

THE GEROULD STATISTICS 1907/08-1961/62

Robert E. Molyneux
Second edition, in pdf, slightly revised
2010

An historical compilation of data from academic libraries
in the United States and Canada

Table of Contents

Dedication	
Acknowledgements	p. 3
Preface to the Second Edition	p. 6
Chapter 1—Introduction	p. 9
Chapter 2—Method Used to Compile the Data	p. 22
Chapter 3—The Gerould Data	p. 42
Appendix 1—Data Sources Consulted	p. 54
Appendix 2—Differences Found in the Values of Variables Reported by the Sources of the Gerould Statistics Consulted	p. 69
Appendix 3—Differences in Data Between the Gerould Statistics and the Princeton Compilation	p. 99
Appendix 4—Differences in Data Between the Gerould Statistics and the Purdue Statistics	p. 120
Appendix 5—The Electronic Version of the Gerould Data Bibliography	p. 144
Bibliography	p. 152

**To Kendon Stubbs,
with thanks**

ACKNOWLEDGEMENTS

For the First Edition

1986

Many people have helped me during the course of the research on the Gerould data and I wish to thank them here. They should not be held responsible for errors in the text or in the analysis presented here.

Before thanking those who helped me in this study, it is appropriate to note that this monograph—and what we will learn from these data—would have been impossible without the pioneering work of James Thayer Gerould (1872-1951), who began this series of library data in 1908 and continued to compile it until he retired in 1938. By that time, the series had been established as making a useful contribution to the understanding of academic libraries. Gerould's work is continued today in the annual issues of the *ARL Statistics* and in the *ACRL Statistics*.

I would particularly like to thank Dr. Robert Broadus and Dr. Martin Dillon, both from the School of Library Science at the University of North Carolina, Dr. Melvin Hinich, Frank Erwin Professor of Government at the University of Texas, and Mr. Kendon Stubbs, Associate University Librarian at the University of Virginia. I would never have been able to contemplate a project of this scope nor to complete it without their support and encouragement. They helped in ways too numerous to list.

Ms. Nicola Daval of the Association of Research Libraries read the manuscript of the first edition and had many helpful suggestions. Ms. Shirley Echelman, Executive Director of ARL, and a good reference librarian, ran down a nagging puzzle in the early history of the ARL data. Dr. Edward G. Holley, former Dean of the School of Library Science at the University of North Carolina, apparently taught me more about historical research than either of us might have suspected and was always ready to listen and to offer good advice. My thanks go to them for their help.

Gathering the data turned out to be substantially more difficult than I had anticipated. Ms. Margaret O'Bryant, Ms. Irene Norvelle, Ms. Peggy Holley, and Mrs. Mary Kessler of the Interlibrary Loan Section at the University of Virginia sent requests for the original Gerould data to every library likely to have a copy. Their good spirits in spite of the

meager returns from that effort was appreciated, as was the work of those working in interlibrary loan departments at the libraries receiving requests.

The Illinois collection of the annual issues of the original Gerould data included data from years before the academic year 1919/20. The consensus in the literature had been that 1919/20 was the first year of the series and, because Interlibrary Loan had been exhausted, a new tactic had to be found to locate all the early data. I asked Dr. Haynes McMullen, of the School of Library Science at the University of North Carolina, if he had any idea where to look. He went to his files and pulled out copies of what appears to be all the early years of the data!

Requests for information on Gerould were also sent to archives or manuscripts departments at a number of university libraries. The response I had from my vague inquiries was remarkable because those responding clearly had made my problem their problem and they supplied me with much useful information. I thank the following people for their help: Ms. Nancy Bartlett, Reference Archivist, Bentley Historical Library, University of Michigan; Mr. Maynard Brichford, University Archivist, University of Illinois; Mr. Earle E. Coleman, University Archivist, Seeley G. Mudd Manuscript Library, Princeton University; Penelope Krosch, Archivist, University Archives, University of Minnesota; and Ms. Jean Preston, Curator of Manuscripts, Princeton University Library. I also thank Mr. Michael Plunkett, Associate Curator of Manuscripts, University of Virginia who introduced me to the files of Harry Clemons, former University Librarian at the University of Virginia.

Producing the tables, plots, and the machine readable version of the Gerould data involved using computers a good deal. I had help from Mr. Jeff Bromm and Mr. Bill Niebel both of whom are at the Administrative Computing Center at the University of Virginia; Ms. Patricia Blanton and Ms. Lynne Petterson, of the Institute for Research in the Social Sciences at the University of North Carolina; and Mr. Marvin McCauley of the Computation Center at the University of North Carolina.

Ms. Marie Carter and Ms. Sada Crismond, both of whom work in the General Office of Alderman Library at the University of Virginia, have been of great help during the work on this and other projects.

I would also like to thank the following organizations for their financial support: the Association of Research Libraries, the Council on Library Resources, the School of Library Science at the University of North Carolina, and the University of Virginia. I could not have finished this work without their help.

For the Second Edition, 1998

Many thanks to Cara Thomisser, Ace GA, and especially Michael Cook, for yeoman work encoding the files into HTML, and compiling bibliographies of Fremont Rider and James Thayer Gerould. Li Zhang and Russ Singletary also helped by checking links, and clarifying infelicities of expression that crept into the text. Any errors in the text are there in spite of their hard work.

For the Second Edition, revised, in pdf, 2010

In the 24 years since the first edition, much has changed in the methods for transmitting data. That function, at least, has gotten easier. The main difficulty in this version of the text was the odd error in English—how I regret never having a copy editor for this work.

The Gerould Statistics

1907/08-1961/62

Second Edition, revised

2010

PREFACE

The first edition to *The Gerould Statistics* was published by The Association of Research Libraries (ARL) in 1986. The ARL data had been compiled and documented in a pioneering work by Kendon Stubbs and David Buxton in 1981 and made available in electronic form by ARL on tape. The first edition of *The Gerould Statistics* presented and documented data that were designed to be compatible with the ARL data as documented in Stubbs and Buxton. Thus, data in *The Gerould Statistics* used the data infrastructure such as variable names, institution numbers, *etc.* in the text and in the electronic version of those data. One change in the structure was necessary: to expand the three digit Stubbs-Buxton institution numbers to four digits in order to accommodate the institutions added to the series in the Gerould data. This point is discussed in section 1.3.3 in Chapter 1.

Thus, this compilation is built on the work of Kendon Stubbs and David Buxton. Both editions of this work are dedicated to Kendon Stubbs for many reasons but the most important one was not properly discussed in the first edition. It was a dreadful oversight not to include their *Cumulated ARL University Library Statistics...*, in the discussion of methodology followed in this work. At the time, I was working with Kendon on a day-to-day basis and was not as aware as I am now of the importance of that work, the contribution to the library literature it represents, and the importance of that work to the warp and woof of this one. Sometimes one's perspective improves with time and distance. That error is corrected here with a new Section 2.4 in Chapter 2.

This work also allows me to correct an error in the Acknowledgements that was unfortunate. Maynard Britchford's full title was omitted. Of course, virtually everyone who would have consulted that edition would know that he was the University Archivist at the University of Illinois but he helped so much and deserved better. At the time of the first edition, the discussion of how libraries grow was clouded by Fremont Rider's exponential growth hypothesis. In 1944 in *The Scholar and the Future of*

the Research Library, Rider stated that academic libraries had doubled in size every sixteen years and:

Research library growth has continued, without any significant change of rate, either upward or downward, for over thirty decades, and a rate so uniform over so many years, and so uniform in so many different libraries, that it might almost seem as though some natural law were at work. (Rider, 1944, p. 15-16)

The impetus to compile these data was to use them to test this hypothesis and the opportunity was taken in the first edition to investigate this question with the long run of data then newly available. Chapter Four of the first edition of *The Gerould Statistics*, then, examined the hypothesis with the new data which allowed studying growth over the largest period of years yet available. That chapter consumed the bulk of the book and concluded that Rider's idea that libraries doubled in size as he had predicted was not correct.

The history of Rider's hypothesis can be found in Molyneux 1986a. In addition to the first edition of this work, the hypothesis is tested empirically in: Molyneux 1984, Molyneux 1986b, and Molyneux 1994a. A related article tried to deduce how Rider made his calculations reported in *The Scholar* (Molyneux 1994b.) To the extent they can be inferred from the clues in various sources where he discussed his thesis, the article concluded that Rider must have used an invalid method for analyzing his data and, hence, these data did not show evidence of exponential growth in the data he seemed to have analyzed.

Rider's hypothesis does not survive the analysis cited here so the question must now be regarded as of historical interest only and repeating the analysis would be of no value. Thus, this second edition presents documentation of the data and the methods of compilation, without concerning itself with the question of academic library growth.

This monograph is a *variorum* edition of a set of data. That is, in addition to making these data available, discrepancies in the record of published data had to be resolved. The text and appendices discuss different published numbers for individual institutions' variables, attempt to arrive at set of principles to resolve the discrepancies, and then use those principles to produce a collection of correct data. The work turned out to be a bit more complex, alas. In any case, the problem of resolving these discrepancies is similar to that of resolving the varying texts of Shakespeare or Chaucer to try to arrive at the

correct text of *Henry V* or the *Canterbury Tales* from the many editions of these works. Thus, part of the problem of the compilation was a problem of textual criticism with its own peculiarities.

Since the publication of the first edition, one error¹ that this writer is aware of have been discovered. The second edition and this revision **make no changes in the data** as documented in the first edition.



The Gerould Statistics

Chapter One:

Introduction

The *ARL Statistics* have been published by the Association of Research Libraries² since 1961/62 but are built on an earlier series begun in 1908 by James Thayer Gerould when he was Librarian at the University of Minnesota. Gerould continued to publish the statistics when he moved to Princeton in 1920 and, after he retired in 1938, they were issued by staff members at Princeton. These data were occasionally referred to as the “Princeton” statistics because of their connection with that university.

The purpose of this edition of this monograph is to make these data available *via* the Internet and to serve as a guide to the data.

Librarians have continued to collect and use statistics in spite of objections to the numbers available. Many issues of the annual Gerould data exist which are annotated with handwritten notes made by people calculating such things as average values of variables or ranks of the libraries according to variables reported for a given year. Another demonstration of the popularity of these data is the number of copies of individual issues of these data apparently not made by Gerould—for instance by reordering the data by Volumes Held. Making a copy of these data from the sheet sent by Gerould would have been difficult because the data would have been retyped, not photocopied or sorted in a spreadsheet and printed as we would do today. The fact that the process of making a copy is complicated implies that the benefits of having any rearranged data were correspondingly great because they outweighed the costs of making it.

There are two important implications to having data which go back to the beginning of the 20th century. Analyzing the behavior of data as a time series mitigates many problems with library data. In addition, time series analysis makes it possible to observe trends in library variables, and to relate trends and changes in them to events external to the library. If we can understand the relationship of the behavior of library variables and events in the economy, for instance, we will have taken a step toward predicting the

behavior of library variables, and the libraries they measure. What problems exist with library statistics?

1.1 Criticisms of Library Data

Librarians have been collecting and reporting statistics as far back in time as the libraries in Alexandria and Pergamon. However, while the statistics have been collected, librarians have also been embarrassed by problems with them. Two kinds of problems that arise with library data will be considered by looking at an early article on library statistics and at a letter written to an academic library director.

1.1.1 Adriano Balbi's Prescient *Statistical Essay*

The 1876 *Public Libraries in the United States of America* included a chapter on "Library Reports and Statistics" (Chapter XXXVII, pp. 745-836). Part of this chapter (pp. 746-761) was a study by "the eminent statistician M. Balbi" who had attempted to discover how many volumes were held by Europe's major libraries by comparing published reports. (Subsequent to the publication of the first edition of *The Gerould Statistics* a full translation of Adriano Balbi's 1835 *A Statistical Essay of the Libraries of Vienna and the World* was published. It is a book well worth reading. Quotes here are from *Public Libraries...*)

The tables included in the article show considerable variation in the estimates of numbers of volumes held at these libraries. Balbi was, as he wrote, "disheartened by this surprising disparity of opinion." (p. 756) Balbi speculated about why this disparity existed and had several theories to explain it, among them:

National writers sometimes repeat, without any criticism whatever, the extravagant verbal or written estimates of an unscrupulous librarian, who thinks he adds to the fame of the collection intrusted (*sic*) to his charge by exaggerating the number of volumes it contains. (p.756)

In addition, Balbi observed that the numbers of volumes held by a library do not measure its quality (pp. 745-746), that different libraries count volumes in different ways (p. 757), and that volume counts are not useful in comparing different libraries (pp. 745-746). Balbi's observations have been echoed in numerous articles since then.

It is debatable whether criticisms of library statistics have advanced beyond Balbi in the 160 years since the publication of his *Statistical Essay*.

1.1.2 Uncomfortable Epistle

Another kind of problem that occurs with library statistics is the awkward question arising from a lack of understanding of what is being measured. Consider the letter from Edmund James, President of the University of Illinois to Phineas Windsor, University Librarian. On December 15, 1914, Windsor sent James the latest Gerould statistics. The next day, James wrote back, in part:

I have received the statement of statistics, etc., prepared by Mr. Gerould relating to university libraries. I note that while Illinois only added 29,000 volumes during the year 1913-14 and Yale added 37,000, we spent for the 29,000, \$86,000 while Yale spent only \$34,000. What accounts for this discrepancy? Is Yale getting better bargains, or are we buying more expensive books, or what?

James also makes similar remarks about salaries: those at Illinois seem higher than at comparable institutions. He concludes his letter by saying:

I should like to have you analyze these things and bring them into harmony with the proposed budget of next year and explain the same. Why does it cost us so much more than these other institutions? (James, 1914)

Windsor was probably disconcerted when he received James's letter—as would a library director who received such a letter today. Yet, James's calculation is not legitimate for three reasons, at least. First, volumes added are not volumes purchased, so amounts spent on materials may or may not reflect volumes actually added to the collection because of the influence of other factors such as gifts, discards, and cataloging backlogs. Second, a volume at Illinois may not be the same as a volume at Yale—they may be counted in different ways that make comparing them in any one year impossible. Third, volumes added fluctuate from year to year and perhaps Yale's figure was unusually high and Illinois's unusually low in this particular year.

1.2 Looking at Data Collected Systematically Over Time

In spite of these problems, we can draw useful information out of these data. One approach is to study library variables over a number of years and observe changes. The time series approach has the advantage of mitigating a few of the bad effects of different ways of counting a variable and allowing more appropriate kinds of comparisons. For example, we have to be careful in making comparisons between two libraries because of questions about definitions of terms. If two libraries count volumes differently, it is not clear what we gain by observing that one library reportedly added 70,000 and another

71,000 in a given year. However, if we observe these two libraries over 20 years and note that one had trebled the size of the collection and the other doubled its size, the fact that each counted volumes in different ways would not deter us from making reasonable observations about the relative growth of the two. Observing variables over time also mitigates the effects of year-to-year fluctuations. In the example in James's letter, if we had the perspective of ten years worth of data instead of one, we would have a better basis to make comparisons.

One advantage which we have over Balbi is our statistics. While the ARL statistics are collected in a systematic fashion, with error checking, and published annually, the statistics which Balbi compiled were not collected systematically nor regularly. Balbi summarized the compilations of others. In fact, they were collected from widely different sources, over different years, and compiled using different methods. The inherent value of collecting data in a systematic fashion over a number of years by one source is dealt with elsewhere (Molyneux, 1984, pp. 48-53) but can be summarized here. Compilers of these data make an attempt to obtain uniform and continuous data and while comparability between the data elements is not assured, it is more likely. Also, when new data items are added to the *ARL Statistics*, they tend to fluctuate in the first few years before the fluctuations are damped. This experience seems to indicate that there is more than editorial control from the compiler of the data but also that adjustments are being made by the reporting libraries. Perhaps the libraries collecting data do a better job as they grow more experienced about how to count a given variable. It is also clear that librarians compiling the data discuss the reported in the *ARL Statistics* and other compilations. This author was present at a meeting where one director spoke forcefully to another about new data he thought had been misreported by his colleague's library and which made his library look bad. Next year, both had adjusted their figures. A collection of data published one time has neither the attempt at uniformity nor the advantage gained by the experience of counting and reporting a variable for a few years. In addition to overcoming certain kinds of problems which arise with these data, observing data over time has the advantage of creating a more dynamic picture of libraries. Observing library variables over time can lead to understanding how the libraries measured behave in relationship to changes in their environments. If we understand the nature of this relationship, we will understand more about the libraries and be able to predict future behavior. Clearly, the better the understanding we have of how libraries respond to factors in their environments such as periods of recession or inflation, the better our predictions and the more years of data we have to examine—as

long as these data are compatible—the better that understanding is likely to be. The ARL data began in 1961/62 and using these data from then through today allows us to study a considerable period. Given that the Gerould data go back further, are a predecessor to the ARL data, and deal with so many of the same libraries, these data should be of considerable value in understanding and predicting the behavior of large academic libraries. Obviously, the question of compatibility of the two series is important. We will turn now to a look at this series so we can understand the general structure of the Gerould data.

1.3 An Overview of the Gerould Data, 1907/08—1961/62

The remarks which follow provide a sketch of the structure of the data based on sources collected in the process of compiling them. In this section, there is a discussion of the years these data were issued, the variables reported, the institutions reported, and the relationship of the Gerould data to the ARL data.

1.3.1 Years the Gerould Data Were Issued

Issues of the data were found for 1907/08, 1909/10, 1911/12 through 1942/43, and 1944/45 through 1961/62. The data were issued on single sheets of paper produced by two major means discussed below. Occasional fugitive *errata* appear on postcards. Note: the last year of the Gerould data overlaps the first year of the ARL data, as discussed below. It appears that the original intention was to issue the data every two years, or so it might be deduced from a 1913 letter from Gerould soliciting information for the next compilation. This letter says, in part:

“I have been requested by a number of libraries to compile and publish annually rather than biennially the statistics of university and college libraries.” (Gerould, 1913)

The statistics solicited in this letter would have been for 1912/13 so it seems reasonable to conclude that previous issues of the statistics were biennial and that there would have been no data for 1908/09 and 1910/11. In addition, there were probably no statistics for 1905/06. Gerould gave a talk at the 1906 ALA meeting at Narragansett Pier entitled “*A Plan for the Compilation of Comparative University and College Library Statistics.*” The printed version of this talk appeared in the November, 1906 *Library Journal*³ and is transcribed for this publication beginning on page 159. In this version, it is reported that a committee was appointed to consider Gerould's recommendations about collecting statistics. (Gerould, 1906 p. 763) There is no record in *Library Journal* or any other source consulted that the committee ever reported.

It seems likely that there was no issue of the data for 1905/06, given that the 1905/06 data would have been collected in the fall of 1906, and that Gerould probably would have allowed time for the committee to report before undertaking the compilation of the data himself.

After 1911/12, the data were issued through 1942/43. 1943/44 was apparently never issued. A letter from the secretary of Julian Boyd, Librarian at Princeton, to Harry Clemons, Librarian at the University of Virginia blames “the war and the shortage of paper” for there being no issue for this year. (Boyd, 1945) These data did finally appear in the 1947 *College and University Library Statistics, 1919/20 to 1943/44* which was a compilation of data from the first 25 years they were issued at Princeton. (Princeton, 1947) The data from 1943/44 which appear in this edition of the Gerould data come from the 1947 compilation. This compilation will be referred to a great deal and in order to avoid confusing it with the annual sheets of data, it will be called the “*Princeton Compilation*” while the annual issues will be referred to as the “Gerould statistics” through their end and the beginning of the *ARL Statistics* in 1961/62—even those issued after Gerould retired in 1938.

1.3.2 Variables Reported in the Gerould Data

The first issue of the Gerould data reports volumes held, volumes added, the dollar expenditures for materials and for salaries, and the number of staff members. The next issue, 1909/10, adds “appropriations,” which is the amount available for materials expenditures in the current fiscal year. The amount paid in wages was added in 1949/50. Appropriations were dropped in 1953/54, the same year that “total expenditures” were added. Total expenditures were defined as the sum of materials expenditures and the expenditures for salaries and wages.

It is argued in section 1.3.4 that the Gerould series is the foundation upon which the Association of Research Libraries built its statistical series. If we consider these series together, we can see a pattern of change in what has been collected. The Gerould data report expenditures for materials and binding while the *ARL Statistics* disaggregate binding expenditures from materials expenditures and report each separately. The volumes added figure reported by Gerould is a gross figure which does not take into account discards while ARL now reports “volumes added, net” and “volumes added, gross.” ARL also added “other operating expenditures,” which have grown so dramatically subsequently (Stubbs, 1985). In addition to the process of disaggregating data into its constituent elements, items have been dropped (for example, Gerould

reports current appropriations for materials) and others added (for example, interlibrary loan figures). Over the years these statistical series have adapted and changed as requirements for data have changed. This process continues today as ARL considers collection of different data items which may be of use to its member libraries.

1.3.3 Libraries Included in the Gerould Data

The institutions which have data reported in any year of the Gerould statistics are given in Table 1, beginning on page 19 with the range of years each reported. There are 60 institutions in this list. A number of these institutions, like Berkeley, report data for every year and two, McGill and Toronto, report data for only two years. There are eleven libraries which are not currently members of ARL in this list (Bryn Mawr, Joint University Libraries, Mt. Holyoke, North Dakota, Oberlin, St. Louis, Smith, South Dakota, Vassar, Wellesley, and Wesleyan). St. Louis was a member of ARL and reported data from 1963 through 1973. The Joint University Libraries have disbanded although one of its parts, Vanderbilt, is an ARL member. Western Reserve merged with Case Institute to form Case-Western Reserve, another ARL member.

There are eight institutions which did not report data for all years in the range of years given. Appendix 5 includes a list of all such missing years because blank lines are inserted in the data for these years.

The institution numbers in the left hand column of Table 1 are necessary in using the machine readable data. For most uses of electronic data, the institution number is easier to use than the institution name. With the forms of the names of the institutions used here, the alphabetical arrangement used in Table 1 results in the same order as that for these institution numbers. Occasionally, however, when space is a consideration, abbreviations are used (for example, "UCLA" for "California, Los Angeles") but the order used here is used throughout the tables and electronic data.

The institution numbers used here are based upon those used in the ARL data to facilitate splicing the Gerould and ARL series together. Three-digit institution numbers were a signal contribution developed by Stubbs and Buxton for the ARL data and are listed in the *Cumulated fARL University Library Statistics, 1962-63 through 1978-79*, which served as the guide to the ARL data in electronic form, among other things. (Stubbs and Buxton, 1981)

Subsequent years of this series added new institution numbers as new institutions joined ARL. By the time of the initial compilation of the Gerould data in 1986, the three-digit numbers had to be expanded to accommodate the added Gerould institutions. In order to give the non-ARL libraries numbers, it was necessary to expand the Stubbs-Buxton numbering system, which was done by multiplying the ARL numbers by 10. All subsequent compilations have followed this custom.

There is one slight exception to this general rule that occurs with the institution numbers of St. Louis (which does not report data after 1973/74) and Saskatchewan (which first reports data in 1979/80). These institutions were inadvertently given the same number (735) for tapes produced from 1981 through 1985, when Saskatchewan's number was changed to 737. There may be copies of these data with this problem uncorrected. Today, of course, St. Louis has the institution number of 7350 in the Gerould data while Saskatchewan's number in the ARL data is 7370.

In addition to the ARL data, compilations of data from the Association of College and Research Libraries (ACRL), a division of the American Library Association, used the same infrastructure with the Stubbs-Buxton institution numbers, with the addition of the fourth digit. This latter ACRL series (there was an earlier one that ended in the 1960s) was based on the ARL series, using the same forms and definitions. The result is that even though the data are compiled by different agencies, a great deal of effort went into making them compatible so that the series of the data could be analyzed together. The related publications from ACRL are Molyneux, 1989 and the compilation of data from the libraries at Historically Black Colleges and Universities (Molyneux, 1991.)

In 1990, Stubbs and Molyneux compiled *Research Library Statistics, 1907/08 Through 1987/88, A Guide to the Machine-Readable Version of the Gerould and ARL Statistics*, to the Gerould/ARL data that included the Gerould libraries which were members of ARL at the time of the compilation. This publication does not discuss details of the compilation of the Gerould data dealt with here.

Appendix 5 lists more recent sources of documentation to these data.

1.3.4 Relationship of the Gerould and ARL Data Series

There are two pieces of evidence which lead to the conclusion that the Gerould data precede the ARL data.

In 1961/62, the two sets of data report the same variables for the same libraries with no discrepancies between the two sets of numbers. This fact indicates that the two series used common definitions and means of collecting data.

The minutes of the sixth meeting of the ARL Board of Directors record that “It was agreed that the Secretariat should assume responsibility for publishing ARL statistics in lieu of the ‘Princeton Statistics’.” (ARL Minutes, January 28, 1963, item 8) How this decision was arrived at is not clear but there are two facts which probably contributed to the decision. After Gerould's retirement in 1938, the person at Princeton who seems to have had the most to do with the compilation was Lawrence Heyl who retired in 1962. If the series had been continued at Princeton, someone else would have had to do the work. The Librarian at Princeton at the time was William Dix who would have made the decision about who was to do the statistics after Heyl's retirement. Dix was also Chairman of ARL. It seems reasonable to speculate that there is a connection between these facts.

A good case can be made that these two series are related but this fact does not mean that splicing the two series of data will yield useful results. The question of the comparability of the two will be dealt with in passing below.

1.4 Organization of This Study

There are three chapters, five appendices, and a bibliography. Chapter 1 includes these introductory matters.

Chapter 2 discusses the methodology used in compiling these data. The Princeton Compilation and the data compiled for the Purdue studies (Dunn, 1965) are discussed in order to show flaws in the approach of both that were taken into account here. The study of (Stubbs and Buxton, 1981) is the one to which this compilation owes the greatest debt and the methodology developed here followed from that work as well as incorporating the infrastructure developed there.

Appendices 1 through 4 are a part of the methodological discussion. Appendix 1 lists the different issues of the Gerould data consulted for this edition and Appendix 2 lists variant values found between these different issues. Appendix 3 lists differences between the values reported by the Princeton Compilation and this edition of the Gerould data while Appendix 4 lists differences between the Gerould and Purdue data.

Chapter 3 is largely made up of tables of the data, arranged by institution and then by year but it begins with a discussion of the variables reported by Gerould with a comparison of the definitions of the Gerould variables and the related ARL variables. Comparability in the end rests on whether the two series can be joined to produce reasonable conclusions. In the First Edition of this work Chapter 4 examined comparability by an experiment where the two sets of data were joined in order to examine the question of academic library growth. As was mentioned above, this section is not repeated here. Those wishing to see this test of comparability can see that monograph. Another examination of the data over more years can be found in Molyneux 1994a.

Appendix 5 contains the documentation for the electronic version of the Gerould data and the data in the Princeton Compilation.

All references in the text can be found in the bibliography.

In these tables and in the electronic editions of the data, blank lines have been added. There are no data for 1908/09 and 1910/11 but institutions which report data before either of those years have a blank line added for the missing year. There are eight institutions (Colorado, Dartmouth, Harvard, Nebraska, Northwestern, Rutgers, Texas, and Wisconsin) which did not report in all years of the Gerould data and lines with the year but no data have also been added for them. This practice was adopted for two reasons, first, to provide a visual cue that data were missing and, second, to make certain data manipulations easier.

These data are produced from the same electronic file which has been used to produce the files of data which are discussed in more detail in Appendix 5. The data are in order by the institution number, which is found in the upper right hand corner of each of the tables in this section and is listed in Table 1 itself which begins on page 19. Missing values are indicated in the data tables by a period ("."). Individual values discussed in Appendix 2, where variant values from different texts are given, are indicated either by a question mark ("?"), when this edition records no value, or by a link to variant texts when there is a value used here. Part 1 of Appendix 2 discusses cases where there is more than one possible value and attempts to decide between these values using textual authority. Chapter 2 deals with textual authority as used in compiling the Gerould data.

Table 1**Institutions Reporting Data in the Gerould Statistics**

Institution Number	Institution	Year Available
800	Brown	1911/12—1961/62
830	Bryn Mawr	1923/24—1950/51
900	California, Berkeley	1907/08—1961/62
1100	California, Los Angeles	1929/30—1961/62
1600	Chicago	1911/12—1961/62
1700	Cincinnati	1940/41—1961/62
1800	Colorado	1917/18—1961/62*
2000	Columbia	1911/12—1961/62
2200	Cornell	1911/12—1961/62
2300	Dartmouth	1913/14—1950/51*
2400	Duke	1928/29—1961/62
2600	Florida	1956/57—1961/62
3100	Harvard	1913/14—1961/62*
3500	Illinois	1907/08—1961/62
3600	Indiana	1907/08—1961/62
3700	Iowa	1907/08—1961/62
3800	Iowa State	1927/28—1961/62
3900	Johns Hopkins	1911/12—1961/62
4000	Joint University	1936/37—1961/62
4100	Kansas	1907/08—1961/62
4300	Kentucky	1951/52—1961/62
4400	Louisiana State	1938/39—1961/62
4500	McGill	1949/50—1950/51

4900	Massachusetts Institute of Technology	1938/39—1961/62
5100	Michigan	1907/08—1961/62
5200	Michigan State	1957/58—1961/62
5300	Minnesota	1907/08—1961/62
5400	Missouri	1907/08—1961/62
5490	Mt. Holyoke	1940/41—1950/51
5500	Nebraska	1907/08—1961/62*
5700	New York University	1935/36—1961/62
5800	North Carolina	1921/22—1961/62
5860	North Dakota	1907/08—1950/51
5900	Northwestern	1912/13—1961/62*
6030	Oberlin	1913/14—1950/51
6100	Ohio State	1907/08—1961/62
6400	Oregon	1921/22—1950/51
6500	Pennsylvania	1911/12—1961/62
6700	Pittsburgh	1942/43—1950/51
6800	Princeton	1911/12—1961/62
6900	Purdue	1956/57—1961/62
7200	Rochester	1925/26—1961/62
7300	Rutgers	1940/41—1961/62*
7350	St. Louis	1942/43—1950/51
7390	Smith	1923/24—1950/51
7430	South Dakota	1907/08—1924/25
7700	Stanford	1911/12—1961/62
8200	Temple	1940/41—1950/51
8400	Texas	1913/14—1961/62*
8600	Toronto	1949/50—1950/51

8860	Vassar	1923/24—1950/51
8900	Virginia	1921/22—1961/62
9100	Washington (Seattle)	1907/08—1961/62
9200	Washington State	1942/43—1950/51
9300	Washington U.—St. Louis	1922/23—1961/62
9420	Wellesley	1925/26—1950/51
9440	Weslyan	1946/47—1950/51
9520	Western Reserve	1942/43—1950/51
9600	Wisconsin	1907/08—1961/62*
9700	Yale	1911/12—1961/62

* Indicates that data are not reported for all years in the range.



The Gerould Statistics

Chapter Two:

Method Used to Compile the Gerould Statistics

2.1 Overview

This chapter will present the methodology used to assemble this edition of the Gerould data from the annual sheets of data issued by Gerould and the Princeton University library from 1907/08 through 1961/62. The data themselves are given in Chapter 3 and are available in the files documented in Appendix 5. The principles followed in compiling the present edition were derived after a consideration of characteristics of the Gerould data and of three other compilations of library data: an edition of the Gerould data done at Princeton and the data used in the studies entitled *The Past and Likely Future of 58 Research Libraries* done at Purdue beginning in 1965, and Stubbs and Buxton's 1981 *Cumulated ARL University Library Statistics...* The latter provides many of the conceptual underpinnings as well as providing the infrastructure upon which this compilation is built. It is a model for any data compilation.

As in Chapter 1, the terms "Gerould data" or "Gerould statistics" are used here to mean the annual sheets of academic library data issued both by Gerould from 1907/08 through 1938/39 and by staff members of the Princeton library through 1961/62 and their *errata*; "Princeton Compilation" refers to the compilation of the annual data for the years 1919/20 through 1943/44 done at Princeton in 1947.

Before dealing with the Princeton and Purdue series, factors which should be considered in compiling data will be discussed briefly in order to make the issues involved clearer.

2.2 Compiling Data

Compiling data means collecting and assembling them, for publication or use. Statistics from libraries should reflect the truth, so that an institution reporting a given number of volumes should report the correct number. However, there is likely to be uncertainty in measuring library variables, with the amount of uncertainty depending on the characteristics of the variable. Consider the variable "Volumes Held." This variable is usually a large number and one collected over a many years by summing

volumes added with the previous year's reported number of volumes held. Recounts are rare. These facts mean that errors tend to be perpetuated. On the other hand, the variables measuring staffing measure a relatively smaller number and are collected each year so that errors are not continued year after year. In addition, this variable measures the number of people employed. These people have names and characteristics and the numbers are more closely audited than volumes held—for example as a part of tax records. Staff should, therefore, have relatively less measurement error than volumes held.

The Association of Research Libraries (ARL) compiles statistics on its member libraries each year and it edits the data because there are errors which are introduced by the process of compilation itself either by the reporting institution or in the publication of the *Statistics*. Errors can sometimes be detected. For instance, if the volumes added at an institution were to increase or decrease by an inordinately large amount, an inquiry can be made about whether the correct number has been reported. More sophisticated means of detecting possible errors, which take into account characteristics of variables and their relationships with other variables, have been developed by ARL. Of course, each year where errors are caught by reporting institutions after the publication of the *ARL Statistics* and these errors are reported by errata or in the *ARL Newsletter* and corrected in the electronic version of the data.

ARL can make inquiries about possible errors because the people who collect and report the statistics annually at the libraries are still there. Detecting possible errors is relatively easy but what happens if an apparent error occurs with a variable reported ten or twenty years ago? A problem faced in compiling this edition of the Gerould data is similar to that faced in the Princeton and Purdue compilations: the data were published years ago so that apparent errors in measurement or reporting cannot be checked with anyone who knows what the facts are. In the case of suspected errors which are more than a few years old, the standards for accepting corrections should be stringent in order to avoid introducing errors which might bias conclusions arrived at using the data in unknown ways. Changes in data which bias results will be discussed below.

2.3 The Princeton and Purdue Compilations of Academic Library Data

This discussion begins with two major compilations of previously existing library data that are examined here to discover the methodologies which have been used to compile library data previously. Unfortunately, the methodological notes made by the compilers

of these series are sparse and the methodologies have to be divined by considering what was said about methodology and what was actually done with the data. The means apparently used to compile these series, to the extent they can be understood, are not adequate for emulation in compiling this edition of the Gerould data. These two compilations are dealt with in detail below. In the years since the first edition of *The Gerould Statistics* no new information on either of these two series has been published. Section 2.4 discusses the Stubbs and Buxton compilation which was the foundation upon which this work was built.

As mentioned, the two series of data examined here for an understanding of methodology are those done at Princeton and at Purdue for a series of studies entitled *The Past and Likely Future of 58 Research Libraries*. The Princeton Compilation was a printed edition of twenty-five years of the Gerould data rearranged to provide data arranged by institution and then in order by year. The Purdue data were assembled from a variety of sources but never formerly published nor documented, although they were made available in a deck of computer cards. The only documentation, such as it is, was in the first edition of this work and it is reprinted here. This documentation deals only with Purdue data that overlap the Gerould data.

There are three kinds of problems which arise with both of these compilations: There are apparent mistranscriptions of data and other minor errors, such as apparent errors in rounding. These kinds of errors are generally not serious not only because they are usually minor—the rounding errors differ from the correct values by only 1—but also because they are apparently not systematic. If the errors are random or if, in a few cases, the mistranscriptions result in large differences from the true values, analyzing the data as a time series will generally allow seeing the trends in the data accurately.

The methodologies used should have been explicitly stated so that it would be possible to know with certainty what was done with the data. Stating the methodology allows others to judge whether the steps taken were proper and to compensate for those steps if they were not appropriate to any analysis at hand. But there is virtually no methodological discussion in the Princeton Compilation and little in the Purdue studies so that we often cannot say unequivocally what was done with the data.

The compilers of the Purdue data explicitly stated that they interpolated data, and the compilers of the Princeton data apparently did also. For reasons given below in passing,

interpolation is not an acceptable practice in compiling data. Because there is little methodological discussion in either it is impossible to compensate for interpolated data or even to identify them.

2.3.1 The Princeton Compilation of the Gerould Data

The data collected and published from 1919/20 through 1943/44 first by Gerould and, after Gerould's retirement, by the staff at Princeton were reissued in an edited form as *College and University Library Statistics*. (Princeton, 1947) This edition appears to be a reproduction of a typescript—that is, the original appears to have been typed.

The Foreword includes a few remarks by Lawrence Heyl, who was Associate Librarian at Princeton, about how the data were assembled for that edition. The data had originally been issued each year, of course, but the *Compilation* is arranged by institution. When the compilers rearranged the data by institution they noted what Heyl called “inconsistencies” from year to year. Letters were sent to various libraries (which ones are not recorded) asking for clarification. Copies of an exchange of such letters between Heyl and Harry Clemons, Librarian at the University of Virginia are at its Alderman Library. (Clemons, 1946)

In his first letter to Clemons, February 6, 1946, Heyl says:

...we are trying to clear up a few questions concerning former years. In your case we do not have a figure for Volumes Added for the year 1921/22. For Expenditures for 1933/34 we have the figure 40,264, although the year before it was 27,827 and the year after it was 29,766. Will you also please check the figures for salaries for the following years:

1941/42	\$81,089
1942/43	70,192
1943/44	93,150

On February 9, 1946, Clemons replied to Heyl. He had been unable to locate figures for volumes added in 1921/22 but the expenditure figures were correct. He also noted that on the salaries, a mistake had been made:

...those for 1942/43 and 1943/44 include full time Staff only while that for 1941/42 includes both full time and part time Staff. Our records are as follows:

Year	1941/42	1942/43	1943/44
Full time	\$78,509.42	\$70,191.84	\$93,150.39
Part time	2,580.00	4,737.50	550.00
Total	81,089.42	74,929.34	93,700.39
Student	11,265.74	12,345.87	11,988.25
Grand total	\$92,353.16	\$87,275.21	\$105,688.64

It appears from Clemons's reply that the value for 1941/42 was calculated in a different fashion from the others. In the Princeton compilation the figure for salaries in 1941/42 is given as \$78,509. The other two values questioned by Heyl are unchanged.

However, two changes in the *Compilation* not discussed in this exchange of letters are made to values from Virginia. It can be seen in Appendix 3, which lists differences between the values reported in the edition of the Gerould data presented here and those reported in the *Compilation*, that Virginia's volumes added for 1933/34 appear in the annual issue as 16,566 while in the *Compilation* this number is given as 16,586; in 1936/37 expenditures are given in the annual issue as \$29,006.74 and in the *Compilation* as \$29,807. While these changes may have been made as a result of a second set of letters between Heyl and Clemons which has not been found, there is reason to be skeptical of this explanation. It seems more likely that errors were introduced in the text of these data during the process of compiling the data. This value is listed in Appendix 2, Part 2 as one which is not used but which may be correct.

Although it is impossible to be sure the *Compilation* introduces errors, it seems likely that many of the differences between the values reported in the annual issues of these data and the values reported in the *Compilation* are the result of inadvertent clerical mistakes in the latter's execution. This judgment is based on circumstantial evidence which is reviewed next in section 2.3.1.1 immediately below. Intentional changes in the data will be examined in section 2.3.1.2. As a result of the evidence presented in these sections, it will be seen that the *Compilation* is flawed.

2.3.1.1 Clerical Errors in the Princeton Compilation

2.3.1.1.1 Rounding Errors

The *Princeton Compilation* reports no fractional values (except for a few cases in the number of staff members) so there are no values to the right of decimal points in the *Compilation* while there are in the most of the annual issues of these data. There are many cases (on the order of several hundred) where there are apparent errors in rounding the values from the annual issues in the *Compilation*. For example, the 1931/32 annual issue reports Brown appropriated \$64,067.65 while the *Compilation* reports \$64,067. The amount expended, however, is correct: \$68,567.58 is rounded to \$68,568.

There is no readily discernible pattern to these many apparent errors. Oddly, apparent rounding errors are also seen in Rider's data (Molyneux, 1986a) and in the Purdue data (see section 2.3.2).

2.3.1.1.2 Transcription Errors

The change in Virginia's volumes added for 1934/35 from 16,566 in the annual issue to 16,586 in the *Compilation* is probably a transcription error. The difference between two numbers is small; they differ by only one digit; and, it is doubtful whether any method used to discover inconsistent values would pick out this year's volumes added for scrutiny (1933/34's volumes added is given as 16,846; 1935/36's as 16,693). On the other hand, it is relatively easy to imagine the difference between these two numbers is a result of typing an "8" instead of a "6" one place. The change in materials expenditures may also be a transcription error. The original value, \$29,006.74, should have been rounded to \$29,007 which again could have been typed as \$29,807 by mistyping one digit.

In an attempt to understand what might have happened, it is worthwhile to consider how this *Compilation* was done. Data in electronic form can be rearranged accurately and easily in a fraction of a second—a fact which can lead us to overlook how difficult it would have been to take data arranged by year and rearrange them by institution without the aid of computers. It can only be speculated about how the process was carried out. The data were probably copied from the individual annual sheets to some kind of work sheet and then typed in a final copy. In such a process, there are many opportunities for mistranscription. (Appendix 3 displays many apparent mistranscriptions.)

2.3.1.1.3 Other Apparent Clerical Errors

There are cases where the discrepancies between the values reported in the annual issues of the Gerould data and the Compilation show a more subtle relationship. Consider these values for volumes added for Washington University at St. Louis as given in the annual issues of the Gerould data and in the Princeton Compilation:

Year	Gerould	Princeton
1926/27	16,584	16,584
1927/28	7,420	16,584
1928/29	8,934	7,420
1929/30	6,256	8,934
1930/31	12,322	6,256
1931/32	12,322	12,322
1932/33	7,445	9,310
1933/34	4,998	7,445
1934/35	7,840	4,998
1935/36	7,426	7,840
1936/37	7,077	7,426
1937/38	6,288	7,077

For most years, the values reported in the annual issues are the same as those reported for the next year in the Compilation. When the values were reported originally, it would have been impossible to report in a given year the number of volumes which would be added in the next so the values in the Compilation which repeat the previous year's values in the annual issues must be wrong. A clerical mistake was probably made in transcribing the data.

Of course, the values reported in the issues of the Gerould data for 1930/31 and 1931/32 are both not likely to be right. That is, it is unlikely that the same number of volumes was added two years in a row and this fact is as obvious now as it would have been in 1946. For this reason, it may be the change in the *Compilation* was intentional and should

be considered with the intentional changes below. However, a change like the one made here, which is logically impossible, could not be the result of a consideration of documents and records at this library. Intentional changes, though, can be a result of a process like that followed by Clemons with the salary figures or for expediency. An expedient change is one made to make the numbers consistent according to some ideal without regard to logic or truth. This kind of change would be inadvisable because, while consistency is a reasonable criterion to use in trying to spot possible wrong numbers with a view to writing letters rectifying errors it is not a sufficient criterion for deciding which numbers should be changed. This point is discussed further below. In any case, what is to be made of the Princeton figures for 1926/27 and 1927/28? It seems inescapable that these numbers can't both be right—any more than the two Gerould values for 1930/31 and 1931/32—and there is an error introduced in the Princeton data at this point.

Apparent clerical errors like those described introduce changes in the text of the data presented in the annual issues. Other changes were intentional—the editorial changes which resulted from exchanges of letters like that between Heyl and Clemons mentioned above. Two of these changes will be discussed next.

2.3.1.2 Intentional Changes

It is impossible to be sure that apparent intentional changes are not the result of clerical errors but two changes made in the data between the annual issues and the Compilation, one minor and one major, do not seem accidental.

A minor change is shown in the figures for volumes held for Indiana from 1929/30 through 1931/32 as given in the Gerould data and the Princeton Compilation:

Year	Gerould	Princeton
1929/30	218,344	218,344
1930/31	218,800	232,800
1931/32	247,320	247,320

The Gerould figure for 1930/31 seems unlikely and it is easy to understand how such an “inconsistency” could have come to an editor's attention, resulted in a letter such as that Heyl wrote to Clemons, and a correction. However, there is no documentation of such a

letter in the *Princeton Compilation*. Which value should a compiler of data today use? One which exists in a known edition or text or one which might be a result of a clerical error?

Changes occur between the Gerould and Princeton data in the values shown for Illinois for all variables but staff members over all years of the data. The variable recording the number of volumes held will be focused on here because it seems to be the key to all the changes. Bear in mind that the annual issue from 1943/44 did not appear (as a result of paper shortages, Boyd, Julian (1945)) so this edition of the Gerould data uses the data from the Princeton Compilation for this year. Consider the following sample of volumes held, according to the issues of the Gerould statistics and the Princeton Compilation:

Year	Gerould	Princeton
1919/20	440,372	549,873
-	-	-
1929/30	841,395	1,006,900
-	-	-
1939/40	1,209,977	1,618,579
1940/41	1,262,046	1,687,847
1941/42	1,306,561	1,765,203
1942/43	1,364,906	1,834,437
1943/44	1,881,432	1,881,432

The few years listed above give an indication of a systematic change in the values for volumes held for each year of the data from Illinois. There is a similar systematic change in the number of volumes added by year.

The big change in volumes held occurs between 1942/43 and 1943/44. We know from the letter from Heyl to Clemons that work on the Compilation was going on in early 1946. Robert B. Downs became librarian at Illinois in 1943. Shortly thereafter, Downs argued that the basis of the volume count should be the bibliographic, not the physical volume. (Downs, 1946; see also Kuhlman, 1960) One effect of this change, arguably, was the dramatic leap in volume count we see.

It is reasonable to speculate that someone decided to make the data more consistent because it seems unlikely that anyone at Illinois recorded the number of bibliographic volumes from 1919/20 until 1943/44, so the numbers were probably interpolated, that is, made up. How it was done is not clear and given that these events occurred almost forty years before the first edition of this work, it could be argued that it was not surprising Professor Downs had no recollection of how the changes were made. (Downs, 1985)

Professor Downs died after the publication of *The Gerould Statistics* and it now seems permissible for aficionados of academic library data to share with others the open speculation that the change in numbers between the two series was the result of a deliberate attempt to boost the volume count at Illinois in the 1940s.

2.3.1.3 The Princeton Compilation: An Assessment

There are fundamental and serious problems with interpolating or making any other kinds of expedient changes in an edition of data. If an incorrect value can be documented, it should be changed, not because it is inconsistent according to some theory, but because it is proven to be wrong and a better value is known. Consistency is a reasonable criterion for examining data for anomalies, but not for changing values; numbers should not be changed unless the changes are documented authoritatively. There are several reasons for advancing this argument:

Until we are certain that we understand the behavior of library variables well enough to know what is consistent behavior and what is not, we should be careful about changing data.

Future research is limited when numbers are altered because potentially useful information is thrown away. Techniques may be developed in the future which will allow others to use data not useable today. We should not use theories to derive numbers; we should derive our theories from numbers.

Consider a result of interpolation. To interpolate data there must be an implicit idea of how the data should be arranged. For example, in 1946 the best evidence about how academic libraries grew indicated they grew exponentially. (Molyneux, 1986a) If data on volumes or volumes added were to be interpolated, this theory of the nature of library growth might be the basis for filling in the missing numbers. The evidence is now unequivocal that library growth is not an exponential phenomenon but subsequent

research on the nature of growth using interpolated numbers would, naturally, discover an exponential series of numbers. This research would conclude growth was exponential (assuming this pattern was the basis of the interpolation) without being aware the results were biased in the editorial process.

The changes in volumes held at Indiana in 1930/31, and those for Illinois also, may be the result of as careful checking the records as Clemons's check of the Virginia records but we do not know, and from what has been shown here about the *Compilation* there is ample reason to doubt whether such care was exercised. Because the suspicion that expedient changes have been made which could bias results and because we cannot distinguish between these kinds of changes or clerical and documented changes, it was concluded the *Compilation* could not be used as a basis for this edition of these data except for 1943/44, year the annual Gerould data were not issued. Unfortunately, in not using the changes in the *Compilation* we will not correct values in the annual issues which were at one time known to be wrong. The cost of not documenting data and changes in them is high.

One problem for compilers of library data and those who use them is that it is relatively easy for fault finders to point to cases of inconsistency—or, as is usually done on articles dealing with library data to repeat the observations of Adriano Balbi limned above in Section 1.1.1. (starting on page 10)—without attribution and without an understanding of the issues involved. It is in the nature of data that there are inconsistencies, missing values, clearly wrong values, dubious values, and the like.

In analyzing data, the problems must be dealt with—which is why we “weigh” evidence. The library field has stumbled over this problem since Balbi’s 1835 discussion, at least. In that time, physicists have learned to make exquisite measurements of electrons in spite of their being invisible by attacking problems with their data and overcoming them. We, on the other hand, can see libraries and hold volumes in our hands but not measure either well.

It was recognized others might disagree with the reasoning which led to excluding the data from the *Princeton Compilation* so they are being made available here. Appendix 5 discusses these data. Appendix 3 lists all values in this edition of the Gerould data that differ from those in the *Compilation*.

2.3.2 The Purdue Data

A series of studies entitled *The Past and Likely Future of 58 Research Libraries* was conducted at Purdue using data assembled for those studies. (Dunn, 1965) Unlike the Princeton Compilation, which was an attempt to present a corrected edition of one series of data, the Purdue data resulted from an amalgamation of several series of data. The resulting text of data overlaps the Gerould data from 1950/51 through 1961/62. Appendix 4 presents a comparison of these two sets of data where they overlap, the sources of the Purdue data, and the discrepancies between the set of the Purdue data analyzed here and the purported sources.

In general, the same kinds of problems arise in this series of data as arose in the Princeton Compilation. There are apparent mistranscriptions and apparent rounding errors (which come from the source of the data used by the compilers). There are also cases of interpolation as the section on Procedures points out:

....In other cases, omitted data could be supplied by interpolation, extrapolation, or other calculations from the data at hand. (Dunn, 1965, p. 5)

In addition to these apparent errors and interpolations, there is the question of what happens when numbers from different series of data derived perhaps from different definitions are intermingled. That is, how was it handled if the two series have different values for the “same” variable or when different definitions were used? It is not clear how these matters were handled. Also, the compilers of the Purdue data report no values as missing. Although instances where values are not reported are a common occurrence in library data, there is none in this series—a fact which, along with interpolating data, leads to curious cases. Consider the figures for volumes added for Indiana for the years 1954/55 through 1962/63 for the Purdue data, the source given by Purdue for its data, the Gerould data, and the ARL data:

Year	Purdue	Purdue Source	Gerould	ARL
1954/55	30,000	*	*	NP
1955/56	83,000	*	*	NP
1956/57	46,500	*	*	NP
1957/58	82,500	*	*	NP
1958/59	69,000	*	*	NP
1959/60	59,000	*	*	NP
1960/61	97,000	*	*	NP
1961/62	102,100	*	*	*
1962/63	0	*	NP	*

(The symbol “*” is used to denote that the source of the data records a missing value for volumes added for Indiana in the given year. “NP” indicates the data were not published in the given year. The value for 1962/63 given here is not noted in Appendix 4 because it does not overlap with a Gerould value. The reader will remember that the last year of the Gerould data is 1961/62.)

Indiana reported values for volumes added to 1953/54 and after 1963/64 but not during this interval. It is not clear where the values reported by Purdue came from. They may have been interpolated, calculated “from data at hand” by an unstated formula, or arrived at by another means. The “0” apparently comes from converting a missing value to zero. It appears that the Purdue compilers substituted zeros for missing values systematically. Appendix 4 shows other cases where this conversion might have been done, particularly the figures for the wages reported at various libraries. Those who use these data should take this peculiarity into account in their analysis.

The objections raised to the means used to assemble the Princeton Compilation apply as well, then, to the Purdue data. Thus, this series is an inadequate source of historical data. No documentation to this series has been found so any discussion of the scope of editorial errors, interpolations, and the like is speculation. In fact, the only published documentation to this series is that published in the editions of this monograph. As is

discussed in Appendix 5, these data are also made available with the publication of this revised second edition of *The Gerould Statistics via the Web*.

Subsequent to the first edition of *The Gerould Statistics*, there was cause to examine aspects of these data more closely. Molyneux, 1990 and Molyneux, 1992 present more detailed discussions of the findings of these examinations. Suffice it to say, there appear to be more troublesome details in this series than were discovered in the earlier edition of this work. For our purposes here, however, these added details only strengthen the case to avoid or to control the kinds of data problems described in Section 2.3 generally and with the Purdue data specifically. Hence, a new method of compiling data has to be substituted for those used before. The next section, 2.4, is concerned with this question.

2.4 Stubbs and Buxton

In 1981, the Association of Research Libraries published the *Cumulated ARL University Library Statistics...* by Kendon Stubbs and David Buxton. This book is a model for anyone compiling data. It made several major changes in methodology from the Princeton and Purdue compilations. The authors attempted to introduce no changes in the data. After a discussion of

“differing perceptions over the years concerning what is to be counted in various categories,” (p. v)

the authors state that

“these vagaries in data collection over the years stand here as they appeared in the original...” (emphasis added).

No editorial changes; no interpolations. Rather, the intent is to make the data available as their original compilers believed them to be.

- They did add corrections to the data that were discovered after the printing of annual issues of *The ARL Statistics*.
- They published the data.
- They stated their methodology in print.
- The data were made available in electronic form with an infrastructure that has been adopted in other library data series.
- There was institutional stewardship so the data were not orphaned. The Association of Research Libraries has continued to collect and maintain control over the data. It publishes corrections and improves collection techniques.
- There is personal stewardship. It is impossible not to recognize that the ARL data are a model for the collection of library data and it is also impossible not to

recognize the contributions of Kendon Stubbs to this effort. His work on these and other data has been a model.

This compilation of the Gerould data attempts to build on the Stubbs-Buxton model.

2.5 Principles Followed in Assembling this Edition of the Gerould Data

A set of principles has been followed in compiling the data presented here that builds on Stubbs-Buxton. This discussion elaborates on those principles while developing new ones that take into account characteristics in the publication of the original Gerould data. One of the most interesting characteristics is that there are variant sources of the Gerould data that have to be consulted and understood. References are made to Appendix 1 where the sources of the Gerould data consulted are given and Appendix 2 which lists variant values of Gerould data found in the different sources used.

2.5.1 The Data Should be Published

The first principle is that the data should be published in both electronic form and hardcopy, if possible. Publication makes the data and the studies derived from them better by bringing to bear the scholarly apparatus which can be regarded as a means of focusing the efforts of people separated by place and time on a problem. To be more specific:

The act of publishing the data will make them better because more care will be taken by the compiler. There is a tendency in using electronic editions of data to ignore the data after a level of confidence in their integrity is reached so results often are considered without reference to the data. Publishing the data is a discipline which forces the compiler to look at the data carefully.

Publishing the data in an accessible medium allows others to examine them thus contributing to the building of an accepted text. If the list of figures of volumes added for Indiana had been published in hardcopy by Purdue, then anyone (and not just those with computers—the data were made available originally only *via* a deck of computer cards) could have seen that no volumes are reportedly added in 1962/63 and this number could have been corrected.

If the data are published and readily available then the research that led to the compiled text of data can be replicated. Replication of research results is important in advancing

our knowledge in most disciplines. If a piece of research cannot be replicated by someone other than the original researcher then there is reason to be skeptical of it. On the other hand, a series of data which is never scrutinized can gain a patina of respectability simply because its “inconsistencies” are never disclosed.

2.5.2 Publish Changes in the Data

Clerical problems like the rounding and transcription errors are inevitable in any edition of data—no matter how strenuous the effort to check for accuracy. In spite of repeated checks of the Gerould data presented here, cases of miskeying will be discovered and this fact must be accepted and dealt with. A problem with the Princeton Compilation, the Purdue data, and the major sources of the Purdue data is that there was no mechanism for correcting known errors because there was no means to publish the changes. As a result, someone using the data today would use data containing errors which at one time were correctable but are not any longer. It should be noted again: the cost of not documenting changes is high.

Corrections to this edition of the Gerould data will be reported in the *ARL Newsletter*, following the practice of reporting errors in the *ARL Statistics* there. As noted, no errors have been discovered since the original publication in 1986.

Providing for corrections seems necessary from the consideration of the data which occurred above. These kinds of changes were made in a number of cases in the Gerould data, as shown in Part 2 of Appendix 2. What emerges from this principle is, while the compilers' intent may be immutable, this edition and our reconstruction of that intent is dynamic. Each change in the text of this edition should bring it closer to its compiler's intent.

2.5.3 Use Authority Rather Than Consistency as a Criterion for Changing Data

In order to avoid the problems inherent in the Compilation, this edition uses textual authority as the guiding principle. Using the textual authority of a document or a number is an attempt to reconstruct what the author (or compiler) of a text had decided. Consistency and authority certainly are not antithetical principles in examining data but consistency is a criterion to be used in examining data for possible incorrect values but not for changing data (for reasons discussed above in section 2.3.1.3). Authority, however, is a criterion which can be used to examine and to change data if a change was the compiler's intent. Consistency could be used to suggest places to look for numbers which are not correct and whose authority should be checked.

The next section (2.5.3.1) will discuss how the criterion of authority was implemented in the construction of this edition of the Gerould data. The following section (2.5.3.2) will discuss a few of the implications of this approach to compiling data.

2.5.3.1 How Principles of Authority Were Implemented

Appendixes 1 and 2 go into detail in the matters addressed here. This section will serve as a general statement of what was done to collect the Gerould data for this edition. The following principles of textual authority have been applied, in reverse order of precedence so the second listed principle takes precedence over the first, and the third over the second and first.

The original text, where there is one, is the source of data.

In cases where there are two or more issues of the Gerould data for a given year the version put out by Gerould is the source. The Gerould data, particularly the early years, exist in a number of different copies done on different typewriters and probably done at different places by different people. Because each retyping increases the possibility of introducing new errors in the text, we should pick Gerould's issue. (The problem of variant annual issues of the data does not exist after Gerould retired.) This question is discussed further in Appendix 2.

In two kinds of cases there is a higher authority than the Gerould text. There are a few errata to individual values in the annual issues of the data listed in the second part of Appendix 1. These errata have precedence over the values in the original copies of the data because they correct mistakes noted in the copy of the data after it was distributed or because they add data reported after publication. In a few cases, the errata have not been found but are presumed to exist because of the nature of the corrections made in individual issues of the data. A second kind of higher authority is the one case where the source Gerould used for his data exists. In this case, which is discussed with the *errata* in Appendix 2 under Texas, 1913/14 (see pages 90-91), Gerould mistranscribed a number.

There are comparatively few cases where there is more than one possible value for a variable and these principles of authority were sufficient to adjudicate all but a handful of them. In the cases where these principles were not adequate, other means were used. All cases where there is more than one possible value are listed in Appendix 2 and it can be seen that these principles were applied rigorously except in one case where internal

evidence in the data make clear that there must have been an erratum which is now missing (see the listing for Louisiana State, 1954/55 in Appendix 2, page 85.)

2.5.3.2 Implications of the Use of Authority as an Editing Criterion

The chief implication of the use of authority as an editing criterion for a series of data is that an edition of data compiled by it is a deliberate attempt to reconstruct what the author intended the numbers to be. We cannot correct errors which arose in the earlier compilations of data, nor will we ever know the true levels of the variables; however, the original compiler's intent is the best guess that can be made today about the values of the variables. In a compilation using authority, there are no interpolations over existing or missing values; the data appear in their raw state with their “vagaries...stand[ing] as they appeared...in the original annual issues of the statistics...” (Stubbs and Buxton, p. v.).

The chief advantage of this approach is that it does not introduce changes in a text for reasons which may be biased in an unclear fashion and which may interfere with future research. Of course, if is research based on these data, the person doing the research can interpolate or change whatever he or she likes (although it should be made clear in reporting results where and how it is done). The restrictions on using data should be less stringent than those on compiling data. A case of undocumented interpolation in the Princeton Compilation—or the Purdue data—cannot now be undone unless we find detailed documentation.

Textual authority has been applied rigorously to the Gerould data and while such rigor limits the kinds of bias which can come from expedient changes, there are disadvantages. In their raw state data may be included which seem extremely suspicious. In addition, in rejecting several possible values as lacking sufficient authority (as indicated in Appendix 2), possibly wrong values were included in the text. Part 2 of this Appendix lists values which may be preferred to the values chosen. Another disadvantage is that techniques which have been used on what might be considered sanitized versions of data may not help elucidate how the libraries or their variables behave with these data. New techniques may have to be developed.

2.5.4 Disclose Variant Data or Texts

If the changes made at Princeton had been documented, the present edition could have built on that work and added to it. Because this earlier work was not documented and

because no one is likely to remember what happened in sufficient detail to be useful, the effort which went into it is lost. Documenting what has been done is another way of permitting different people at different places and times to build on and, if necessary, correct earlier work.

What should be disclosed and documented?

- The sources of data consulted.
New copies of the raw data may be found which may shed light on a value for which there is more than one possibility.
As noted, Appendix 1 lists the sources consulted.
- Errata consulted.
Any new erratum is a useful discovery and it is important to alert anyone who discovers one that it is new.
Part 2 of Appendix 1 lists errata consulted.
- Variant values with good authority.
By disclosing variant values that have good authority, places where there are questions are made explicit. This practice has the great advantage of alerting others to these questions and thereby bringing other views or even other copies of data to bear on them. Also, in applying the principle of textual authority, a conclusion was arrived at that one of the two main Gerould texts was the correct one. This conclusion was based on circumstantial evidence and may prove to be wrong. In that case, the choices made between the values which differ between the editions arranged by name and by size will have to be changed. Disclosing those differences makes correcting this potential problem easy.
Appendix 2 lists the variant values.
- Variant values without good authority but which may be correct.
Applying the principles used here can lead to cases where the authority for a decision is impeccable but the value chosen is likely to be wrong. These few cases are listed in Part 2 of Appendix 2.

In addition to the items which should be disclosed in presenting the data, an implication of the principles of authority and of disclosing as much as possible about the data affects how changes in this edition of the Gerould data should be reported in the *ARL Newsletter*. The authority for the change should be given. For example, if the change results from the discovery of a typographical error that fact should be noted.

2.6 Conclusion

This chapter has outlined the method used to compile the Gerould data presented here. The two appendices referred to should be consulted for more detail. Chapter 3 lists the data by the name of the institution sorted by year. As will be discussed in Appendix 5, the Gerould data, data from the Princeton Compilation, and the Purdue data are being made available.



The Gerould Statistics

Chapter Three:

The Gerould Data

This chapter contains a discussion of the variables in the Gerould data. The documentation for the electronic edition of these data is in Appendix 5, beginning on page 144.

An important question about these data is whether the Gerould variables are comparable to similar variables reported by the Association of Research Libraries in compiling its annual statistics. Examining data stretching back to 1907/08 should give us a means to understand libraries better than we do now so the comparability of definitions of variables has to be considered. Preliminary studies mentioned above seemed to indicate the data are comparable.

A general discussion of library variables by Stubbs and Buxton (pp. vii-xiii) in 1981 provides background on the variables found in the ARL data and is helpful in understanding variables in the Gerould statistics. More recently, Molyneux, 1989 and Stubbs and Molyneux, 1990 discuss library variables and their characteristics.

3.1 The Variables Recorded in the Gerould Data

Eight variables are reported in the Gerould statistics: volumes held, volumes added, expenditures for materials, staff, salaries, appropriations (for materials expenditures in the current year), total expenditures, and wages. The latter three are not recorded in all years. These variables will be discussed in the pages which follow.

3.1.1 Volumes Held

The number of volumes a library holds is probably the oldest and most frequently referred to library statistic. It is reported for all years of these data. Although librarians generally recognize that the number of volumes held is not a measure of a library's quality, discussions of libraries as far back as those about the Alexandrian library often sound like bragging about how big the library is which, presumably, is a sign of how good the library is.

One of the problems with using volumes to compare libraries is that there is no accepted definition of what a volume is. The two most widely accepted definitions are the physical volume and the bibliographic volume, a distinction discussed by Downs, (1946) and Kuhlman, (1960). Whatever kinds of volumes are being measured, they have increased greatly over the years of these data. The twelve university libraries which have reported data over the entire period:

The twelve institutions are:

California, Berkeley	Minnesota
Illinois	Missouri
Indiana	Nebraska
Iowa	Ohio State
Kansas	Washington
Michigan	Wisconsin

During the period, these institutions grew from an average of 107,425 volumes in 1907/08—or “Books in Library,” as that year's statistics called this variable—to 1,772,831 by 1961/62. The 1995/96 total volumes held for these 12 libraries is 64,015,448! It is often forgotten that while Fremont Rider's math was wrong, his understanding of the implications of library growth was prophetic. This general point is discussed in Molyneux, (1996)

Volumes held should increase steadily without any declines, unless there is a disaster, such as a fire, which destroys volumes. However, this statistic, as reported in these data, occasionally declines from one year to the next for no apparent reason. This kind of drop in volumes held which is also occasionally seen in the data in section 3.2 may be a result of mistakes in reporting, a change in a library's definition of a volume, or new and more accurate estimates of volumes held. While these kinds of fluctuations do not seem logical, they do occur, although grouping libraries will obscure fluctuations such as a decline in volumes at one library.

3.1.2 Volumes Added

While the number of volumes held at an academic library is an indicator of the past commitment of its parent institution, the number of volumes added is an indicator not only of the health of the library but of a continuing commitment by its institution. If two libraries of the same size were observed over a decade and one added twice as many

volumes as the other, it would be reasonable to conclude the one which added the larger number of volumes had more institutional support and, in addition, a newer collection than the other.

The *ARL Statistics* report volumes added gross and volumes added net. The Gerould figure probably corresponds to volumes added gross, that is, volumes added without taking into account volumes which were discarded or lost. The gross figure was not reported in the first year of the ARL data (1961/62). However, the distinction was drawn between the two kinds of volumes added in the next year the data are reported, 1962/63, when only 25 of the 63 libraries reporting had figures for volumes added net while 60 reported figures for volumes added gross. (Molyneux, 1986b) These figures indicate that there was more familiarity with the gross figure, probably because it had been collected longer. In fact, of the three libraries which did not report two, Syracuse and Wayne State, are not included in the Gerould data. The third library was Indiana which, as was noted in section 2.3.2, did not report volumes added from 1954/55 through 1962/63.

The mean number of volumes added for the twelve libraries listed above is given here:

Table 3.1

Year	Mean Volumes Added	Year	Mean Volumes Added	Year	Mean Volumes Added
1907/08	8,689	1926/27	18,638	1945/46	33,841
1908/09	*	1927/28	22,104	1946/47	31,168
1909/10	10,976	1928/29	19,915	1947/48	46,369
1910/11	*	1929/30	22,604	1948/49	37,839
1911/12	12,245	1930/31	21,941	1949/50	37,199
1912/13	11,420	1931/32	22,414	1950/51	42,677
1913/14	13,771	1932/33	19,748	1951/52	42,262
1914/15	14,090	1933/34	18,475	1952/53	46,191
1915/16	13,831	1934/35	18,915	1953/54	52,215
1916/17	13,276	1935/36	21,660	1954/55	51,159

1917/18	11,670	1936/37	24,379	1955/56	52,948
1918/19	11,780	1937/38	27,014	1956/57	54,343
1919/20	11,614	1938/39	25,400	1957/58	57,517
1920/21	13,144	1939/40	25,988	1958/59	56,701
1921/22	16,747	1940/41	28,445	1959/60	61,988
1922/23	15,357	1941/42	28,319	1960/61	66,465
1923/24	15,508	1942/43	24,588	1961/62	76,068
1924/25	17,518	1943/44	25,066		
1925/26	17,398	1944/45	24,551		

The Gerould data were not published in 1908/09 and 1909/10.

The number of volumes added for these libraries fluctuates from year to year but it can be seen that number of volumes added annually increased over the period. There are drops in the annual levels of volumes added during the period of World Wars I and II. There is also a drop in the early 1930's which is probably a result of the Depression. Molyneux, 1994a is the most recent treatment of the subject of library growth. It includes graphs and a recent bibliography on the subject.

3.1.3 Expenditures for Library Materials and Binding

The first issue of the Gerould statistics reports the amount "Spent for Books, periodicals & Binding (*sic*), and this amount is reported for all subsequent years of the data. The *ARL Statistics* currently separate expenditures of books, serials, and binding. Binding expenditures were first separated in 1962/63 and expenditures for serials in 1975/76. In order to make the data from the Gerould series comparable to those from ARL, the disaggregation of these data will have to be taken into account.

Table 3.2 shows mean expenditures for materials for the 12 libraries which reported data from 1907/08 through 1961/62. There has been a substantial increase in the levels reported for this variable over the period. Materials' expenditures, as with the other expenditure categories, have increased with the decline of the value of the dollar. It would be worthwhile to know the relationship between increases in expenditures and in

price levels in general in order to know if gains in library expenditure categories have usually led or followed inflation. There are probably periods where expenditures have increased faster than inflation and periods where they have failed to keep up with inflation, but an understanding of what happened over the course of these years might allow us to form a better understanding of library funding.

Table 3.2

Year	Mean Dollars Expended	Year	Mean Dollars Expended
1907/08	\$16,561	1935/36	\$68,242
1908/09	*	1936/37	69,284
1909/10	20,535	1937/38	74,752
1910/11	*	1938/39	85,862
1911/12	23,668	1939/40	75,940
1912/13	23,386	1940/41	80,049
1913/14	28,049	1941/42	76,789
1914/15	24,046	1942/43	88,307
1915/16	24,893	1943/44	83,699
1916/17	23,860	1944/45	95,025
1917/18	20,954	1945/46	102,075
1918/19	24,123	1946/47	130,030
1919/20	26,466	1947/48	163,931
1920/21	30,533	1948/49	184,812
1921/22	38,195	1949/50	194,381
1922/23	43,180	1950/51	212,134
1923/24	54,240	1951/52	224,729
1924/25	55,407	1952/53	246,286
1925/26	57,229	1953/54	239,605
1926/27	61,557	1954/55	260,499

1927/28	64,913	1955/56	269,050
1928/29	71,037	1956/57	309,885
1929/30	76,399	1957/58	362,738
1930/31	80,208	1958/59	369,542
1931/32	73,335	1959/60	415,960
1932/33	69,163	1960/61	498,644
1933/34	60,662	1961/62	539,913
1934/35	61,148		

The Gerould data were not published in 1908/09 and 1909/10.

3.1.4 Appropriations

Appropriations are reported from 1909/10 through 1953/54. There is no comparable figure currently reported by ARL. The figure reported for materials expenditures shows how much was spent last year. A possible reason for keeping track of the appropriations is to show how much is to be spent in the current year. It is not clear why this variable was dropped after 1953/54.

Table 3.3 lists the mean values of appropriations at the 12 libraries. The mean values for appropriations are:

Table 3.3

Year	Mean Dollars Appropriated	Year	Mean Dollars Appropriated
1907/08	*	1935/36	\$60,261
1908/09	*	1936/37	65,022
1909/10	\$21,174	1937/38	73,518
1910/11	*	1938/39	72,806
1911/12	20,547	1939/40	79,523
1912/13	25,761	1940/41	73,123
1913/14	22,853	1941/42	77,749

1914/15	26,717	1942/43	75,212
1915/16	27,326	1943/44	89,545
1916/17	25,134	1944/45	89,125
1917/18	24,879	1945/46	105,943
1918/19	28,589	1946/47	152,710
1919/20	26,154	1947/48	157,112
1920/21	34,026	1948/49	181,207
1921/22	36,475	1949/50	211,357
1922/23	51,131	1950/51	220,160
1923/24	52,925	1951/52	211,614
1924/25	51,733	1952/53	231,865
1925/26	51,982	1953/54	241,984
1926/27	58,435	1954/55	*
1927/28	58,525	1955/56	*
1928/29	65,514	1956/57	*
1929/30	69,426	1957/58	*
1930/31	69,744	1958/59	*
1931/32	65,212	1959/60	*
1932/33	51,527	1960/61	*
1933/34	54,183	1961/62	*
1934/35	56,013		

3.1.5 Total Expenditures

A figure for the total expenditures was first reported in 1954/55 and replaced appropriations in the annual statistics. The figure is calculated by adding expenditures for materials, salaries, and student salaries. For this edition, the total figure was entered from the annual data, not calculated from the formula.

The Gerould figure for this variable is not directly comparable to the ARL figure which is calculated by adding expenditures for materials, binding, salaries (which include student salaries), and other operating expenditures. There is no Gerould equivalent of operating expenditures, and student salaries are not reported before 1949/50. A total figure which would be comparable for both sets of data can be assembled from 1949/50 by using the formula from 1954/55 to 1961/62 from the Gerould data back to 1949/50 and forward from 1962/63 through 1983/84.

Total expenditures for the 12 libraries for the eight years this figure exists show that it increased substantially. The mean values for this variable for the years it was published are:

Year	Mean Total Expenditures
1954/55	\$832,823
1955/56	873,547
1956/57	973,933
1957/58	1,127,834
1958/59	1,169,559
1959/60	1,284,647
1960/61	1,458,109
1961/62	1,594,947

3.1.6 Total Staff

A count of staff is reported each year of the data. The issue of the data from 1909/10 defines this number as the “Number of members of the staff (excluding pages and student assistants)” while the letter Gerould sent out in 1916 defines it as the “Number of members of staff (excluding janitors, pages and student assistants).” This definition was apparently followed throughout the period. ARL collects data on the numbers of professional and nonprofessional employees. Adding the two figures will yield the number of full-time staff which should be comparable to the Gerould variable. See the discussion on salaries in section 3.1.7 for a related matter.

The average number of staff for the 12 libraries is listed in Table 3.4 The steady increase in staff members is readily apparent as are the sharp rise and then drop after World War II.

Table 3.4

Year	Mean Total Staff	Year	Mean Total Staff	Year	Mean Total Staff
1907/08	13.5	1926/27	41.5	1945/46	73.9
1908/09	*	1927/28	45.1	1946/47	86.7
1909/10	14.3	1928/29	48.6	1947/48	99.6
1910/11	*	1929/30	50.4	1948/49	139.7
1911/12	16.9	1930/31	50.3	1949/50	124.4
1912/13	18.7	1931/32	49.3	1950/51	125.7
1913/14	21.4	1932/33	47.6	1951/52	124.4
1914/15	22.1	1933/34	45.8	1952/53	130.6
1915/16	23.2	1934/35	49.3	1953/54	131.8
1916/17	24.5	1935/36	50.0	1954/55	136.2
1917/18	23.8	1936/37	51.5	1955/56	138.2
1918/19	27.0	1937/38	55.4	1956/57	153.0
1919/20	27.2	1938/39	57.1	1957/58	154.1
1920/21	27.5	1939/40	63.3	1958/59	158.6
1921/22	29.8	1940/41	63.6	1959/60	159.7
1922/23	31.9	1941/42	64.0	1960/61	166.6
1923/24	34.5	1942/43	67.8	1961/62	177.7
1924/25	37.2	1943/44	70.7		
1925/26	41.2	1944/45	66.8		

3.1.7 Staff Salaries

This variable is the “total salary account during current fiscal year” (1909/10 issue) and seems to have been interpreted as the money paid full-time employees—presumably the same people measured under total staff.

Salaries as measured in the Gerould statistics appear to differ from the salary figure in the ARL data because this latter figure reports the previous year's salaries. When the change of definitions occurred is not clear. ARL's salary figure includes full-time employees, students, and part-time employees. For this reason, the variable name for these salaries is different from that for the ARL salaries. The ARL variable name is TOTSAL while the Gerould variable name is FTESAL. Future research may show these variables are the same, but at this time, this fact is not established.

The problem of comparability inhibits one intriguing calculation, average salary per library employee, which could be calculated by dividing total salaries by total staff. However, interpreting the results of this calculation would require care not only because of the question of comparability but also because both numbers mix professional and non-professional employees. A drop or rise in the average amount paid each employee might be a result of a change in the salaries paid or a change in the ratio of professionals to non-professionals. There is evidence from the period covered by the ARL data to indicate this ratio is dynamic.

The average figures paid in staff salaries at the 12 libraries for the years of the Gerould data are given in Table 3.5:

Table 3.5

Year	Mean Total Salaries		Year	Mean Total Salaries		Year	Mean Total Salaries
1907/08	\$11,256		1926/27	\$86,760		1945/46	\$162,445
1908/09	*		1927/28	91,967		1946/47	213,292
1909/10	14,023		1928/29	97,588		1947/48	255,195
1910/11	*		1929/30	105,321		1948/49	335,470
1911/12	17,548		1930/31	105,557		1949/50	343,281

1912/13	19,511		1931/32	104,274		1950/51	364,119
1913/14	20,624		1932/33	93,070		1951/52	403,035
1914/15	22,283		1933/34	87,132		1952/53	435,513
1915/16	23,691		1934/35	92,761		1953/54	461,680
1916/17	26,230		1935/36	97,961		1954/55	488,458
1917/18	27,686		1936/37	96,491		1955/56	519,638
1918/19	34,290		1937/38	103,409		1956/57	561,228
1919/20	41,744		1938/39	96,843		1957/58	650,817
1920/21	44,665		1939/40	100,726		1958/59	683,445
1921/22	51,681		1940/41	109,667		1959/60	742,276
1922/23	56,494		1941/42	112,549		1960/61	811,330
1923/24	61,309		1942/43	115,302		1961/62	896,349
1924/25	71,702		1943/44	130,735			
1925/26	79,897		1944/45	134,947			

3.1.8 Student Salaries

The salaries paid students were added to these data in 1949/50. It can be speculated that this variable was added because student wages were becoming an increasing part of library budgets, but there is no way to be sure at this time. The ARL data report total salaries which are the sum of salaries for full-time employees, students, and part-time employees.

The Appendices give detailed documentation to the data themselves and to related data series.

Appendix 1 lists the data sources consulted begins on page 54.

Appendix 2 discusses cases where there is reason to doubt a value chosen and presented in the data tables and in the electronic data, in spite of its good authority. See page 69.

Appendix 3 discusses the Princeton Compilation and compares these data to the Gerould Statistics. See page 99.

Appendix 4 discusses the Purdue Data and compares these data to the Gerould Statistics. See page 120.

Appendix 5 gives details on the Electronic version of the data documented here. It begins on page 144.

The Bibliography begins on

These tables also include blank lines which are not included on the electronic forms of the data strictly as a visual cue to a special set of cases which occur as a result of the data collected by Phineas Winsor at Illinois in 1909/10. These data are discussed in Appendix 2 (see Chicago, 1909/10 for a general explanation) and include data comparable to the Gerould data from ten institutions (Chicago, Columbia, Cornell, Harvard, Johns Hopkins, Northwestern, Pennsylvania, Princeton, Stanford, and Yale). In order to include a reference to the data in Appendix 2, there will be a link in these ten cases for 1909/10 in each of these institutions' lists to the variant values. These cases therefore, are not included in the data but there is a good variant text as discussed in Part 2 of Appendix 2. These variants may be preferred in some applications. A blank line is then inserted in the data in these tables for 1910/11 and for each year until data are reported.



The Gerould Statistics

Appendix One:

Data Sources Consulted

The Gerould statistics were issued by James Thayer Gerould at Minnesota in 1907/08, 1909/10, and subsequently from 1911/12 through 1918/19 or 1919/20. In 1920, Gerould moved to Princeton. From then through 1938/39 Gerould continued to publish the statistics. Gerould retired in 1938 and publishing the data was continued by the staff at Princeton from 1939/40 until 1961/62. In that year the Association of Research Libraries began to publish its statistical series which continues the series started by Gerould through today as the *ARL Statistics*.

In the latter years, the issues of these data available from the various sources are virtually identical but in the early years there are different editions of the various years in which the data were rearranged, edited, or retyped. It is not clear where this work was done, although one edition of the data in which the data were put in order by the size of the institution appears to have been done at Illinois. In spite of these different editions, there are few major discrepancies in data between these different editions. Appendix 2 lists the variant items of data found in the sources consulted.

There have been two ways used to order the Gerould data over the years: alphabetically by the name of the institution and by size of the library in descending number of volumes held. In the early years the issues organized by institution name were entitled simply "Statistics of University Libraries" with a date; while those in order of the size of the institution had a title followed by: "Compiled by J. T. Gerould, University of Minnesota Library."

In addition to these annual issues there are errata which were sent by library directors or by the compiler in postcards or letters. These errata are noted with the source and also in a separate section at the end of this Appendix.

The copies of the Gerould statistics gathered for this study came from university libraries, archives, and a private collection of these data. Originally, interlibrary loan requests were sent to all institutions which reported data in the 1921/22 issue, including

Princeton. The only universities which could locate copies were Illinois, Texas, and Virginia. Dr. Haynes McMullen of the Library School of the University of North Carolina subsequently contributed what ultimately proved to be a copy of issues of all the early years' data from his dissertation notes. These copies came from the archives of the University of Chicago in the mid-1940's. The Manuscripts Department at the University of Virginia was subsequently checked. Letters were also sent to other manuscripts departments or archives of large university libraries which seemed likely to have early data. Penelope Krosch at the University of Minnesota sent early issues of the Gerould statistics and Maynard Brichford at the University of Illinois sent early issues of the Gerould statistics but also correspondence of Gerould and Phineas Winsor, the University Librarian at Illinois. These letters were most useful in sorting out facts about the early years of this data series. The University Microfilms collection containing data from 1944/45 to 1961/62 was also consulted and is available from UMI as Special Film No. S-338. The University of Minnesota was also the source of the photograph of Gerould that appears in the first edition.

To summarize, the issues of the Gerould statistics consulted in this compilation come from the following sources:

- Illinois Archives
- Illinois stacks (by ILL)
- (Haynes) McMullen (originally from Chicago)
- Minnesota Archives
- Virginia stacks
- Texas (by ILL)
- University Microfilms (1944/45 to 1961/62). The earlier edition published by Xerox-University Microfilms, one of UM's Books on Demand (AG1-30634) is actually the compilation of the data for 1919/20-1943/44 done at Princeton and referred to in the text as the *Princeton Compilation*.

No attempt will be made here to discuss differences in the data between the various issues for given years. This task is left to Appendix 2.

A list of the sources consulted by year follows. In order to make the remarks clearer, a distinction will be drawn between the types of copies of the Gerould statistics which exist. Gerould and subsequent compilers would have generated what will be called a "master" copy which reflected their views about what the numbers were. This copy is

the corrected text assembled by the compiler from which the distribution copy—the copy which is to be sent to subscribers—is to be made. A distribution copy can exist as an “original,” that is, the actual piece mailed by the compiler or as a copy (used here to mean either a photographic copy or one made by the Xerox, or similar, process). Since the invention of the Xerox process, a distribution copy is often a direct copy of the master. That is, in such a copy the text itself has not been touched or altered and there is little chance of a change in the text's being introduced in the process of making the distribution copy. This process of reproduction of documents has tended to obscure a distinction that is useful in considering these early data. When Gerould compiled his master copy, he could not Xerox the master and distribute the resulting issues, he had to have a copy made. By as early as 1913/14, these issues were made by a process which allowed reproduction but before then the distribution copies appear to have been retyped. Each retyping, of course, introduces the possibility of a change in the text.

Year	Source	Remarks on Sources
1907/08	Minnesota	Both are in alphabetical order and appear to have been done on the same typewriter but at different times.
	McMullen	
1909/10	Illinois Archives	Both are in alphabetical order and appear to have been done on the same typewriter but at different times.
	McMullen	
1911/12	McMullen	In alphabetical order.
1912/13	McMullen	In alphabetical order.

Year	Source	Remarks on Sources
1913/14	Illinois stacks	In order by size. ¹
	McMullen	In alphabetical order. ²
	Minnesota	In alphabetical order. ³

Notes on 1913/14:

1. The Illinois stacks issue is blurred and unclear.
2. The McMullen and Minnesota issues may have come from the same distribution copy.
3. There is a letter from Winsor to Gerould at Illinois and an erratum from Gerould in the Illinois Archives.

Year	Source	Remarks on Sources
1914/15	McMullen	Both are in alphabetical order and may have come from the same distribution copy.
	Minnesota	
1915/16	Illinois Archives	In order by size. ¹
	Illinois stacks	
	McMullen	In alphabetical order. ²
	Minnesota	

Notes on 1915/16:

1. These two issues appear to have been done on the same typewriter but were done at different times.
2. These two issues appear to be from the same distribution copy.

Year	Source	Remarks on Sources
1916/17	Illinois Archives	All three are in alphabetical order. ¹
	Illinois stacks	
	McMullen	

Notes on 1916/17:

1. The issue in the stacks is probably a carbon of the Archive copy.

Year	Source	Remarks on Sources
1917/18	McMullen	In alphabetical order.
1918/19	McMullen	In alphabetical order.
1919/20	McMullen	In alphabetical order.
1920/21	Illinois Archives McMullen	All in alphabetical order. ¹

Notes on 1920/21:

1. There are two issues in the Archives, one of which appears to be a carbon of the other.

Year	Source	Remarks on Sources
	McMullen	
1921/22	Texas Virginia	All in alphabetical order. ¹

Notes on 1921/22:

1. All three issues are from the same distribution copy. The Texas issue is a photocopy.

Year	Source	Remarks on Sources
	McMullen	
1922/23	Texas Virginia	All are in alphabetical order. ¹

Notes on 1922/23:

1. Professor McMullen had two issues, an original and a photocopy.
2. All four issues are from the same distribution copy.

Year	Source	Remarks on Sources
1923/24	McMullen	All are in alphabetical order. ¹
	Texas	
	Virginia	

Notes on 1923/24:

1. These three issues are from the same distribution copy.
2. There is an undated erratum in the Illinois Archives the first line of which reads: "Northwestern University reports:"

Year	Source	Remarks on Sources
1924/25	Illinois Archives ¹	Illinois issues in order by size.
	Illinois stacks ²	
	McMullen Texas Virginia ³	McMullen, Texas, and Virginia in alphabetical order. ⁴

Notes on 1924/25:

1. The Illinois Archives issue includes a typed erratum dealing with data from Iowa labeled "MEMORANDUM; Concerning; Statistics of University Libraries, 1924-1925, Recently Distributed." This erratum is date-stamped "Dec 21 1925." It is not clear where this erratum came from. There is an erratum date stamped "Dec 12 1925" reporting data from Colorado. The first line says: "University of Colorado reports:". This correction appears to have been typed on the same typewriter as the original data and may have come from Gerould.
2. The issue in the Illinois stacks is apparently a carbon of the Archives copy.
3. The Virginia issue includes three errata. Two are other issues of the errata found in the Illinois Archives although neither has been stamped with a date. The erratum with data from Iowa appears to be identical to that in the Illinois Archives while the erratum with data from Colorado appears to have been done on the same typewriter as the issue of this erratum in the Illinois Archives but at a different time. The data in the two issues are the same. The third erratum is in the form of a typed letter from William Henry, Librarian at the University of Washington (Seattle), and is dated "February 25, 1926." The latter three issues appear to be from the same distribution copy.

Year	Source	Remarks on Sources
1925/26	Illinois Archives	Both Illinois issues in order by size.
	Illinois stacks ¹	
	McMullen	McMullen, Texas, and Virginia in alphabetical order.
	Texas ²	
Virginia		

Notes on 1925/26:

1. The issue in the Illinois stacks appears to be a carbon of the Archives issue.
2. Texas appears to be a photocopy. The latter three issues appear to be from the same distribution copy.

Year	Source	Remarks on Sources
1926/27	Illinois stacks ¹	In order by size.
	McMullen	McMullen, Texas, and Virginia in alphabetical order. ²
	Texas	
	Virginia	

Notes on 1926/27:

1. The Illinois issue is not clear.
2. The latter three issues appear to be from the same distribution copy.

Year	Source	Remarks on Sources
1927/28	Illinois stacks ¹	In order by size.
	McMullen	McMullen, Texas, and Virginia in alphabetical order. ²
	Texas	
	Virginia	

Notes on 1927/28:

1. Some numbers on the Illinois stacks' issue are unclear because of the quality of the copy.
2. The latter three are from the same distribution copy.

Year	Source	Remarks on Sources
1928/29	McMullen	All three in alphabetical order. ¹
	Texas	
	Virginia	

Notes on 1928/29:

1. These copies are from the same distribution copy.

Year	Source	Remarks on Sources
1929/30	Illinois stacks	In order by size
	Texas	Texas and Virginia in alphabetical order ¹
	Virginia	
1930/31	Illinois stacks	In order by size
	Texas	Texas and Virginia in alphabetical order ¹
	Virginia	
1931/32	Illinois stacks	In order by size
	Texas	Texas and Virginia in alphabetical order ¹
	Virginia	
1932/33	Illinois stacks	In order by size
	Texas	Texas and Virginia in alphabetical order ¹
	Virginia	
1933/34	Illinois stacks	In order by size
	Texas	Texas and Virginia in alphabetical order ¹
	Virginia	
1934/35	Illinois stacks	In order by size
	Texas	Texas and Virginia in alphabetical order ¹
	Virginia	

Year	Source	Remarks on Sources
1935/36	Illinois stacks	In order by size
	Texas	Texas and Virginia in alphabetical order ¹
	Virginia	
1936/37	Illinois stacks	In order by size
	Texas	Texas and Virginia in alphabetical order ¹
	Virginia	
1937/38	Illinois stacks	In order by size
	Texas	Texas and Virginia in alphabetical order ¹
	Virginia	
1938/39	Illinois	In order by size.
	Texas	Texas and Virginia in alphabetical order. ¹
	Virginia ²	
1939/40	Illinois	In order by size.
	Texas	Texas and Virginia in alphabetical order. ¹
	Virginia	

Notes:

1. The latter two are from the same distribution copy.
2. There is an erratum (an undated, printed postcard) from M. Llewellyn Raney, Director, at Chicago.

Year	Source	Remarks on Sources
1940/41	Illinois	Illinois and Virginia Archives in order by size. ¹
	Virginia Archives	
	Texas	Texas and Virginia stacks in alphabetical order. ¹
	Virginia stacks	

Notes on 1940/41:

1. All four copies appear to have been done on the same typewriter from two different distribution copies, one arranged by size and the other alphabetical. Both sets were apparently issued at the same time as there is an explicit reference on the alphabetically ordered list to footnotes on the “sheet attached...arranged by size.”

Year	Source	Remarks on Sources
1941/42	Illinois	All three in order by size. ¹
	Texas	
	Virginia	
1942/43	Illinois	All three in order by size. ¹
	Texas	
	Virginia	
1943/44	Princeton compilation	²
1944/45	Illinois	All three in order by size. ¹
	Univ. Microfilms	
	Virginia	

Notes:

1. These three issues are from the same distribution copy.
2. The data from this year were not distributed because of war time paper shortages.

Year	Source	Remarks on Sources
1945/46	Illinois	All three in order by size. ¹
	Univ. Microfilms	
	Virginia	

Notes on 1945/46:

1. These three issues are from the same distribution copy. However, there are two different sets of footnotes. Illinois and Virginia have what appear to be identical sets of footnotes on three pages while UM's copy has two pages of footnotes. The texts of these two sets of footnotes appear to be the same.

Year	Source	Remarks on Sources
1946/47	Illinois	All four in order by size. ¹
	Texas	
	Univ. Microfilms	
	Virginia	
1947/48	Illinois	All four in order by size. ¹
	Texas	
	Univ. Microfilms	
	Virginia ²	
1948/49	Illinois	All four in order by size. ¹
	Texas	
	Univ. Microfilms	
	Virginia	
1949/50	Texas	All three in order by size. ³
	Univ. Microfilms	
	Virginia	
1950/51	Illinois	All four in order by size. ¹
	Texas	
	Univ. Microfilms	
	Virginia	

Notes:

1. All four issues are from the same distribution copy.
2. Virginia issue contains an erratum (an undated, mimeographed postcard) from the "Princeton University Library" for Kansas.
3. All three issues are from the same distribution copy.

Year	Source	Remarks on Sources
------	--------	--------------------

	Texas	
1951/52	Univ. Microfilms Virginia ²	All three in alphabetical order. ¹
	Texas	
1952/53	Univ. Microfilms Virginia	All three in alphabetical order. ¹
	Texas	
1953/54	Univ. Microfilms Virginia ³	All three in alphabetical order. ¹
	Texas	
1954/55	Univ. Microfilms Virginia	All three in alphabetical order. ^{1,4}
	Texas	
1955/56	Univ. Microfilms Virginia	All three in alphabetical order. ¹
	Texas	
1956/57	Univ. Microfilms Virginia	All three in alphabetical order. ¹
	Texas	
1957/58	Univ. Microfilms Virginia	All three in alphabetical order. ¹
	Texas	
1959/60	Univ. Microfilms Virginia	All three in alphabetical order. ¹

Year	Source	Remarks on Sources
1960/61	Texas	All three in alphabetical order. ^{1,3}
	Univ. Microfilms	
	Virginia	

Notes 1951/52-1960/61:

1. All three issues are from the same distribution copy.
2. Virginia contains an erratum for Washington, St. Louis.
3. The Virginia issue is a photocopy. Both the UM issue and the Virginia issue have identical handwritten notes on them suggesting that both were copies from an unknown original or that Virginia's copy was made from the UM copy.

Year	Source	Remarks on Sources
1961/62	Univ. Microfilms Virginia	Both in alphabetical order. ¹

Notes on 1961/62:

Both issues are from the same distribution copy.

ERRATA

What follows is a list of the errata used in this compilation. Individual values will be discussed in Appendix 2.

Year	Source	Institution Dealt With
1913/14	Illinois Archives	Texas

Notes:

In the Illinois Archives, there is a letter, dated November 17, 1914, from Phineas Windsor to Gerould saying that he had received a letter from the University of Texas Library with data similar to the statistics distributed by Gerould about other libraries. The letter of reply, dated November 19, 1914, records that Gerould was sending these figures to “the

institutions which have received copies of the table of university library statistics.” The November 19th letter from Gerould giving the values from Texas and addressed “Dear Sir” is also included in the Archives. As noted in Appendix 2, Gerould made a transcription error from the Winsor letter to the erratum in the number of volumes reported for Texas. Winsor's letter records the number of volumes at Texas as 100,774 while Gerould's erratum has this number as 100,744.

Year	Source	Institution Dealt With
1923/24	Illinois Archives	Northwestern

Notes:

This erratum begins “Northwestern University reports:” and resembles the erratum for Colorado for 1924/25 in the format displaying the data and in the typewriter used. This typewriter appears to be the same one used to produce the alphabetically ordered distribution copy of the Gerould data. This erratum bears no date and has not been date stamped. The distribution copies of the data from 1922/23 and 1923/24 list no data from Northwestern so the assignment of this erratum to 1923/24 is based on circumstantial evidence which is discussed in Appendix 2.

Year	Source	Institutions Dealt With
1924/25	Illinois Archives	Colorado ¹ Iowa ²
	Virginia	Colorado Iowa ² Washington ³

Notes:

1. The Colorado erratum might have been sent by Gerould. In the distribution copy there is space for Colorado's values but nothing was typed in. Colorado might have sent the reply late and Gerould might have passed these additions on. As discussed above, the two issues of this erratum appear to have been done on the same typewriter although at different times. These errata resemble the erratum for Northwestern for 1923/24 in that they use the same format to display the data and they also use the same typewriter. The typeface on this typewriter closely resembles that on the typewriter used on the distribution copy of the alphabetically ordered Gerould data.

2. The two Iowa errata are identical, with the exception that the issue from Illinois is stamped with a date. It is not clear where they came from.
3. The Washington erratum is in the form of a letter dated February 25, 1926 from William E. Henry, Librarian.

Year	Source	Institutions Dealt With
1938/39	Virginia	Chicago ¹
1947/48	Virginia	Kansas ²
1951/52	Virginia	Washington, St. Louis ³

Notes:

1. This erratum is in the form of an undated printed postcard from M. Llewellyn Raney, Director.
2. This erratum is an undated mimeographed postcard from the "Princeton University Library."
3. This erratum is a mimeographed postcard dated "13 January 1953" from Lawrence Heyl, Associate Librarian at Princeton.



The Gerould Statistics Appendix Two:

Differences Found in the Values of Variables Reported by the Sources of Gerould Statistics Consulted

This appendix gives the cases where there is uncertainty in the values which might be used in this edition of the Gerould data. Part 1 lists the cases of uncertainty with the different possible numbers and the decision made in this edition between these possibilities. Part 2 lists the cases where the decision made might be wrong.

Uncertainty in the context of this appendix occurs because of a conflict between the values given by different sources of these data. The most common kind of conflict arises when two sources record a different value for a variable at an institution in a given year. For example, there are two different issues of the Gerould data for 1909/10, one of which reports that there were 16 staff members employed at Minnesota, while the other reports that Minnesota had 18 staff members. This appendix will discuss these kinds of conflicts as well as two others, both of which have only one example. Below, at Michigan, 1945/46 (see page 85), there is a discussion of a value a part of which is illegible and at Illinois, 1960/61 (page 82) there is a discussion of a case where a wrong value has been corrected but it is not clear what the corrected value should be. While for most of the cases of conflict an attempt is made to develop a set of principles to follow in adjudicating disputes between possible values, these two cases are handled in an *ad hoc* manner.

Part of the problem in assembling these data is sorting out which number is correct when there is more than one possibility. This process is made difficult by the fact that there are many places at which error can creep into any text, whether a play of Shakespeare or a set of numbers. In the case of these data, an error might be introduced before the data reached Gerould, of course, but we will concern ourselves with errors which arise after Gerould received the data.

Before the data for 1913/14 were published (when these data were apparently first distributed in mimeographed form) Gerould would have assembled the data and generated what might be called a master copy. In order to distribute the text, he would have to duplicate the master copy, which would have been done by having a copy typed

from the master. Before massive reproduction was possible, each distribution copy would have been produced separately and each new copy could vary from the master copy and from other distribution copies.

Making copies by mimeograph allowed the production of numerous identical copies of a text. As a result, deciding which of several numbers is correct should be easier after 1913/14 but is still complicated by three facts:

1. There are editions of these data made by people other than Gerould and at other institutions. Each new edition requires that the data be rearranged or retyped, processes which increase the chances for error.
2. Changes have been made in the distribution copies sent to the various libraries. For example, suppose identical distribution copies were sent to twenty institutions and suppose further that all of these distribution copies have been changed in the same way. Is the change in the text we see correct? What about cases where identical changes are made on half of the distribution copies?
3. There are instances where the text being used is barely legible.

In cases where there are more than one possibility for a given number, deciding which to choose has been guided by the attempt to establish which of the different candidates has the best authority. That is, what did the master text look like or what would it have looked like after a correction was sent out? In this context authority will be established by:

1. Using the earliest generation of the data possible. This principle follows from the notion that each new generation of a text allows (without necessarily promoting) the introduction of errors. This matter is serious enough to deserve separate treatment below.
2. Preferring changes in data from errata over the principle that the earliest generation should be used. In compiling data in the days of Gerould or now, mistakes are inevitable and corrections issued either by the compiler or by the library which reported the data should have priority over the master copy or the distribution copy made from it. For the lone exception to this principle see the listing below for Texas, 1913/14 (see pages 90-91.) The one change in the text of the

data made without an erratum or clear authority is reported at Louisiana State, 1954/55 (page 85.)

3. In cases where neither of the first two principles allows the selection between two possible values, use other variables or data from other years for clues about which value is correct.

There have been two ways used of presenting the Gerould data each year. One is to arrange the data alphabetically by the name of the institution and the other is to arrange the data for each institution in order by the number of volumes held. In the early years of the Gerould series there are two editions issued about the same time each year. There are stray examples of data arranged each way but there are two editions of these data which are readily identifiable because each appears to be consistently done on its own typewriter. Which of these editions is closer to the master copy is not absolutely clear but the alphabetical set has been treated as the earlier edition. In cases where the alphabetical text and the text ordered by size differ, the alphabetical text is assumed to have precedence. The case for this view is circumstantial:

1. There are instances of variant forms given below where it seems clear that the list ordered by size was done later than the alphabetical list. For example, in the cases of Texas, 1913/14, Harvard, 1924/25 and Colorado, 1924/25, there are handwritten changes made in various distribution copies of the alphabetical list which are included in the typed list ordered by size.
2. There is an issue of an alphabetically arranged edition of all years of the data from the beginning until Gerould retired but there is not an issue ordered by size of institution for all of those years.
3. There is a copy of "Preliminary statistics" from 1913/14 in the Illinois Archives which has what appear to be notes in Gerould's hand. This issue of these statistics is a predecessor to the master copy and is arranged alphabetically. The expenditure categories include cents. One note says: "Arrange by size....omit - cents-." It appears from this issue of the data that a copy ordered alphabetically was produced first.
4. A series of these data apparently done on the same typewriter and arranged alphabetically are in several libraries from about 1918/19 through the 1930's while

a series of these data arranged by size (and done on a different typewriter) are only at Illinois.

A fairly conservative view is taken here of making changes in what is presumed to be the text closest to the master copy. While the treatment of changes has been conservative in the text of the data, the inclusion of variant values has been fairly liberal in this appendix. For instance, the following variant values are included in the appendix but not in this edition of the data:

1. Handwritten changes to the typed text for which there is no authority, e.g. errata.
2. Values for the volumes held at Wisconsin's Historical Library. These values were reported for a number of years in the Gerould data as a footnote to the reports for the University of Wisconsin's library.
3. Values for 1909/10 from Chicago, Columbia, Cornell, Harvard, Johns Hopkins, Northwestern, Pennsylvania, Princeton, Stanford, and Yale which were collected by Phineas Winsor at Illinois using the same definitions as those used by Gerould.

Excluding these kinds of values in this edition of the Gerould data helps insure the integrity of the Gerould text but causes other kinds of problems. For instance, the values collected by Winsor using Gerould's definitions are probably as good as those collected by Gerould and handwritten changes may reflect an erratum noted at one institution but not another. Occasionally, the changes which lack authority seem more likely to be correct than the values chosen here. In a few cases, therefore, there is an apparent conflict between insuring the integrity of the Gerould text and insuring that the numbers are always correct. While using authority as the criterion leads to a dilemma in a handful of cases, it avoids accepting values which seem right because they fit some preconceived notion of how the data should lie. Future research may resolve many of these conflicts but it seems reasonable to note those values which meet the criteria used here, that is, there is good authority for the values chosen but the values not chosen might seem better in some circumstances. A list of these values is included at the end of this appendix.

The symbol "*" is used to indicate no value is given for the variable.

PART 1

Variant Values From the Various Sources

CALIFORNIA, BERKELEY

Year	Variable	Value Used in this Edition
1911/12	Appropriations	\$24,000

The Minnesota issue reports that appropriations were \$24,000 while the McMullen issue reports \$34,000. \$24,000 is assumed to be the correct value because this is the value on the Minnesota copy which must be either a later copy than the McMullen copy or closer to the master copy for reasons discussed under Chicago, 1911/12 and Wisconsin, 1911/12 in this Appendix.

There is reason to be suspicious of the value from the Minnesota copy, however, as the following list should make clear:

Year	Expenditures	Appropriations
1910/11	Not published	
1911/12	\$33,800	\$24,000 or \$34,000
1912/13	36,538	33,950

\$34,000 seems to fit in this series of numbers better, although it is not yet clear how the variable Appropriations behaves. That is, not enough is known about the behavior of this variable to know if a jump from an appropriation of \$24,000 to an expenditure of \$36,538 is an anomaly.

This value is listed in Part 2 as one which is not used but which may be correct.

CALIFORNIA, LOS ANGELES

Year	Variable	Value Used in this Edition
1929/30	Volumes Added	17,220

The edition of these data ordered by size records the number of volumes added as 17,200 while the edition ordered alphabetically records 17,220. The latter edition is presumed to be closer to the master copy so 17,220 is used here.

CHICAGO

Year	Variable	Value Used in this Edition
1909/10	All	*

In November, 1910 Phineas Winsor, the Librarian at Illinois, sent a letter to the directors of the libraries at Chicago, Columbia, Cornell, Harvard, Johns Hopkins, Northwestern, Pennsylvania, Princeton, Stanford, and Yale asking for information on "the six questions which Mr Gerould asked of the state universities." Data from the replies to these letters are included in this appendix but not in the edition of the Gerould data which is presented here. Data from Chicago first appear in the Gerould reports beginning in 1911/12, the next year the data appear. The values for Chicago, according to Winsor's survey are:

Year	Volumes	Volumes Added	Expenditures	Appropriations	Staff	Salaries
1909/10	310,000	22,871	\$38,484.54	\$41,679.95	60	\$25,860.66

The figure for appropriations is broken down into appropriations: \$25,265 and "Balances carried forward:" \$16,414.95.

CHICAGO (continued)

Year	Variable	Value Used in this Edition
1911/12	Salaries	\$43,216

The Minnesota issue of these data reports the total salaries as \$43,216 while that from McMullen reports this variable's value as \$52,570. This latter issue has a footnote indicating "there has been an additional appropriation of \$14,484 for student service." As was noted in the main text, Salaries was apparently regarded as distinct from wages. Salaries were paid to full-time employees while wages were paid to part-time workers like students. The language of this footnote and the inclusion of wages with salaries seem to suggest that the McMullen issue of this year's Gerould statistics was a retyping done at Chicago. The value on the Minnesota copy, \$43,216, is used here.

CHICAGO (continued)

Year	Variable
1922/23	Salaries

In the first edition, there was an error in the text that incorrectly documented this value. The error is discussed elsewhere. The errant value reported for this institution for this year was actually for Columbia and this has been corrected in this edition. This error affected only the text of the monograph, not the data.

CHICAGO (continued)

Year	Variable	Value Used in this Edition
1927/28	Salaries	\$230,857

The three copies of the alphabetical edition and the copy from the edition ordered by size have the figure for salaries for Chicago as \$230,857. The McMullen copy has a handwritten change of \$233,041.35. The McMullen copy comes from the University of Chicago so there is a strong temptation to regard this new figure as correct. There is a problem and that is: when was the change made, by whom, and was the definition followed in making the change the same as that used by Gerould in compiling the original figure? The argument advanced in the main text is that changes in the text of a series of data, such as the Gerould data, should be timely and widely distributed, and done for a clear reason. This change done at Chicago meets none of these criteria and is suspect until the reasoning behind the change is clear. Therefore, the original value, \$230,857, is used in this edition.

CHICAGO (continued)

Year	Variable	Value Used in this Edition
1938/39	Salaries	\$189,578.36

An erratum found attached to the Virginia issue of these data corrects the value reported for salaries from \$157,732.66 to \$189,578.36. This change is made on the Texas issue and on that from Illinois, but not on the Virginia issue.

COLORADO

Year	Variable	Value Used in this Edition
1920/21	Appropriations	\$13,000

The alphabetically ordered edition gives the figure of \$13,000 while the edition ordered by size gives this figure as \$15,000. \$13,000 is used in this edition of the Gerould statistics because it comes from a copy of the data deemed to be closer to the master copy.

COLORADO (continued)

Year	Variable	Value Used in this Edition
1924/25	All	*

The alphabetically ordered distribution copy of the data from 1924/25 has a blank line instead of reporting values for Colorado's library. Both the Virginia and McMullen issues have the values handwritten in the blank line while the Texas issue has no values written in. There is an erratum with the Virginia copy giving the following values, which have been included in this edition:

Year Volumes Volumes Expenditures Appropriations Staff Salaries

Added						
1924/25	167,000	9,000	7,500	7,700	15	24,085

These values are included in the distribution copies which are ordered by size.

COLUMBIA

Year	Variable	Value Used in this Edition
1909/10	All	*

Phineas Winsor collected data from Columbia in 1909/10. More detail on his effort is found above at Chicago, 1909/10. Data from Columbia first appear in the Gerould reports beginning in 1911/12, the next year the data appear. The data from Columbia for this year, according to Winsor are:

Year	Volumes	Volumes Added	Expenditures	Appropriations	Staff	Salaries
1909/10	446,112	16,090	\$24,279.85	\$37,424.34	42	\$41,700

COLUMBIA (continued)

Year	Variable	Value Used in this Edition
1911/12	Volumes Added	23,528

The Minnesota issue reports the number of volumes added as 23,528 while that from McMullen has this number as 23,522. 23,528 is used here because the Minnesota issued is presumed to be closer to the master copy.

COLUMBIA (continued)

Year	Variable	Value used in this edition
1922/23	Salaries	*

All four issues appear to be from the same distribution copy and they all have the original figure for salaries at Columbia lined out. The figure which has been lined out is not clear but it appears to be \$28,300. Consider the figures for this variable for these years:

Year	Salaries
1921/22	\$125,821.00
1922/23	28,300(?)
1923/24	*

Given the values for this variable from the year before and given the fact the same change was made on all four copies, this variable is treated here as if there were no value reported.

COLUMBIA (continued)

Year	Variable	Value used in this edition
1929/30	Staff	163

The alphabetically ordered edition reports Columbia employed 163 staff members while the edition ordered by size reports 183. Because the alphabetical list is presumed to be closer to the master copy the value from it is used here. This number appears to be correct as a comparison of the numbers of staff and salaries at Columbia from 1928/29 through 1930/31 shows that 163 appears to fit the related data better than 183 would.

Year	Staff	Salaries
1928/29	157	\$285,568.00
1929/30	163	308,202.00
1930/31	166	337,569.00

CORNELL

Year	Variable	Values used in this edition
1909/10	All	*

Phineas Winsor collected data on Cornell in 1909/10. More detail on his effort is found at Chicago, 1909/10. Data from Cornell first appear in the Gerould reports beginning in 1911/12, the next year the data appear. The data from Cornell for this year, according to Winsor are:

Year	Volumes	Volumes Added	Expenditures	Appropriations	Staff	Salaries
1909/10	383,696	14,645	\$20,116	\$21,500	18	\$22,660

CORNELL (continued)

Institution	Year	Variable	Value used in this edition
CORNELL	1931/32	Expenditures	\$38,638.00

The edition of these data ordered by size records a figure for Expenditures in 1931/32 as \$38,636.00 while the alphabetically ordered edition records a value of \$38,638.00. The latter edition is presumed here to be closer to the master copy and the value from it is used.

HARVARD

Institution	Year	Variable	Values used in this edition
HARVARD	1909/10	All	*

Phineas Winsor collected data from Harvard in 1909/10, as discussed above at Chicago, 1909/10. Data from Harvard first appear in the Gerould reports in 1913/14 although they appear only sporadically before 1919/20. Data from Harvard for this year, according to Winsor are:

Year	Volumes	Volumes Added	Expenditures	Appropriations	Staff	Salaries
1909/10	543,439	20,616	\$30,422	\$19,100	47	\$42,000

These figures apply “to the central or College library only.” Expenditures are further broken down into: \$21,962 “from regular funds” and \$8,460 “from current gifts for immediate use.”

HARVARD (continued)

Institution	Year	Variable	Values used in this edition
HARVARD	1912/13	All	*

The McMullen issue of this year's Gerould data has the following values handwritten on it:

Year	Volumes	Volumes Added	Expenditures	Appropriations	Staff Salaries
1912/13	1,083,750	40,318	\$106,965	*	116 \$99,745

It is not clear where these values came from. It can be speculated that Gerould hoped to have something to report for Harvard when the master copy was made and that he received a reply from Harvard late, reported the new values to those who were sent copies, and someone at Chicago wrote in these values for Harvard and did not keep the letter from Gerould. This case is a bit tenuous and we cannot be sure of the source for these numbers so they are not included in this edition.

This value is listed in Part 2 as one which is not used but which may be correct.

HARVARD (continued)

Institution	Year	Variable	Values used in this edition
HARVARD	1917/18	All	*

The McMullen issue of this year's Gerould data has the following values typed in with a different typewriter from that used in making the distribution copy:

Year	Volumes	Volumes Added
1917/18	1,854,800	61,800

As with the variables for 1912/13, it is not clear where these values came from and they are not included in this edition.

This value is listed in Part 2 as one which is not used but which may be correct.

HARVARD (continued)

Institution	Year	Variable	Value used in this edition
HARVARD	1924/25	Expenditures	\$157,040

The alphabetical distribution copy, the one presumably sent by Gerould, has no value for Expenditures for 1924/25 but the issues from Virginia and McMullen have been amended to include the value \$157,040. The Texas issue has not been changed. The issues from the Illinois stacks and Archives are arranged by size and are dated "Dec. 1925" and both have the same figure for Expenditures (i.e. \$157,040) on the distribution copies. It appears that someone sent an erratum after the distribution of the alphabetical copy and before the compilation of the Illinois copy. This erratum is missing but was accepted as genuine by the libraries which had received distribution copies of the Gerould data. The amended Expenditures figure is used in this edition.

HARVARD (continued)

Institution	Year	Variable	Value used in this edition
HARVARD	1936/37	Expenditures	\$292,924.27

The edition of these data ordered alphabetically records the expenditures for Harvard in 1936/37 as \$292,924.27 while the edition ordered by size records a value of \$292,924.26 for this variable. The value from the alphabetical list is presumed to be closer to the master copy and is used here.

ILLINOIS

Institution	Year	Variable	Value used in this edition
ILLINOIS	1936/37	Volumes	1,086,212

The Illinois stacks issue of these data has a handwritten correction to the value recorded in this edition. The value recorded in the distribution copies of both the edition ordered by size and the alphabetically ordered edition is 1,086,212 while the handwritten correction is 1,087,286. Even though this correction was done at Illinois on data from Illinois the value in the originals is used here for reasons given above under Chicago, 1927/28.

ILLINOIS (continued)

Institution	Year	Variable	Value used in this edition
ILLINOIS	1960/61	Volumes Added	95,246

The distribution copy records that Illinois added 66,789 volumes but this figure has been altered with a handwritten change indicating that either 95,266 volumes were added (on the Texas issue) or 95,226 volumes were added (on the Virginia and University Microfilms issues. The Virginia issue is either a Xerographic copy of the UM issue or they have both been copied from the same original.) The two changes appear to have been made by the same person which suggests the change was made at Princeton before the copies were distributed. The change, therefore, is acceptable by the criteria used here but which value should be used? There is no clear choice and the decision was made to take the average of the two values until some reason emerges for choosing one of the two. Following this procedure means that a value will be used here which does not exist in any of the texts consulted. While this solution to the conflicting values is not easy to defend on intellectual grounds, it is practical because the difference between the two values is so small. If a similar case arose where the differences were substantial, some other solution would have to be found for it and this case.

JOHNS HOPKINS

Institution	Year	Variable	Values used in this edition
JOHNS HOPKINS	1909/10	All	*

Phineas Winsor collected data from Johns Hopkins in 1909/10 as discussed at Chicago, 1909/10. Data from Johns Hopkins first appear in the Gerould reports beginning in 1911/12, the next year the data appear. The data from Johns Hopkins for this year, according to Winsor are:

Year	Volumes	Volumes Added	Expenditures	Appropriations	Staff	Salaries
1909/10	148,242	5,928	\$13,990.88	\$12,000	12	\$9,088.35

JOHNS HOPKINS (continued)

Institution	Year	Variable	Value used in this edition
JOHNS HOPKINS	1920/21	Salaries	\$21,713.50

The two issues from Illinois report \$21,713.50 spent for salaries. The McMullen issue has this same value but it has been crossed out and had \$24,763.50 handwritten beside it. There is no other authority for this change and it is not used here.

KANSAS

Institution	Year	Variable	Values used in this edition
KANSAS	1907/08	All	*

The McMullen issue has a blank line for Kansas while the Minnesota issue has variables for this year included. The values given in the Minnesota issue are included in this edition of the Gerould data.

KANSAS (continued)

Institution	Year	Variable	Value used in this edition
KANSAS	1921/22	All	*

The three issues of this year's data appear to have come from the same distribution copy but two of them have been amended. The issue from Texas has a blank line beside Kansas. Virginia's issue has been amended to include all six variables. These changes have been typed in on the original. An asterisk is beside the figure for expenditures referring to a note at the bottom of the page: "Binding is not a charge on the budget." The McMullen issue has the values for the variables volumes, volumes added, and expenditures handwritten on it. These three values are the same as those on the Virginia version of these data. In addition, the McMullen issue has an asterisk beside the expenditure figure with this note at the bottom of the page: "Not incl. binding". It appears that an erratum was sent out with the Kansas values and that there are no copies of this erratum. Virginia received and noted all changes but for some reason Kansas noted only three changes and Texas did not receive or did not note any changes. The changes recorded on the Virginia copy have been included in this edition. These values are:

Year	Volumes	Volumes Added	Expenditures	Appropriations	Staff	Salaries
1921/22	150,104	5,675	\$20,538.00	\$17,700.00	18	\$28,430.00

KANSAS (continued)

Institution	Year	Variable	Value used in this edition
KANSAS	1947/48	Salaries	\$60,924

The three distribution copies consulted here recorded \$660,924 as the amount of money spent on salaries at Kansas in 1947/48. An erratum corrects this figure to \$60,924.

LOUISIANA STATE

Institution	Year	Variables	Values used in this edition
LOUISIANA STATE	1954/55	Volumes	656,639
"	"	Volumes Added	32,108

The originals of the data for this year record the number of volumes as 32,108 and the number of volumes added as 656,639. The University Microfilms issue and the Virginia issue have the same change made on the text with arrows showing the two figures reversed. The Virginia issue is either a xerographic copy of the UM issue or they are both made from the same issue. The issue from Texas makes no note of this change. Although the change seems obvious there is no erratum. The values used here are the changed ones. A comparison of the values for these two variables for 1953/54 through 1955/56 is given below in order to show the context of this change in the distribution copy:

Year	Volumes	Volumes Added
1953/54	625,813	34,683
1954/55	656,639	32,108
1955/56	690,529	34,204

MICHIGAN

Institution	Year	Variable	Value used in this edition
MICHIGAN	1945/46	Volumes Added	*

In the issues of the data from 1945/46 the middle digit for volumes added is illegible, appearing as "28, 87." The middle digit, of course, can be any number which creates a conflict which is not easily resolvable because there are no good choices in this circumstance. A solution offered at the end of this Appendix, which lists values which are excluded from this edition of these data but which may be may be useful in some circumstances, suggests that substituting a 5 for the blank may be useful in some calculations, as was done in Molyneux, 1994a. The solution adopted here is to treat this

value as if it were not reported at all. This value is listed in Part 2 as one which is not used but which may be correct.

MINNESOTA

Institution	Year	Variable	Value used in this edition
MINNESOTA	1909/10	Staff	18

The McMullen copy gives the number of staff as 16, while the copy from the Illinois Archives gives the number as 18.

These two copies appear to be two separate distribution copies done from the same master. In the process of typing one of the copies an error was made. The number of staff and salaries at Minnesota for 1907/08, 1909/10, and 1911/12 are as follows:

Year	Staff	Salaries
1907/08	13	\$12,490
1908/09	<i>Not published</i>	
1909/10	16 or 18	18,890
1910/11	<i>Not published</i>	
1911/12	24	24,245.00

Over the period, the figure for salaries increased approximately \$3,000 a year and in a regular fashion; that is, it hasn't increased by a small amount from 1907/08 to 1909/10 and then a larger amount from 1909/10 to 1911/12. It is likely that the staff follow a similar regular progression. 18 is 5 more than 13 and 6 less than 24, a more orderly series than 3 (from 13 to 16) and 8 (16 to 24).

NEBRASKA

Institution	Year	Variable	Values used in this edition
NEBRASKA	1925/26	Appropriations	\$41,000
"	"	Salaries	81,000

The alphabetically arranged distribution copy records \$41,000 for appropriations and \$81,000 for salaries. The issue from Texas has had the appropriations figure changed to \$40,000 and the figures for salaries changed to either \$40,000 or \$41,000. The source for these changes is not given. The other copy of these data, the one arranged by size gives the same values as the alphabetical list. The changes made to the issue from Texas are not used in this edition because there is no authority for them.

Even though these changes are not accepted they may be correct. Consider these data from Nebraska for 1924/25 through 1926/27:

Year	Staff	Salaries
1924/25	23	\$33,940.00
1925/26	30	81,000.00 or 40,000
1926/27	32	41,000.00

It is difficult to see how the 1925/26 figure could be correct although the criteria used here to decide correct values force the acceptance of \$81,000.

This value is listed in Part 2 as one which is not used but which may be correct.

NORTHWESTERN

Institution	Year	Variables	Values used in this edition
NORTHWESTERN	1909/10	All	*

Phineas Winsor collected data from Northwestern in 1909/10, as discussed above at Chicago, 1909/10. Data from Northwestern first appear in the Gerould reports beginning in 1912/13. Data from Northwestern in this year are:

Year	Volumes	Volumes Added	Expenditures	Appropriations	Staff	Salaries
1909/10	123,595	9,531	\$8,642.42	\$10,973.31	10	\$7,820

The figures for volumes and volumes added include departmental libraries. A footnote to the figure for expenditures says: "Besides the Law School expends as much again." A note to appropriations says the law school "has about \$8-10,000."

NORTHWESTERN (continued)

Institution	Year	Variables
NORTHWESTERN	1923/24	All

The three issues of this year's data appear to come from the same distribution copy. In two of these issues, those from McMullen and from Virginia, there is a blank line beside the value for Northwestern. The Texas copy, however, has a set of values handwritten in. There is an erratum in the Illinois Archives for this set of data. This erratum is undated but, given the note on the Texas issue for this year, and the surrounding values, it seems probable this erratum is for 1923/24. These values, along with those for Northwestern from a few years before and after are reproduced here:

Year	Volumes	Volumes Added	Expenditures	Appropriations	Staff	Salaries
1921/22	219,448	11,401	\$15,688.43	\$17,319.00	13	\$23,810.22
1922/23			<i>Not reported</i>			
1923/24	236,946	12,283	27,963.31	28,000.00	21	31,594.87
1924/25	247,768	10,154	38,744.00	28,700.00	21	49,130.00

The values for volumes and volumes added which are handwritten on the Texas issue are slightly different from those on the erratum. The number of volumes held on this issue is 236,945 (not 236,946) and the number of volumes added is 12,285 (not 12,283.)

OHIO STATE

Institution	Year	Variables	Values used in this edition
OHIO STATE	1923/24	Expenditures	\$47,718.33
"	"	Appropriations	61,376.67

The three issues for this year record \$54,547.50 as the amount expended for books in 1923/24 and appropriated for next year's book purchases but in all three cases there is a change made on the original. In each case, a line is drawn through the two values and

the values used here typed in on the bottom of the page by the same typewriter. It looks as if an error was caught after the distribution copy was mimeographed but before the issues were mailed.

PENNSYLVANIA

Institution	Year	Variables
PENNSYLVANIA	1909/10	All

Phineas Winsor collected data from Pennsylvania in 1909/10 as discussed above at Chicago, 1909/10. Data from Pennsylvania first appear in the Gerould reports beginning in 1911/12, the next year the data appear. The data for Pennsylvania for this year, according to Winsor are:

Year	Volumes	Volumes Added	Expenditures	Appropriations	Staff	Salaries
1909/10	313,772	15,246	\$25,412.34	*	23	\$20,000

A note under appropriations says this figure is about the same as expenditures “but probably more. Appropriations vary each year.”

PRINCETON

Institution	Year	Variables
PRINCETON	1909/10	All

Phineas Winsor collected data from Princeton in 1909/10 as discussed above at Chicago, 1909/10. Data from Princeton first appear in the Gerould reports beginning in 1911/12, the next year the data appear. The data for Princeton for this year, according to Winsor are:

Year	Volumes	Volumes Added	Expenditures	Appropriations	Staff	Salaries
1909/10	264,242	14,258	\$14,510.81	*	35	\$20,300

A note under appropriations reports this figure is the “same approximately” as the expenditure figure.

STANFORD

Institution	Year	Variables	Value used in this edition
STANFORD	1909/10	All	*

Phineas Winsor collected data from Stanford in 1909/10 as discussed above at Chicago, 1909/10. Data from Stanford first appear in the Gerould reports beginning in 1911/12, the next year the data appear. Data for Stanford for this year are:

Year	Volumes	Volumes Added	Expenditures	Appropriations	Staff	Salaries
1909/10	143,881	15,118	\$29,300	\$32,000	19	\$19,240

TEXAS

Institution	Year	Variables
TEXAS	1913/14	All

The Illinois Archives has a copy of a letter from Phineas Winsor, University Librarian at Illinois, to Gerould dated November 17, 1914. In this letter, Winsor says he has “received a personal letter from Mr. Godwin at the University of Texas, in which he gives me the corresponding figures for his own library. You may be interested in them.” He lists the values and notes “I am adding these Texas figures to the copies of the summary which I am using here in Urbana.” There is also a letter in the Archives from Gerould to Winsor (dated “November 19, 1914”) saying, in part: “I am sending on to the institutions which have received copies of the table of university library statistics, the figures from the University of Texas Library.” A letter of the same date, with the heading “UNIVERSITY LIBRARY STATISTICS (*sic*)” from Gerould and addressed “Dear Sir” asks that data from Texas be added “to the table of University Library Statistics recently sent you.”

Unfortunately, this letter mistranscribes the number of volumes held by Texas from the value reported by Winsor to Gerould (100,774) to 100,744. This letter must be a copy of the general mailing which might have been retyped for each institution receiving the data so this mistake may not have been repeated on other issues of this erratum. However, the McMullen issue has these values from Texas written in after Yale (that is, in the first blank line) and for volumes the figure given is 100,744. The Minnesota issue

has the Texas values typed in after Yale and also has 100,744 volumes. The Illinois stacks issue of these data, which unlike the others is arranged by size, has Texas typed in, in proper order, and with 100,774 volumes. The Texas data are included in this edition of the Gerould data and the number of volumes used here is 100,774, the value in the Winsor letter. In this case, the value in this letter must take precedence over the erratum. Data from Texas are included on subsequent reports from Gerould. The data from 1913/14 are:

Year	Volumes	Volumes Added	Expenditures	Appropriations	Staff	Salaries
1913/14	100,774	11,694	\$17,000	\$21,000	18	\$15,040

TEXAS (continued)

Institution	Year	Variables
TEXAS	1926/27	All

The alphabetical edition of this year's data has blanks for Texas and Wisconsin. The issues at Texas and Virginia were left blank while values for all variables for these two libraries were typed in on the McMullen copy. Both of these sets of values have footnotes saying "See attached slip," although no slip is attached. The same values were handwritten on the edition of these data ordered by size. The values for Texas and Wisconsin are included here because they appear to have come from an erratum which has not been found. The values from the years on either side of this year are worth considering, though, because the numbers are not in good order. The values for Texas for 1925/26 through 1927/28 are:

Year	Volumes	Volumes Added	Expenditures	Appropriations	Staff	Salaries
1925/26	416,523	17,041	\$32,862.83	\$30,000.00	30	\$66,130.00
1926/27	388,016	31,362	74,818.54	35,000.00	26	52,530.00
1927/28	411,310	20,161	59,912.67	35,000.00	26	52,900.00

TEXAS (continued)

Institution	Year	Variable	Value used in this edition
-------------	------	----------	----------------------------

TEXAS 1937/38 Volumes 564,360

The edition of these data ordered by the size of the institution records the number of volumes held at Texas in 1937/38 as 564,350 while the value recorded in the alphabetical list is 564,360. Because this latter edition is deemed closer to the master copy the value from that edition is used here.

WASHINGTON (SEATTLE)

Institution	Year	Variables
WASHINGTON (SEATTLE)	1924/25 and 1925/26	Volumes, Volumes Added

The Virginia issue of this year's data has a letter dated February 25, 1926 from William E. Henry, the Librarian at the University of Washington's library, correcting the figures for Volumes and Volumes Added given in the Gerould report. The distribution copy reports 156,416 volumes held and 26,030 Volumes Added while the erratum asks that these values be changed to 143,840 and 15,391, respectively. These changes have been made in this edition.

The issue from Texas does not include this correction, nor do the issues from Illinois's stacks or Archives. The Illinois issues are dated "Dec. 1925" and were, therefore typed before Henry's letter was sent. The issue from the Illinois stacks for 1925/26, dated "December, 1926" has had a handwritten change made to the volumes held and volumes added figures. Volumes has been changed to 145,840 (or 145,640—the change is not legible) and Volumes Added has been changed to 15,391. It is clear from the dates that the changes made to the 1925/26 issue were made on the wrong year's data. The Volumes figure, in any case, was mistranscribed.

WASHINGTON (SEATTLE) (continued)

Institution	Year	Variable	Value used in this edition
WASHINGTON (SEATTLE)	1929/30	Volumes	207,069

The number of Volumes reported for Washington in the alphabetically ordered edition is 207,069 while the edition ordered by size reports 207,067 volumes held. Because the alphabetical edition is deemed to be closer to the master copy, the value from that edition is used here.

WASHINGTON (ST LOUIS)

Institution	Year	Variable	Value used in this edition
WASHINGTON (ST. LOUIS)	1951/52	Staff salaries	\$137,106
"	"	Student salaries	10,290

An erratum dated "13 January 1953" from Lawrence Heyl at Princeton University corrects the values in the original issues (\$59,190 and \$17,570, respectively) to those used here.

WISCONSIN

Institution	Year	Variable	Value used in this edition
Wisconsin	1911/12	All	
"	1913/14	Volumes	210,000

The McMullen issue of the 1911/12 data says “(Figures not yet received)” for Wisconsin. The Minnesota issue has values for all six variables. The values from the Minnesota issue are used here.

The McMullen and Minnesota issues from this year report 210,000 volumes held while the alphabetically ordered issue in the Illinois stacks reports another number which is barely legible. What is visible may be either “393,000” or “398,000.” This number is probably 393,000 because the footnote says the count “includes state Hist. Soc. Lib. 183,000 vols. in same building.” 210,000 plus 183,000 equals 393,000. For reasons given at the beginning of this Appendix, the alphabetically ordered list is deemed not to be a distribution copy from Minnesota and might be a copy made (presumably at Illinois) from a distribution copy. This issue of these data includes a change in the data which exists in no other issue of these data, namely the inclusion of the volumes held at the Wisconsin Historical Society Library. In most of the Gerould data the number of volumes held by Wisconsin is reported with a note similar to the one found in this year: “Figures do not include State Hist. Library numbering 183,000 vols. in same building.”

The Wisconsin library's statistics, as they are now published in the *ARL Statistics*, include the holdings of Wisconsin's Memorial Library and the State Historical Library. The holdings of the main library at Wisconsin are included in this edition of the Gerould statistics but footnotes often record the number of volumes held at the State Historical Library. These figures gleaned from the various years' editions are:

Year	Volumes	Year	Volumes
1911/12	175,000	1923/24	234,000
1912/13	<i>Not reported</i>	1924/25	<i>Not reported</i>
1913/14	183,000	1925/26	242,000
1914/15	185,000	1926/27	248,000
1915/16	193,000	1927/28	255,000
1916/17	201,000	1928/29	<i>Not reported</i>
1917/18	201,000	1929/30	<i>Not reported</i>
1918/19	212,275	1930/31	271,000
1919/20	218,000	1931/32	271,000

1920/21	220,000	1932/33	<i>Not reported</i>
1921/22	226,000	1933/34	282,000
1922/23	228,000	1934/35	286,000

There is no explanation of the repeat in 1917/18 of the figure from 1916/17 or of the repeat in 1931/32 of the figure from 1930/31. There are no other data on the State Historical Library until 1949/50 when it is noted there are “720,246 physical volumes and about 850,000 manuscripts.”

WISCONSIN (continued)

Institution	Year	Variable
WISCONSIN	1926/27	All

Values for 1926/27 were added after the alphabetically ordered edition was distributed and before the edition ordered by size was distributed. See TEXAS, 1926/27 for a fuller explanation. The values have been included in this edition of these data:

Year	Volumes	Volumes Added	Expenditures	Appropriations	Staff	Salaries
1926/27	377,452	18,385	\$61,673.55	\$53,700.00	35	\$79,765.00

YALE

Institution	Year	Variable	Values used in this edition
YALE	1909/10	All	*

Phineas Winsor collected data from Yale in 1909/10, as discussed above at Chicago, 1909/10. Data from Yale first appear in the Gerould reports beginning in 1911/12, the next year the data appear. Data from Yale for this year according to Winsor, are:

Year	Volumes	Volumes Added	Expenditures	Appropriations	Staff	Salaries
1909/10	600,000	30,000	\$26,843	\$27,400	47	\$37,233

The volumes figure is reported as “about 600,000.” The reply to volumes added says “Total accessions, bound + unbound, 30,00 (*sic.*)” This mistake may be a result of carelessness on the part of Dr. J.C. Schwab, who apparently collected the data at Yale, or it may be the figure on the original sheet was, in fact, 30,000 but the last zero was cut off when the copy which was used in this compilation was made. This copy was not aligned properly on the copying machine and, as a result, part of this number may be missing.

YALE (continued)

Institution	Year	Variable	Value used in this edition
YALE	1911/12	Volumes	900,000

The McMullen issue reports 600,000 while Minnesota reports 900,000. For reasons given at Chicago, 1911/12 and Wisconsin, 1911/12 the Minnesota issue is believed to be later and closer to the master copy.

In 1912/13 Yale is reported to have 1,000,000 volumes. 1911/12 is the first year data from Yale appears in the Gerould data. Of course, there is a 1909/10 figure (600,000) from the Winsor data (see Yale, 1909/10).

PART 2

Values Not Used in This Edition But Which May Be Correct

The biggest group of the values not included in this edition of the Gerould data which may be correct is the set of data collected by Phineas Winsor in 1909/10. These data are not, strictly speaking, Gerould data but Winsor made an effort to collect these data using the definitions used by Gerould. There may be circumstances where these data are useful. See above the listings under Chicago, 1909/10; Columbia, 1909/10; Cornell,

1909/10; Harvard, 1909/10; Johns Hopkins, 1909/10; Northwestern, 1909/10; Pennsylvania, 1909/10; Stanford, 1909/10; and Yale, 1909/10 for these values. Other values which may be correct are given below. These values are discussed above at the linked text.

CALIFORNIA, BERKELEY, 1911/12

Institution	Year	Variable	Value used in this edition	Value not used in this edition
CALIFORNIA, BERKELEY	1911/12	Appropriations	\$24,000	\$34,000

HARVARD, 1912/13

Institution	Year	Variables
HARVARD	1912/13	All

The values given in the McMullen issue for 1912/13 for Harvard may be correct.

HARVARD, 1917/18

Institution	Year	Variable
HARVARD	1917/18	All

The values given in the McMullen issue for 1917/18 for Harvard may be correct.

MICHIGAN, 1945/46

Institution	Year	Variable
MICHIGAN	1945/46	Volumes Added

The distribution copy of these data give the value of volumes added at Michigan in 1945/46 as "28, 87" with the middle digit illegible. The decision was made to exclude this

value from these data and treat it as a missing value because the missing digit can be any number from 0 through 9. In many circumstances, treating the missing digit as a “5” and reading the number as “28,587” will give a usable approximation of the number of volumes added. As mentioned above, this practice was followed in Molyneux, 1994a.

NEBRASKA, 1925/26

Institution	Year	Variable	Value used in this edition	Value not used in this edition
NEBRASKA	1925/26	Salaries	\$81,000	\$40,000

The value used does not fit the context of related variables for Nebraska as well as the value for which there is no authority. \$40,000 seems to fit better than \$81,000. See [above](#) for the related values.

VIRGINIA, 1941/42

Institution	Year	Variable	Value used in this edition	Value not used in this edition
VIRGINIA	1941/42	Salaries	\$81,089.22	\$78,509.42

These values are discussed at some length in Chapter 2, Section 2.3.1



The Gerould Statistics Appendix Three:

Differences in Data Between the Gerould Statistics and the Princeton Compilation

This appendix lists differences between the 1947 compilation of the data (*College and University Library Statistics*) done at Princeton and the data reported in the original issues of the Gerould statistics as they appear in this edition of these statistics. This compilation is discussed in Chapter 2, beginning in section 2.3.1. (See page 25 and please see Appendix 5 (page 146) for more information on obtaining these data.

The values for dollar variables have been rounded to the nearest whole dollar in the Princeton compilation. In a large number of cases the compilers apparently did not round the values correctly. It is possible, though unlikely, that these many cases may have actually resulted from the correction of errors in the data issued by Gerould. This list of differences between the two sets of numbers has ignored those cases where there are apparent rounding errors. They are numerous and not large. Those interested can find them with the machine-readable data.

Data from the library at the University of South Dakota which appeared in the issues of the Gerould statistics from 1907/08—1924/25 were not included in the Compilation. The symbol “*” will be used to indicate that there is no value for the given variable in that year and source. The variable recording expenditures for books, periodicals and binding is labeled “Expenditures.”

BROWN

Year	Variable	Gerould	Princeton
1922/23	Salaries	*	28,300
1930/31	Volumes Added	21,485	21,435

BRYN MAWR

Year	Variable	Gerould	Princeton
1934/35	Expenditures	13,604.00	18,604

CALIFORNIA, BERKELEY

Year	Variable	Gerould	Princeton
1935/36	Appropriations	*	102,000
1936/37	"	*	113,000
1937/38	"	*	160,700
1938/39	Salaries	119,831.47	179,062

CALIFORNIA, LOS ANGELES

Year	Variable	Gerould	Princeton
1932/33	Salaries	56,750.00	66,100

CHICAGO

Year	Variable	Gerould	Princeton
1925/26	Salaries	188,337.89	181,338
1936/37	Volumes Added	34,259	34,269

CINCINNATI

Year	Variable	Gerould	Princeton
1940/41	Salaries	54,634.02	54,632

COLORADO

Year	Comment
1922/23	Not reported in the original.
1923/24	Not reported in the original.
1930/31	Not reported in the original.

The Princeton compilation reports the following values for these years:

Year	Volumes	Volumes Added	Expenditures	Appropriations	Staff	Salaries
1922/23	143,885	9,200	10,500	31,045	13	16,070
1923/24	153,289	9,404	14,230	31,095	14	17,940
1930/31	235,000	13,163	25,000	45,260	16.5	24,550

Other differences between the two sets of data:

Year	Variable	Gerould	Princeton
1924/25	Volumes	167,000	163,470
"	Volumes Added	9,000	10,181
"	Expenditures	7,500.00	15,650
"	Appropriations	7,700.00	31,785
"	Staff	15	15.5
"	Salaries	24,085.00	19,150
1925/26	Staff	24	14

COLUMBIA

Year	Variable	Gerould	Princeton
1921/22	Appropriations	*	75,238
1922/23	"	*	73,337
"	Salaries	*	138,844

1923/24	Appropriations	*	71,855
"	Salaries	*	257,098
1924/25	Appropriations	*	87,337
"	Salaries	*	168,000
1925/26	"	*	183,584

CORNELL

Year	Variable	Gerould	Princeton
1923/24	Appropriations	*	40,206
"	Staff	36	32
"	Salaries	52,437.00	46,640
1924/25	Volumes	724,453	724,452
"	Appropriations	*	39,514
"	Staff	22	34
"	Salaries	58,215.00	51,560
1925/26	Appropriations	*	41,767
"	Staff	32	33
"	Salaries	57,255.00	52,040
1926/27	Volumes	787,127	761,868
"	Appropriations	*	42,048
"	Staff	20	33
"	Salaries	71,125.00	53,706
1927/28	Appropriations	*	44,375
"	Staff	20	35
"	Salaries	58,700.00	55,540

DARTMOUTH

Year	Variable	Gerould	Princeton
1922/23	Appropriations	*	27,440
1926/27	Salaries	39,600.00	34,435
1927/28	“	57,000.00	39,577
1928/29	“	62,000.00	54,140
1929/30	Volumes	331,146	306,496
1930/31	“	330,000	328,905
1939/40	Staff	56	58
1940/41	Appropriations	90,536.00	56,300
“	Staff	49.5	61
1941/42	Appropriations	97,954.00	54,250
“	Staff	56	59
1942/43	Appropriations	98,900.00	53,547

DUKE

Year	Variable	Gerould	Princeton
1933/34	Appropriations	102,500.00	102,000
1937/38	Salaries	99,519.24	112,161
1940/41	“	103,584.98	118,773
1941/42	“	111,989.33	126,160
1942/43	Appropriations	*	96,245
“	Salaries	98,238.87	117,686

HARVARD

Year	Variable	Gerould	Princeton
1919/20	Salaries	*	67,390
1920/21	"	*	89,377
1921/22	"	*	91,557
1924/25	Expenditures	157,040.00	*
1931/32	Staff	131	118

ILLINOIS

Year	Variable	Gerould	Princeton
1919/20	Volumes	440,372	549,873
"	Volumes Added	21,423	36,884
"	Salaries	52,384.00	54,757
1920/21	Volumes	456,503	581,193
"	Volumes Added	17,352	31,924
"	Salaries	55,184.00	57,984
1921/22	Volumes	480,595	628,595
"	Volumes Added	28,873	48,006
"	Salaries	77,795.00	80,933
1922/23	Volumes	541,942	677,534
"	Volumes Added	28,506	49,543
"	Salaries	108,407.00	101,002
1923/24	Volumes	567,787	720,697
"	Volumes Added	25,863	43,767
1924/25	Volumes	655,139	762,227
"	Volumes Added	32,678	42,134
"	Salaries	121,898.00	105,382

ILLINOIS (continued)

1925/26	Volumes	687,345	810,948
"	Volumes Added	32,206	49,325
"	Salaries	110,494.00	106,644
1926/27	Volumes	708,850	854,947
"	Volumes Added	25,522	44,603
"	Salaries	132,838.00	116,086
1927/28	Volumes	768,329	903,137
"	Volumes Added	59,579	48,794
"	Expenditures	127,006.00	127,007
"	Salaries	143,302.00	123,841
1928/29	Volumes	800,330	955,931
"	Volumes Added	32,001	53,398
"	Expenditures	117,599.00	117,600
1929/30	Volumes	841,395	1,006,900
"	Volumes Added	35,086	51,573
"	Salaries	158,409.00	133,819
1930/31	Volumes	887,884	1,060,733
"	Volumes Added	36,489	54,437
"	Salaries	164,769.94	135,770
1931/32	Volumes	916,452	1,113,733
"	Volumes Added	38,568	53,604
"	Expenditures	103,634.00	109,560
"	Salaries	183,868.00	155,995
1932/33	Volumes	947,043	1,167,978
"	Volumes Added	30,591	54,849
"	Expenditures	94,456.20	97,414
"	Salaries	169,179.00	145,535

ILLINOIS (continued)

1933/34	Volumes	978,212	1,220,746
"	Volumes Added	31,169	53,372
"	Expenditures	94,648.00	98,927
"	Salaries	134,389.00	127,586
1934/35	Volumes	1,011,938	1,273,446
"	Volumes Added	34,155	53,304
"	Expenditures	103,535.86	106,497
"	Salaries	161,776.00	133,554
1935/36	Volumes	1,052,684	1,335,456
"	Volumes Added	40,746	62,614
"	Expenditures	113,910.70	116,988
"	Salaries	171,742.00	142,266
1936/37	Volumes	1,086,212	1,395,691
"	Volumes Added	33,528	60,888
"	Expenditures	117,557.89	128,908
"	Salaries	151,318.00	145,061
1937/38	Volumes	1,130,584	1,468,684
"	Volumes Added	43,298	73,749
"	Expenditures	139,565.00	151,734
1937/38	Volumes	1,175,692	1,545,344
"	Volumes Added	45,108	77,393
"	Expenditures	138,859.99	154,170
1939/40	Volumes	1,209,977	1,618,579
"	Volumes Added	34,285	73,998
"	Expenditures	135,189.39	141,923
"	Salaries	204,552.00	208,407

ILLINOIS (continued)

1940/41	Volumes	1,262,046	1,687,847
"	Volumes Added	44,371	70,002
"	Expenditures	133,724.68	146,674
"	Appropriations	137,500.00	151,200
"	Salaries	244,124.00	212,903
1941/42	Volumes	1,306,561	1,765,203
"	Volumes Added	44,515	78,154
Year	Variable	Gerould	Princeton
"	Expenditures	156,682.00	151,229
"	Salaries	244,124.00	235,719
1942/43	Volumes	1,364,906	1,834,437
"	Volumes Added	58,455	69,733

INDIANA

Year	Variable	Gerould	Princeton
1927/28	Staff	19	10
1930/31	Volumes	218,800	232,800
1941/42	"	436,183	*

IOWA

Year	Variable	Gerould	Princeton
1923/24	Salaries	33,760.00	28,633
1924/25	"	59,014.00	50,607
1925/26	"	73,719.80	59,698
1926/27	"	73,696.95	56,165
1927/28	"	77,785.00	58,922

IOWA (continued)

1928/29	"	78,445.00	61,020
1929/30	"	83,640.00	65,186
1930/31	"	96,174.11	75,440
1931/32	Appropriations	*	50,000
"	Salaries	97,476.00	72,834
1932/33	"	90,096.00	66,675
1933/34	"	82,789.00	59,923
Year	Variable	Gerould	Princeton
1934/35	"	83,760.00	60,488
1935/36	"	83,823.00	61,613
1936/37	"	87,473.78	65,708
1937/38	Volumes Added	14,595	14,593
"	Salaries	98,869.95	74,353

JOHNS HOPKINS

Year	Variable	Gerould	Princeton
1920/21	Salaries	21,713.50	21,715

KANSAS

Year	Variable	Gerould	Princeton
1921/22	Volumes	150,104	149,140
"	Volumes Added	5,675	5,359
"	Expenditures	20,538.00	21,500
"	Appropriations	17,700.00	33,545
"	Staff	18	19
"	Salaries	28,430.00	22,116

MICHIGAN

Year	Variable	Gerould	Princeton
1936/37	Expenditures	133,673.00	133,676

MINNESOTA

Year	Variable	Gerould	Princeton
1920/21	Salaries	65,030.00	63,030
1936/37	Appropriations	93,885.00	93,835
1937/38	Staff	87	97
1937/38	"	88	98
1939/40	"	126	95
1941/42	"	88	98
1942/43	"	94	98

MISSOURI

Year	Variable	Gerould	Princeton
1932/33	Salaries	40,506.00	36,175

NEBRASKA

Year	Variable	Gerould	Princeton
1922/23	Expenditures	27,000.00	27,400
1924/25	Volumes	185,575	195,575
1925/26	Appropriations	41,000.00	40,000
"	Salaries	81,000.00	41,000
1935/36	Staff	25	35

NEW YORK

Year	Variable	Gerould	Princeton
1941/42	Staff	104	99
1942/43	"	60	67

NORTH CAROLINA

Year	Variable	Gerould	Princeton
1930/31	Salaries	55,922.36	46,025
1931/32	"	45,317.00	37,035
1932/33	"	25,806.00	35,437
1933/34	Volumes	254,552	264,552
"	Salaries	33,006.00	28,423
1934/35	"	44,537.00	30,412
1937/38	Staff	30	20
1937/38	Appropriations	40,000.00	25,000
1939/40	"	25,500.00	25,000
1940/41	Volumes	403,051	405,051
1942/43	Expenditures	38,437.01	38,369

NORTH DAKOTA

Year	Variable	Gerould	Princeton
1927/28	Volumes	80,000	95,215
1939/40	Appropriations	*	14,862
1940/41	"	8,500.00	14,026

NORTHWESTERN

Year	Comment
1922/23	Not reported in the original.

The Princeton compilation reports the following values for this year:

Year	Volumes	Volumes Added	Expenditures	Appropriations	Staff	Salaries
1922/23	225,061	8,320	18,332	20,915	*	26,392

Other differences between the two sets of data:

Year	Variable	Gerould	Princeton
1923/24	Volumes	236,945	237,821
"	Volumes Added	12,285	12,760
"	Expenditures	27,963.31	*
"	Appropriations	28,000.00	*
"	Staff	21	*
"	Salaries	31,594.87	*
1924/25	Staff	21	*
1929/30	Volumes	423,744	354,601
1930/31	Volumes	575,245	388,055
1931/32	"	605,288	418,098
1932/33	"	652,689	452,272
"	Volumes Added	41,121	34,194
1933/34	Volumes	719,412	483,949
"	Volumes Added	43,933	31,657

OBERLIN

Year	Variable	Gerould	Princeton
1925/26	Salaries	*	34,647
1929/30	Staff	27	32.8
"	Salaries	55,972.00	51,051
1930/31	Staff	26	32
"	Salaries	49,480.00	54,451
1931/32	Staff	42	37
"	Salaries	49,070.00	55,538
1932/33	Staff	30	36
"	Salaries	58,737.00	53,887
1933/34	Staff	36	33
"	Salaries	47,830.00	45,036
1934/35	Staff	30	33
"	Salaries	49,978.00	43,673
1935/36	Staff	33	32.6
"	Salaries	51,100.00	46,252
1936/37	Staff	34	33
"	Salaries	54,420.00	54,092
1937/38	Staff	35	34
"	Salaries	56,750.00	54,492

OHIO STATE

Year	Variable	Gerould	Princeton
1923/24	Expenditures	47,718.33	54,547
"	Appropriations	64,376.67	54,547
1930/31	"	*	36,935
1935/36	"	*	45,859

OREGON

Year	Variable	Gerould	Princeton
1921/22	Salaries	26,965.41	20,176
1922/23	Staff	16	13
"	Salaries	32,085.84	21,359
1923/24	"	33,556.52	24,400
1924/25	Staff	19	16
"	Salaries	36,099.00	24,413
1926/27	"	41,987.40	29,555
1927/28	"	46,915.00	30,849
1928/29	Staff	27	25
1929/30	"	27	25
"	Salaries	55,349.00	35,082
1930/31	Staff	26	24
"	Salaries	59,630.70	39,363
1931/32	Volumes	235,645	235,643
1935/36	Staff	25	24
"	Salaries	53,058.00	36,475
1932/33	Staff	19	23
"	Salaries	44,680.00	34,108
1933/34	Staff	21	23
"	Salaries	40,621.00	28,587
1934/35	Staff	24	23
"	Salaries	32,874.00	31,911
1935/36	Staff	25	24
"	Salaries	41,758.00	33,048
1936/37	Staff	22	24
"	Salaries	39,921.24	31,530

OREGON (continued)

1937/38	Staff	23	24
1937/38	"	23	24
1940/41	"	24.5	24
"	Salaries	42,457.64	42,790
1941/42	Staff	28	24
1942/43	"	24	27

PENNSYLVANIA

Year	Variable	Gerould	Princeton
1930/31	Expenditures	83,877.50	93,877
1942/43	Appropriations	*	62,592

PRINCETON

Year	Variable	Gerould	Princeton
1925/26	Volumes	593,954	595,954

ROCHESTER

Year	Variable	Gerould	Princeton
1925/26	Volumes	140,380	125,380
"	Volumes Added	13,826	8,900
1926/27	Volumes	163,700	133,488
"	Volumes Added	8,101	8,108
1927/28	Volumes	163,015	142,165
"	Volumes Added	8,647	8,677
1928/29	Volumes	175,839	153,791
1929/30	"	191,266	167,514
"	Volumes Added	15,427	13,723

STANFORD

Year	Variable	Gerould	Princeton
1921/22	Staff	34	32
"	Salaries	58,900.00	54,000
1927/28	Appropriations	*	53,615
1942/43	"	55,000.00	54,000

TEXAS

Year	Comment
1922/23	Not reported in the original.
1924/25	Not reported in the original.
1929/30	Not reported in the original.

The Princeton compilation reports the following values for these years:

Year	Volumes	Volumes Added	Expenditures	Appropriations	Staff	Salaries
1922/23	294,308	25,746	*	*	27	40,960
1924/25	325,873	13,404	*	*	25	41,210
1929/30	407,536	14,507	*	*	29	43,205

Other differences between the two sets of data:

Year	Variable	Gerould	Princeton
1919/20	Volumes	183,263	226,479
"	Volumes Added	17,483	17,662
1920/21	Volumes	205,097	248,312
1921/22	Volumes	225,347	268,562

TEXAS (continued)

Year	Variable	Gerould	Princeton
1923/24	"	282,048	312,469
"	Volumes Added	18,536	18,161
1925/26	Volumes	416,523	339,458
"	Volumes Added	17,041	13,585
1926/27	Volumes	388,016	358,821
"	Volumes Added	31,362	19,363
1927/28	Volumes	411,310	379,165
"	Volumes Added	20,161	20,344
1928/29	Volumes	430,082	393,029
"	Volumes Added	12,692	13,864
1930/31	Volumes	459,873	426,076
"	Volumes Added	17,425	18,540
1931/32	Volumes	471,515	436,224
"	Volumes Added	10,642	10,148
1932/33	"	12,235	14,025
1933/34	"	20,158	23,588
1934/35	"	17,686	16,508
1935/36	"	17,782	18,782
1936/37	"	22,985	22,327
1937/38	Volumes	564,360	563,697
"	Volumes Added	32,906	32,243
1937/38	Volumes	618,856	613,615
"	Volumes Added	55,159	49,918

VASSAR

Year	Variable	Gerould	Princeton
1931/32	Salaries	36,709.00	36,700
1937/38	"	41,529.00	37,529

VIRGINIA

Year	Variable	Gerould	Princeton
1934/35	Volumes Added	16,566	16,586
1936/37	Expenditures	29,006.74	29,807
1941/42	Salaries	81,089.22	78,509

WASHINGTON

Year	Variable	Gerould	Princeton
1922/23	Volumes	120,964	120,946
1924/25	"	143,840	156,416
"	Volumes Added	15,391	26,030
1931/32	Volumes	257,853	316,136
"	Appropriations	*	45,000
1932/33	Volumes	265,492	327,431
1933/34	"	275,820	344,220
1941/42	Salaries	108,747.00	71,693

WASHINGTON UNIVERSITY-ST. LOUIS

Year	Variable	Gerould	Princeton
1923/24	Volumes Added	12,176	9,926
1925/26	"	15,233	16,301
1926/27	Appropriations	*	50,950

WASHINGTON UNIVERSITY-ST. LOUIS (continued)

1927/28	Volumes	243,221	294,469
"	Volumes Added	7,420	16,584
1928/29	Volumes	250,699	301,889
"	Volumes Added	8,934	7,420
1929/30	Volumes	168,752	310,823
"	Volumes Added	6,256	8,934
1930/31	Volumes	271,808	317,079
"	Volumes Added	12,322	6,256
"	Expenditures	36,502.00	36,302
1931/32	Volumes	271,808	329,401
1932/33	"	189,632	338,711
"	Volumes Added	7,445	9,310
1933/34	Volumes	295,249	346,156
"	Volumes Added	4,998	7,445
1934/35	Volumes	303,089	351,154
"	Volumes Added	7,840	4,998
1935/36	Volumes	305,515	358,994
"	Volumes Added	7,426	7,840
1936/37	Volumes	305,515	366,420
"	Volumes Added	7,077	7,426
1937/38	Volumes	167,556	373,497
"	Volumes Added	6,288	7,077
1937/38	"	14,384	25,322
"	Staff	34	16
"	Salaries	*	25,186
1939/40	Volumes	409,872	413,203
"	Volumes Added	11,063	14,384

WASHINGTON UNIVERSITY-ST. LOUIS (continued)

1940/41	“	14,915	67,934
1942/43	Volumes	276,390	536,819
“	Volumes Added	15,941	35,196

WELLESLEY

Year	Variable	Gerould	Princeton
1942/43	Appropriations	*	22,000

WESTERN RESERVE

Year	Variable	Gerould	Princeton
1942/43	Volumes	565,500	565,000

WISCONSIN

Year	Comment
1928/29	Not reported in the original.
1929/30	Not reported in the original.

The Princeton compilation reports the following values for these years:

Year	Volumes	Volumes Added	Expenditures	Appropriations	Staff	Salaries
1928/29	404,386	12,624	50,313	57,500	37	62,544
1929/30	417,467	13,081	55,791	57,500	39	65,065

YALE

Year	Variable	Gerould	Princeton
1928/29	Staff	*	81
“	Salaries	*	164,015



The Gerould Statistics Appendix Four:

Differences Between the Gerould and Purdue Statistics

This appendix presents the values of five variables, (volumes held, volumes added, materials expenditures, salaries for full time employees, and wages for student employees) which differ between the edition of the Gerould statistics being published here and the statistics used in the Purdue studies in the years these two series overlap, 1950/51-1961/62. These variables are those which are directly comparable with each other in the two series. There are 46 institutions common to both sources. Appendix 5 lists the other variables in this series and makes the data available for these years.

The Purdue data were not published in hardcopy but were made available in a deck of computer cards. Dr. Warren Seibert located one of the Purdue decks. These data were originally obtained from him in 1985. While it seems likely that the set of data analyzed here is the same as those data analyzed at Purdue, there is no guarantee because, as has been mentioned, these data never have been published nor documented. A copy of this set of these data is made available here. Appendix 5 provides the source for these data. The comparison between the Gerould and Purdue statistics made in this Appendix uses the Purdue values from a tape supplied from Dr. Seibert which transcribed values for the 58 Purdue libraries from 1950/51 through 1971/72 from the deck of cards he obtained and will, therefore, reflect any changes in those data made through 1985 when this tape was obtained.

The Purdue values which differ from the values reported in the Gerould statistics have been checked with the original sources indicated in the Appendix to *The Past and Likely Future of 58 Research Libraries, 1951-1980* and references to these sources are included when necessary. According to this Appendix, the Purdue values for these early years come from three sources:

Values for these five variables from 1950/51 through 1958/59 come from an annual cumulation published in *College and Research Libraries*, volumes 13-21.

Values for 1959/60 and 1960/61 come from *Library Statistics of College and Universities* for those two years. This title was produced by the Office of Education at the U.S. Department of Health, Education, and Welfare and published by the Government Printing Office.

Data for 1961/62 came from the first issue of the annual statistical series published now by the Association of Research Libraries statistical series under the title *ARL Statistics* but known in 1961/62 as *Statistics for the Fiscal Year 1961/62*.

In the references given below to these three sources, *C&RL* refers to the issues of *College and Research Libraries*. In that series, values for a given year will be found in the next year, that is, data from 1950/51 will be in volume 13, 1952. The statistics appear in the first issue each year. *LS* refers to *Library Statistics* and *ARL* refers to the statistics published by the Association of Research Libraries for 1961/62. According to the Appendix in *The Past and Likely Future*, the Purdue values given are said to have been taken from one of these three sources with exceptions as noted in the next paragraph.

The compilers of the Purdue data also used other sources of data and other means of obtaining data. In a section on PROCEDURES, the authors mention they discovered omissions in the data and “steps were taken to locate and insert the missing figures. In some cases, data initially missing were found in reports which had not served as principal data sources. In other cases, omitted data could be supplied by interpolation, extrapolation, or other calculations from the data in hand.” (p. 5)

As was noted in the discussion of the Princeton compilation, the fact that a number in the Purdue tape differs from the value in one of the sources does not indicate the Purdue value is wrong because the compilers of the data may have discovered a mistake and corrected it. However, Appendix 5 cites sources that suggest there is a prima facie case that the data made available there are the same as the original data.

In many of the cases which follow the values differ by one and may result from rounding errors. Purdue and the sources it consulted do not include amounts below a dollar; so there are no numbers to the right of decimal points. The Gerould data during this period used decimals in 1950/51 with library materials expenditures and did not thereafter. These values less than a dollar are reported for the Gerould data for 1950/51.

Values for 1950/51 are not regarded as different for purposes of this appendix if the rounded Gerould value equals the Purdue value.

In 1959/60 and 1960/61, the Purdue studies used *Library Statistics* which did not record the value for library materials expenditures and binding: these two variables were separate. In order to derive figures which were comparable to those reported for other years, the compilers of the Purdue data apparently added the figure given for "Books and other library materials" expenditures with those for "Binding." In two cases (Colorado, 1959/60, and Rochester, 1959/60), these figures were apparently not added correctly.

The notes following the differences between the values that differ for each institution include comparisons from the stated sources of the Purdue data and the data themselves. There are discrepancies which may be the result of editorial changes or other sources. There are a number that appear to be transcription errors or mathematical errors mentioned for Colorado and Rochester above. On the other hand, they may be corrections. There is no way to know because of the lack of documentation.

There are cases where values for student wages are given by Purdue as \$0. In all such cases (except Yale's student wages in 1950/51) the original source did not report a value, that is, the space was blank. There were a number of reasons for the practice for reporting no value as evidenced by footnotes, the most common of which are either that the data are "not reported or not available" or that student wages are included with staff salaries. In 1950/51 Yale is reported to have spent \$0 because the student budget did not come out of the library's budget.

There are no missing values in this version of the Purdue data and zeroes occur often where other sources report missing values. *In doing computations on these data, the zeroes must be taken into account.* For a further discussion of these zeros, and their effect on computations, see Molyneux, 1992, p. 117-118

Cases where the source suggested for data gives no values for an institution for a given year are also given. There are a number of institutions for which there are no data in *C&RL* in given years but there are data in the Purdue statistics. For example, Johns Hopkins has no data in *C&RL* before 1954/55, yet Purdue reports data for these years. In six (Iowa before 1950/51, Johns Hopkins before 1954/55; Nebraska before 1953/54; Notre Dame

before 1954/55; Pittsburgh before 1954/55; and Stanford before 1959/60 there are few, if any, discrepancies between the Purdue and Gerould data. It is possible that in these cases the Gerould data was one of the “reports which had not served as principal data “sources” which were used by the compilers of the Purdue data as sources of data.

The symbol “*” indicates that the value for the given variable is missing in the source.

BROWN

Year	Variable	Gerould	Purdue
1950/51	Volumes added	17,014	18,534
1951/52	“	17,220	19,429
“	Wages	14,926	14,936
1952/53	Volumes added	19,741	21,461
“	Expenditures	92,090	92,089
1953/54	Volumes	807,902	817,514
“	Volumes added	18,056	20,363
1954/55	Volumes	825,797	836,159
“	Volumes added	17,895	19,425
1955/56	Volumes	844,383	857,660
“	Volumes added	18,586	23,858
1956/57	Volumes	862,674	891,033
“	Volumes added	18,291	37,510
1957/58	Volumes	884,247	922,077
“	Volumes added	21,584	32,937
“	Salaries	226,961	223,523
“	Wages	29,699	33,587
1958/59	Volumes	906,415	956,472
“	Volumes added	25,993	37,953

BROWN (continued)

Year	Variable	Gerould	Purdue
"	Salaries	265,897	264,807
"	Wages	33,093	34,183
1959/60	Salaries	321,968	319,887
"	Wages	35,313	37,392 ¹
1960/61	Salaries	369,064	363,619
"	Wages	43,394	48,839

Notes:

1. The 1959/60 figure for Wages according to *LS* was 37,394 instead of 37,392.

CALIFORNIA, BERKELEY

Year	Variable	Gerould	Purdue
1950/51	Salaries	877,194	864,614
"	Wages	196,190	193,950
1951/52	Expenditures	412,570	412,571
"	Wages	191,654	191,655
1953/54	Volumes added	79,276	79,279
"	Salaries	1,183,811	1,192,004
"	Wages	294,202	241,027
1958/59	Wages	303,166	330,166

CALIFORNIA, LOS ANGELES

Year	Variable	Gerould	Purdue
1950/51	Expenditures	353,094.76	353,093
1957/58	Wages	281,084	281,064

CHICAGO

Year	Variable	Gerould	Purdue
1950/51	Volumes added	45,459	49,457 ¹
1950/51	Wages	*	0
1951/52	Salaries	483,369 ²	216,737 ²
"	Wages	*	266,633 ²
1954/55	Salaries	394,003	394,004
1956/57	"	456,381	456,382

Notes:

1. The 1950/51 figure for Volumes Added according to C&RL is 53,971 instead of 49,457.
 2. The 1951/52 figures for Salaries according to C&RL are \$483,369 for but there is no value for Wages. According to a footnote these are included with salaries. The Purdue figures for Salaries, \$216,737, plus the figure for wages, \$266,633, equals \$483,370. The source of these numbers is not documented.
-

CINCINNATI

Year	Variable	Gerould	Purdue
1950/51	Expenditures	51,188.78	51,639
1952/53	Volumes	661,227	661,572
1955/56	Expenditures	95,441	94,646
1957/58	"	102,531	102,331

COLORADO

Year	Variable	Gerould	Purdue
1950/51	Wages	33,685	33,686
1951/52	Salaries	107,778	107,779
1952/53	Volumes	805,090	758,190
"	Volumes Added	36,809	34,529

COLORADO (continued)

Year	Variable	Gerould	Purdue
"	Expenditures	91,727	79,327
"	Salaries	148,013	133,033
"	Wages	36,733	34,133
1953/54	Expenditures	86,846	86,847
1956/57	Wages	40,901	40,900
1959/60	Expenditures	187,488	187,448 ¹

Notes:

1. The 1959/60 figure for Expenditures according to *LS* gives \$165,018 for "Books and other library materials" and 22,470 for "Binding." The sum of these two numbers is \$187,488, not 187,448.

COLUMBIA

Year	Variable	Gerould	Purdue
1950/51	Wages	*	0
1951/55	"	*	0
1955/56	Expenditures	330,483	329,483
"	Wages	*	0
1956/57	"	*	0
1957/58	"	*	0

CORNELL

Year	Variable	Gerould	Purdue
1952/53	Wages	53,734	53,735
1954/55	Salaries	561,126	561,127
"	Wages	70,512	70,513
1957/58	"	97,663	97,623

DUKE

Year	Variable	Gerould	Purdue
1950/51	Expenditures	166,679	166,680
"	Salaries	204,811	204,812
1955/56	Volumes	1,244,880	1,243,691
"	Volumes Added	47,748	46,500
1957/58	Expenditures	250,104	250,105
1959/60	Salaries	403,799	390,031
"	Wages	34,415	48,183

FLORIDA

Year	Variable	Gerould	Purdue
1957/58	Salaries	479,446	479,447
1960/61	Wages	75,590	75,591

HARVARD

Year	Variable	Gerould	Purdue
1951/52	Salaries	890,000	889,702 ¹
1952/53	Volumes Added	54,153	54,123 ²
1953/54	Volumes	5,832,912	5,833,116
"	Volumes Added	129,965	130,169
"	Salaries	1,057,510	1,107,510 ³
1953/54	Wages	50,000	0 ⁴
1954/55	Salaries	1,125,154	1,100,154
"	Wages	70,000	95,000
1956/57	Volumes	6,225,447	6,225,444
1957/58	Volumes Added	162,186	124,723 ⁵
"	Salaries	1,552,752	1,401,385
1958/59	Volumes	6,492,124	6,492,111

HARVARD (continued)

Notes:

1. The 1951/52 *C&RL* does not record any values for Harvard.
2. *C&RL* records 112,262 instead of 54,123.
3. *C&RL* records this variable as “not reported or not available.”
4. *C&RL* records this variable as “not reported or not available.”
5. *C&RL* records 214,723 instead of 124,723 volumes added. The Purdue number appears to be a result of transposition of digits from the source.

ILLINOIS

Year	Variable	Gerould	Purdue
1950/51	Expenditures	385,444.00	385,445
“	Wages	96,215	96,216
1951/52	Expenditures	378,006	378,007
“	Wages	100,635	100,636
1952/53	Volumes Added	85,825	92,146
1953/54	“	133,760	109,377
1954/55	“	98,694	113,294
“	Expenditures	447,240	448,995
1955/56	Volumes Added	90,040	98,427
1960/61	“	95,246 ¹	95,266 ¹

Notes:

1. This pair of values is discussed in Appendix 2 (page 82.)

INDIANA

Year	Variable	Gerould	Purdue
1950/51	Volumes added	40,000	34,958
1952/53	“	40,271	44,300
“	Expenditures	252,764	253,164
	Salaries	328,896	329,996
1954/55	Volumes Added	*	30,000 ¹

INDIANA (continued)

1955/56	Volumes Added	*	83,000 ¹
1956/57	Volumes Added	*	46,500 ¹
1957/58	Volumes	1,190,566	1,188,827
"	Volumes Added	*	82,500 ¹
1958/59	Volumes Added	*	69,000 ¹
1959/60	Volumes Added	*	59,000 ²
1959/60	Salaries	440,340	444,340
1960/61	Volumes Added	*	97,000 ³
1961/62	Volumes	1,828,992	1,678,992 ⁴
1961/62	Volumes added	*	102,100 ⁵

Notes:

1. *C&RL* records these variables as "not reported or not available."
2. *LS* does not record a value for this variable.
3. *LS* records this value as an "estimated figure."
4. *ARL* records 1,828,992 instead of 1,678,992 volumes held.
5. *ARL* does not report a value for this variable.

IOWA

Year	Variable	Gerould	Purdue
1951/52	Volumes ¹	773,205	664,814
"	Volumes Added ¹	22,205	18,098
"	Expenditures ¹	129,385	113,335
"	Salaries	233,480	236,375 ²
1953/54	Wages	40,318	45,525
1954/55	Salaries	259,605	251,628
"	Wages	39,395	47,372
1955/56	Salaries	272,530	270,910
"	Wages	44,405	46,025

IOWA (continued)

Notes:

1. *C&RL* records no values for Iowa in 1950/51.
2. *C&RL* records 236,675 instead of 236,375.

IOWA STATE

Year	Variable	Gerould	Purdue
1950/51	Wages	14,421	14,422
1951/52	"	18,053	18,054
1952/53	Expenditures	88,724	88,725
"	Salaries	136,382	136,383
"	Wages	18,414	18,415
1953/54	Expenditures	95,098	95,099
"	Wages	21,393	21,394
1955/56	Salaries	154,684	154,685

JOHNS HOPKINS

Year	Variable	Gerould	Purdue
1951/52	Wages	8,597	8,598 ¹
1958/59	Expenditures	141,069	141,070
1959/60	Wages	13,758	15,530
1960/61	Expenditures	189,471	162,237
"	Salaries	335,777	319,087
"	Wages	20,050	22,280

Notes:

1. *C&RL* records no values for Johns Hopkins before 1954/55.

JOINT UNIVERSITY LIBRARIES

Year	Variable	Gerould	Purdue
1950/51	Expenditures	104,685.03	104,684
1951/52	Salaries	109,438	109,439
1953/54	Volumes Added	23,183	24,839
"	Wages	30,965	30,966
1954/55	Expenditures	89,343	89,344
"	Salaries	125,083	125,084
"	Wages	30,768	30,769
1955/56	Expenditures	96,035	96,036
1956/57	Volumes	708,752	708,952
1957/58	Volumes Added	24,616	27,914
1958/59	"	26,591	29,436

KANSAS

Year	Variable	Gerould	Purdue
1950/51	Expenditures	121,761.00	122,761

KENTUCKY

Year	Variable	Gerould	Purdue
1951/52	Volumes	558,442	558,896
"	Expenditures	80,641	80,642
1952/53	"	108,319	108,320
"	Salaries	144,457	144,458
1953/54	"	145,055	145,056
1954/55	Volumes Added	31,470	31,009
1956/57	Expenditures	133,486	133,487
"	Wages	15,178	15,179
1957/58	Expenditures	176,170	195,170

KENTUCKY (continued)

1958/59	Salaries	272,827	272,828
"	Wages	24,847	24,848
1959/60	Expenditures	322,222	322,223
"	Salaries	298,058	298,059
1960/61	Volumes	904,757	925,904
"	Volumes Added	38,064	41,180
"	Salaries	327,613	354,449
"	Wages	26,336	28,550

LOUISIANA STATE

Year	Variable	Gerould	Purdue
1950/51	Salaries	183,680	183,681
"	Wages	20,157	20,158
1951/52	Expenditures	118,340	118,341
"	Salaries	189,805	189,806
"	Wages	18,022	18,023
1953/54	Expenditures	188,190	188,191
"	Salaries	252,333	252,334
1954/55	"	281,759	281,760
1956/57	Expenditures	228,317	228,318
"	Salaries	321,635	321,636
"	Wages	39,295	39,296
1959/60	Expenditures	410,196	410,197
"	Salaries	481,594	481,118

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Year	Variable	Gerould	Purdue
1950/51	Volumes Added	15,623	15,625
"	Salaries	151,000	130,850
"	Wages	14,900	16,132
1951/52	Volumes Added	16,809	16,876
"	Expenditures	54,736	48,700
"	Salaries	158,340	141,400
"	Wages	16,491	16,500
1954/55	Expenditures	51,340	51,339
1955/56	"	74,825	74,826
"	Salaries	177,200	177,201
"	Wages	19,074	19,075
1958/59	"	27,384	27,383
1960/61	Volumes Added	40,812	40,910

MICHIGAN

Year	Variable	Gerould	Purdue
1951/52	Expenditures	244,708	251,500 ¹
"	Salaries	557,971	570,432 ¹
"	Wages	67,468	74,855 ¹
1953/54	Volumes Added	58,211	58,018
1954/55	Salaries	783,945	783,946
1956/57	Expenditures	523,304	523,305

Notes:

1. C&RL records no values for Michigan in 1951/52.

MICHIGAN STATE

Year	Variable	Gerould	Purdue
1957/58	Expenditures	363,791	336,309
1960/61	Salaries	463,630	463,730
"	Wages	94,131	94,000

MINNESOTA

Year	Variable	Gerould	Purdue
1957/58	Salaries	643,866	643,876
1958/59	Expenditures	388,342	388,341

MISSOURI

Year	Variable	Gerould	Purdue
1951/52	Expenditures	140,388	140,389
1952/53	"	133,624	133,622
"	Salaries	154,023	154,203
"	Wages	18,303	18,302
1954/55	Volumes Added	17,467	17,500
"	Expenditures	115,367	116,368
1955/56	Salaries	147,292	147,293
1956/57	Expenditures	190,263	190,264
"	Salaries	164,206	164,207
"	Wages	36,937	36,938

NEBRASKA

Year	Variable	Gerould	Purdue
1951/52	Salaries	176,158	176,159
"	Wages	27,334	27,335
1953/54	Salaries	212,528	212,529 ¹
1954/55	Volumes Added	20,504	18,485
"	Expenditures	123,653	123,654
1956/57	Wages	34,547	34,548
1958/59	"	35,870	35,871
1959/60	"	39,997	36,997

Notes:

1. C&RL records no values for Nebraska before 1953/54.

NEW YORK UNIVERSITY

Year	Variable	Gerould	Purdue
1950/51	Expenditures	144,048	144,038
"	Salaries	410,846	410,847 ¹
1950/51	Wages	*	0
1951/52	Salaries	365,557	365,558
"	Wages	*	0
1952/53	Expenditures	112,761	122,201
"	Salaries	411,372	419,784
"	Wages	*	0
1953/54	Expenditures	125,585	91,795
"	Salaries	436,071	389,306
"	Wages	*	0
1954/55	"	*	0
1955/56	"	*	0

NEW YORK UNIVERSITY (continued)

1956/57	Expenditures	132,039	132,040
"	Salaries	*	216,166
"	Wages	*	0
1957/58	"	*	0
1958/59	"	40,284	40,285
1960/61	Salaries	561,933	275,343
"	Wages	45,437	330,028

Notes:

1. C&RL records that this variable is "not reported or not available."

NORTH CAROLINA

Year	Variable	Gerould	Purdue
190/51	Expenditures	157,389.92	157,389
"	Wages	37,859	37,860
1951/52	Expenditures	131,091	131,092
"	Wages	43,348	43,349
1956/57	Volumes Added	67,295	69,043
"	Wages	49,549	49,550
1959/60	Volumes	1,025,144	1,025,944
1960/61	Salaries	522,450	522,451

NORTHWESTERN

Year	Variable	Gerould	Purdue
1951/52	Volumes	1,089,295	1,089,458
"	Expenditures	174,450	174,151
"	Salaries	279,483	279,484
1954/55	"	319,962	310,962

NORTHWESTERN (continued)

1956/57	Volumes added	43,364	47,940
1957/58	Volumes	1,339,218	1,322,040
"	Volumes added	43,364	57,783

PITTSBURGH

Year	Variable	Gerould	Purdue
1950/51	Volumes	617,000	616,216 ¹
"	Volumes Added	38,700	21,919 ¹
"	Expenditures	105,576	87,676 ¹
"	Salaries	138,263	119,250 ¹
"	Wages	12,633	0 ¹

Notes:

1. *C&RL* records no values for Pittsburgh before 1954/55

PRINCETON

Year	Variable	Gerould	Purdue
1950/51	Salaries	263,064	266,683
"	Wages	9,522	9,523
1951/52	Salaries	271,939	275,925
"	Wages	12,156	12,157
1952/53	Salaries	280,848	291,860
1953/54	Expenditures	151,685	151,686
"	Wages	13,750	25,853
1954/55	Expenditures	154,375	154,357
"	Salaries	329,940	320,826
"	Wages	17,177	26,291

PRINCETON (continued)

Year	Variable	Gerould	Purdue
1955/56	Expenditures	169,632	169,631
"	Salaries	362,676	352,417
"	Wages	18,408	28,667
1956/57	Salaries	376,371	364,462
"	Wages	21,562	35,554
1957/58	Volumes added	53,134	53,146
"	Expenditures	243,707	243,706
"	Salaries	418,633	418,753
"	Wages	20,539	35,291
1958/59	Salaries	474,897	531,416
"	Wages	23,060	39,933
1959/60	Salaries	462,625	518,845
"	Wages	22,267	43,256
1960/61	Salaries	568,777	568,778
"	Wages	25,670	52,132

PURDUE

Year	Variable	Gerould	Purdue
1956/57	Volumes	417,258	416,637
"	Volumes Added	23,890	23,909

ROCHESTER

Year	Variable	Gerould	Purdue
1950/51	Expenditures	97,619.51	97,619
"	Salaries	136,699	136,700
1951/52	Wages	14,264	14,265

ROCHESTER (continued)

Year	Variable	Gerould	Purdue
1952/53	Expenditures	96,644	96,662
1953/54	"	101,010	100,892
"	Salaries	158,082	158,083
1954/55	Expenditures	96,544	96,545
"	Salaries	170,549	170,550
"	Wages	21,322	21,332
1955/56	Expenditures	105,028	100,317
1956/57	Wages	26,474	26,475
1959/60	Expenditures	143,874	143,742 ¹

Notes:

1. LS gives \$121,442 for "Books and other library materials" and 22,432 for "Binding." The sum of these two numbers is \$143,874.

RUTGERS

Year	Variable	Gerould	Purdue
1950/51	Expenditures	109,468.00	103,200

STANFORD

Year	Variable	Gerould	Purdue
1950/51	Expenditures	184,430.50	184,430 ¹
"	Wages	*	0 ¹
1951/52	Expenditures	179,625	179,626 ¹
"	Salaries	312,842	312,843 ¹
1960/61	Volumes	1,615,740	1,691,008
"	Volumes Added	92,945	98,721

Notes:

1. C&RL records no values for Stanford before 1959/60.

TEMPLE

Year	Variable	Gerould	Purdue
1950/51	Wages	16,693	16,694

TEXAS

Year	Variable	Gerould	Purdue
1950/51	Wages	68,277	68,278
1951/52	Expenditures	139,778	139,779
"	Wages	67,564	67,565
1953/54	Salaries	260,343	260,344
"	Wages	82,154	82,155
1956/57	Expenditures	197,625	197,626
1958/59	"	633,749	633,748

VIRGINIA

Year	Variable	Gerould	Purdue
1950/51	Salaries	201,412	201,413
"	Wages	25,761	25,762
1951/52	Volumes Added	40,305	40,315
"	Expenditures	97,516	97,517
1952/53	Salaries	227,476	227,447 ¹
1952/53	Wages	22,800	22,801
1957/58	Volumes added	45,748	45,745
"	Expenditures	170,369	170,370

VIRGINIA (continued)

Year	Variable	Gerould	Purdue
1960/61	Volumes	1,104,610	1,111,642
"	Volumes Added	58,795	60,407
	Salaries	340,744	347,464
"	Wages	36,470	38,751

Notes:

1. C&RL records 227,477 instead of 227,447.

WASHINGTON

Year	Variable	Gerould	Purdue
1950/51	Expenditures	181,127.00	152,770
"	Salaries	303,742	303,505
1951/52	Expenditures	143,615	120,479
"	Salaries	317,046	298,596
"	Wages	83,824	75,971
1952/53	Volumes	802,681	795,442
"	Volumes added	35,463	35,488
"	Expenditures	212,039	189,257
"	Salaries	355,465	334,585
"	Wages	87,584	81,584
1953/54	Volumes Added	42,779	49,225
"	Expenditures	188,281	165,594
"	Salaries	415,032	368,224
"	Wages	87,416	82,801
1954/55	Volumes Added	23,545	23,547
"	Salaries	471,209	417,209

WASHINGTON (continued)

Year	Variable	Gerould	Purdue
1957/58	Volumes Added	43,711	45,252
1959/60	"	45,918	47,977

WASHINGTON STATE

Year	Variable	Gerould	Purdue
1950/51	Volumes added	11,299	12,148
"	Expenditures	60,300	49,337
"	Salaries	143,935	140,231
"	Wages	36,000	34,573

WASHINGTON UNIVERSITY-ST. LOUIS

Year	Variable	Gerould	Purdue
1950/51	Salaries	136,460	136,461
1951/52	"	137,106 ¹	59,190 ¹
1951/52	Wages	10,290	17,570
1955/56	"	36,623	36,628 ²

Notes:

1. For a discussion of this pair of values see Appendix 2.
 2. C&RL records 36,623 instead of 36,628.
-

WISCONSIN

Year	Variable	Gerould	Purdue
1950/51	Expenditures	240,829.00	237,629
1953/54	Salaries	433,414	433,410 ¹

Notes:

1. C&RL records 433,414 instead of 433,410.
-

YALE

Year	Variable	Gerould	Purdue
1950/51	Wages	*	0
1951/52	Expenditures	303,375	303,275
"	Wages	*	0
1952/53	"	*	0
1953/54	"	*	0
1954/55	"	*	0
1955/56	"	*	0
1957/58	Salaries	817,331	860,784
"	Wages	43,453	0
1958/59	Expenditures	478,663	478,633 ¹
1958/59	Wages	*	0
1959/60	"	*	0
1960/61	Volumes	4,490,420	4,478,040
"	Volumes added	95,432	83,052
"	Wages	*	0
1961/62	"	*	0

Notes:

1 . C&RL records 478,663 instead of 478,633.



The Gerould Statistics Appendix Five:

The Machine Readable Version of the Gerould Data

5.1. Introduction

The purpose of this chapter is to make the Gerould data available, as well as data from the Princeton Compilation and the Purdue studies. In addition, this publication is made available in compressed form.

For related documentation, see:

ARL, 1992-1996

Stubbs (1996)

Stubbs and Molyneux (1990), particularly p. 12.

The usage here is that of ARL. Variable names, including "Region," "MEMBYR," and "Type" are included to help the Gerould, Princeton, and Purdue files conform to ARL practice.

These data will be the same as those found in the ARL data for the years before 1961/62 for the ARL libraries. However, there are institutions which reported to Gerould or the compilers at Princeton which are not ARL members. These data are only available in the Gerould data.

These institutions are:

Institution Number	Institution	Years
830	Bryn Mawr	1923/24-1950/51
4000	Joint University	1936/37-1961/62
5490	Mt. Holyoke	1940/41-1950/51
5860	North Dakota	1907/08-1950/51
6030	Oberlin	1913/14-1950/51

Institution Number	Institution	Years
7350	St. Louis	1923/24-1950/51
7340	Smith	1923/24-1950/51
7430	South Dakota	1907/08-1924/25
8860	Vassar	1923/24-1950/51
9420	Wellesley	1925/26-1950/51
9440	Wesleyan	1946/47-1950/51

Data from these three series have been encoded with Institution Numbers as listed in Table 1 following the system used by ARL. These institution numbers are discussed in Chapter 1. In addition, the variable names in these series follow ARL practice. As a result of these two facts, these data can be merged, analyzed, or compared with each other. In addition, data from two related series, that from the Association of College and Research Libraries (ACRL) documented in Molyneux, 1989 and that from the Historically Black College and University Libraries both of which were sponsored by ACRL but used the ARL forms and definitions in order to ensure comparability between these various series. Note the dates on this later ACRL series (1987/88+) because they are not to be confused with the earlier series published in the ALA Bulletin and later College and Research Libraries. The HBCU data were published for one year (1988/89) and no subsequent attempt to recompile these data from this important set of libraries has attracted interested until recent work by SOLINET (now Lyris).

A lesser-known but related series of library data deserves attention. This series was published through the auspices of the Louisiana State University from 1928/29 through the late 1980s when they were compiled by D.W. Schneider. It is known by various titles, most commonly, *Statistics of Southern Colleges and University Libraries*. The first year was known by the title *Statistics of Southern University Libraries, 1928-29* and consists of a mimeographed sheet with data from 25 libraries. Letters soliciting data obtained from the LSU archives make clear that James McMillen, Librarian, had a copy of the Gerould form and sent it out to Southern institutions in an attempt to start a series comparable to the Gerould series. Subsequently, local exigencies created a separate series, related but separate from the ARL and its tributaries: those from ACRL and HBCU.

As a result of funding from OCLC, the LSU data are in electronic form—in the possession of the author—and await documentation by anyone interested. These data use the same infrastructure pioneered by Stubbs and Buxton with the four digit institution numbers that this monograph introduced as a modification of the Stubbs-Buxton institution numbers.

The following variable names are used in these series:

Variable Name	Comments	Years Variable is Reported
APPROP	Appropriations (Chapter 3)	1909/10-1953/54
EXPLMB	Materials Expenditures and Binding (Section 3.1.3)	All
FTESAL	Staff salaries (Section 3.1.7)	All
INAM	Institution Name (Table 1)	All
INSTNO	Four digit institution numbers (Section 1.3.3)	All
MEMBYR	Year the institution joined ARL. The non-ARL members listed above do not have a MEMBYR.	All
REGION	01 = New England 02 = Middle Atlantic 03 = East North Central 04 = West North Central 05 = South Atlantic 06 = East South Central 07 = West South Central 08 = Mountain 09 = Pacific 10 = Canada	All
SALSTUD	Student salaries (Section 3.1.8)	1949/50-1961/62
STOTEXP	Library expenditures (Section 3.1.5)	1954/55-1961/62
TOTSTF	Total staff (Section 3.1.6)	All
TYPE	C = Canadian university P = Private US university S = State controlled or public US university Each institution is typed by its current status	All
VOLS	Volumes Held (Section 3.1.1)	All
VOLSADG	Volumes Added, Gross (Section 3.1.2)	All
YEAR	Expressed as two digits. For example "17"=1916/17 academic year.	All

5.2. The Gerould Data

Data are available as SAS Datasets and comma delimited ASCII files (.csv).

A few important details about the data:

1. The data are arranged by institution number and then by year.
2. Blank lines have been inserted in cases where data are missing from institutions that have reported data before or after the missing year(s). (These blank lines are often helpful in time series applications.) Thus, a blank line is inserted for Northwestern in 1922/23, when data are missing for this one year but not for Pittsburgh after 1950/51 because no data are reported after that date in the Gerould statistics.
3. Eight institutions (Dartmouth, McGill, Oregon, Pittsburgh, St. Louis, Temple, Toronto, and Washington State) are not available in the Gerould data after 1950/51 and subsequently report data to ARL. The biggest class of these inserted lines is for institutions that reported data in 1907/08 or 1909/10. No data exist for 1908/09 or 1910/11 so a blank line is inserted for all such institutions. The blank lines are included in the tables of the Gerould data listed in Table 1. The inserted lines are, (in addition to those for 1908/09 and 1910/11):

Institution	Years Blank Lines Are Inserted
Colorado	1922/23—1923/24, 1930/31
Dartmouth	1914/15—1919/20
Harvard	1914/15, 1917/18, 1918/19
Nebraska	1943/44
Northwestern	1922/23
Rutgers	1951/52—1955/56
Texas	1922/23, 1924/25, 1929/30
Wisconsin	1928/29—1929/30

5.2.1 The Data

File Name	Contents/Comments
geriy.csv (138K)	The Gerould data in an ASCII "text" file The file format is documented immediately below
geriy.sas7bdat (301K)	Gerould data in a zipped SAS Dataset This is the master file.

5.2.2 Variables Included in the Gerould Data

The .csv file of the Gerould data consists of 2,091 records. Each record consists of information on one institution for one year. Variable names are detailed above on page 146.

Variable Name	Width of Field	Special Format	Years in File
YEAR	2		All
INSTNO	4		All
INAM	27 Character		All
TYPE	1 Character		All
REGION	2		All
MEMBYR	2		All
VOLS	8		All
VOLSADG	8		All
TOTSTF	6 Includes decimals		All
EXPLMB	10 Includes decimals		All
FTESAL	10 Includes decimals		All
SALSTUD	10		1949/50-1961/62
STOTEXP	10		1954/55-1961/62
APPROP	10 Includes decimals		1909/10-1953/54

5.3. The Princeton Compilation Data

Data Files	Contents
Princeton.csv (59K)	Data from the Princeton Compilation in .csv file.
prisas.sas7bdat (121K)	The unzipped SAS dataset of the Princeton data.

The files of the Princeton data consist of 992 records. The variables SALSTUD and STOTEXP do not exist in these data. Non-ARL libraries listed above do not have a MEMBYR because they did not join ARL.

5.4. The Purdue Data

This compilation is discussed in general early in Chapter 2 and in more detail in Section 2.3.2. of that Chapter. The edition here is that given the author by W.F. Seibert in the mid-1980s and comprises data from 1950/51-1970/71. He said that he obtained the data from Maurice Marchant.

There is at best a *prima facie* case that these data are the same as those used by Seibert later in Molyneux, 1990 and Molyneux, 1992. It seems probable that these data are the same as the Purdue data but there is uncertainty because the data have never been published, nor documented; indeed their only documentation is in the two editions of this work.

The data have been modified to include ARL variable names, institution names, and institution numbers—the same infrastructure developed by Stubbs and Buxton and used here with the Gerould and Princeton data and, more generally, with the ARL data and derivative data such as the ACRL and HBCU series.

As was discussed in Appendix 4, the Purdue data overlap the Gerould for five variables, across forty six institutions and for the years 1950/51-1961/62. In addition to the Gerould variables listed above, Purdue records six others:

Dataset Name	Contents
PRFSTF	Number of professional staff
NPRFSTF	Number of non-professional staff
TOTEXP	Total Expenditures
BEGSAL	Beginning professional salary
TOTSTU	Total Students
GRADSTU	Number of graduate students
	Original Purdue library numbers.
LIBNO	These range from 1 to 60. 10 and 46 were not used.

NPRFSTF, PRFSTF, TOTEXP, TOTSTU are found in the ARL data. In order to accommodate these variables, the ASCII format for these data is altered. The ASCII format is given after the listing of the data files.

5.4.1 The Data

Data Files	Contents
Purdue.csv (114K)	The zipped ASCII file of Purdue data. The unzipped file is purascii.txt and it is 191K.
purdue.sas7bdat (225K)	A zipped SAS dataset. The unzipped file is pursas.sd2 and it is 213K. It is written with engine V612.

5.4.2 Variables included in the Purdue data

Variable Name	Width of Field	Special Format	Years in File
YEAR	2		1951-1972
INSTNO	4		1951-1972
INAM	27 Character		1951-1972
TYPE	1 Character		1951-1972
REGION	2		1951-1972

MEMBYR	2	1951-1972
GRADSTU*	6	1951-1972
VOLS	8	1951-1972
VOLSADG	8	1951-1972
TOTSTU*	6	1951-1972
EXPLMB	10	Amounts in whole dollars 1951-1972
FTESAL	10	Amounts in whole dollars 1951-1972
SALSTUD	10	Amounts in whole dollars 1951-1972
TOTEXP	10	Amounts in whole dollars 1951-1972
PRFSTF*	10	Includes decimals 1951-1972
NPRFSTF*	10	Includes decimals 1951-1972
BEGSAL*	10	Amounts in whole dollars 1951-1972
TOTSTU*	10	1951-1972
LIBNO*	2	1951-1972

* Variables found in Purdue but neither Gerould nor Princeton.



The Gerould Statistics: Bibliography

Association of Research Libraries. <http://www.arl.org/>

Balbi, Adriano (1835). *A Statistical Essay of the Libraries of Vienna and the World*, translated by Larry Barr and Janet L. Barr (Jefferson, North Carolina, McFarland, 1986). This is a translation of *Essai Statistique sur les Bibliothèques de Vienne* (Vienne, Frederic Volke, 1835). ISBN 0-89950-149-4. LC 84-43235

Boyd, Julian (1945). Julian Boyd, Librarian at Princeton to Clemons "July 23, 1945." This letter says, in part:

"Due to the war and the shortage of paper, the Princeton University Library did not publish statistical totals for the year 1943-44..."

The letter is signed by the "Secretary to Mr. Boyd" whose signature is not clear [Anne Goldsmith?]. Boyd's letter is probably in response to a letter to him from Clemons to him dated "21 July 1945" which says, in part:

"Apparently we have not received a copy of the statistical totals of university libraries for the year 1943-44."

Manuscripts Department, Alderman Library, University of Virginia, Charlottesville.
Manuscript number 8836. Archives Records Group RG 12/1.

Clapp, Verner (1951). A note beginning "The 'Princeton Statistics'". *Library of Congress Information Bulletin* 10, no. 17, pp. 15-16.

Clemons, Harry (1946). Lawrence Heyl to Clemons, "6 February 1946". Letter begins: "In order to bring out a compilation..." Reply from Clemons to Heyl dated "9 February 1946" (begins "We have been able to locate no figures whatever for volumes added to the University of Virginia Libraries for the year 1921/22.") Manuscripts Department,

Alderman Library, University of Virginia, Charlottesville. Manuscript number 8836. Archives Records Group RG 12/1.

Downs, Robert B. (1946). "Uniform Statistics for Library Holdings." *Library Quarterly* 16, 63-69.

Downs, Robert B. (1985). Letter to Robert Molyneux dated October 22, 1985.

Dunn, O. C.; Seibert, W. F.; and Scheuneman, J. A. (1965). *The Past and Likely Future of 58 Research Libraries, 1951-1980: A Statistical Study of Growth and Change.* (University Libraries and Audio Visual Center, Purdue University, West Lafayette, Indiana, 1965.) This was the first of nine studies with the same title. The last was: Dunn, O. C.; Tolliver, D. L.; and Drake, M. A., *The Past and Likely Future of 58 Research Libraries, 1971/72, "Ninth Issue."* (Instructional Media Research Unit, University Libraries and Audio Visual Center, Purdue University, 1973).

Gerould, James Thayer (1906). "A Plan for the Compilation of Comparative University and College Library Statistics." *Library Journal* 31, 761-763.

Gerould, James Thayer (1913). Letter from J.T. Gerould to G.W. Harris (at Cornell), October 25, 1913. University Archives, University of Minnesota, Minneapolis. Bears the handwritten note: "Saved as a sample."

This letter appears to be a form letter with the name of the person to whom the letter is being sent filled in. The letter was probably typed (not printed) and reproduced in some fashion. In this copy of the letter, Harris's name and address were typed in a different typeface from that of the rest of the letter and without allowing enough space for Harris's name and address.

Gerould, James Thayer (1932). *The College Library Building: Its Planning and Equipment.* (New York, Scribners, 1932), p. 31.

Kuhlman, A. F. (1960). "Two ARL Approaches to Counting Holdings of Research Libraries." *College and Research Libraries* 21, 207-211.

Molyneux, Robert E. (1984). "An Examination of the Growth of Academic Libraries in the United States, 1972/73-1981/82." Unpublished dissertation, University of North Carolina, Chapel Hill.

Molyneux, Robert E. (1986a). "Patterns and Processes of Growth, and the Projection of Library Size: A Critical Review of the Literature on Academic Library Growth." *Library and Information Science Research* 8, 5-28.

Molyneux, Robert E. (1986b). "Growth at ARL Member Libraries, 1962/63-1983/84." *Journal of Academic Librarianship* 12, 211-216.

Molyneux, Robert E. (1989). *ACRL Academic Library Statistics, 1978/79-1987/88; A Guide to the Machine Readable Version of the ACRL Statistics*, (Chicago: Association of College and Research Libraries, 1989.)

Molyneux, Robert E. (1990). Review of *Research Library Trends...* in *Libraries & Culture*, 25, no. 4, 1990, pp. 619-620.

Molyneux, Robert E. (1991), compiler, *ACRL/Historically Black Colleges & Universities Library Statistics, 1988-89* (Chicago: Association of College and Research Libraries, 1991)

Molyneux, Robert E. (1992). "Communication." *Libraries & Culture*, 27, no. 1, 1992, pp. 113-118.

Molyneux, Robert E. (1994a). "More Hortatory Than Factual: Fremont Rider's Exponential Growth Hypothesis—and the Context of Exponentialism." In *For the Good of the Order: Essays in Honor of Edward G. Holley*, Delmus E. Williams, et al. eds. (Greenwich, JAI Press, 1994), pp. 85-117. Updates "Patterns" and presents a more rigorous analysis of the subject of academic library growth.

Molyneux, Robert E. (1994b). "What Did Rider Do? An Inquiry Into The Methodology of Fremont Rider's 'The Scholar and The Future of The Research Library.'" *Libraries & Culture*, 29, no. 3, 1994, pp. 297-325. Discusses how Fremont Rider arrived at his exponential growth hypothesis and calls his methods into question.

Molyneux, Robert E. (1996) "Fremont Rider's Legacy," *ARL: A Bimonthly Newsletter of Research Library Issues and Actions*, 188, (October 1996): pp.6. Internet edition is at: <http://arl.cni.org/newsltr/188/fremont.html>.

Princeton University Library (1947). *College and University Library Statistics, 1919/20 to 1943/44.* (n.p. Princeton University Library). This title is available from University Microfilms as one of its Books on Demand: AG1-OP30634.

Rider, Fremont (1944). *The Scholar and the Future of the Research Library.* (New York: Hadham Press, 1944)

SAS. SAS Institute. <http://www.sas.com/>

Stubbs, Kendon and David Buxton (1981). *Cumulated ARL University Library Statistics 1962-63 through 1978-79* (Washington, Association of Research Libraries, 1981).

Stubbs, Kendon and Robert Molyneux (1990). *Research Library Statistics, 1907/08 Through 1987/88, A Guide to the Machine-Readable Version of the Gerould and ARL Statistics* (Washington: Association of Research Libraries, 1990.)

Stubbs, Kendon (1996). *ARL Statistics, 1992-96: A Guide to the Machine-Readable Version of the ARL Statistics* (Washington: Association of Research Libraries, 1996) and Internet edition: *ARL Statistics, 1992-96: A Guide to the Machine-Readable Version of the ARL Statistics* at: <http://viva.lib.virginia.edu/socsci/arl/1994/96doc.html>

Stubbs, Kendon (1985). "Introduction." *ARL Statistics, 1983-84.* pp 4-7

U.S. Bureau of Education (1876). *Public Libraries in the United States of America* (Washington, Government Printing Office, 1876).

FOOTNOTES

First Edition Error

There is an error in the text of the first edition that appears in two places. This error does not affect earlier editions of the data.

Appendix 2 reversed part of the text for Salaries for Columbia and Chicago in 1922/23. Chicago should have a value of \$126,015.32 (as it correctly appears in the data) and Columbia should have noted the value as missing. This error was perpetuated because Chicago's value is noted as Missing with a reference to Appendix 2. Appendix 2 had no entry for Columbia's salaries for this year but for Chicago, the entry on page 223 of the first edition was:

Year	Variable	Value used in this edition
1922/23	Salaries	*

All four issues appear to be from the same distribution copy and they all have the original figure for salaries at Chicago lined out. The figure which has been lined out is

not clear but it appears to be \$28,300. Consider the figures for this variable for these years:

Year	Salaries
1920/21	\$121.742.11
1921/22	28,300(?)
1922/23	165.962.96

Given the values for this variable from the year before and after and given the fact the same change was made on all four copies, this variable is treated here as if there were no value reported.

The entry for Columbia in 1922/23 is corrected in this edition.

1. **Association of Research Libraries.** <http://www.arl.org/>
2. **SAS.** <http://www.sas.com/>