the establishment of objective measurements but rather in the manner in which they are interpreted. Use of comparative statistics does not eliminate the possibility of misinterpretation but it does tend to minimize it.

The comparative studies selected for the cataloging operation, in addition to those already cited, were: a Colorado study based on average cataloging times for eleven librarians from six cooperating libraries (2), and a study of ordering, cataloging, and preparations in several Southern California libraries (6). The catalog cost information for these five studies is summarized in Table 2.

Table 1. *Cataloging Cost Elements**Columbia University Science*

(With LC information)

1. Assign class number
2. Compare book and card, check entries in general catalog, establish subjects, etc.
3. Make necessary changes in LC proof slip, or type temporary slip giving brief descriptive information and class number
4. Completed books revised and sent for shelf listing

(Without LC information)

1. Supply descriptive cataloging
2. Subject analysis, classification and authority work
3. Type workslip for processing section.

*New York Public*

1. Review work done by searcher. Reconcile conflicts and approve new entry forms
2. Full descriptive cataloging
3. Assign subject entries
4. Assign divisional catalog designators
5. Check authority files and establish new authorities and cross references
6. Determine classmark

*Lockheed Research Laboratory*

1. Get book and analyze for subject. Obtain Dewey and Cutter numbers
2. Check shelf list for duplicates and copy number
- 3a. (With LC information)  
Insert and type copy slip and temporary catalog card, check LC subject headings and other references. Descriptive and subject catalog book. Pencil call number on title page
- 3b. (Without LC information)  
Insert and type descriptive part only on copy slip and temporary catalog card. Write subject data only on catalog card. Pencil call number on title page
4. Tear and separate copy slips and temporary cards. Proof and correct as necessary.
5. Take report to reports cataloging
6. Travel to library, check national union catalog or other reference book
7. Count and tally titles cataloged

Table 2. Comparative Costs of Cataloging

<i>Library Source</i>	<i>Date</i>	<i>Average Time, min.</i>	<i>Cataloging Cost</i>	<i>Implied Avg. Salary (per hour)</i>
Lockheed	1967	10.0	—	—
Colorado	1969	28.6	\$ 2.07	\$ 4.34
New York	1968	39.8	6.30	5.25
So. Cal.	1961	44.8	2.23	2.98
Columbia	1967	84.0	5.85	4.17

In the Lockheed and Colorado studies, basic times of each operation were studied and then "standard" time factors added to allow for non-productive time. The standard factors increased the Lockheed times by 13 percent and the Colorado times by 48 percent. (The times in the table include these allowances.) The figures for New York were derived from their reported statements that they processed 65,000 books using 49 catalogers at a total cost of \$409,500 (not including fringe benefits). The Columbia figures have been reduced by 20 percent to eliminate fringe benefits. The implied average salary for each source was obtained by dividing the total cataloging cost by the average time and multiplying by 60 to convert to cost per hour.

The simplest conclusion to reach from a study of Table 2 is that cataloging costs vary widely from one library to another. Average times differ by more than 8 to 1 and total cost varies by more than 3 to 1. The low salary for the Southern California study is presumably explained by the fact that that study was done in 1961. Adjustment of this figure for average salary increases from 1961 to 1968 would undoubtedly bring their total cost more directly in line with the other studies (Bureau of Labor Statistics shows hourly wages increased approximately 30 percent over this period). It would be interesting to know if the presumed increased salaries of the Southern California catalogers has led to a decrease in the average time they spend on cataloging. The more recent data on Colorado and New York suggest that this might be expected.

The Columbia and Lockheed time data represent, perhaps not unreasonably, the extremes in this table. The Lockheed research library is small compared to the others, and Lockheed is, of course, a private corporation, whereas the other sources represent public and university libraries. Columbia, on the other hand, is a large university library; however, the figures given are from a study of cataloging of science monographs, which may be more time-consuming.

As these cataloging cost figures will be used only as a point of comparison with subsequent operations, it is not necessary to further resolve the apparent differences. The average time for the five sources is 41.4 minutes. Assuming that a cataloger currently earns \$4.50 an hour, the average cost for the five sources would be \$3.11 for the unit cost of cataloging.



Table 3. *Processing Cost Elements*

Columbia University	New York Public	Sacramento State
<ol style="list-style-type: none"> <li>1. Card production</li> <li>2. Card set completion</li> <li>3. Sorting and preliminary filing</li> <li>4. Shelf listing</li> <li>5. Typing of book pockets</li> <li>6. Filing</li> </ol>	<ol style="list-style-type: none"> <li>1. Receive and distribute planning sheets</li> <li>2. Type headings for added entries and subject entries</li> <li>3. Mark designators and sort completed cards</li> <li>4. Distribute cards to filing section</li> <li>5. Paint edges of cards when required</li> <li>6. Glue and separate batches</li> <li>7. Type masters for offset printing</li> <li>8. Prepare copy for Itek masters</li> <li>9. Check format of entry on masters</li> <li>10. Check letter for letter on planning sheet</li> <li>11. Gather statistics and keep log of card preparation</li> <li>12. Prepare Itek masters and print cards on offset</li> <li>13. File</li> </ol>	<ol style="list-style-type: none"> <li>1. Type master cards from handwritten slips</li> <li>2. Produce subject cross reference cards</li> <li>3. Maintain guide cards</li> <li>4. Card production and purchase</li> <li>5. Complete card sets</li> <li>6. Proof</li> <li>7. Alphabetize</li> <li>8. File and revise</li> <li>9. Card shifting</li> <li>10. Update existing cards</li> <li>11. Correction of problems</li> <li>12. Withdrawals</li> <li>13. Weed order slips</li> <li>14. Assembly of statistics</li> <li>15. File temporary slips</li> <li>16. File permanent slips</li> <li>17. Shelf list shifting</li> <li>18. Blank catalog card stock</li> </ol>

## CARD PROCESSING COSTS

If cataloging is the least likely part of the library operation to be automated in the near future, the procedures that immediately follow cataloging are precisely opposite in character. Card preparation, production, and filing all involve time-consuming routine operations that can be done automatically, thus relieving the library community of a significant proportion to man-hours to apply to problems of greater intellectual content. Cost factors must nonetheless be considered.

As with cataloging, description of basic cost elements will vary from one library to another. For the detailed breakdown in Table 3, use is again made of the Columbia and New York Public studies previously cited. Added to them is data from an unpublished study made available by Neil Barron of Sacramento State College Library. Barron's cost elements are given in finer detail than those in the other studies reported in this section.

In Table 4, data from the New York Public Library and from the Sacramento State College Library have been grouped into three categories (preparation, production and filing) to achieve maximum compatibility with data from other sources reported in the table. These sources are: a study (7) at the University of Toronto of manual costs made in conjunction with early machine methods; a comparative study (8) of manual methods and a special-purpose machine procedure at the Air Force Cambridge Research Laboratory Library; and results of three years of computerized card production at the Yale Medical Library (9).

Costs shown in Table 4 are on a "per-card" basis, rather than on a title basis, as differing library requirements show averages ranging from 4.6 cards per title at Sacramento to 9.8 cards per title at New York Public.

Most significant in Table 4 is the extraordinary agreement between two of the studies: the total processing costs amount to 23.2c per card and 23.5c per card for these two sources, even though the reports were prepared over a six-year period and include significant changes in the cost of labor and materials. Furthermore, these costs are reasonably constant for the individual categories in all three sources: card preparation varies from 11.4c per card to 11.6c per card; card production varies from 6.4c per card to 7.9c per card; and card filing varies from 4.2c per card to 5.2c per card. In one sense this close agreement should not be surprising. If it is indeed true that cataloging involves relatively high intellectual content that is difficult to automate, and card processing involves straightforward operations that are relatively easy to automate, it is reasonable to argue that the latter should show much less variability from one operation to the next.

The fact that the New York Public operation has significantly higher costs can be partially explained by the following observations. The NYPL costs are based on the supposition that all cards are locally produced. The

*Table 4. Comparative Costs of Card Processing*

Date	1968	1969	1965	1963			1968	
Library	NYPL	SSC	ONULP	AFCRL		CHY		
Cards per title	(9.8)	(4.6)	(~9)	(local)	(LC cat.)	(machine)	(9.3)	
Preparation	0.140	0.116	0.114	0.233	0.166	0.075	0.088	
Production } Filing }	0.186	0.064	0.079		0.043	0.043	0.043	—
Totals	0.336	0.232	0.235		0.276	0.209	0.118	—
				0.228				

other libraries indicate that a significant proportion of their work is based on the acquisition of LC cards. The breakdown for the AFCRL study is shown in Table 4 and the breakdown for Sacramento is approximately the same. Secondly NYPL is clearly the largest of the operations under consideration here, and it is not unreasonable to expect that the size of the file will have an effect on the cost of filing. In fact, assuming that the NYPL cost of preparation and production is the same as that for the AFCRL's locally produced cards (27.6c) and assigning the rest of the NYPL cost to filing, the latter figure becomes 10.3c per card, or a little more than twice the average for the other three sources (4.8c per card). If this is the case, it would be of interest to know whether the problem is one of sheer size of the catalog or rather one of increased density that naturally occurs in larger files. E.g., is it more costly to file "Smith, Adrian J." in a file with 100 Smith's or 1000 Smith's?

Finally, in the two cases of partial automation (AFCRL and Yale) the cost of card preparation and production is significantly lower (7.5c and 8.8c) than that indicated for LC cards (16.6c), or the average for the three closely agreeing sources (23.2c). This observation alone should point the library community strongly towards automation of the card processing function. Nor is this observation new; both authors of the preliminary studies at AFCRL and Yale made the point more than adequately. Furthermore, as will be demonstrated shortly, the cost of filing is also reduced in an automated system.

Several factors may be contributing to the slowness of the library community to introduce changes to achieve such cost savings. First, there is inevitably a substantial initial cost involved in any automation project. Second, although the potential cost saving is a substantial proportion of the processing cost, it is still small when compared to the cost of cataloging; a librarian under pressure to reduce costs could gain more by cutting back on the time allowed for cataloging without the initial investment necessary for automation. Third, there is a persistent difficulty in finding trained personnel in the automation field. Finally, librarians are certainly aware of the rapid changeover in equipment in the computing field with the concomittant costs of adapting programs to new equipment.

#### CASE AND SPACE

The preceding discussion has provided some notion as to the cost of obtaining the required cataloging information, encoding it on catalog cards, and entering those cards in a catalog file. These costs can be compared with other possible approaches to the problem, including those that involve some degree of automation. There are, of course, a number of associated costs that must be taken into account to obtain a full picture of the cost of card cataloging. They would include, at a minimum, the cost of the space occupied by the catalog, the purchase price of catalog filing

cases, the cost to the user of consulting the catalog, and the cost to the library of maintaining the catalog in usable form.

The allocation of capital expenditure costs to a form comparable to the costs per title and the costs per card used in the earlier sections of this report raises certain difficulties. Accounting procedures vary from one institution to another. Further there is the real but difficult-to-measure problem of comparing funds of various types in a particular situation. Nonetheless, it is useful to know whether under any reasonable accounting system the cost of space and cabinets is of sufficient magnitude to make it worthwhile to consider these costs in the overall evaluation. Assuming, therefore, that a filing case capable of storing 72,000 cards fully packed costs \$800 and occupies approximately 30 square feet of space, including room for aisles and access area, and further assuming that land and construction costs are approximately \$30 per square foot, the total cost of the cabinet and the space it occupies would be approximately \$1,700. Finally if it is assumed that on the average a catalog is approximately 60 percent full, the initial cost of space and case is approximately 4c per card. Four cents a card is not negligible, but it is only about 15 to 20 percent of the cost of producing the cards and an even smaller fraction of the total cost when cataloging is included. Hence, it seems reasonable to put this cost for space and case in the category of a secondary cost item that will favor book catalogs, microfilm catalogs, and other high-density forms. It is unlikely to be a determining factor unless other cost factors are very closely balanced.

#### BOOK AND CARD CATALOGS: SOME RELATIVE ADVANTAGES

Among the various cost factors involved in cataloging, the most difficult to assess objectively is the cost to the user. The problem is that no one really knows what a user does in a library, nor what impact a given change will have on its utility to him. Whether they like a card catalog or not, library users do consult it and it is thus a usable device for providing access to library materials. Equally, many libraries in times past, and again more recently, have had book catalogs; and they also are viable devices. But which is better?

A card catalog is updated by the simple expedient of entering recently obtained cards in the file. A book catalog is updated by periodically printed revisions. Hence any search for a particular item will in general require fewer specific searches in the card catalog than in the book catalog, if the proper information is available to the searcher. Card catalogs are large and costly and there are few savings over the original cost in producing a second copy. Reproducing books after the first copy is relatively inexpensive. Libraries with many branches, or a decentralized set of users, will provide better service with book catalogs. The added cost of maintaining more than a few files is heavy with cards and light with books. Whether card or book catalogs are used, the existence of a machine

Table 5. Comparative Conversion Costs Per Title

	Mar. 68 LC 446 char.	1968 LACP ~450 char.	1964 ONLUP 400 char.	1968 NYPL 300 char.	1966 UC/B 317 char.	1964 CHY 243 char.	1966 SUL 180 char.
Coding/editing	\$0.169	—	—	—	\$0.080 <sup>1</sup>	—	\$0.044
Keying	0.207	} \$0.480	\$0.307	} \$0.450	0.188	\$0.198	} 0.183
Re-keying	0.033		} 0.259		} 0.030	} 0.117	
Proofing	0.125	0.127					0.085
Rental	0.156	0.084	0.650 <sup>2</sup>		—	0.036	0.037
Conversion & List	} 0.359	0.020	} 0.096	0.046	0.020	0.024	} 0.104 <sup>3</sup>
Edit List		0.084		0.141	—	—	
Sort & Merge		—	—	0.165	—	—	
Supplies	0.080	0.036	0.508 <sup>4</sup>	—	—	—	0.033
Supervision	0.183	—	0.580	—	—	—	—

<sup>1</sup>Includes provision for keypunch rental, and supplies

<sup>2</sup>Full keypunch rental absorbed by pilot project

<sup>3</sup>Includes use of automatic error-detection routines

<sup>4</sup>Includes cost of magnetic tapes and other supplies

readable catalog provides much greater flexibility as time goes on. Revisions of cataloging practice become much simpler if the revisions can be programmed on a computer.

In sum, machine readable book catalogs appear less advantageous than card catalogs only when immediate updating is the primary criterion for comparison.

### COMPARATIVE COSTS OF CATALOG CONVERSION

Table 5 (an extension and revision of Table 7 appearing on page 42 of Reference 1) gives comparative conversion costs for three public libraries (Library of Congress (10), New York Public Library and Los Angeles Public Library), the Library of the University of California at Berkeley, the Stanford Undergraduate Library (11), the Ontario New Universities Library Project, and the Columbia-Harvard-Yale study. Although the data was gathered for the most part independently over a four-year period, it is worth making a number of internal comparisons to test for consistency.

The most outstanding comparison is between the encoding costs for the Library of Congress and those for the Los Angeles Public Library. For records of essentially the same average length (446 characters versus 450 characters) the coding costs agree to the penny! Yet the methods of production are significantly different. The Library of Congress invested heavily in the coding and editing operation and used paper tape typewriters with their relatively high rental. As a result its costs in this area are significantly higher than those for LACP. On the other hand these procedural changes resulted in significantly lower keying costs, so that the overall cost for encoding was the same.

The encoding costs of UC/B, CHY, and SUL are all very close (within three cents per title) even though there is a fair range of record size (from 180 for SUL to 317 for UC/B). These three studies probably provide a more reasonable picture of the underlying variation in cost than the unusually close figures for LC and LACP.

As a further test of consistency, average cost is plotted against average record length (in characters per record) in Figure 1. The rightmost points are for LC and LACP, and the line is simply drawn through the origin (zero dollars, zero cost) and those points. The points of UC/B, CHY and SUL cluster about the center of the line. Following is an interpretation of the other points charted.

The NYPL point of \$.45 for a 300-character record is not based on actual NYPL experience, but rather on a study of information from other investigations. Its proximity to the line suggests that NYPL's analysis of existing information reaches a conclusion similar to that of this paper.

The average encoding cost used to plot the ONULP point does not contain the full rental charge reported in the ONULP study, because the entire cost of keyboard rental was charged against the project although

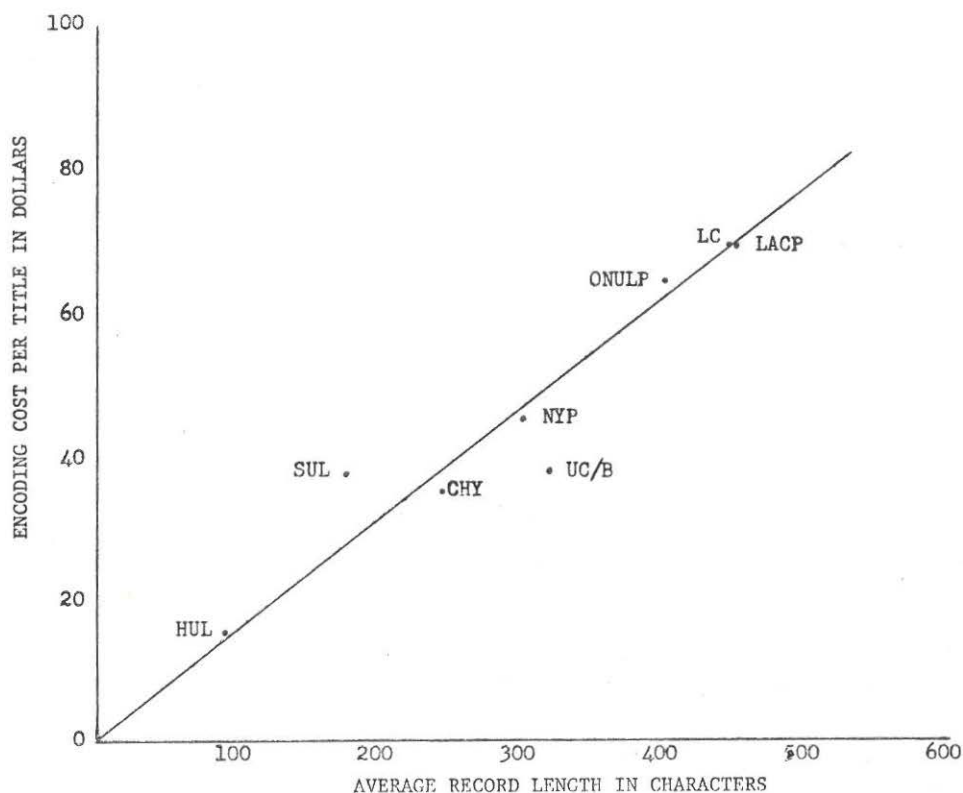


Fig. 1. Encoding Costs per Title as a Function of Average Record Length.

the machines were only partially utilized. The point for Harvard University Library (HUL) is based on information received in a private communication.

Although there is a significant amount of variation from one study to another it seems reasonable to conclude that the cost of encoding is approximately \$.15 per title per hundred characters.

The cost of computation is not as well-documented as the cost of conversion. Studies that reported computer costs all include the following three operational costs: The first is the cost of conversion and listing. This cost includes the cost of converting the original machine readable form (be it cards or paper tape) to magnetic tape form. In most cases a by-product of this operation was a listing (all-caps only) of the material on the tape.

The second is the cost of an edit run, including a listing in upper- and lower-case. The latter was eschewed in a number of cases because of the added costs. However, many libraries would require a proper edit run and many librarians would prefer to edit from an upper/lower-case printout than from an all-caps printout.



The third is the cost of sorting and merging the tapes. Many of the early studies did not explicitly report on this cost because they were primarily concerned with the cost of converting the retrospective list. However, in an on-going operation this would be a continuing cost of some magnitude.

The available information points to a uniform cost of approximately \$.02 per record for conversion and list, and approximately \$.08 per record for editing. The two studies where both these costs are given indicate that a ratio of 4 to 1 is appropriate. The only study giving a ratio between the sort and merge operation and the edit operation is the NYPL study and this is based on before-the-fact-information only; the ratio is approximately 8 to 7. For convenience, one can assume that this ratio is unity, giving an overall ratio of 4-4-1. The most complete history of total computer cost is given by LC: a total of \$.36 per record for 446 character records. Applying the above ratio to the LC total yields a breakdown of \$.04 for conversion and list, \$.16 for editing, and \$.16 for sort and merge. Extending the Stanford cost of \$.12 for conversion and list and editing gives a total cost for SUL of \$.22 for its 180 character records. This figure is considerably more than 180/446 parts of the LC cost.

One other pertinent piece of information is available from the SUL data. In the production of the annual catalog, Stanford estimates a cost of \$.121 per title for what is roughly comparable to the cost of sort and merge. This cost is then roughly 1.2 times the SUL cost for conversion and list and editing, verifying the notion that the cost of "sort and merge" is of the same general magnitude as the cost of editing.

The ratios of SUL costs to LC for encoding are  $.367/.690 = .532$  and  $.225/.359 = .625$  for computer time. This suggests that the means of computing average record length may be different for the two institutions. Taking the LC figures as the standard and assuming that both computing and encoding costs are strictly a function of record length, the SUL record length should be between  $.532 \times 446 = 238$  and  $.625 \times 446 = 279$ . This discrepancy may be a result of one source (presumably LC) counting all delimiter and other non-printing characters while the other does not. NYPL indicates that the ratio of printed characters to total characters is approximately 3:4. If the SUL figure of 180 is expanded by one third, one obtains the figure of 240 which agrees well with the lower limit (based on encoding costs) given above.

The cost of sort and merge is a function of the size of the data base, not the amount of material being put into it. The Library of Congress points this out in its study (11) and report on an average month (where the data base grows for a period and then is reduced to zero.) Stanford Undergraduate Library figures are based on its second year of operation, in which 16,000 titles were added to form a total base of 41,000 titles. The actual cost of this step in the operation will therefore depend strongly on the operating strategy employed. Clearly, the number of times one

has to sort and merge the entire data base should be minimized, particularly taking into account the fact that sorting costs go up faster than linearly. If the master file is arranged in  $n$  orders (author, subject, title, class number, etc.), it will generally be less expensive to sort the updating material into those  $n$  orders and make  $n$  merge runs with the sorted master files than to make a single merge with a single ordering of the master file and then sort the master file  $n$  times to obtain the required updated orderings of the master file.

#### MANUAL AND COMPUTER PROCESSING: COMPARATIVE COST

One objective of this paper is to define factors whose costs enter into calculations of relative costs of manual and computer processing of catalog information and to report these factor costs. The following paragraphs present a simplified comparison of actual costs of manual and machine processing for a "typical" library characterized by average costs approximating those in the preceding tables.

Table 5 yields average figures for two cases: catalogs with approximately 425 characters per entry and catalogs with approximately 250 characters per entry; they may be called "full entries" and "short entries," respectively.

From Table 4, it is possible to compute similar figures for "full catalogs" and "short catalogs" by clustering the three larger cases (those having 9.8, 9.0, and 7.0 cards per title) and the three smaller cases (those having 3.0 and 4.6 cards per title). For the full catalogs the average cost of processing is 26.7c per card and 8.6 cards per title, or a total cost of \$2.29 per title. For the short catalogs the average cost of processing is 20.3c per card and 3.8 cards per title, or \$0.78 per title. Combining these two sets of figures gives the results in Table 6.

Table 6. Comparative Costs of Manual and Computerized Processing

	<i>Short Entries</i>	<i>Full Entries</i>
Manual	\$0.78	\$2.29
Computer	\$0.84	\$1.31

Table 6 shows that an hypothesized "typical" library would be slightly better off with manual methods if it chose the short form entries, and noticeably better off with the machine if it chose the full form of the entry.

In making this quick comparison, consideration has not been given to several factors that should obviously be taken into account even in this simple example. First, there is not included either the initial cost of programming or the initial cost of converting the retrospective records. Either or both of these costs could be substantial, but as they are one-time costs and as libraries are basically long-term institutions, such costs should be written off over a relatively long period, even though they must be financed out of a given year's budget.

Second, the cost of printing the catalog is not included (assuming a book catalog is in fact to be used in the computerized system). Thus the comparison in Table 6 is between a card catalog and a catalog in machine readable form. Such a comparison is complicated by the fact that a card once filed stays in the catalog indefinitely, subject only to long-term wear and tear and a certain rate of attrition due to unauthorized removal, misfiling, and so forth, whereas the machine readable catalog must be updated periodically and supplemented by interim publications. And, of course, the comparison is also complicated by the corresponding low cost of producing a number of copies of the book catalog where this is useful for a given system.

However, to put the printing cost in some degree of perspective, one may make a quick calculation based on the production of a single book catalog using a standard upper- and lower-case print chain. At present commercially available prices this would cost between 35c and 50c per 10,000 characters, or approximately 9c per entry for the full form entries and 5c per entry for the short form entries (assuming four complete listings for author, title, subject, and class number listings). This added cost would make the comparison between manual and computerized methods even less favorable for the short form, but still substantially better for the long form entries (\$1.40 to \$2.29).

## CONCLUSION

It may be concluded that the card-processing operations in typical libraries can be automated economically in many situations today. Libraries using the short form of a catalog and having no immediate need for multiple copies of the catalog may find it desirable to wait a year or two, depending upon their local situation, the availability of trained personnel and, of course, the availability of capital to finance the initial cost of programming and retrospective conversion.

However, libraries using the full form in their catalogs, or those needing multiple copies of their catalogs, will almost certainly find that there is a substantial economic advantage to computerization at the present time. Even when allowance is made for substantial departures from the "typical" costs found in this study, it is difficult to visualize any library using full form information not finding significant economic gains in computerization.

Considering the further advantages of the greater flexibility available in machine readable records, the increased services that can be offered to the user, and the fact that machine costs are decreasing while labor costs are increasing, one is led to the conclusion that more and more libraries will move towards catalog automation.

Tables 7 to 11 appearing on the following pages are reference tables for calculating costs.

## ACKNOWLEDGMENTS

The work reported in this paper was supported by the U. S. Office of Education under Contract Number OEC-9-8-00292-0107.

Mrs. Henriette Avram (Library of Congress) and Mr. Neil Barron (Sacramento State College, Sacramento, California) made important contributions of cost figures and other technical data used in this report. Various State Libraries supplied detailed cost information.

## BIBLIOGRAPHY

A 400-item bibliography on cost and automation is available from the National Auxiliary Publication Service of ASIS (NAPS 00696).

## REFERENCES

1. Dolby, J. L.; Forsyth, V. J.; Resnikoff, H. L.: *Computerized Library Catalogs: Their Growth, Cost and Utility* (Cambridge, Mass.: M.I.T. Press, 1969).
2. Dougherty, Richard M.: "Cost Analysis Studies in Libraries: Is There a Basis for Comparison," *Library Resources and Technical Services*, 13 (Winter 1969), 136-141.
3. Kozumplik, William A.: "Time and Motion Study of Library Operations," *Special Libraries*, 58 (October 1967), 585-588.
4. Henderson, J. W.; Rosenthal, J. A.: *Library Catalogs: Their Preservation and Maintenance by Photographic and Automated Techniques* (Cambridge, Mass.: M. I. T. Press, 1968).
5. Fasana, Paul J.; Fall, James E.: "Processing Costs for Science Monographs in the Columbia University Libraries," *Library Resources and Technical Services*, 11 (Winter 1967), 97-114.
6. MacQuarrie, Catherine: "Cost Survey: Cost of Ordering, Cataloging, and Preparations in Southern California Libraries," *Library Resources and Technical Services*, 6 (Fall 1962), 337-350.
7. Bregzis, Ritvars: "The ONULP Bibliographic Control System: An Evaluation," In University of Illinois Graduate School of Library Science: *Proceedings of 1965 Clinic on Library Applications of Data Processing* (Urbana: University of Illinois, 1966), pp. 112-140.
8. Fasana, Paul J.: "Automating Cataloging Functions in Conventional Libraries," *Library Resources and Technical Services*, 7 (Fall 1963), 350-365.
9. Kilgour, Frederick G.: "Costs of Library Catalog Cards Produced by Computer," *Journal of Library Automation*, 1 (June 1968), 121-127.
10. Avram Henriette: *The MARC Pilot Project* (Final Report on a project sponsored by Library Resources: Chapter VIII: "Cost Models" (Washington, D. C.: Library of Congress, 1968).
11. Johnson, Richard D.: "A Book Catalog at Stanford," *Journal of Library Automation*, 1 (March 1968), 13-50.

Table 7. *Cost/Card—Library of Congress Catalog Cards (July 1968)*

<i>LC Cards Ordered by/for</i>	<i>1-2 cds only</i>	<i>1st cd of 3 or more order</i>	<i>Add'l copies same cd ordered same tm.</i>	<i>All titles specific subject</i>	<i>Subsc for all cds</i>	<i>Extra chgs/title all orders lacking req info</i>
1) LC #	\$ .22	\$ .10	\$ .06	\$ ___	\$ ___	
2) Author & Title	.27	.15	.06	—	—	
3) Series	—	.10	.06	—	—	
4) Subject	—	.10	.06	—	—	
5) Chinese/Japanese/Korean	.22 - .27	.10 - .15	.06	—	.04	\$ .04
6) Motion Pictures & Filmstrips	.22 - .27	.10 - .15	.06	.10	.04	
7) Phonorecords	.22 - .27	.10 - .15	.06	.10	.04	
8) Revised & Cross Ref.	—	—	—	—	.04	
9) Anonymous		\$ .04				

Source—LC cds, July 1968

Table 8. Catalog Card Costs

<i>Cards</i>	<i>Cost/Card</i>	<i>Cost/Hour</i>	<i>Time</i>
LC Cards	\$.22-.27 (min order 1-2 cds) .10-.15 (1st cd—3 or more order) .04-.06 (add'l copies same cd-same order)	\$04 extra chg all orders lacking req. info.	
Blank Cards	< 3-< 4 for \$.01		
Original Card Preparation	\$.20-2.34		\$2.40 - 4.70
Card Checking Before Filing	\$.21	\$4.20	3 min/cd
Correcting Detected Errors	\$.12	\$2.40	3 min/cd
File	\$.024 .03 .047	\$2.40 3.00 4.71	100 cds/hr 100 cds/hr 100 cds/hr
Store	\$.01		
Reproduce	\$.0023-.00208 (AB Dick Offset Press =\$.125/bk(54-60 cds) .045 (Xerox—1K-100K cds)		

Table 9. (Estimated) Annual Cost of 1000 Sq Ft of Storage Space

1)\* Minnesota State Dept. of Education (1968)—\$520

\*Source—Private communication

2) R&D Estimate\*\*

$$\frac{1968 \text{ Construction Cost}}{\$30 \text{ sq ft} \times 1000 \text{ sq ft}} = \frac{\$30,000}{100 \text{ yrs (life of bldg)}} = \$ 300/\text{yr}$$

$$+ \text{Maintenance Costs, clean up, etc. } (\$1 \text{ yr/sq ft}) = \frac{\$ 1000}{\$ 1300/\text{yr}}$$

$$\frac{1974 \text{ Construction Cost}}{\$50 \text{ sq ft} \times 1000 \text{ sq ft}} = \frac{\$50,000}{100 \text{ yrs (life of bldg)}} = \$ 500/\text{yr}$$

$$+ \text{Maintenance Costs, clean up, etc. } (\$1 \text{ yr/sq ft}) = \frac{\$ 1000}{\$ 1500/\text{yr}}$$

\*\*Source—E. Graziano, Univ. Calif. at Santa Barbara

Table 10. Card Catalog Cost/Year

Given the following variables, 1 card catalog case with a maximum card capacity of 72,000 cards (purchase price—\$789)—the cost/card to store would be \$.01.

	<i>Estimated Cost sq ft Rental @ \$.42 sq ft/mo</i>	<i>Construction Cost \$30/sq ft ÷ 100 yrs life bldg</i>	<i>Maintenance Est. @ \$1/sq ft</i>	<i>Cost/Yr</i>
Cabinet (6 sq ft)	\$30.24	\$1.80	\$ 6.00	\$ 38.04
Room for Users (16 sq ft)	80.64	4.80	16.00	101.44
Aisles (3 sq ft)	15.12	.90	3.00	19.02
Catalog Table (5 sq ft)	25.20	1.50	5.00	31.70
				<u>\$190.20</u>
			+ 72,000 cards @ \$ .01 (to store)	720.00
			<b>TOTAL COST/YR</b>	<u><b>\$910.20</b></u>

Table 11. Card Catalog Maintenance Costs

	<i>Requirement Space</i>	<i>Estimated Cost/Sq Ft</i>	<i>Cost/Mo</i>	<i>Cost/Year</i>
Card Catalog Cabinet	— 6 sq ft	\$ .42	\$ 2.52	\$ 30.24
Room for Users	—16 sq ft		6.72	80.64
Aisles	— 3 sq ft		1.26	15.12
Catalog Table	— 5 sq ft		2.10	25.20
	30 sq ft		\$12.60	\$151.20

Source—E. Graziano, Univ. Calif. at Santa Barbara and R&D Consultants Co.