Library Evaluation: The Shanghai Experience

Kuei Chiu

Follow this and additional works at: https://scholarsarchive.byu.edu/jeal

BYU ScholarsArchive Citation

Available at: https://scholarsarchive.byu.edu/jeal/vol2000/iss122/6
LIBRARY EVALUATION: 
THE SHANGHAI EXPERIENCE

Kuei Chiu
University of California Riverside

Library statistics are an important source for library management since they provide library administrators with a tangible reference for decision making. They are also useful for library science research. In North America, there are statistical resources such as ARL (Association of Research Libraries) statistics to compare library standings. In China, however, no statistics equivalent to ARL data are currently available. Historically, Chinese libraries have not been active in library data collection. As a part of the overdue reform sweeping through the Chinese education system, an "Evaluation of Libraries of Higher Learning" was conducted in China from 1992 to 1994. Even if this evaluation has its Chinese characteristics of a strong political tone and some of the evaluation criteria may be considered "interesting" at least from a Western point of view, it is a breakthrough in collecting Chinese library statistics as well as improving libraries and their services. This paper will try to investigate the evaluation, its methodology, its significance, and to see how it differs from the ARL statistics.

The Evaluation

Since its implementation in 1978 of the "open door" policy, which aimed to modernize China and to boost her standing among the nations, the Chinese government has been trying vigorously to reform its education system, especially what is at the peak of the pyramid: the institutions of higher education. More and more funding has been allocated to higher education. The libraries of higher learning institutions, in turn, have also benefited from this trend and have been given much more attention in terms of such things as funding allocation, facility expansion, staff benefits, and training. In order to have a clear view of the status of the libraries and to help future development, the State Education Commission requested libraries of higher learning in China to launch an evaluation in 1991. This evaluation was to aim at ascertaining the "overall level of librarianship so as to enable libraries at all levels to play necessary roles, improve efficiency and get involved in and serve the socialist modernization drive in a more effective manner."¹ The evaluation emphasized three main aspects: 1. Internal conditions and external environment of library development. 2. Library operating conditions. 3. Library performance capability.

There are four levels of evaluation: national, regional, internal within the school, and internal within the library. So far, the highest level achieved is the regional evaluation in which libraries of certain regions have been evaluated. The guidelines of the library evaluation were the "Rules on Libraries of Higher Learning" enacted by the State Education Commission in 1987. Originally, the Rules were introduced to regulate college and university libraries, so that the libraries would have a state-issued manual for development. In many ways, the Rules have helped the libraries gain the share they deserve from school administrations. For example, the Rules require that each school of higher learning allocate at least 5% of its whole budget to its library system.
Under the auspices of the Chinese State Education Commission, the evaluation of libraries of higher learning was launched in different regions of the country. In Shanghai, it was initiated and administered by the Bureau of Higher Education of the Shanghai municipal government. Being at the threshold of the new millennium, the Shanghai Bureau of Higher Education found it extremely crucial to determine the strengths and the weaknesses of college and university libraries in Shanghai, hoping that the results of the evaluation could help future planning. The library evaluation was also part of the overall performance evaluation of higher learning institutions in Shanghai.

**Methodology**

It must be pointed out that although it is a nationwide effort, the evaluation was conducted in each region in its own way. In Shanghai, the "Library Evaluation Index System" was developed after "considering the spirit of the Memorandum on Library Evaluation of Higher Learning Institutions’ issued by the State Education Commission... and with reference to the evaluation index systems compiled by library evaluation committees of other regions". The Shanghai evaluation covered all higher education institutions in Shanghai: fifty universities and colleges in total. Specifically, there were eight key state universities affiliated to the State Education Commission, twenty-one universities affiliated to state ministries, eleven local universities, and eighteen post-secondary colleges. Due to the recent mergers of Chinese universities and colleges, some of them no longer exist as independent schools. Although the evaluation was conducted in 1992-1994, it was based on the statistics collected during the four years of 1988-1991.

a. Evaluation administration

In order to carry out the evaluation effectively and accurately, an evaluation administration system was created. On the top of the evaluation administration structure, there was *Shanghai* *Putong Gaodeng Xuexiao Tushuguan Pinggu Weiyuanhui* (Evaluation Committee on Libraries of Higher Learning in Shanghai), which oversaw the whole evaluation operation. The key mission of this 15-member committee was to develop the evaluation index system as well as application manuals following the evaluation guidelines set by the State Education Commission. In addition, it also appointed the evaluation expert team.

The *Pinggu Zhuanjiazu* (Evaluation expert team) consisted of nine to eleven library experts. Its basic responsibilities were to review the statistics and self-evaluation reports submitted by libraries being evaluated and to canvass these libraries to verify if all evaluation guidelines had been strictly followed. It was also the duty of the expert team to rate the libraries at the end of the evaluation. The expert team was also assisted by an evaluation secretariat that managed the daily administrative work for the expert team including collecting and sorting self-evaluation reports submitted by libraries and scheduling library field trips for expert team members.

In each individual library, there was a library evaluation task force whose responsibilities were to collect library statistics according to the guidelines, to conduct self-evaluation, and to present reports on self-evaluation as well as plans for library's future improvement based on the results of self-evaluation.

b. Evaluation standards

The centerpiece of the evaluation is the *Shanghai* *Putong Gaodeng Xuexiao Tushuguan Pinggu Zhun Tixi ji Pinggu Biaozhun* (Evaluation Index System and Standards for Libraries of Higher
Learning in Shanghai, referred to as EIS hereafter). When statistics were collected from each library, they were used as raw data that would be used eventually to rate libraries into four different ranks: A (A-), B, C, and D.

The EIS in Shanghai was created according to the Putong Gaodeng Xuexiao Tushuguan Guicheng (Rules on Libraries of Higher Learning, referred to as Rules hereafter) issued by the State Education Commission. Another document that helped develop the EIS was the Putong Gaodiao Tushuguan Pinggu Zhibiao Tixi Dagang (Outlines for Evaluation Indexing System of Libraries of Higher Learning, referred to as Outlines hereafter) by the National Commission of Libraries of Higher Learning.

There are six main categories of evaluation with a number of subcategories. At the bottom of this 3-tier evaluation system there are 48 indexes. With this evaluation system, the Evaluation Committee hoped to measure the standings of these libraries in a fair and scientific way. More details on EIS are available in the appendix of this article.

The biggest challenge in developing EIS in the Shanghai evaluation was that the Rules on which the evaluation was based were too abstract to be used directly in EIS. In the meantime, although the Outlines tried to make some indexes quantifiable, they were not adequate in practice. To solve this problem, supporting documents were used including library-related government rules and regulations. Even so, it was still difficult to quantify many elements such as accomplishments in promoting library resource sharing, observation of library rules, regulations by the staff, work on political consciousness and thought, etc. Although it is now more than twenty years after the Cultural Revolution ended, and political ideology plays a smaller role in the Chinese society today, it is still perceived as an essential part of job measurement even if it is not measurable. Among the 48 third-tier indexes, only 14 or 29.2% were quantifiable.

c. Evaluation procedure
Using EIS, the evaluation was conducted with the combination of self-evaluation and inspection of the expert team. As soon as the evaluation was launched, it was the job of the library evaluation task force of each individual library to collect statistics on the library and evaluate them according to the guidelines.

The main benefit of self-evaluation was that it provided the libraries with the opportunity to find out its weaknesses first so that they could try to correct them if possible. All the collected statistics were then sorted according to EIS, before later review by the expert team. At the end of self-evaluation, each library self-evaluation task force completed a self-evaluation report with a score sheet. The report narrated what the library’s current standing was in terms evaluation categories specified in EIS and how it would improve if weaknesses were found.

After self-evaluation, an expert team consisting of library specialists from outside the library canvassed the library, inspecting it in comparison with the self-evaluation report and statistics compiled by the library. This field inspection by the expert team was the core of the whole evaluation. First, the inspection proved the validity of the statistics collected by individual libraries. In self-evaluation, many libraries might lose their objectivity by giving themselves higher scores. The inspection of the expert team offered a chance for proofreading the results of the preliminary
evaluation. Second, since the expert team visited all libraries being evaluated, and reviewed all documents submitted, they served as a controlling mechanism in the evaluation that made the whole process standardized. Although it was difficult to eliminate completely the personal preferences of the team members, the inspection was an effective way to place all libraries under the same standards of measurement as well as rule interpretation. Finally, because of the expertise and experience of the experts, they were able to offer valuable advice for improvement to the libraries being inspected. Many problems in library management and operations neglected by some libraries were fixed due to the involvement of the expert team.

After field inspection, the expert team conducted an overall review of the library and filed a final report of the evaluation. The report included all six aspects of evaluation with a single score as well as the ranking of the library at the end. Since the essential goal of this library evaluation was to have a precise view of the current standing of libraries of higher learning in Shanghai and to make improvements in the future, many libraries took necessary steps for library service improvement and reorganization right after the evaluation.

Conclusion

According to James Krikelas, the uses of library statistics "may be grouped under one of two categories: 1) to support administrative decisions and 2) to describe various types of library activities. A third category, seldom encountered in the literature but frequently implied (and perhaps of more concern in the future) is the use of statistical data in library research, that is, in attempting to establish general principles and relationships concerning library organization, administration, and use."

Although the U.S. is a relatively young country compared with China, it has a long history of collecting nationwide statistics on libraries. The collection of library statistics in the U.S. has been developed into a routine process. Statistics are regularly collected from all kinds of libraries, and reports are published. According to library historians, "the practice of collecting statistics on "public" (i.e. not private) libraries can be traced back as far as 1837 when Charles Bower started to include library statistics in his American Almanac and Repository for Useful Knowledge."

For academic libraries in North America, the ARL statistics are the most authentic and important statistical source for library administrators and researchers on library science to compare the status of ARL member libraries. "ARL libraries are a relatively small subset of all academic libraries in North America, but they account for a large portion of academic library resources in size of assets, budgets, and number of users they serve. Although limited to describing only ARL member institutions, the ARL data are indicators of library trends for North America in general because of the size of resources and users they represent and serve."

The Shanghai evaluation surely has its Chinese characteristics. However, despite its political and ideological tone, the general goals of the Shanghai evaluation meet Krikelas' description and are similar to what the ARL statistics do every year. Due to the lack of previous evaluation, the unprecedented attempt in Shanghai was very valuable. The evaluation offered the libraries an opportunity to discover their deficiencies in all aspects of library administration, services, equipment, and so on. Many problems were found and fixed in the evaluation. The evaluation also served as an alert to the school administrators who did not pay enough attention to the importance of library
resources and their development. A number of schools offered substantial library budget increases as soon as the evaluation started. The libraries of higher learning in Shanghai are definitely in much better shape after the evaluation.

In spite of few similarities between the evaluation in Shanghai and the ARL statistics, there are many distinct features that separate the two.

1. A one time attempt vs. annual reporting
It was the first time and only time in Shanghai. There is no plan to conduct a similar evaluation in the near future. The ARL statistics, however, started as Gerould statistics in 1908 when James T. Gerould, librarian at the University of Minnesota started collecting academic library data. It was later know as "Princeton series", when Gerould moved to Princeton. The Gerould/Princeton/ARL statistics are the longest continuous series of data on libraries in the United States. The most obvious benefit of collecting statistics annually and systematically is its value for comparison studies and future projection. In Shanghai's case, the data collected for the four years under evaluation would have been more meaningful if there were at least a second set of data for comparison.

2. Comprehensive data collection vs. self reporting
Although the Shanghai evaluation was a one-time effort, it was well organized. The whole evaluation mechanism was carefully planned and developed. The self-evaluation, inspection of the expert team and re-inspection after the evaluation truly helped the libraries to take advantage of the evaluation and improve themselves. Another benefit is the presence of the expert team that evaluated all libraries using the same guidelines. Here lies the biggest drawback of the ARL statistics. Over the years, the ARL data "have been based on voluntary submissions, by the individual libraries, of answers to specific questions in the light of rules and definitions, tight or loose, provided by the compilers. There has been little or no policing of these submissions to ensure compliance with the rules and conformity with the definitions and instructions... It is evident that all libraries do not interpret the rules in the same way and that individual libraries have made major changes in their own methods of gathering and reporting over the years, resulting in otherwise inexplicable discontinuities in their data." However, we have to realize that given the geographical coverage and number of ARL members, it would be cost-prohibitive to have an inspection team canvass all member libraries on an annual basis.

3. Scoring system vs. direct statistics report
One of the goals of the Shanghai evaluation was to rate libraries in several ranks. The raw data was collected and then filtered through EIS in a way that was somewhat subjective due to the numerous weighting coefficients assigned. In the final evaluation report, all the data were translated into a text with a single score attached at the end. In the ARL statistics, statistical data are the main body of the report and each library ranks differently in each category depending on how it does in that category. Explanation of statistics is only secondary, and it is hidden in the notes section. The ARL statistics "do not reveal whether this or that library meets accepted quantitative standards. But in certain ways the data can tell us where university libraries are." The Shanghai evaluation, however, tried to tell us where the libraries should be. From a library research point of view, the ARL statistics are more useful because they let the readers draw conclusions themselves based on the statistics per se. Also, it is difficult to have a "one-size-fits-all" approach for some libraries since the libraries may perform and serve their patrons well in their own way even if they don't meet standards in certain categories.
This is the kind of problem the Shanghai evaluation encountered while assessing some medical university libraries.10

4. Confidential report vs. open publication
While all the ARL statistics are available to the public with its annual publication, the results of the Shanghai evaluation are only for internal circulation. The official publication of the Shanghai evaluation is: Gaodeng Xuexiao Tushuguan Pinggu Shijian yu Yanjiu (Practice and Research on Evaluation of Libraries of Higher Learning), which includes much information on evaluation background and methodology as well as numerous forms and analysis. However, except for a list at the beginning of the book of libraries being evaluated, no specific names of libraries are mentioned thereafter.

5. Local vs. national
Although the Shanghai evaluation was a nationwide effort under the auspices of the Chinese State Education Commission, its evaluation criteria and standards were tailored to the local specifications. Such being the case, it is difficult to have a line-to-line comparison with similar evaluations conducted in other regions of China during the same period. The ARL statistics, however, cover libraries located all over North America, and it is ready for comparison assuming that all libraries interpret the categories in the same/similar way.

6. Activities covered
This is what sets the Shanghai evaluation apart from the ARL statistics. When they are compared side by side, only a fraction of them are similar. There should be no surprise for this finding. The reason is that while the ARL statistics are mainly a statistical report, the Shanghai evaluation is not. It is only an evaluation that tries to rank libraries. In the Shanghai evaluation, data collection is only a method used to facilitate the evaluation. It is not the result. Also, as mentioned earlier, in the EIS of Shanghai only 14 or 29.2% among the 48 third-tier indexes are quantifiable. Many of the questions in the Shanghai evaluation are of a subjective type, i.e. How are the reference services in your library?

There is no simple answer to the big difference between the two, but we may find some hint from James Krikelas: “No obvious measurements can be made to determine how effective the library is in providing the services necessary to meet its objectives, and there is no obvious way of applying uniform measurements to libraries which are serving diverse objectives.” At least, it is a difficult if not impossible mission to compare two different kinds of libraries. The ARL statistics do not happen overnight. It is a product of efforts over several generations. Even today, it is still evolving to meet the challenges of new library services that still cannot be recorded in the current reporting system (such as how to measure the use of electronic databases). While American academic libraries are in their full maturity after so many years of development, their Chinese counterparts are just starting to modernize themselves. Despite its weaknesses and drawbacks, the Shanghai evaluation is a groundbreaking effort in the Chinese library statistics collection. It will be a very useful statistical resource to support library administrative decisions, to improve library services, and to facilitate library research if more future data can be collected in a timely and open fashion.
Appendix: Evaluation Index System (EIS)

There are six main categories of evaluation with a number of subcategories (number of subcategories listed in parenthesis): A. Library administration (6); B. Patron Services (3); C. Collection Development and Technical Services (3); D. Application of Library Technology (2); E. Application of Scientific Management (3); F. Awards for Research (2). At the bottom of this 3-tier evaluation system there are 48 indexes. While the main categories and their subcategories are assigned with different weighting coefficient depending on the importance of the category, each of the 48 indexes is assigned a different value (the highest possible score one library might get).

In EIS, category A shows the current conditions of library development, while categories B-F demonstrate the functions and performance of library and its services. In category A, there are basically three elements: a. Library personnel and their educational background and professional capabilities; b. Budget for library material purchases; c. Size of the library, its furniture and equipment, as well as number of library volumes per student.

Category B has been assigned a weighting coefficient of 0.3, the highest among all categories. It covers library opening hours, circulation numbers, percentage of open stacks, as well as performance in the areas of circulation, reading guidance, bibliographic instruction, and information retrieval. Category C is for collection development, cataloging and weeding. Category D is for the application of computer, audio-video equipment and the use of copy machines in the library. Although Category D stands alone, it has the smallest weighting coefficient among all categories: 0.08. Categories E and F cover library management and awards and research activities of library staff. Like Category D, almost all items covered in Category E and F are not quantifiable.

3 Ibid., p. 104.
7 Krikelas, p. 467.
10 Gaodeng Xuexiao Tushuguan Pinggu Shijian yu Yanjiu, p. 151.