

The Yale Survey: A Large-Scale Study of Book Deterioration in the Yale University Library

Gay Walker, Jane Greenfield, John Fox,
and Jeffrey S. Simonoff

A large-scale survey of the physical condition of books and the nature of the collections in the Yale University Library system that evaluated more than 36,500 volumes was carried out. Results have been tabulated, compared by computer, and analyzed to provide statistical information on the fifteen distinct collections surveyed in thirty-six separate strata. Environmental conditions were also monitored. These studies, along with the analyses of binding materials and methods, were used to formulate probable reasons for deterioration levels as well as documenting these levels. Several questions of particular interest were compared in two-way intersections, and a brief analysis was made of publication dates in relation to age and condition of a selected group of books. It was found that 37.1 percent of the books sampled overall had brittle paper (i.e., broke after two double folds) and that 82.6 percent of the books overall had acidic paper (i.e., a pH of below 5.4). These and other results should help Yale and libraries elsewhere to identify their preservation needs and develop appropriate programs.



One of the most serious problems facing research libraries today is the preservation of the materials that comprise their collections—materials that are deteriorating because of their chemical composition, the mechanics of their construction, and the effects of uncontrolled environmental conditions. Deterioration is a particularly

critical problem in large libraries, where the age and size of the collections make evaluation and corrective action difficult. It has been estimated that more than six million volumes in the collections of the Library of Congress have deteriorated so badly they cannot be given to users without risk of irreparable damage;¹ at the New York Public Library, it is estimated

R. Gay Walker is head of the Preservation Department and curator, Arts of the Book, Yale University Library, New Haven, Connecticut 06520. Jane Greenfield, now retired, was past conservator and head of the Conservation Studio, Yale University Library. John Fox is a graduate student, Statistics Department, Yale University. Jeffrey S. Simonoff was a graduate student in the Statistics Department, Yale University, at the time of this survey. He is presently on the faculty of the Graduate School of Business Administration at New York University, New York 10012. The authors would like to acknowledge financial support for the design and implementation of this survey from the National Endowment for the Humanities (#RC-32954-79-1585) and the Mellon Foundation. Those interns involved in the survey were Paul Beck, David Boardway, Charlotte Brown, Candace Brugmann, Diane Burke, Margaret Byrnes, Ellen Chin, Anne Dutlinger, Patricia Gladys, Sharlane Grant, Nancy Grussing, Michael Holland, Cheryl Jones, Lynne Keller, Margaret Madison, Jan Merrill-Oldham, Marion Munzer, Carolina Portela, Sharon Pugsley, Mary Schlosser, Pamela Spitzmueller, Gregor Trinkaus-Randall, Karen Walsh, and Virginia Wisniewski-Klett. Special thanks go to Jan Merrill-Oldham for her careful editorial review and suggestions. Thanks also go to Rutherford D. Rogers and the Yale Library's administration for local support, to Louis J. Volpi, and to the staffs of the Preservation and Conservation divisions for their participation and support.

that as much as half of the collection has reached a similarly advanced state of disintegration.²

During the past decade, several attempts have been made to explore this problem, and a variety of responses to it have been initiated. The first major study was conducted by the Association of Research Libraries; it attempted to "identify specific steps that might be taken . . . to work towards resolution of the many problems . . . brought on by the physical deterioration of books and journals."³ Authors treating this subject included Darling⁴ and Walker,⁵ both of whom urged more complete documentation of preservation activities. Several major libraries, including New York Public, Columbia University, Stanford University, University of California at Berkeley, University of Michigan, Newberry, and Yale University libraries, have formed preservation units for the purpose of repairing and replacing damaged books and journals. Outstanding recent developments have been initiated by the Library of Congress, where a pilot project on the storage of information in digital form is under way and where the Preservation Research and Testing Office is conducting extensive research on the permanence of materials, practical methods for mass treatment of deteriorated books, and the effects of buffering agents on papers and inks.

Those institutions that have attempted to address the preservation problem have been hampered by the lack of a detailed study to determine its scope. Although small, limited surveys were conducted at Stanford University⁶ and at several other academic libraries, a large-scale study had never been attempted. In 1979, the Preservation/Conservation group at Yale applied for and received a three-year grant from the National Endowment for the Humanities (NEH) to survey the Yale collection, evaluate the results, develop educational tools, and provide interns with advanced training in preservation/conservation procedures and theories. Additional support was provided by the Andrew W. Mellon Foundation. The grant was administered by project codirectors, Jane Greenfield and Gay Walker.

The proposed survey was undertaken to determine the extent and nature of the deterioration of books in the Yale University Library system. This was a job of major proportions; Yale Library has the second largest collection of any academic library in the nation. In 1982, 7,725,424 volumes were held in forty separate library units. The records for that year show that more than one million volumes had circulated. This figure does not reflect in-house use of materials. In order to obtain results in which we could be confident, a very large sample—more than 36,500 volumes—was surveyed. Fifteen of the sixteen major libraries were divided into thirty-six subunits, each of which was treated separately in terms of its statistical framework and the generation of results. The surveyed libraries varied greatly in size, age, and nature of buildings and collections; environmental conditions; reader access; and circulation patterns. The following descriptions of some of the surveyed units illustrate this point.

The Sterling Memorial Library (the main library) houses approximately four million volumes, including some that date back to 1600. The collection has grown steadily since 1701, when the university was founded. Opened in 1931, the building has fifteen floors of stacks accessible to the Yale community, a centrally controlled heating system, and no air-conditioning. Among the subunits within the Sterling Library that were sampled separately was the Preservation Division. The 13,000 books held there represent the work flow of materials regularly sent to Preservation for repair, replacement, or reproduction. The 3,359-volume Statistics Library (the smallest library sampled) is located in one room of what was once a private house, built in 1849. The Cross Campus Library, a two-floor, air-conditioned underground structure built in 1971, houses 150,000 volumes. This undergraduate collection receives the heaviest use of any within the library system. The Kline Science Library is comprised largely of twentieth-century periodicals and recent scientific texts (the older science materials are in the Sterling stacks). Because it was felt that rare books should not be tested for pH and brittle-

ness, none were surveyed. The major unit omitted was the Beinecke Rare Book and Manuscript Library. Folios were also omitted because they are awkward to handle and are easily damaged.

SURVEY GOALS

The survey was designed to yield a detailed description of the collections in the discrete units of the Yale system; to examine the complex relationships between the nature of materials, their condition, and the environment in which they are housed; and to estimate how many volumes require immediate attention, how many will need attention soon, and what kind of attention will be needed. In order to gather the requisite data, the project co-directors devised a series of questions that could be used to evaluate books. Eight of these questions helped to establish the size of the preservation problem:

1. Is the primary protection (binding, box, or protective cover) intact?
2. Is the leaf attachment (sewing, gluing, or stapling together of pages) intact?
3. Is the paper very brittle (does the corner of a page break off after two double folds—i.e., after being folded in one direction, then in the opposite direction, twice)?
4. Is the paper very acidic (i.e., does a test using an archivist's pen filled with bromocresol green show the paper to be pH 5.4 or below)?
5. Is the printed area of all pages intact?
6. Is the book mutilated (i.e., damaged by humans or animals)?
7. Is the book damaged by environmental factors (i.e., are there signs of fading or water damage)?
8. Does the volume require immediate treatment (replacement, reproduction, repair, or rebinding)?

Other questions were devised to expand the profile of the deteriorated volumes and to suggest reasons for their deterioration:

9. What is the country of publication?
10. What is the date of publication?
11. Is the book circulating or noncirculating?
12. What kind of primary protection

(binding, box, wrapper) does the book have?

13. What kind of material covers the joint (the outer hinges)?

14. How are the leaves of the book attached?

15. What is the width of the gutter (inner) margin?

Most questions are of particular interest as they intersect with others. Among those sets analyzed were questions 1 and 12, 2 and 14, 3 and 4, 3 and 8, 3 and 15, 4 and 14, 4 and 15, 8 and 12, 8 and 13, and the three-way intersection of 3, 9, and 10. See appendix A for the sampling methodology and procedures.

THE PILOT STUDY

An important step in the construction of the Yale survey was a pilot survey, that is, a preliminary run-through on a small subset of the total number of volumes to be sampled. The pilot helped identify and eliminate problems in the sampling design (for instance, it was discovered that certain questions were worded ambiguously). A pilot study of 1,000 books in one stratum was carried out. This was a number large enough to achieve the desired objectives but small enough so that the study could be done quickly and analyzed inexpensively.

The pilot study emphasized the need for the following: (a) a consistent method of locating books, e.g., by always moving clockwise around a range when counting sections; (b) detailed instructions on how to fill out questionnaires and guidelines for answering the questions; (c) a knowledge of book structure and the ability to recognize different methods of leaf attachment and the various materials used for book covering.

The educational program for the surveyors and the instructions for locating books and evaluating them (see appendix B) were evolved during the trial run. A truck of books containing various binding styles and covering materials was assembled for each group of NEH intern surveyors to study. The surveyors also spent time in the stacks practicing evaluation techniques to standardize findings, and they attended a discussion session at which the

statisticians explained statistical theory. After a surprisingly short period of practice, each group was able to work smoothly and efficiently.

IMPLEMENTATION

Six groups of four interns each carried out the survey over the course of two and one-half years. Each group stayed at Yale for five months and spent close to half of each day surveying. The total time spent evaluating books was about thirty-eight hundred hours.

College Board form IBM-H45352 was used to record findings in a machine-

readable format. This procedure eliminated errors that are sometimes introduced when data must be input into a computer manually. An overlay of thin cardboard (figure 1) with windows cut out to expose areas where answers were to be recorded was placed on the survey form. The form and overlay were supported in correct relative position by a jig (figure 2) that also held finished forms, the thin cardboard strip used to measure gutter margin, #2 pencils needed to fill in the form, and an archivist's pen used to check pH. A short list of abbreviations for names of countries (appendix C) was taped on

Call number

Floor

Range

Section

Shelf

Book

Random number

Date

Country

Date

Questions

Intern code letter

1. Yes No Tagged

2. Yes No Circulating

3. Rigid binding
Limp binding
Acidic paste
Acid-free paste
Acidic box/env.
Acid-free box/env.
None
Other

Primary protection

4. Cloth
Leather
Paper
Vellum
Other

Covering of joint

5. Yes No N/A Primary prot. functional

6. Sem/stapled thru fold
Oversewn/Glued-sewn
Adhesive
Stapled

None
Unknown
Other

Leaf attachment

7. Yes No N/A Leaf attachment intact

8. Breaks after 2 folds
Breaks after 4 folds
Does not break
Not tested

Brittle paper

9. Blue
Not blue
Not tested

pH of paper

10. Less than 1 cm
1 cm or greater

Gutter margin width

11. Yes No Text intact

12. Yes No Mutilation

13. Yes No Environmental damage

14. Yes No Repair needed

FIGURE 1

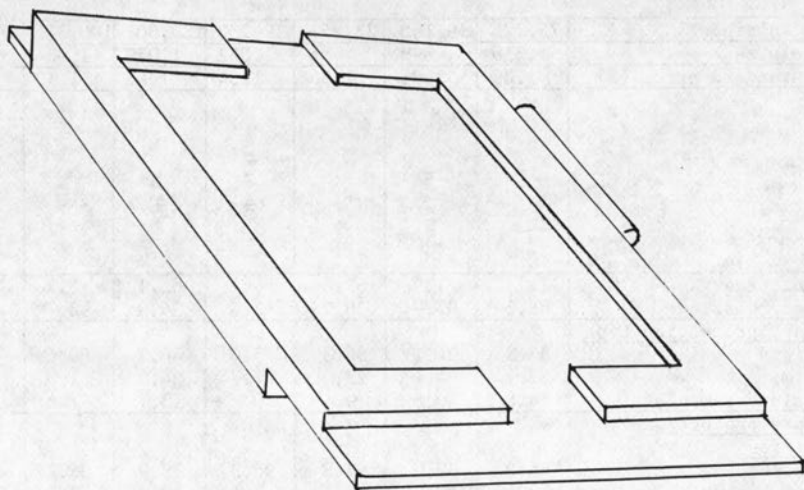


FIGURE 2

Jig to Hold Overlay and Evaluation Form in Alignment

the back of the jig. In addition, a full list of abbreviations, translations of Latin and Cyrillic place-names, an atlas, and shelf-lists were available.

RESULTS

Appendix D gives the statistical analysis of the data gathered. The sample results for each of the fifteen libraries surveyed are presented in the tables in figure 3. Findings for thirteen questions are expressed as percentages; answers to the question about tagging are omitted (this figure was always under 10 percent, as mentioned before). Great care was taken to obtain sufficiently large sample sizes to ensure that our estimates were accurate to within a few percentage points. The *maximum* standard error observed for each of the thirteen questions listed in the tables is given at the top of each column. Although the highest standard error in an entire unit was 2.58 percent (in the Statistics Library, where the smallest sample was taken), most were well under 1 percent. For questions with several possible answers, e.g., "How are the leaves attached?" only those answers comprising at least 2 percent of the total response in any stratum were included in the tables. The percentages, therefore, do not always add up to 100 percent in each category. Libraries

having air-conditioning are indicated in the tables.

The following example illustrates how survey results should be interpreted. Refer to the first table in figure 3, Art Library column. In a sample of 1,336 volumes, 17.2 percent were found to be in need of treatment. (Given the calculated standard error of 1.03 percent for the question about repair, a 99 percent confidence interval for the actual percentage of books in the Art Library in need of treatment is 14.53 percent to 19.87 percent.)

The survey findings were very interesting—some because they confirmed previous estimates of the scope of the preservation problem, and others because they provided new data with which to analyze the problem. Salient aspects of these results are discussed below, question by question. (Because the Sterling sample was the largest, statistics from that collection are most often cited. Findings from other strata are mentioned when they are of particular significance.)

Is the Primary Protection Intact?

The data gathered in response to this question can help identify those collections that would be good candidates for rebinding projects. Findings also suggest levels of use and maintenance of a given

Total holdings	75,775	19,165	22,357	29,250	130,001	100,315	170,116
Sample size	1,336	898	965	1,032	1,079	1,964	1,588
Max. standard error (%)	1.36	1.66	1.56	1.54	1.52	1.13	1.28
Library	Art	Classics	Drama	Engineering*	Forestry	Geology*	Kline Science*
Circulating	11.1	1.9	62.6	56.7	65.1	3.2	46.1
Non-circulating	88.6	97.9	36.9	43.3	34.9	96.7	53.9
<u>Primary protection type</u>							
Rigid	55.2	94.1	66.3	75.0	46.7	89.6	85.9
Limp	37.7	4.1	22.0	24.7	34.9	1.0	9.8
Acidic pamphlet	4.8	0.5	9.4	0	13.7	7.6	3.8
<u>Primary protection condition</u>							
Intact	86.2	89.0	92.8	96.3	79.9	96.7	90.8
Not intact	13.4	10.6	6.9	3.4	19.7	3.1	9.0
<u>Joint covering</u>							
Cloth	60.4	90.0	69.8	74.4	58.7	89.5	85.2
Paper	33.5	5.6	23.7	24.5	35.0	3.4	9.7
Leather	1.4	1.6	2.0	0	0.6	0.7	2.4
<u>Leaf attachment type</u>							
Sewn-through-fold	44.5	56.2	61.9	57.4	44.8	56.0	46.1
Over/cleat sewn	28.6	37.9	13.9	33.1	24.6	27.4	40.0
Stabbed	10.7	1.6	10.0	2.3	13.6	1.7	4.8
Adhesive	8.4	2.8	8.7	5.1	9.5	12.3	7.9
<u>Leaf attachment condition</u>							
Intact	94.6	89.1	91.6	98.2	90.6	96.3	97.5
Not intact	3.1	9.6	6.0	1.5	4.9	2.1	2.0
<u>Text condition</u>							
Intact	96.1	95.3	94.5	98.6	94.0	97.8	99.0
Not intact	1.5	3.8	2.4	1.2	1.5	0.7	0.5
<u>Brittle</u>							
Does not break	81.5	64.0	77.8	93.3	71.1	89.4	73.4
Breaks at 4 folds	7.1	11.9	10.3	3.0	8.3	3.5	6.0
Breaks at 2 folds	9.2	23.2	7.7	3.3	16.0	5.5	20.1
<u>pH</u>							
pH above 5.4	22.6	16.9	33.2	33.5	17.1	48.0	32.0
pH below 5.4	75.1	82.1	63.9	66.2	78.4	50.3	67.4
<u>Gutter margin</u>							
1 cm. or more	66.5	70.2	76.7	69.5	68.2	60.9	75.4
Less than 1 cm.	31.4	29.1	21.1	30.1	27.4	37.5	24.1
<u>Mutilated</u>							
No	99.0	97.9	98.0	98.6	98.4	95.5	98.4
Yes	1.0	2.1	2.0	1.4	1.5	4.5	1.5
<u>Environmental damage</u>							
No	92.4	83.5	96.9	94.5	89.4	98.7	98.8
Yes	7.5	16.5	2.9	5.5	10.2	*1.3	1.1
<u>Treatment needed</u>							
No	82.7	83.3	88.7	95.5	75.8	95.3	89.7
Yes	17.2	16.7	11.2	4.5	24.1	4.7	10.2

*Air-conditioned libraries

FIGURE 3
Sample Results (%)

Total number	3.1 mil	18,000	1,500	13,000	5 mil.
Sample size	17,096	300	222	528	36,512
Max. standard error (%)	0.38	2.70	3.35	2.17	0.30
Library	Sterling stacks (15 floors)	Reference	Periodical stacks	Preservation	Yale Overall
Circulating	97.2	0	0	0	87.2
Non-circulating	2.5	100	100	100	12.5
<u>Primary protection type</u>					
Rigid	75.3	91.7	96.0	86.0	76.0
Limp	7.5	4.7	1.4	10.0	9.7
Acidic pamphlet	12.8	0.7	2.7	3.2	11.0
<u>Primary protection condition</u>					
Intact	91.8	95.3	94.1	20.1	91.4
Not intact	7.6	3.7	5.9	79.6	8.1
<u>Joint covering</u>					
Cloth	75.3	74.0	96.4	54.7	76.9
Paper	9.8	6.0	1.4	10.4	11.3
Leather	6.8	7.7	2.3	33.1	5.4
<u>Leaf attachment type</u>					
Sewn-through-fold	51.7	67.7	23.0	78.6	52.4
Over/cleat sewn	22.5	21.3	69.8	10.6	24.6
Stabbed	17.7	3.0	5.9	8.9	13.7
Adhesive	4.8	7.3	1.4	1.3	6.3
<u>Leaf attachment condition</u>					
Intact	92.5	94.0	94.6	54.9	93.6
Not intact	4.4	5.3	5.0	44.5	4.0
<u>Text condition</u>					
Intact	94.4	95.7	88.3	52.1	95.4
Not intact	2.2	1.7	11.3	47.4	2.0
<u>Brittle</u>					
Does not break	50.8	74.0	52.3	13.8	59.4
Breaks at 4 folds	8.8	7.0	11.3	8.0	8.2
Breaks at 2 folds	35.9	18.3	36.0	69.3	28.9
<u>pH</u>					
pH above 5.4	8.3	24.7	15.3	2.1	13.7
pH below 5.4	87.0	74.3	83.8	89.0	82.6
<u>Gutter margin</u>					
1 cm. or more	75.9	79.3	65.3	88.3	75.2
Less than 1 cm.	21.2	20.0	34.2	11.0	22.5
<u>Mutilated</u>					
No	96.5	97.7	85.6	90.0	95.9
Yes	3.1	2.3	14.4	9.7	3.8
<u>Environmental damage</u>					
No	95.3	100	98.7	81.3	95.8
Yes	4.3	0	1.4	18.4	3.9
<u>Treatment needed</u>					
No	86.5	90.7	82.4	3.2	87.0
Yes	13.2	9.3	17.6	96.6	12.8

*Air-conditioned libraries

FIGURE 3 (Continued)
Sample Results (%)

Total number	652,896	24,604	272,215	117,021	115,416	3,359	150,000
Sample size	2,923	1,026	1,507	1,266	1,129	364	1,289
Max. standard error (%)	1.03	1.55	1.29	1.30	1.48	2.58	1.30
Library	Law	Mathematics	Medicine	Music	Social Science*	Statistics	Cross Campus*
Circulating	99.1	97.3	56.9	81.6	76.9	94.0	98.3
Non-circulating	0.7	2.0	43.0	18.2	23.1	4.1	1.2
<u>Primary protection type</u>							
Rigid	87.2	84.8	67.2	69.3	45.2	74.7	94.8
Limp	7.3	8.0	28.1	1.2	14.5	23.6	0.1
Acidic pamphlet	4.7	6.8	2.3	27.8	36.9	0.6	4.3
<u>Primary protection condition</u>							
Intact	92.5	97.2	88.1	97.2	92.6	97.5	93.2
Not intact	7.1	1.9	11.5	2.4	7.2	1.4	5.8
<u>Joint covering</u>							
Cloth	84.4	82.4	67.6	95.7	72.5	71.7	90.7
Paper	7.2	12.6	27.7	2.0	18.3	25.8	3.1
Leather	6.4	3.0	0.5	1.2	0	0.8	0.2
<u>Leaf attachment type</u>							
Sewn-through-fold	53.3	40.0	50.5	75.4	34.6	67.6	71.5
Over/cleat sewn	32.0	44.9	31.7	13.2	12.9	19.0	12.3
Stabbed	7.7	5.6	4.8	4.6	21.3	2.5	5.7
Adhesive	5.7	8.4	10.0	4.2	27.6	9.6	10.0
<u>Leaf attachment condition</u>							
Intact	96.9	98.1	93.7	94.2	97.3	98.1	97.3
Not intact	2.4	1.7	3.8	4.6	2.6	0.6	2.0
<u>Text condition</u>							
Intact	98.1	99.1	96.4	95.4	98.9	97.8	98.0
Not intact	1.2	0.4	1.1	3.8	0.9	0.6	0.9
<u>Brittle</u>							
Does not break	54.2	74.7	87.7	82.2	96.6	95.6	91.1
Breaks at 4 folds	9.9	9.8	6.0	7.4	2.0	3.6	3.5
Breaks at 2 folds	33.7	15.0	3.9	9.6	1.3	0	4.9
<u>pH</u>							
pH above 5.4	9.5	10.3	40.5	8.6	25.0	40.9	30.3
pH below 5.4	88.3	89.4	57.1	90.4	74.8	58.8	67.9
<u>Gutter margin</u>							
1 cm. or more	80.4	61.8	67.1	73.8	63.2	86.0	84.6
Less than 1 cm.	18.8	37.6	30.3	25.5	36.7	13.5	14.7
<u>Mutilated</u>							
No	96.3	95.1	98.0	90.1	97.9	99.7	74.6
Yes	3.5	4.7	1.9	9.9	2.1	0.3	25.0
<u>Environmental damage</u>							
No	97.0	97.2	97.5	98.4	96.8	99.2	98.4
Yes	2.8	2.6	2.3	1.5	3.2	0.8	1.2
<u>Treatment needed</u>							
No	90.4	96.8	85.3	92.8	88.8	98.1	85.2
Yes	19.5	3.0	14.6	7.0	11.2	1.9	14.4

*Air-conditioned libraries

FIGURE 3 (Continued)
Sample Results (%)

collection. More than 7 percent of the sample surveyed in the Sterling stacks had broken bindings (which represents around 235,000 volumes if extrapolated to the entire stack holdings). Surprisingly, the percentages of volumes with broken bindings in the high-use Periodical and Reference collections were lower (5.9 percent and 3.7 percent, respectively), which is perhaps a function of the Reference staff's prompt processing of materials in need of rebinding. Predictably, nearly 80 percent of volumes in the Preservation Division had bindings that were not intact. More than 10 percent of the volumes in the long-established Forestry, Medicine, and Classics libraries had broken bindings—the Forestry Library having the largest percentage (19.7 percent); while the new collection in the Geology and Engineering libraries had very few broken bindings. However, the low numbers at Music (an old collection), and the high numbers at Kline and Art (relatively new collections), suggest that level of usage and care may also be important factors for predicting binding condition.

Is the Leaf Attachment Intact?

The condition of the leaf attachment suggests the levels of use and maintenance of a collection, as does the condition of the primary protection, but the implications of leaf-attachment problems can be more serious. The text blocks of books with broken leaf attachments must be re-sewn or reglued, procedures that are not possible when margins are narrow or paper is brittle. The condition of leaf attachments was reasonably good throughout the library system. As might be expected, 44.5 percent of the volumes in the Preservation Division had leaf-attachment problems, but percentages in the remaining units ranged from 0.6 percent to 9.6 percent.

Is the Paper Very Brittle?

The test for paper embrittlement produced the most significant results of the survey. Brittle volumes cannot be easily rebound or repaired, cannot withstand photocopying or heavy use, and would not benefit appreciably from deacidifica-

tion. The test for embrittlement is fairly objective: the corner of a page was folded back and forth four times (two double folds). If the corner broke off after one double fold, the paper was considered extremely brittle; after two double folds, brittle. Of the books surveyed in the main Sterling stacks, 44.7 percent did not survive the four-fold paper test—a percentage that represents between 1,351,600 and 1,420,420 books. As might be expected, more of the paper in the Preservation Division collection was brittle (77.3 percent). Most of the older collections contained a high percentage of brittle books. Of the materials in the Periodical stacks (where back files of the 100 most heavily used titles are shelved), 47.3 percent were embrittled. The two collections that had the lowest percentage of brittle books, Social Science (3.3 percent) and Statistics (3.6 percent), are relatively new. When results for the entire Yale Library system's holdings were weighted and combined, a total of between 1,796,100 and 1,879,377 volumes were estimated to have brittle paper. These findings signal the need for expanded replacement and reproduction programs.

Is the Paper Very Acidic?

The pH is important because of the established correlation between paper acidity and longevity. In general, the more acidic the paper, the more short-lived it is. Determining the percentage of acidic materials in a collection is useful for predicting long-range preservation needs. We measured acidity using a simple pH indicator. A small mark was made in the gutter margin of each book using a felt-tipped pen filled with bromocresol green. The chemical is green when applied but turns blue within about thirty seconds if the pH of the paper is above 5.4. Since a pH of 5.4 or below is *very* acidic (i.e., well below neutral), the percentage of acidic books identified in the survey is conservative. We approached these results with some caution because color changes are sometimes hard to read in dim stack areas; however, the results *did* corroborate similar findings in other studies. In no library unit did more than 48 percent of the paper

tested have a pH higher than 5.4. Those collections with the largest percentages of better paper were generally the newer ones. In the Sterling Library, 87 percent of all papers tested had a pH of under 5.4. We can estimate, then, that about 2.57 million volumes in Sterling have acidic paper and are either brittle or will become brittle. In the Cross Campus and Kline Science libraries, 68 percent and 67 percent, respectively, of the paper tested was highly acidic. We had expected this lower percentage, since the bulk of both collections is late-twentieth-century material. When the estimates for each stratum in the survey are weighted and the results added together, the total number of items having highly acidic paper in the five-million-volume target population is between 4,065,192 and 4,128,542 (82.6 percent).

Is the Printed Area of All Pages Intact?

The identification of pages with tears or breaks extending into the text is significant because of the immediate potential for loss of information. The Periodical stacks showed severe damage; 11.3 percent of the books sampled had broken or torn paper. As might be expected, 47.4 percent of the volumes sampled in the Preservation Division had damaged paper. However, the results from all other units were low, e.g., only 2.2 percent of the books surveyed in Sterling showed damage to printed areas of pages.

Is the Book Mutilated?

Extensive mutilation suggests heavy reader use. The problem was moderate throughout the library system, with the exception of the Cross Campus Library, where 25 percent of the books sampled were mutilated. (The Cross Campus Library contains all class texts including items placed on reserve.) The Periodical stacks, the Preservation Division collection, and the Music Library showed mutilation rates of 14.4 percent, 9.7 percent, and 9.9 percent, respectively, while all other collections showed rates of under 5 percent. These findings identify sites where augmented programs for reader

education are necessary.

Is the Book Damaged by Environmental Factors?

Damage to books caused by environmental factors (e.g., water, sunlight, mold, insects) indicates problems with physical housing, including building construction, environmental control, and housekeeping practices. As might be expected, the Preservation Division collection showed the highest rate of environmental damage (18.4 percent). Other collections showing high rates were the Classics Library (16.5 percent) and the Forestry Library (10.2 percent). The median for all strata was 2.9 percent. Some collections have been repeatedly damaged by water from leaking pipes or windows, overflow from sinks, condensation from steam heating units, or rainwater seeping through walls and ceilings. In some undetected cases, mold had grown, exacerbated by high heat and humidity.

Does the Volume Require Immediate Treatment?

The results of this question identified those library units that have the greatest numbers of deteriorated volumes in need of immediate attention—that is, volumes with broken bindings, missing or damaged text, and/or broken leaf attachments. (Intact brittle materials were not included; only those already damaged were identified here.) Surprisingly, the percentage of books needing immediate treatment was much lower than we had believed. (We had estimated that roughly 30 to 40 percent of all items in Sterling would fall into this category.) Although 96.6 percent of the books surveyed in the Preservation Division collection needed immediate treatment, in no other library was this figure higher than 25 percent. In the Sterling stacks, only 13.2 percent of the materials fell into this category. It should be noted, however, that this percentage represents more than 400,000 volumes. More than 10 percent of the collections sampled in Sterling's Periodical stacks, and the Cross Campus, Art, Classics, Drama, Forestry, Kline Science, Law, Medicine, and Social

Science libraries, were also identified as needing immediate treatment.

Is the Book Circulating or Noncirculating?

The question regarding circulation was included to determine whether there is a relationship between the condition of the books and circulation outside the library. Surprisingly, no clear correlation was found—need for treatment being more closely related to age and nature of the collection. For example, the Classics Library is a noncirculating collection but showed a high number of volumes needing treatment, while other circulating collections showed a low rate of damage.

What Kind of Primary Protection Does the Book Have?

Identifying and quantifying the types of primary protection (including bindings, boxes, envelopes, and wrappers) are useful for estimating the number of volumes in need of first-time binding and the number of acidic pamphlet binders that must be replaced by alkaline binders. The most common primary protection was the rigid binding (hard covers that provide firm support). The percentage of rigid bindings varied throughout the library system, from 45 percent to 96 percent. Also common were limp supports (paper or other flimsy covers) and acidic pamphlet binders. The percentage of limp bindings varied from 0.1 percent to 38 percent. In those few libraries where the number of limp bindings is high, a review of binding policies may be appropriate. The percentage of acidic pamphlet binders ranged from 0 percent to 36.9 percent throughout the system. The estimated number of these binders in Sterling alone was 396,800 (12.8 percent).

What Kinds of Materials Cover the Joint?

Identification of the materials covering the joints (the outer hinges) of books helps to describe library collections, particularly when it is coupled with information about condition. Because of the degree to which joints must flex, they are extremely vul-

nerable to failure. The nature of the material covering the joint is therefore critical to the durability of the binding. Thirty-three percent of the books awaiting treatment in the Preservation Division were bound in leather. Since the highest percentage in any other collection was 7.7 percent and the median was 1.7 percent, this finding suggests that leather is more fragile than other binding materials. (It is also difficult to repair.) The large number of books with paper-covered joints may represent a future binding problem, although many of the pre-nineteenth-century paper bindings have held up extremely well.

How are the Leaves of the Book Attached?

The method by which the leaves of a book are held together (e.g., sewing-through-the-fold, oversewing, gluing) is an important factor in determining whether the book can be rebound if necessary. In all but two collections, volumes bound by sewing-through-the-fold (i.e., through the folds of the signatures) outnumber those bound by any other method. Volumes that are sewn through the folds not only open easily, but can usually be rebound provided the paper is still flexible. In the Mathematics Library and the Periodical stacks, oversewing and cleat sewing were more common. Both methods require trimming away the folds and some of the inner margin and can make it impossible to rebind a volume successfully. The percentage of adhesive-bound volumes in collections with large holdings of new books proved the current popularity of this binding method. The Social Science Library had the highest percentage of adhesive-bound volumes (27.6 percent). Stab sewing, a method long used in pamphlet binding at Yale, was also widespread—ranging from 1.5 percent to 21.3 percent in the various collections, with a median of 5.7 percent.

What Is the Width of the Gutter Margin?

The width of the gutter margin was examined in order to estimate the percentage of books that could not be rebound

readily by commercial methods, regardless of paper quality. A small strip of heavy paper was marked with a one-centimeter line and used to measure the narrowest inner margin in the book. (One centimeter was judged to be the minimum width required to rebind a volume easily. When margins are narrower, special care must be taken either to retain the original sewing structure, and thus all of the margin, or to trim and bind the pages with great care so as not to obscure text. Sometimes neither method is possible.) The highest percentage of books with margins less than one centimeter wide was found in the Mathematics Library (37.6 percent). Between 34.2 percent and 37.5 percent of the books sampled in the Social Science and Geology libraries and in the Periodical stack collection had very narrow margins. These statistics suggest that when bindings fail, reproduction and replacement of materials will in many cases be the only available options.

The last column in figure 3 is a statistical description on the entire sample for the target population of five million volumes. The percentages were derived by weighting the results from the various strata based on their proportion of the whole and adding these results together. It should be remembered, however, that these systemwide statistics must be used with caution because they do not describe a coherent collection. Yale's library units are very different from one another, and the statistics gathered in each are probably best understood as separate studies.

INTERSECTIONS

Some of the most useful statistics resulted from determining the percentage of books that exhibited two or more of the characteristics isolated in the survey (e.g., how many books had *both* broken bindings and paper covers). The results of many of the most important intersections are given in figure 4. Only statistics for the main research collection are given here due to space limitations. These results are shown in a mileage table format where the figures are read from both the horizontal and the vertical axes. For instance in the

first column, 66.7 percent of the books did not need treatment and had rigid bindings, 4.9 percent did not need treatment and had limp bindings, and 11.2 percent did not need treatment and had acidic pamphlet bindings; 8.4 percent *did* need treatment and had rigid bindings, 2.6 percent needed treatment and had limp bindings, and 1.6 percent needed treatment and had acidic pamphlet bindings. The figures for any given intersection may not add up to 100 percent since findings were purposefully omitted when a category (e.g., vellum joint coverings) made up less than 2 percent of the sample and when data were missing because of human error (e.g., the surveyor skipped a question).

All of the intersections that were analyzed proved interesting. Some of the more significant ones are discussed here. Not surprisingly, it was found that acidity and brittleness were directly related; although approximately 80 percent of the nonbrittle books were acidic, more than 99 percent of the brittle books were acidic. Similarly, while only 6 percent of the nonbrittle books needed treatment, more than 20 percent of the brittle books needed treatment. This latter group of books (around 285,000 when extrapolated to the entire Sterling collection) will probably need replacement or reproduction, rather than repair; the books that need treatment but are not brittle can probably be repaired or rebound. An estimated 592,000 volumes, or 18.8 percent of the sample, were brittle *and* had been oversewn, cleat sewn, or stab sewn. These volumes are particularly vulnerable to damage; even gentle use could easily result in broken pages. At the time of the survey, however, the leaf attachments in oversewn and cleat-sewn volumes were generally intact; only 2 percent of those sampled were broken. This contrasts with the other leaf-attachment types, where more than 5 percent were broken. Limp bindings were more prone to failure than other forms; 33.3 percent of all limp bindings were not intact, while 4.8 percent of rigid and 7.8 percent of acidic pamphlet bindings were not intact (around 49,000 of the total number of acidic pamphlet binders in Sterling

Library: Sterling Memorial		Sample size: 17,096				Size of target population: 3.1 million							
	Treatment Needed		Leaf Attachment		Gutter Margin		Primary Protection		pH		Brittle		
	No	Yes	Intact	Not intact	1 cm or over	Less than 1 cm	Intact	Not intact	Above 5.4	Below 5.4	Does not break	Breaks 4 folds	Breaks 2 folds
pH													
pH above 5.4					6.3	1.9					7.9	0.2	0.2
pH below 5.4					68.2	18.7					42.7	8.3	35.6
Primary protection type													
Rigid	66.7	8.4					71.5	3.6					
Limp	4.9	2.6					5.0	2.5					
Acidic pamphlet	11.2	1.6					11.8	1.0					
Joint covering													
Cloth	67.5	7.6											
Paper	7.0	2.8											
Leather	4.9	2.0											
Leaf attachment type													
Sewn-through-fold			48.4	3.1							25.0	4.8	20.6
Over/cleat sewn			21.9	0.5							11.1	2.3	8.9
Stab sewn			17.0	0.7							9.9	1.6	6.0
Adhesive			4.7	0.1							4.4	0.1	0.2
Brittle													
Does not break	47.7	3.0			39.2	11.5			7.9	42.7			
Breaks at 4 folds	7.5	1.2			6.9	1.9			0.2	8.3			
Breaks at 2 folds	27.9	8.0			28.6	7.2			0.2	35.6			

FIGURE 4
Significant Intersections for Questions in Sterling Library

needed immediate replacement). There was also a relationship between binding type and need of attention; while 11.2 percent of the rigid bindings and 12.5 percent of the acidic pamphlets needed attention, 37.4 percent of the limp bindings needed attention. It should be noted, however, that the rigid ones needing attention formed the largest group in absolute terms (about 260,000 needing attention). A relationship between joint covering and need of treatment was also evident; only 10.1 percent of the books with cloth-covered joints required treatment, while 28.6 percent of volumes with paper-covered joints and 29 percent of volumes with leather-covered joints were in need of immediate treatment.

Figure 5 shows the percentage of books sampled in the Sterling Library that were

brittle (paper broke after two or four folds), broken down by date and by country of publication. Statistics from three geographic areas are plotted: the United States, Great Britain and Ireland, and Germany. The results were surprising in that papers older than expected are now embrittled. Paper from the early 1800s through the 1950s is now brittle, with the peak extending from 1860 to 1930. The sharp decline in brittleness by the end of the 1950s is probably not due to a major improvement in paper quality since high levels of acidity continue to be found, but because flexibility has not yet been lost.

ENVIRONMENTAL DATA

During the grant period, five hygrothermographs were placed throughout the library system to document environmental

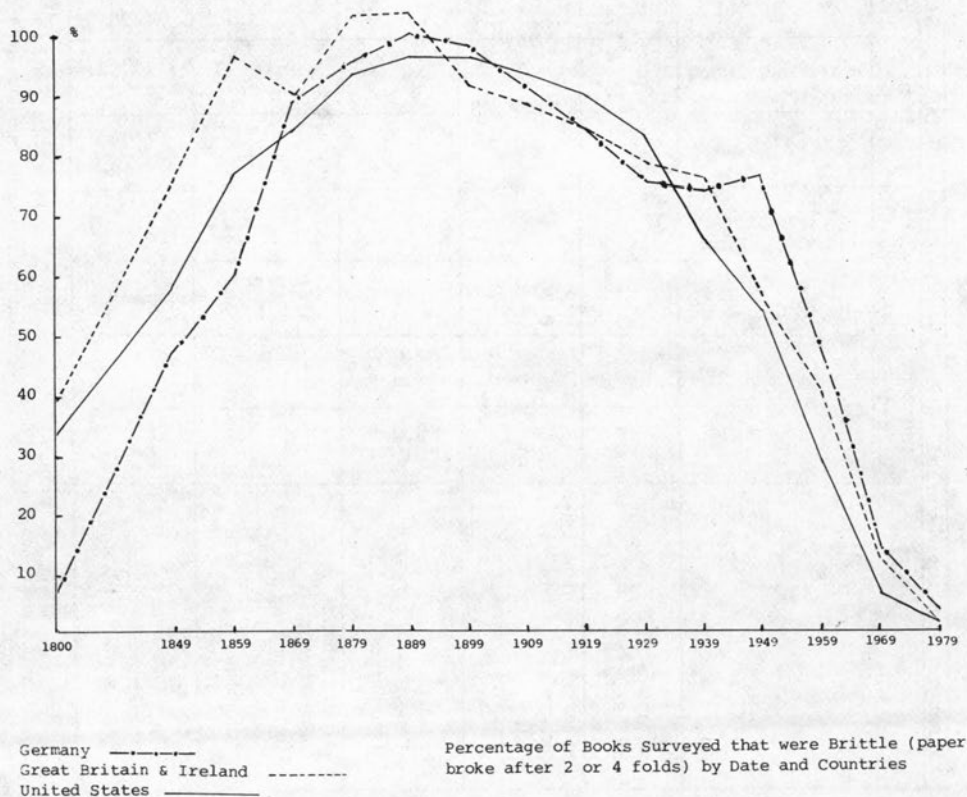


FIGURE 5
Percentage of Books Surveyed that Were Brittle
(Paper Broke after Two or Four Folds) by Date and Countries

conditions. These instruments continuously recorded (onto paper charts) levels of temperature and relative humidity. The hygrothermographs were moved at least once a year so that as many sites as possible could be monitored, but they were left in place at each site long enough to record seasonal extremes. Findings from one site (the top floor of the Sterling stacks) is shown in figure 6. Rapid fluctuations in climate, and a significant deviation from ideal conditions, is apparent. Because of the established link between air pollution and deterioration of paper and other library materials, the U.S. Environmental Protection Agency, Connecticut Air Quality Monitoring Division, was approached regarding pollution in New Haven. It is an urban area and has a bad problem, and though Connecticut is presently (1984) meeting the annual federal standards, for instance, for both sulfur dioxide and nitrous oxide, these pollutants tend to be concentrated in the city streets, including those surrounding the various library units, and there are still many days during

the year when levels of pollutants are so high as to be "unhealthy," a category that does not meet federal health standards.

In addition to the monitoring program and communication with the EPA, one group of NEH interns conducted in-depth environmental studies of each of the fifteen library units surveyed. Using information gathered during site visits and discussions with unit heads, they developed detailed descriptions of each unit and proposed solutions to problems. Although the climate in the five air-conditioned units (the Cross Campus, Engineering, Geology, Kline Science, and Social Science libraries) was fairly good, in general the environment both inside and outside the library buildings at Yale was found to be inhospitable to the storage of library materials.

CONCLUSIONS

The results of the Yale survey provide a detailed description of each of the surveyed collections. This profile includes

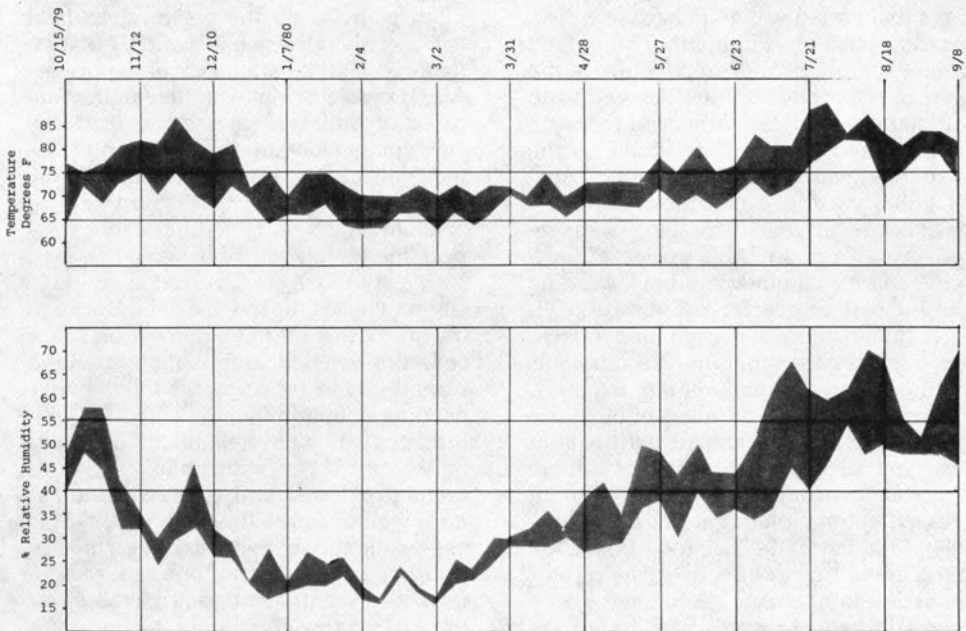


FIGURE 6
Weekly Fluctuations in Temperature and Relative Humidity in SML, Stack Floor 7M
Zeta Collection Hygrothermograph Readings 10/15/79-9/8/80

the physical condition of each unit's holdings (e.g., the number of volumes requiring immediate treatment) and an analysis of the physical composition (e.g., the number of volumes bound in paper) and publishing history (e.g., the number of volumes published in the United States during the 1890s) of each collection. The data are of practical interest to local planners and of more theoretical interest to selection officers, administrators, and students of book publishing and collecting. The amount of statistical information available from such a large survey is tremendous, and other analyses in both areas of physical condition and collection description may be carried out in the future.

A large survey that details the collection's composition and preservation problems is a powerful planning tool. This study documents the preservation needs of the Yale University Library system, and our present efforts must be reviewed in this new light. The survey was large, with a sample of more than 36,500 volumes, quite accurate, and included library collections that varied widely in age, size, use, location, and environment. The overall size of the problem should influence the level of effort and funding devoted to it; budgeting increases over several years can be proposed and justified based on the hard data available from this survey. Specific library units of the fifteen surveyed have been identified as problem areas requiring either onetime projects (e.g., screening for commercial library binding candidates) or greater preservation efforts, for instance, an expanded search-and-replacement program. We can focus on those units as our program expands. We know that large numbers of items are irreparable and threatened with loss of text, and survey results support a more aggressive program—especially in increased control of the storage environment for many of the Yale collections. The most interesting results in terms of program development are the percentages of books in need of immediate treatment (12.8 percent overall), with broken bindings (8.1 percent overall), with brittle pa-

per (37.1 percent overall), and, for future planning, with acidic paper (82.6 percent overall).

Although this survey does not take the place of a local preservation survey that identifies environmental problems or specific candidates for treatment, it may be useful in estimating levels of deterioration at other libraries. The fact that many units of varying natures were surveyed as separate strata should allow libraries elsewhere with similar collections to identify results relevant to their own situations for local planning.

Furthermore, the survey design, the questions, and the implementation procedures described here may be particularly useful as working tools adaptable to different situations and needs at other libraries. The survey methodology was based on random sampling techniques; and the use of mapping, a presampling strategy, tagging, machine-readable forms, consistent surveyor training, and computer analysis increased the accuracy and efficacy of the actual surveying process.

For many years the preservation field has cherished those few statistics that attempt to set the parameters of the preservation problem. Many of these figures are based on educated guesses or small surveys combined with experience and common sense. For instance, the commonly quoted figures for both the Library of Congress and the New York Public Library appear in the introduction here, although both institutions are involved in new survey activities. Our rough estimate prior to the study that 30 to 40 percent of Yale's collection needed preservation attention turned out to be correct if those books needing attention include all books with brittle paper, an overestimate if it means only those books with immediate treatment problems, and conservative if it means all books with acidic paper. This large-scale survey provides one set of hard statistics in a number of different categories that presents a statistical microcosm of the various preservation problems and a sobering picture of book deterioration in a large research library.

REFERENCES

1. Library of Congress, *Preserve* (Washington, D.C.: Library of Congress, 1978?), p.1. Also quoted by Merrily Taylor in the Yale University Library NEH Preservation Grant application, *A Proposal to the National Endowment for the Humanities from the Yale University Library to Survey, Evaluate and Develop Methods of Correcting Preservation Problems in a Large Academic Research Library*, Apr. 1978, p.1.
2. Robert DeCandido, "Preserving Our Library Materials. Part 1: The Nature of Paper and Causes of Its Deterioration," *The Library Scene* 7: 10-12.
3. Warren J. Haas, *Preparation of Detailed Specifications for a National System for the Preservation of Library Materials* (Washington, D.C.: Association of Research Libraries, 1972), p.2.
4. Pamela W. Darling, "A Local Preservation Program: Where to Start," *Library Journal* 101:2343-47 (Nov. 15, 1976).
5. Gay Walker, "Preservation Efforts in Larger U.S. Academic Libraries," *College and Research Libraries* 36:39-44 (Jan. 1975).
6. Sarah Buchanan, and Sandra Coleman, "Deterioration Survey of the Stanford University Libraries Green Library Stack Collection" (Stanford June 1979).

APPENDIX A: SAMPLING
METHODOLOGY AND PROCEDURES

Two statisticians from the Yale Statistics Department served as consultants for the project. They helped to plan and implement all aspects of the survey, carried out the pilot survey, analyzed the data, and wrote the statistical narration. Since time and monetary constraints made it impossible to examine every volume in the Yale system, a random sample representative of the entire collection was constructed. (For a simple random sample, every volume in the target population must have an equal chance of being selected for sampling.)

The basic principles of sampling within the library framework have already been treated in the literature.* However, the structure of the sample needed to examine the books in a large academic library is more complicated than those of the surveys previously conducted. Although the Yale survey, like others, used a random sampling technique, the sampling also took into account particular attributes of each library unit being studied. A stratified sampling design[†] was used whereby the entire library system was divided into strata. Each departmental and area library comprised a different stratum, and reference and reading rooms were often broken away from the main collections because of their peculiar characteristics. The Sterling Memorial Library was subdivided so that each floor, and several of the special units (the Periodical stacks, the Main Reading Room, and the Preservation Division), were studied as separate strata. Thus, the location, environment, reader access, and general level of maintenance within strata were similar, while the characteristics of one stratum could be quite different from those of another.

In order to sample the collections in these strata randomly, a sampling frame was built. A sampling frame is a systematic, usually hierarchical method for giving each member of a target population a unique label, the label usually being a number or a series of numbers.[‡] Our sampling frame was based on floor plans of each stratum showing all stack ranges and the number of sections in each range. Random numbers based on these plans were generated by computer. Each nine-digit code identified a particular book by stratum (two digits), range number (two digits), section number (two digits), shelf number (one digit), and book number (2 digits). All random numbers were sorted in hierarchical order

*See, for example, M. Carl Drott, "Random Sampling, a Tool for Library Research," *College & Research Libraries* 30:119-25 (Mar. 1969); also, Marianne Goldstein and Joseph Sedransk, "Using a Sample Technique to Describe Characteristics of a Collection," *College & Research Libraries* 38:195-202 (May 1977).

[†]See, for example, W. G. Cochran, *Sampling Techniques*, 3d ed. (New York: Wiley, 1977), p.89-146.

[‡]A good primer on sampling concepts and terminology is F. J. Anscombe's "Some Principles of Sampling," unpublished manuscript, 1975.

so that the surveyor could move in a logical progression through a given stack area.

The numbers identifying ranges, sections, and shelves had identifiable maximums based on stack configurations. For example, stratum fourteen may have had forty ranges with no more than twelve sections per range, each section having no more than seven shelves. The number of books on a shelf, however, tended to vary from 0 to as high as 150 in some cases. For reasons of efficiency, an artificial book-per-shelf maximum of 30 was chosen. This was found to be usually as large as, or larger than, the average number of books per shelf in each stratum. The computer-generated random numbers corresponding to book position, then, went no higher than 30. Mapping (i.e., annotating floor plans so that empty ranges and sections were not included in the sampling frame) was used to make the sampling procedure more efficient. So that books with shelf positions exceeding 30 were allowed to enter the sample, a tagging procedure was devised. The sample that resulted from this design was an approximate simple random sample of the target population.

If the random numbers designating book positions went up to, say, 150, the frame would have encompassed the entire target population and resulted in an exact simple random sample of the target population. However, it also would have enlarged the sampling frame to such a degree that the hit rate would have been drastically lowered. That is, only rarely would one find books in the positions identified by the higher random numbers designating book positions, since most shelves contained thirty or fewer items. The enlarged sampling frame would also increase the surveyors' work load and introduce more opportunity for errors.

Tagging (i.e., systematic subsampling) worked in this way: every time a book was sampled, the thirtieth book on the same shelf beyond the one identified by the random number was also sampled (if it existed). For instance, if the book identified by the computer printout was the seventh from the left-hand edge of the shelf, the thirty-seventh, sixty-seventh, and ninety-seventh books were also sampled if present. This procedure ensured that unusually full shelves were not undersampled. We tried to keep the proportion of tagged books in the sample at 10 percent or less as a safeguard against any possible "long-shelf" effects. In some subunits we could do this by adjusting the book-per-shelf limit to a more appropriate value for that unit. Tagged books were identified as such on survey forms in order to track percentages.

Once the frame was completed, we needed to determine the required sample size and how many random numbers to generate. In this survey, the chosen sample size was a function of the desired accuracy. In a simple dichotomous situation, for example when determining the percentage of books with an intact text, the standard error associated with the sample proportion of books with intact text is no greater than $1/(2\sqrt{n})$, where n is sample size, and we can ignore the finite population correction (fpc).^{*} Hence, a sample of size 1,600 would give us an estimate of the proportion of books in the library with intact text and a standard error no greater than 1/80 or 1.25 percent. This would mean that we could be fairly confident about placing the proportion of books with intact text in an interval of the form: sample proportion \pm 2.5 percent. We also needed to select enough books to analyze the condition of library holdings in terms of several variables, e.g., what proportion had brittle paper and also needed repair? In general, this necessitated taking large samples, usually between 1,000 and 2,000 books in a stratum.

Once sample size (n) was determined, we could calculate how many random numbers (k) had to be generated in order to sample the desired number of books by solving $(k)(r) = n$ (where r equals the hit rate). For example, if the sample size required was 1,600 and the hit rate was known to be 50 percent, then one would generate 3,200 numbers inside the sampling frame. Unfortunately the hit rate can only be guessed at beforehand, and it tends to vary from stratum to stratum. An underestimate of the hit rate would produce an unnecessarily large sample, while an overestimate would produce too small a sample. This problem was addressed at Yale by developing a presampling strategy. For each stratum the statisticians generated between 200 and 500 random numbers inside the respective frames. By recording whether a book was to be found at each location specified by the number, and by recording any tagged books that would result, they were able to determine the presample hit rate. This served as an estimate of the hit rate (r) that could be expected in the actual survey sample and was used in the equation $(k)(r) = n$ to determine k . Because presample hit rate is not an exact predictor of r , and because r affects the ultimate size of the sample, the actual sample sizes we observed were close but never equal to the desired sample sizes. The actual sample size usually fell within 50 to 100 books above or below the desired sample size.

^{*}See Cochran, *Sampling Techniques*, p.24-25.

APPENDIX B: SURVEY INSTRUCTIONS

The following instructions are a synthesis of those given to the Yale surveyors.

Supplies

Gather the following materials:

1. support jig
2. heavy paper overlay
3. supply of forms marked with code letter and date
4. #2 pencil
5. pH indicator pen
6. random number printout
7. heavy paper strip to measure gutter margin width

Finding a Book to Evaluate

A computer printout of random numbers gives the location of a book. If the random number does not locate a book, mark a zero after the number. If a book is located, check off the random number. The random numbers appear on the printout in units, e.g., 2 49 8 3 9. Each unit comprises two-digit numbers for the floor, range, section, and book and a one-digit number for the shelf. Enter the number on the form with zeros added where appropriate to look like this:

Floor or	{	0
Library	{	2
Range	{	4
		9
Section	{	0
		8
Shelf		3
Book	{	0
		9

Floor: The first two digits can refer to a floor or can be a number arbitrarily assigned to a specific library subunit.

Range: In some libraries, ranges are numbered, in others they are not. The floor plan used to map out the sampling frame showing the range numbers can be used.

Section: A pattern of movement must be designed for counting sections. Always start from a designated point and move around or along the range in the same direction. The pattern will, of course, change to meet specific conditions.

Shelf: Shelves are counted down from the top.

Book: Books are counted from left to right. In counting, any material tied together is counted as a unit as is material with the same call number. Dummies (boards left in place of books) are not counted.

Filling out the Form

1. Enter the random number on the form.
2. Enter the call number of the book and underline the letters s, n, o, and b to distinguish them from 5, h, zero, and 6, respectively.
3. Enter country and date of publication.
4. Answer questions 1 and 2.
5. Evaluate the book. Answer questions 3-14. For boxes and envelopes answer questions 1-3, 5, and 12-14 only.
6. Put the finished form in the pocket underneath the jig.
7. Fill in ovals for date and country at the end of the survey period.
8. All completed forms should be returned to a central location each day.

Guidelines for Answering Questions

Country: Record country of publication or reprint if given. Record any lack of information in the ques-

tions box. If a city is given but the country is unknown, consult an atlas and/or the shelflist.

Date: Record latest date shown or date of reprint. If no information is given, consult the shelflist. Record any lack of information in the questions box.

1. *Tagged:* Every thirtieth book after the first book located on a shelf is considered tagged.
2. *Circulating:* Noncirculating books are stamped "noncirculating" on the inside of the book cover or are in a special collection. Also, periodicals published within the last ten years are noncirculating.
3. *Primary protection* (outer protective covers):
 - a. *Rigid* includes rigid attached covers and limp vellum bindings.
 - b. *Limp* includes paper bindings (oriental or occidental) and attached paper covers.
 - c. *Acidic pamphlets* are all old ones with cloth tape on spine (pre-1977).
 - d. *Acid-free pamphlets* are all new ones known to be acid-free (post-1977).
 - e. *Acid boxes and envelopes* include all old ones (pre-1977).
 - f. *Acid-free boxes and envelopes* are new and known to be acid-free (post-1977);
 - g. *None* includes unbound materials such as pamphlets without attached covers of any kind. For single units, answer all questions if possible. Although material tied together, with or without boards, is considered a unit, answer questions 1-3, 5, 12-14 only for this material.
 - h. *Other* includes rigid unattached covers such as slipcases and oriental cases.
4. *Covering of joint* (material covering outer hinge): If question 3 is "none" answer N/A.
5. *Primary protection functional:* Not functional includes inner hinge torn more than 25 percent, book block loose enough in case to extend beyond book covers, board broken, and limp binding that does not support text.
6. *Leaf attachment:*
 - a. *Sewn/stapled through the fold:* Look at head of book for rounded gatherings. Look for sewing thread or staples in gutter.
 - b. *Oversewn/cleat sewn:*
Oversewn: Book opens only to sewing thread visible at intervals of about one-fourth of an inch;
Cleat sewn: Adhesive looks like rubber bands visible at intervals of about three-fourths of an inch in gutter.
 - c. *Adhesive:* Backs of gatherings are cut off or sawn in at intervals. This can look somewhat like sewing but no threads are visible.
 - d. *Stabbed:* These usually show three to six holes. Thread or staples are at a right angle to plane of book.
 - e. *None:* Record unlinked materials here.
 - f. *Unknown:* The binding method cannot be ascertained without damaging the book.
 - g. *Other:* This includes spiral, ring binders, accordion bindings.
7. *Leaf attachment intact:* This is self-explanatory.
8. *Brittle paper:* A page is chosen toward the middle of the book, its corner folded four times, and the crease pinched on each fold. Do not test books printed before 1800. Note when the paper breaks after two folds, after four folds, and if it does not break.
9. *pH of paper:* A short line of indicator fluid is drawn in an inner margin. If it turns blue within thirty seconds, the paper has a pH of 5.4 or above. If the mark is not blue (blue-green, green, yellow), the paper has a pH of less than 5.4.
10. *Gutter margin width:* Record narrowest margin visible in brief inspection as less than 1 cm. or as 1 cm. or greater. Use the 1 cm. marked tag to measure where necessary.
11. *Text intact:* Not intact includes tears into text, pages entirely detached, pages missing, and parts of pages missing. Do not include torn blank leaves as not intact.
12. *Mutilation:* Thumb quickly through the text to find leaves cut out, underlining, Scotch tape, food stains, or obvious evidence of mutilation by people, animals, or machines.
13. *Environmental damage:* At the same time look for fading, mold, pest damage, water stains, or charring.
14. *Immediate treatment needed:* Treatment is needed if the primary protection is *not* functional (#5 = No), the leaf attachment is *not* intact (#7 = No), or the text is *not* intact (#11 = No). Do not record cosmetic damage, such as torn headcaps, frayed corners, or loose labels, as needing immediate treatment.

**APPENDIX C: ABBREVIATIONS FOR
FREQUENTLY USED COUNTRIES OF PUBLICATION**

ARGE	Argentina	IREL	Ireland
AUSL	Australia	ISRL	Israel
AUST	Austria	ITAL	Italy
BELG	Belgium	JAPN	Japan
BOLV	Bolivia	LUXG	Luxembourg
BRZL	Brazil	MEXI	Mexico
BULG	Bulgaria	NETH	Netherlands
CANA	Canada	NWZL	New Zealand
CHIL	Chile	NICA	Nicaragua
CHNA	China (People's Republic)	NIRE	Northern Ireland
TAIW	China (Nationalist)	NORW	Norway
CLMB	Colombia	PAKI	Pakistan
CSTR	Costa Rica	PANA	Panama
CUBA	Cuba	PRGY	Paraguay
CZEC	Czechoslovakia	PERU	Peru
DENM	Denmark	PHIL	Philippines
DOMR	Dominican Republic	PORT	Portugal
ECUA	Ecuador	SCOT	Scotland
ENGL	England	SAFR	South Africa
FINL	Finland	SPAN	Spain
FRAN	France	SWDN	Sweden
GERM	Germany	SWTZ	Switzerland
GREC	Greece	TAIW	Taiwan
GUAT	Guatemala	USSR	Russia
HOND	Honduras	USAM	United States
HUNG	Hungary	VNZL	Venezuela

APPENDIX D: STATISTICAL ANALYSIS

The data recorded on the survey forms were scanned electronically and transferred to magnetic tapes. Prior to analysis, the data were cleaned and reordered onto master volumes. From the total target population of five million volumes, more than 36,500 records were gathered from libraries throughout the university. Each record contained information about a single book, including answers to the fourteen preservation-related questions, and five other pieces of information: the place and date of publication, a code identifying the surveyor who evaluated the book, the random number used to locate the book, and the call number. Call numbers were recorded so that the condition of any item could be checked again at a later date.

Our main software tool was the package Table Producing Language (TPL), version 4.0, developed at the Bureau of Labor Statistics, U.S. Department of Labor. TPL is designed for efficient construction of tables from large data sets.

For each of the thirty-six strata, straight tabulations were run that gave raw counts of the number of books that fell into each category for each of the fourteen main questions. Attaching standard errors to these allowed us to produce confidence intervals for the proportions of books in each stratum with the characteristics of interest, e.g., condition of text, embrittlement, mutilation, etc. In all cases where a large unit was broken into subunits for analysis, the data gathered were weighted and grouped together for presentation in the tables reproduced in this paper. We also produced most of the possible two-way tabulations, e.g., the numbers of books that had both brittle paper and a broken binding.

In addition, frequency tables were produced for the date and place of publication of the books surveyed in each stratum. Information about books published in small geographic units was grouped together within larger geographic areas so that sample sizes would be large enough to give reliable results. Dates were grouped into one period from 1801 to 1850, and thereafter by decades. Such batching made it possible to compare critical variables, e.g., embrittlement, across time and place.

In order to estimate the precision of the data gathered, the standard error was calculated for all sur-

vey results. Standard error is a measure of the accuracy to be expected in making statistical estimates, and is calculated thus:

$$SE(\bar{X}) = \sqrt{p(1-p)/n}$$

where \bar{X} is the sample proportion with characteristic X , p is the population proportion with characteristic X , and n is the sample size. Hence, the standard error of \bar{X} decreases as \sqrt{n} increases. Increasing the sample size by a factor of four will halve the standard error of \bar{X} . For example, a sample size of 900 will give a standard error for \bar{X} of

$$\sqrt{p(1-p)/30}$$

which is less than or equal to 1/60 or about 1.67 percent. A more realistic standard error is

$$\sqrt{\bar{X}(1-\bar{X})/n},$$

where \bar{X} is used to estimate p .

Another essential factor in interpreting the survey results was establishing the confidence interval, i.e., our estimate for p and a range for the estimate. The formula for determining the confidence interval is:

$$\bar{X} \pm (z)(\text{standard error of } \bar{X}),^*$$

where \bar{X} is the sample proportion with characteristic X and can be approximated as a normal variate, and z is the appropriate critical value for a normal. A 99 percent confidence interval would be interpreted as follows: if this study were repeated under the same conditions many times and if confidence intervals were constructed in this way each time, 99 percent of the intervals should contain the true value of p , the unknown proportion of books in the population with characteristic X . All results given here are based on a 99 percent confidence interval.

In the Sterling Memorial Library, stratification by floor yielded a wealth of information. The results from each floor were weighted according to the size of that floor in relation to the whole. To obtain an overall summary, we aggregated the weighted results from each of the fifteen floors. The resulting overall frequency estimates and associated standard errors were computed as follows:

Denote the fifteen floor weights as

$$c_1, \dots, c_{15} \text{ where } c_1 + \dots + c_{15} = 1.$$

Suppose the sample frequencies of characteristic X from each floor are

$$\bar{X}_1, \dots, \bar{X}_{15}.$$

Then the overall estimate for the frequency of characteristic \bar{X} in Sterling is given by:

$$\bar{X} = c_1(\bar{X}_1) + \dots + c_{15}(\bar{X}_{15}),$$

and the standard error for \bar{X} is

$$SE(\bar{X}) = \sqrt{c_1^2(\text{var}(\bar{X}_1)) + \dots + c_{15}^2(\text{var}(\bar{X}_{15}))},$$

where var means variance.[†] A similar approach was used to produce the weighted estimates for the entire library system. These appear in the column labeled "Yale Overall" in the third table of figure ^a.

*See W. G. Cochran, *Sampling Techniques* 3d ed. (New York: Wiley, 1977), p.27.

†*Ibid.*, p.92, 107-8.