Obsolescence of Books in College Libraries

Dr. Gosnell is librarian of Queens College, Flushing, N.Y., and associate in library service, School of Library Service, Columbia University. This report is based on a thesis submitted at New York University in September 1943.

The problem of maintaining efficient college library book collections requires consideration of the effects of obsolescence and planning for the segregation or withdrawal of obsolete materials. Heretofore obsolescence has been considered principally with reference to individual and specific titles rather than to groups of books, but no general planning can be done except on a quantitative basis. For evaluation of libraries by accrediting agencies, it is important to have a fairly simple and precise basis for comparison of libraries and lists of books with regard to obsolescence or the presence of obsolete material. By quantitative analysis, primarily of three lists of recommended books and secondarily by analysis of library collections, the writer believes that he has presented a clearer picture of obsolescence than has previously been available. Tentative conclusions in the form of comparative rates of obsolescence for various groups of material are reported as suggestive for further study.

Several years ago the writer called attention to the fact that Shaw and Mohrhardt showed marked preference for newer books in their lists of books for college libraries.¹ He pointed out that this preference fell into a rough pattern of distribution which was essentially the same in each of the lists and suggested several questions as to the nature and significance of these curves of distribution. He now presents a progress report of his efforts to answer some of these questions.

The fundamental assumption in this study is that the compilers of the three select lists consciously or unconsciously followed certain principles of choice which are significant without reference to the particular titles chosen and that these principles may be discovered and studied by a form of documentary analysis of the lists. The principle of immediate interest is that of preference for the newer or more recent titles. “Newer” and “more recent” are of course relative terms referring in each case to the particular time the list was compiled. Conversely, this is the principle of obsolescence—that older books in general have less value for use in the college library.

Statistical Bibliography

The second assumption is that mere masses of books (or titles) may be

analyzed for certain characteristics without reference to their individual titles. Statistical bibliography is a relatively new or uncultivated field. But the astronomical proportions to which some of our libraries, their catalogs, and bibliographies in general are growing must force librarians to consider collections of books as populations. A typical example of what can be done is the study of "certain biological properties" of literature by Wilson and Fred. 

This assumption demands that individual titles, especially the exceptional (and therefore more noteworthy) ones, be forgot; that attention be to the group as a whole—its averages and general trends. The larger the groups, the greater the play of many factors and the more reliable the conclusion. Life insurance mortality tables are not primarily based on infant deaths or the longevity of octogenarians but on these and all others of the population. A study of book obsolescence or mortality will not depend primarily upon ephemera or classics but upon all kinds of books. The very multiplicity of causes of death seems to lend stability to the human mortality rate.

The causes of book mortality or obsolescence are many, varying from pure fad through extension of scientific knowledge, technological changes, to fundamental changes in our civilization. The object of the present study has been, not to discover or classify these causes, but to analyze their total effect. Deterioration or destruction of books is not true obsolescence, and no consideration has been given to the physical properties or conditions of books.

Distribution of Imprint Dates

The fundamental data of the study were the dates of publication of titles included in the three lists. The significant features of the distribution curves were: (1) There is a rapid rise in number of titles per year, going backward for the first few years preceding the date of issue of each respective list; (2) The maximum number of titles per year occurs within three years preceding the publication of each list; (3) After this maximum is reached, the number of imprints per year drops rapidly at first and then more slowly, approaching the base line of zero asymptotically in the early years. The initial rise may be ascribed to lag in selection of books. This lag is due to delay in appearance of reviews and acceptance by scholars. The drop, rapid at first but slower as the age of remaining titles increases, is expressive of a preference for newer books over old or rejection of older titles for the newer, because the older ones have become obsolete.

This pattern is repeated independently for each of the three lists. The latter part does not show in the Shaw supplement for 1931-38, because, as a supplement, titles over nine years old were automatically excluded. The pattern is not markedly affected by the rate of book production as derived from annual tables in the Publishers' Weekly. It is a function of the age of the titles at the time each list was compiled rather than of any other observed variables.

From the fall in the curves it is immediately evident that the older a title is at a given time of selection, the less

---


3 Illustrated in College and Research Libraries 2:218, June 1941.

are its chances of being selected or of surviving the selective process. In general, each preceding year back from the date of selection is represented by fewer titles. That is, if in a given subject group there are one hundred titles 23 years old and ninety titles 24 years old, the rate of obsolescence between age 23 and 24 is 10 per cent.

If a similar select list were to be prepared one year later, the group of one hundred would have become one year older (from 23 to 24) and, suffering the same rate of obsolescence, be reduced to ninety. And if the rate from age 24 to 25 were approximately the same, the group of ninety would drop to approximately eighty-one. And similarly, the drop from eighty-one would be to seventy-three. Conceivably the process continues indefinitely, with the older groups becoming practically zero.

Several curves were considered for fitting to the data, including the normal curve, the logarithmic normal curve, Pearson's types I, III, and IX, and others. But a satisfactory fit was not secured and logic indicated that the opposing forces of lag and obsolescence should be isolated. By dropping data for the rising portion of the curve for the first few years and limiting consideration to the rate of decline after the early maximum was reached, the effect of lag was substantially eliminated.

The curves thus remaining are presented in Diagram 1. Data for the Shaw and Mohrhardt lists have been grouped in five-year periods for simplicity, but this could not be done for the Shaw supplement. A similar curve for book circulation at Hamilton College Library is included and will be described later.

These curves, plotted on a logarithmic scale, or the data, plotted in logarithms, tend to form a straight line. Pearson's criteria of moments indicate his Type X, the exponential curve. Thus the curve of organic decay appears as the expression of obsolescence:

\[ y = y_0 b^x \]

where \((y)\) is the number of titles, \((y_0)\) the number of titles at the maximum or initial point, with lag eliminated, and \((x)\) is time elapsed. When

\[ \omega = 1 - b \]

\((\omega)\) omega, becomes the annual rate of decrease in the curve, or the rate of obsolescence. Per cent of decrease is expressed by \((100\omega)\).

This rate is almost the exact opposite of the rate of compound interest. It indicates the rate at which the principal or capital is decreased instead of increased.

By shifting the equation to logarithmic form:

\[ \log y = \log y_0 + x \log b \]

an equation of the type:

\[ Y = A + BX \]

the curve takes the form of a straight line. By the short method of least squares this is easily fitted to the data in logarithmic form. For simple graphic analysis and illustration, the data of imprint dates may be plotted directly on semi-logarithmic grids, as already indicated.

For easy computation and quick graphic checking, the data of imprint dates may be grouped into five-year periods and immediately transferred into logarithms. Plotting of logarithms is illustrated for four subject sections in Diagram 2.

**Rates of Obsolescence**

By means of the formulae and methods outlined above, the rates of obsolescence have been computed for nineteen subject
groups in the select lists; they are presented in Table I. The number of groups has been reduced to nineteen for two reasons. First, Shaw and Mohrhardt did not use exactly the same classifications, e.g., "Romance Languages" as

tion of the list the number of titles chosen was 8.1 per cent smaller than for the preceding year. That is, if there were 500 titles ten years old, there were some 8.1 per cent fewer, or approximately 460 in the eleven-year-old group; and

used by Shaw was split by Mohrhardt into "French" and "Spanish." Second, some subject classes such as "Botany" and "Zoology" were too small for effective separate treatment and similar enough to combine readily into larger groups.

Table I reads as follows: For the entire Shaw List the rate of obsolescence is 8.1 per cent. For "Physical Education and Hygiene" it is 21.6 per cent. This rate is highest in rank of all subject groups in the Shaw List. In the number of titles which comprise it, this subject ranks eighteenth. The rate of 8.1 indicates that in general for each year back from the compila-

8.1 per cent of 460 fewer, or 421, in the twelve-year-old group.

These rates may be compared by rank but as ratios they cannot be added or averaged.

It is evident that there is substantial agreement in (ω) for each subject in the three lists. The rank correlation between the rates for the Shaw List and the supplement is +.53; between Shaw and Mohrhardt, +.84; between the Shaw supplement and Mohrhardt, +.66. Further detailed analysis reveals no direct relationship between the size or total number of titles in a subject group and its (ω) nor between the number of

**Diagram I**

**Imprint Dates of Titles in the Select Lists and Circulation at Hamilton College**
titles in the maximum year and \((\omega)\). The rate of obsolescence seems to be a property peculiar to each subject and not a variable dependent upon some other continuous variable. No other ordinal relationship of the respective subject sections has been found to correlate definite-

Education” and “Chemistry and Physics.” But “Geography,” one of the smallest groups, has a medium rate. And “Mathematics,” “Music,” and “Philosophy” are small groups with low rates.

The Mohrhardt List, as a whole, is the smallest of the three and has the highest rate. The Shaw supplement is small in total but limited to only ten years. The totals for each of these years are much nearer to those of the original Shaw List for titles of the same age.

In some fields, such as “Chemistry and Physics,” “Psychology,” and “Mathematics,” a large proportion of current research is published in professional journals. Many of the books are either textbooks or handbooks and reference books. The bearing of this fact on obsolescence is difficult to evaluate. From one point of view the need for current publication is met by the journals and the output of new book

**Diagram 2**

**The Rate of Obsolescence**
titles is consequently reduced. Conversely the rapid developments characterized by the journal articles must occasion frequent revisions and changes in textbooks and handbooks and should be reflected in a high rate of obsolescence. The problem of such obsolescence in the field of technical books has been dramatically stated by Thompson.  

Fluctuations in (ω)

It is quite possible that the (ω) for a given subject might change over a span of years. This change might be due to a general change in approach or in any such changes within the limits of the present study. The entire process of selection for all three of the lists took place within eleven years (1929-40). This period is not long enough to show clearly any fundamental changes.

The possibility of a general shift in the rate during the period 1890-1920 is suggested by the yearly totals for the Shaw and Mohrhardt lists and in many subject sections. In the process of fitting the straight lines, it was found that nearly all curves rose above the straight lines for the period of 1900-10. In the succeeding ten-year period they fell below.

### Table 1

<table>
<thead>
<tr>
<th>Subject Section</th>
<th>Shaw ’31</th>
<th>Shaw Supplement</th>
<th>Mohrhardt</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ω Rank</td>
<td>Size ω Rank</td>
<td>ω Rank</td>
</tr>
<tr>
<td>Total</td>
<td>8.1</td>
<td>21.6 1 18</td>
<td>8.4</td>
</tr>
<tr>
<td>Physical Education, Health</td>
<td>15.3</td>
<td>13.2 3 5</td>
<td>12.9</td>
</tr>
<tr>
<td>Education</td>
<td>12.4</td>
<td>11.7 6 9</td>
<td>10.9</td>
</tr>
<tr>
<td>Economics</td>
<td>11.7</td>
<td>10.9 7 8</td>
<td>9.6</td>
</tr>
<tr>
<td>Chemistry, Physics</td>
<td>10.9</td>
<td>9.6 8 19</td>
<td>8.1</td>
</tr>
<tr>
<td>Psychology</td>
<td>9.6</td>
<td>7.9 9 13</td>
<td>6.7</td>
</tr>
<tr>
<td>Sociology</td>
<td>8.1</td>
<td>7.7 10 12</td>
<td>6.7</td>
</tr>
<tr>
<td>Political Science</td>
<td>7.7</td>
<td>6.7 11.5 2</td>
<td>6.7</td>
</tr>
<tr>
<td>Geography</td>
<td>6.7</td>
<td>6.5 13 1</td>
<td>6.5</td>
</tr>
<tr>
<td>Botany, Zoology</td>
<td>6.7</td>
<td>6.3 14 3</td>
<td>6.3</td>
</tr>
<tr>
<td>Religion</td>
<td>6.3</td>
<td>6.0 15 15</td>
<td>6.0</td>
</tr>
<tr>
<td>History</td>
<td>6.0</td>
<td>5.7 16 4</td>
<td>5.7</td>
</tr>
<tr>
<td>Fine Arts</td>
<td>5.7</td>
<td>4.9 17 14</td>
<td>4.9</td>
</tr>
<tr>
<td>English</td>
<td>4.9</td>
<td>4.2 18 11</td>
<td>4.2</td>
</tr>
<tr>
<td>General, Astronomy, Geology</td>
<td>4.2</td>
<td>4.0 19 6</td>
<td>4.0</td>
</tr>
</tbody>
</table>

methodology in a subject or to a sudden increase of interest or expansion in the field. It has not been possible to isolate

This pattern or cycle is a function of contemporary conditions, not of the age of the titles.

The rise may be ascribed to a number of factors: (a) an increased production

120 COLLEGE AND RESEARCH LIBRARIES
of books in the period; (b) an increased proportion of desirable books in the period; and (c) survival of a larger proportion of these titles because enough stronger titles did not appear in the war years immediately following. The drop may be due to a reverse of these conditions. Certainly it was not due to drop in production alone, for there was a wealth of publications in history and political science during the war years. It is more likely that the very urgency and immediacy of these publications in a critical period condemned them to short life.

A much larger volume of data, covering a longer span of time, would be required to demonstrate conclusively the nature and extent of fluctuations in \( (\omega) \) for given subjects.

In this connection Sorokin has declared that "purely quantitative [astronomical] time cannot replace sociocultural time and is inadequate for a study of sociocultural phenomena." As an example he points out that "One year of existence of a modern social group is packed with more numerous and greater changes than are fifty years of existence of some isolated primitive tribe." Some investigators have studied a similar problem in the process of forgetting, where the nature and intensity of activities between the learning period and the retention test are factors.

**Mortality and Life Expectancy**

Fitting of the exponential equation to the data and computation of \( (\omega) \) makes possible certain generalizations regarding the life expectancy and mortality of books. Of course no definite statements can be made in terms of individual titles. But, regarding a given group, predictions can be made with the same justification as they are in similar situations in life insurance and annuities.

The annual mortality rate for any given group is expressed by \( (\omega) \), or, in terms of per cent, \( (100\omega) \). The number of titles remaining in the group \( (y_r) \) after the lapse of a time \( (t) \) is given by:

\[
y_r = y_0 (1 - \omega)^t
\]

It is possible to compute the time required to reduce the group to any given remainder and the average life or life expectancy. There are many applications of the exponential equation which have been fully explored by workers in other fields and they need not be mentioned here. Some of the simplest and most satisfactory analogies may be found in the field of radioactivity and the decay of radioactive substances.

**Comparison with Library Holdings**

The lists have been generally accepted as practical standards for college library collections. But no general principles which emerge from analysis of these lists can be generally accepted until they are compared and tested against actual library situations. To do this, samples were taken from the shelflists of five local libraries for four subject sections. The results of computation are presented in Table 2. Libraries B, C, and D have acquired the bulk of their collections within the past twenty-five years. Libraries A and E are much older but each has followed a fairly systematic policy of weeding and discarding.

Although there is strong agreement between the libraries and the lists in relative ranks of the coefficients for each
subject, it is evident that, in the aggregate, the coefficients or rates are lower for the libraries than for the lists.

The interpretation of this difference between the libraries and the lists depends upon the point of view. A defender of the lists might say that the libraries have failed to add new material rapidly enough or in sufficient quantity and have failed to discard to a great enough extent.

### Table 2
Rates of Obsolescence in College Libraries

<table>
<thead>
<tr>
<th>Subject</th>
<th>Library</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemistry and physics</td>
<td></td>
<td>4.1</td>
<td>6.1</td>
<td>12.8</td>
<td>12.9</td>
<td>5.7</td>
</tr>
<tr>
<td>Sociology</td>
<td></td>
<td>2.5</td>
<td>11.0</td>
<td>10.5</td>
<td>10.5</td>
<td>2.1</td>
</tr>
<tr>
<td>History</td>
<td></td>
<td>0.1</td>
<td>3.1</td>
<td>6.1</td>
<td>5.2</td>
<td>1.1</td>
</tr>
<tr>
<td>Classics *</td>
<td></td>
<td>1.1</td>
<td>6.0</td>
<td>3.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Denotes negative results, i.e., a preponderance of older material rather than new.

Conversely, it may be said that the libraries are successfully meeting the demands upon them—pragmatic proof of the validity of lower rates—and that the rates derived from the lists are too high, especially for the larger and older libraries. It is obvious that in selecting for a list or in buying for a new library, certain preferences will accrue to the latest publications. Yet the balance in favor of the new may be so slight that the older library is not often justified in discarding an older title it already has for a new one which may be only slightly better. This may well be the situation in classics in libraries A and E, where almost no new titles have been added in recent years and where the demand is small.

The absolute size of a collection may also be a factor in defining obsolescence. For one direction in which the scope of a collection may be broadened is in greater emphasis on the historical background and development of a subject. Thus there is more reason for retaining examples of books which have become obsolete for ordinary daily use.

The truth seems to lie somewhere between the two extremes. Only a careful analysis of each individual library and its functions will indicate the directions in which the emphasis should be. It is essential, however, that the problem be redefined and restudied at intervals of not more than ten years. And the librarian must campaign ceaselessly against the inevitable lag and inertia which, coupled with all too frequent lack of funds, prevents his attaining the ideal.

Logsdon's Study

Further evidence from the point of view of compilers of lists has been presented by Logsdon. In his study of instructional literature of sociology, he includes an extensive analysis of obsolescence. His basic list of titles was gathered from reading lists in eighty-eight textbooks. He describes obsolescence in terms of the citation record of titles in three periods. From his table (p. 32) a rate of obsolescence (60) has been computed as .148 or 14.8 per cent. This would indicate that interest in the titles he found listed declined nearly 15 per cent each year. This is comparable with the rates listed in Table 1 for "Sociology," viz.: Shaw, 11.7; supplement, 19.7; Mohrhardt, 13.7. Logsdon's findings thus appear to be in substantial agreement with those already presented.

Circulation Experience

The rates of obsolescence which have been presented so far have been computed on the basis of select lists, approaching ideal situations, and actual library collections. Even actual library collections have something of the ideal about them, since they are gathered in anticipation of need. Actual experience in the use of books is desirable. Such experience is the daily life of any busy library but it is seldom recorded. An important exception is Stieg's study of circulation at Hamilton College.9

In his Table 4 (p. 40) Stieg records the imprint dates of titles circulated. A curve for the data of 1938-39 has been plotted as a part of Diagram 1. The strong similarity between distribution of imprint dates in the three select lists and the distribution of imprint dates of titles circulated is apparent. The rate of obsolescence (ω) may be computed from each column in Stieg's table and expressed as follows:

For 1938-39 5.0 per cent
For 1939-40 4.8 per cent
For 1940-41 4.9 per cent

A rough mean may be taken at 4.9 for comparison with 8.1 for the Shaw List and 8.4 for the supplement.

The lower rate at Hamilton for circulation may be accounted for in part by several circumstances. First, the collection ("about 190,000 volumes") is much larger than the lists and perforce must contain more older material. Second, the instructional program of the college may be planned to take advantage of greater resources by provision for use of older materials beyond that contemplated by the compilers of the select lists. These compilers made virtually no provision (less than 0.2 per cent) for titles published prior to 1850, while at Hamilton 1.24 per cent of the circulation was in titles published prior to 1850, including even seventeenth and sixteenth century titles. Some of the older titles may have been withdrawn because of interest in their physical form rather than for ordinary reading or reference use. On the other hand, these rates are closer to those derived from checking of the five library collections, as shown in Table 2.

Suggestions for Further Study

The writer is conscious of having scratched only the surface of one corner of a vast field. He hopes that he has turned up a few useful ideas and facts. It seems worth while to list some of the possibilities for future research and practical application. Much of the practical application must await further research in clarification and standardization of many phases.

The coefficients of obsolescence presented here are by no means final. Their computation has been presented in support and illustration of the method. It remains to standardize them by application to large groups of libraries and by further study.

What happens when a library does not discard old books is itself a problem for further study. By what curve or law the growth of libraries takes place and what may be expected in the future, are open questions.

Analysis of Collections

The writer believes that a powerful tool can be built for the analysis of book

collections. Heretofore libraries have been compared with respect to size of their collections and annual additions or with respect to their holdings of specific titles. Now it should be possible to evaluate in quantitative terms at least one factor in the quality of collections.

The extent to which the \( w \) computed for a given library may differ from a generally accepted standard may be taken as an objective indication of its deviation from the norm. It must be remembered, however, that such figures as may be derived are no substitute for special expressed purposes or policies of a given library. If a library chooses to be different and has good reasons, therein lies the justification for deviation. But the deviation can still be measured.

Comparison of the \( w \) for a given subject or collection in a library with a generally accepted standard can be the basis for an evaluation or for corrective action. For example, if in a study of its collection of sociology a college library finds an \( w \) of .01 whereas the generally approved figure for comparable libraries is at least .08, there is an obvious deviation from the norm. The cause for this deviation may be that (a) the library is not adding new books in sufficient quantities and (b) that it is not discarding older material. If the generally accepted \( y_a \) for sociology is forty titles within five years (i.e., the maximum point on the curve as described) and the library has only twenty titles at maximum, then at least part of the difficulty is evident. It is failure to add sufficient new material each year.

But if the library has been adding an average of forty titles per year, the slight slope of the curve is due to failure to discard older materials. This failure may be examined in the light of the purposes of the library. If there is a deliberate policy of retaining otherwise obsolete material for some definite purpose, no further justification is needed. But it must be remembered that the costs of housing and caring for the older material can legitimately be chargeable only to this special use.

**Administrative Problems**

New light on many administrative problems in the gathering and maintenance of library collections can be cast by computations from the exponential formula or by simple graphic illustration.

An offer of a large gift of older material can be weighed with regard to what it will do to the present distribution of a library. It might fill in a gap left in the past but, more likely, it will increase the proportion of obsolescent material. And, if the library is to maintain the same standards in the future, the pace of adding new books and discarding old ones will have to be accelerated.

For the college library faced with the necessity of storing or setting aside the lesser-used portions of its collections, as suggested by many and practiced by few, the formula and rates may be used in planning what to segregate and in estimating the demand for segregated material.

Assuming an \( w \) of .05 for an entire collection, as suggested by the experience at Hamilton College, the half-life will be approximately fourteen years. That is, half of the useful collection will be in titles fourteen years old or less. Likewise, three-fourths of the demand will be for books less than twenty-eight years old or certainly less than thirty years old. The average life or life expectancy would be about twenty years.
Titles over thirty years old in many college libraries comprise at least half of the collection. But only a small proportion are in active or potential demand, probably not more than 10 per cent. Thus 50 plus 10 per cent, or 60 per cent of the collection, may account for 90 to 95 per cent of the demand or circulation. Because the remaining 40 per cent would be used so little, these volumes might be removed to a less accessible place of storage with very little inconvenience. Or they might be removed from service entirely with a loss of only 5 to 10 per cent of the total use of the library.

This loss of use would certainly be counterbalanced by a corresponding increase in efficiency in the use of the live material retained.

Within a library it may be desirable to compare several sections, with a view to determining relative need for book funds. In budgeting departmental purchases it is important to know approximately how many titles are required each year in each subject and to know that replacements in such fields as classics are needed less urgently than in certain of the sciences. Conversely, the professor of classics may claim that money is more wisely spent on books with longer life expectancy.

One of the least explored areas of cost accounting is that for public and semi-public nonprofit institutions. Understanding and statement of the rate of obsolescence of library book collections in financial terms would be an important contribution toward the solution of this problem. Such information should also provide a useful lever in annual budget requests. The librarian can readily demonstrate how many titles must be purchased annually to keep his collection up to the standards set for it.

The publisher has other phases of the accounting problem. He must decide when to dispose of unsold inventories of older books. His editors must decide when a book needs to be revised or even when a new book is needed in a given field. Some help in these problems should come from knowledge of the rates of obsolescence for various types of book.

Among the more general problems to the solution of which this study may make some contribution is that of the general rate of cultural evolution. Books represent one of the higher forms of culture and the rate at which they are discarded and replaced may give some suggestion as to the rate of evolution of the general culture of which they form a part. Perhaps some future investigator with the means and spirit may find it possible to compare various culture groups in speed of evolution by study of their books. One might compare some of the relatively stable and permanent Oriental groups with the relatively dynamic state in Europe and America.

---