RATES OF PROCESSING MANUSCRIPTS AND ARCHIVES

KAREN TEMPLE LYNCH
and
THOMAS E. LYNCH

Archivists must constantly balance the amount and kind of material taken into their repositories against the staff available to arrange, describe, preserve, and service it. In order to coordinate a program of arrangement and description and balance the time devoted to processing against that spent on other elements of an archival program, it is useful for the archivist to have a sense of how long it takes to arrange and describe papers. Without this knowledge, the archivist is unable to set realistic priorities and project the completion of processing assignments. For budget purposes, too, it is important to understand processing time requirements, since staff time is the largest expense involved in processing.[1] Although the role of the archivist's intuition is enormous, good planning should also draw on whatever concrete evidence is pertinent. The project on which this paper is based was designed to explore estimates of processing rates that appeared in fifty-five grant proposals submitted to the National Historical Publications and Records Commission and the National Endowment for the Humanities.[2] Grant proposals are a useful source of processing estimates, because the archivist requesting funding is obliged to justify requests for staff.

Both the NHPRC Records Program and the NEH Research Collections Program fund projects to process and make available historical records. The range of projects dealing with manuscripts and archives which the two agencies fund is, of course, much broader, but for this study only arrangement and description projects were examined. Thirty active and closed files for grants funded by NHPRC and twenty-five active grant proposals funded

by NEH were examined. All of the proposals had been funded. Every effort was made to make the figures from the various proposals and reports consistent, but since each document was written from a different perspective, these figures are not entirely reliable.

Neither NEH nor NHPRC has guidelines for what constitutes an appropriate processing rate. Rather, each agency relies on the comments of reviewers, panelists, and sometimes consultants to determine whether a project can realistically be carried out as described in the proposal. Therefore, even if the figures obtained from the proposals are not strictly empirical, they should represent a consensus from the applicants, the various professional archivists who reviewed the proposals, and the staffs of NEH and NHPRC that will give a range into which appropriate processing rates might fall. The fact that all the proposals examined had been funded should be an indication of the perceived feasibility and appropriateness of their work plans.

The following information was taken from each proposal: the total project staff, the total amount of material to be processed, the amount of time scheduled for the project, the number of separate collections to be processed (if indicated), the level of control to be achieved, the type of records involved, and the dates of the records. The average processing rate was found by dividing the total number of processing weeks into the linear feet of records to be processed. For grants funded by NHPRC, proposed average processing rates ranged from 0.1 linear foot to 72.5 linear feet per processor per week. For NEH-funded grants, the range was from 0.1 linear foot to 13.9 linear feet per processor per week.

Although the range of data is wide, it is not as wide as it seems at first. The largest rate was apparently based on an incorrect initial estimate of bulk. The project proposed to process 20,000 linear feet of records, but only 2,250 were actually processed, while 3,750 feet were discarded. Using the 2,250 figure, the recalculated processing rate is about 18.8 feet per week. Similarly, the bulk listed in another proposal as 5,000 linear feet was later estimated by a consultant to be closer to 500 linear feet.

The distribution of the data is represented in Figure 1. The data can be seen to be in roughly normal form—approximately a bell-shape curve. In order to compress the horizontal dispersion of the
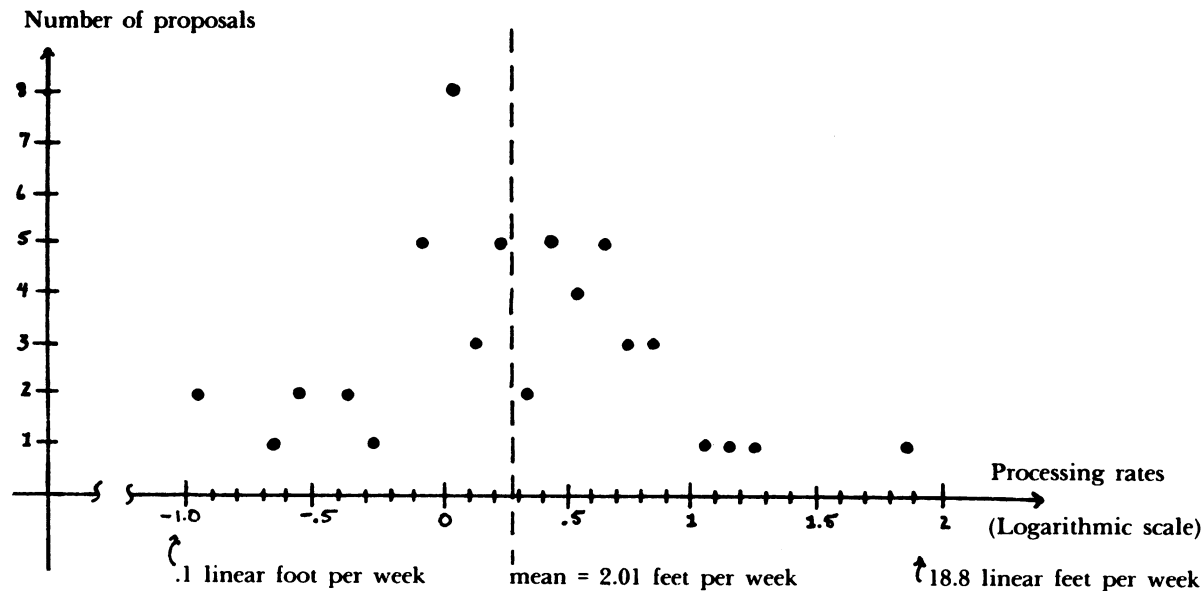
Figure 1: Frequency distribution of processing rates based on proposals.

data, the logarithms of the processing rates are shown on the horizontal axis. The appropriate mean for this logarithmic scale is a geometric mean, shown in Figure 1 at the point corresponding to a processing rate of 2.01 linear feet per week. This mean was calculated using 52 data points. The two lowest and one highest values were dropped. In this sort of statistical analysis, dropping these two outlyers is legitimate, largely because the data are so highly subjective. In summary, if the statistics based on initial miscalculations of bulk are omitted, the applicants expected processing projects to be completed at a rate of 0.1 to 18.8 linear feet per week.

These expectations, especially those at the upper end of the scale, were not always realistic. Of the twelve completed NHPRC grants, five were completed as proposed and on schedule. Their processing rates were 0.8, 0.9, 1.3, 2.3, and 3.4 feet per week per full-time processor. The other seven projects took longer than esti-mated. In no instance was there any indication that a grant had been completed early or that initial estimates of processing time had been too large. Often there were signs that processing esti-mates had been too low, even for grants that had proposed a relatively slow rate of processing.

The progress reports to NHPRC and NEH often described factors that slowed progress. Preservation often proved to take much longer than estimated, in one case five times as long. Even pulling out staples and removing paper clips required a more substantial time investment than some grantees had expected. Other factors that slowed progress included addenda to collections that had to be incorporated into an existing arrangement, foreign language materials, and confidential material that was sprinkled throughout the collection rather than concentrated in one series. Some grantees also admitted that personnel problems and poor supervision had contributed to their failure to complete projects on schedule.

Aside from miscalculations, processing rates varied with the dates of the records processed, the types of records, and the level of intellectual control that was to be established. An attempt was made also to analyze the influence of the size of collections and different staff compositions.

To see the effect of the dates of records on processing rate,

consider a distribution of the processing rates of only those projects that dealt with pre-1800 records. The geometric mean rate for those projects was 0.98 linear feet per week, considerably lower than the aggregate rate. Processing sixteenth, seventeenth, and eighteenth century records apparently takes longer than average. If all grants dealing with records dated before 1900 are included, the geometric mean is 1.55 feet per week, still lower than the aggregate mean rate of 2.01, but higher than the mean rate for earlier records. For twentieth century records, however, the geometric mean rate is 2.51 feet per week, indicating that twentieth century records can generally be processed faster than average. In part, the longer processing time for older records is due to the necessity for more careful handling and more extensive preservation work. Then too, often when twentieth century records need cleaning, deacidification, and other preservation work, their volume disallows extensive work.

The records to be processed were designated as belonging to one of five categories: personal papers, institutional records, business records, government records, and records of mixed types. Only one designation was used for each grant project. The data suggest that business records (with a mean rate of 3.07 feet per week) and government records (3.11) are processed more quickly than institutional records (1.87), which in turn can be processed at a slightly faster rate than personal papers (1.51) or papers of mixed types (1.54). It is interesting that the four most rapidly progressing projects involving personal papers were all concerned with processing papers of twentieth century political figures. Because of the large bulk of records produced by politicians and the large amount of mail they receive, their papers tend to be more like institutional, business, or government records than they are like other personal papers.

The intellectual controls proposed were classified into four levels beyond basic accessioning. Partial control was defined as the stage when the collection has been examined, series have been delineated, and possibly a container list or series list has been generated. The most usual level of control was the inventory level, where the collection has been arranged, usually to the folder level, and described in an inventory or register consisting of a biography or agency history, a scope and content note, and a container list or

folder list. Limited indexing may also have been done. Special controls were tailored to each collection and included item lists, extensive indexes, bibliographies, lists of speeches, and other lists. For some proposals, the archivist stated the intention of using whatever level of control was necessary to make the collection usable. These variable controls usually ranged from the inventory level up to more specialized controls. Almost all of the projects proposed to go to the inventory level and often to include special controls as well. Of the projects that created special controls, thirteen had processing rates slower than the mean rate, while five had faster rates. It is interesting to note that the two proposals that used variable controls both had relatively fast processing rates of 3.4 feet per week.

Average size of the individual collections was also computed in an effort to determine whether large collections or small ones take proportionately more processing time. It appears that large collections are processed at a faster rate, but it should be noted that most of the large collections consisted of institutional, business, or government records. Therefore, it was impossible to tell whether the faster rate was due to the size of the collection, the type of records, or the initial order. It might be argued that doing background research and writing an inventory take about the same amount of time regardless of the size of the collection. Sometimes, though, processing a small collection of papers created by a person for whom little biographical information is available requires that more time be spent on research and writing. On the other hand, it has been argued that a large collection in disarray requires a much more intricate job of arrangement than any small collection can.

The full-time equivalent processing staff as computed from the proposals was broken down by function whenever possible. Project activities including administration, decisions on appropriate arrangement and description methods, routine sorting, preservation work, and typing finding aids and catalog cards were sometimes all undertaken by a single processor, but more often the different activities were divided among staff members. Staff was classified into four categories: administrative staff (concerned with budgetary matters and sometimes advising processors), processors, clerical staff, and processors' assistants (often students). From the information gathered from the fifty-five proposals, it was not

possible to tell whether a more specialized staff—that is, a staff divided by function—is more efficient than a single processor working alone on all phases of the work. Some archivists have found that the use of assistants is helpful. Delegating routine tasks to students or paraprofessional assistants frees the processor's time for decision-making, research, doing more complicated arrangement, and the like. On the other hand, some archivists encounter problems with such division of labor, chiefly in the large amount of time required to supervise the assistants and in the difficulty of maintaining the morale of junior staff members as they become bored with routine tasks.

The basic results of this study are summarized below in Figure 2, which shows the geometric mean for each category into which the projects were divided, the number of data points used to calculate each mean, and the 90% confidence intervals.[3] These confidence intervals indicate that if a new sample were taken, there is a 90% certainty that the means of the new data would fall into these ranges.

Proposals by Records Grouped	Geometric Mean*	Number of Data Points	90% Confidence Interval
All proposals	2.01	52	1.592-2.541
Pre-1800	0.98	7	0.676-1.545
Pre-1900	1.55	31	1.13-2.11
Post-1900	2.51	19	1.92-3.26
Personal	1.51	11	1.06-2.17
Institutional	1.87	9	0.97-3.58
Government	3.11	5	insufficient data
Business	3.07	4	insufficient data
Mixed types	1.54	19	0.99-2.36

* Processing rate in linear feet per processor per week.

Figure 2: Processing rate as affected by time period and record type.

Conclusion

The statistics taken from the proposals suggest that a rule-of-thumb rate for processing personal papers might fall into the range of 0.5 to 2 linear feet per full-time processor per week. Government and business records take somewhat less processing time and might reasonably be processed at a rate ranging from 2 to 4 linear feet per full-time processor per week. In between, institutional records might be done at a rate of 1 to 3 feet per week. The archivist might expect these figures to be about right for nineteenth and twentieth century records processed to the inventory level. Earlier records would probably take longer. Control beyond the inventory level would probably require a substantially larger investment of processing time.

It must be stressed that these estimates are only—at best—ballpark figures. By definition grant proposals include only estimates of how long processing will take. Further, as any processor knows, a multitude of factors can slow or speed progress on a given project. Inevitably some collections will be processed at a rate different from that expected, but the archivist should try to analyze the factors that cause the variance and incorporate them into future estimates.

A few of the grants submitted to the agencies included processing estimates based on experience. Grants submitted by the State Historical Society of Wisconsin usually included the statement that the processing estimates in the grant were based on current performance statistics in the repository. Similarly, a Temple University grant included this statement:

> The total of the combined unprocessed collections amounts to approximately 400 cubic feet. Work involving weeding, arrangement, description at the box and folder level and preparation of a complete inventory for each collection will involve an eighteen month project. This estimate of work progress is based on experience with very similar records undertaken during the 1975–76 grant period above.

The fact that statistics for the same type of records were the basis for the estimate increases the probability that an accurate estimate of processing time will have been made.

More controlled analysis of processing rates should greatly

increase our planning capabilities. Suppose that an archivist is able to determine from an initial look at a collection how much material is present, what the inclusive dates are, what types of records are involved, and maybe even a degree of disorder. Then it should be possible to devise a rough formula to indicate approximately how long processing the collection to a given level of control might require. There would still be factors that would change the estimate—discovery of foreign language materials, restricted or sensitive documents, more preservation needs than expected—but the archivist would develop some sense of the time required to deal with these additional problems. Adjustments in the estimated overall processing time could be made as processing proceeded.

The implications of such planning are not easy to assess. Not all of the variables in processing time are easy to control. The types of records, their dates, and the size of the collections can be observed and analyzed but not changed. Two other factors, however, are different: it is possible to change the composition of processing staff, and it is possible to vary levels of control of collections. Such changes are not mandated by statistics. Efficiency, after all, is not the only standard archivists use. Even if they were to be convinced that using students or paraprofessional assistants was most efficient, many archivists would still be unable or unwilling to use them. But knowing the options increases archivists' ability to plan.

An interesting by-product of increased planning might be a renewed exploration of methods of processing at various levels, particularly at the series or box level. For many collections, the so-called preliminary inventory provides quite adequate access to the collection, and it might well be accepted as permanent. It is crucial always to remember the reason that intellectual controls are created—to serve the needs of the research public. Faced with a clear choice between having a few collections processed to inventory level and having minimal controls for all holdings, archivists may find a new appeal in partial and variable level processing.

FOOTNOTES

1. In a survey of grants funded by NHPRC, Robert W. Coran found that processing costs ranged from $61 to $321 per linear foot, and that labor costs accounted for about 90% of the total cost.
2. We are grateful for the assistance of people at both agencies, and especially for the kindness of Larry Hackman of NHPRC and John Fleckner of NEH.
3. We are grateful to Holy Cross College for the use of the Statistical Package for the Social Sciences in computing these data.