

# LIBRARY RESOURCES & TECHNICAL SERVICES

Vol. 23, No. 1

Winter 1979

## CONTENTS

	Page
Some Practical Observations on the Writing, Implementation, and Revision of Collection Development Policy. <i>Charles B. Osburn</i>	7
Collection Evaluation in Research Libraries: The Search for Quality, Consistency, and System in Collection Development. <i>Paul H. Mosher</i>	16
Selection Methodology in Academic Libraries. <i>Hendrik Edelman</i>	33
The Necessity for a Collection Development Policy Statement. <i>Y. T. Feng</i>	39
Allocation of Funds in Support of Collection Development in Public Libraries. <i>Ann Bender</i>	45
Collection Development: A Summary of Workshop Discussions. <i>Norman Dudley</i>	52
Specificity in Subject Headings. <i>Hilda Steinweg</i>	55
Examining the Library of Congress <i>Subject Catalog</i> . <i>David F. Kohl</i>	69
Resources and Technical Services Division: Annual Reports, 1977/78	75
RTSD Nominees—1979 Election	88

---

AMERICAN LIBRARY ASSOCIATION  
RESOURCES AND TECHNICAL SERVICES DIVISION

# Collection Evaluation in Research Libraries: The Search for Quality, Consistency, and System In Collection Development\*

PAUL H. MOSHER  
*Assistant Director for Collection Development*  
Stanford University  
Stanford, California

*The history, literature, and methodology of collection evaluation or assessment in American research libraries are reviewed; current problems, tools, and methodology of evaluation are discussed; and an ongoing collection evaluation program at the Stanford University Libraries is described.*

ROGERS AND WEBER remind us that

a distinguished collection of books is the *sine qua non* of a great research library. Librarians and nonlibrarians alike may forget this because of the current preoccupation with computers, the enormous fiscal problems associated with larger and larger library buildings, and the vexatious bibliographic problems of controlling huge bodies of materials in increasingly diverse formats. However important these other factors may be, we should not forget that they are subsidiary to the root element that gives the library its name—*liber* the Latin word for book.<sup>1</sup>

We would by now agree, I am sure, that the generic term "book" includes an impressive array of formats which are not strictly books but which serve the needs of our readers and scholars in ways still analogous to the rolls and codices which filled the libraries of Alexandria or Rome.

"Evaluate" is an English word coming from the French term *evaluer*—to determine the value of something. In the case of collection evaluation, the term has to do with measuring or assessing the *quality*

\*Edited version of a paper presented at the Preconference Institute on Collection Development, sponsored by the Collection Development Committee, Resources Section, RTSD, Detroit, June 1977.

of a collection or collections. Let me avoid a possible early pitfall by defining quality as "the utility or benefit of library collections to library patrons, their needs and work, and to institutional programs"; in the case of academic libraries, for example, the utility of collections to the requirements of academic programs of teaching and research.

Evaluation of collections is most often considered a function of collection development and should be related to the planning, selection, and pruning of collections. The benefits of a well-planned program of collection evaluation are many, including the following important ones:

1. A more accurate understanding of the scope, depth, and utility of collections.
2. A guide and basis for collection planning.
3. An aid to the preparation of a collection development policy.
4. A way to measure the effectiveness of a collection development policy.
5. A method to ascertain collection adequacy or quality.
6. A means to rectify inadequacies in and improve library holdings.
7. An opportunity to focus human and monetary resources on collection areas most needing attention.
8. A justification for book budget increases.
9. A demonstration to administrators that something is being done to change the "bottomless pit" of library acquisitions budgets.

Collection evaluations cannot only provide answers to questions like "How large is our collection in, say, art history?" or "How good is our collection in French linguistics?" They should also be able to tell us how well we are doing and how rapidly we are achieving specified goals.

Now that I have introduced the topic and described some of its benefits, I would like to do some other things: first to give a brief overview of the history and progress of the art of collection evaluation in American academic libraries, then deal with some common problems of recent scholarship and practice—particularly the issues of quality versus quantity, subjective versus objective evaluation, and the degree to which collection evaluation can be considered an art or a science at this stage of its development, and finally to describe the theory and practice of some recent methodology.

## History

Surveys were undertaken at a number of university libraries between 1933 and 1950, including Chicago, Harvard, Pennsylvania, Indiana, Stanford, Cornell, New Hampshire, Texas A & M, and Alabama Polytechnic.<sup>2</sup> Robert Downs presented the results of a statistical survey of research materials in U.S. libraries (primarily academic) in 1942 and completed a similar study for Canada in 1967.<sup>3</sup> With the exception of Raney's 1933 study of Chicago, Orr and Carlson's report on Texas A & M, and to a lesser extent Wilson, Downs, and Tauber's

examination of Cornell, all of these were descriptive surveys rather than analytic or systematic evaluations. The distinction between a survey as a descriptive overview and an evaluation as a systematic analysis is important and should be kept in mind as we proceed.

The principal body of collection evaluators, describers, and selectors during these decades consisted of the department faculties of each institution, and they were responsible for the descriptive or evaluative portions of nearly all of these studies. The most comprehensive of the three surveys that attempted evaluations of collections is the Chicago study of 1933, which involved 200 faculty members checking some 400 bibliographies and a schematic master list compiled from the *Union List of Serials*; a separate check of lists or bibliographies was made by each department or school of the university. While the methods and products were uneven, the overall result was very impressive. This mammoth enterprise, which was never again undertaken in the same scope and which may first have given collection evaluation its reputation of being inordinately costly and time-consuming, was orchestrated, though not carried out, by M. Llewellyn Raney, the director of libraries. In retrospect, its achievement may possibly have been worth the work involved. It resulted in justification of retrospective needs of 1,400,000 volumes, and expenditure of the then staggering sum of \$4,000,000, with a "fallback" core of 713,000 volumes costing \$2,700,000 which Raney urged upon Chicago's trustees as the effort necessary to bring Chicago's collections to the level of those at Harvard or Yale. The study also resulted in a vastly improved understanding of Chicago's library collections and needs and set the university library on a path of development that created one of the greatest library collections of the central United States.

Library surveys have remained a common form of simple evaluative tool—especially in the area of accreditation, which I will not cover here—and the 1930s had a distant echo in 1961 when the University of Michigan conducted a faculty survey appraisal of the University Library collections, which appears to have had little more practical result than most of its predecessors.<sup>4</sup>

All of the surveys before 1950, both the analytic one by Chicago and the purely descriptive ones elsewhere, were carried out by faculty rather than librarians. During this period collection development was regarded as a faculty preserve. Harry Bach reports that librarians did not participate as "active agents in the selection of a majority of books that went into the library," and in 1930 a study of U.S. land grant universities and colleges revealed that in thirty-three institutions librarians functioned in the collection development sphere only to prevent duplication. In ten other institutions, books were ordered by departments without any supervision by librarians.<sup>5</sup> This pattern of faculty control of collection development in American academic libraries continued into the 1960s, as Periam Danton has noted.<sup>6</sup> It is still regarded as the natural order of things in some libraries. In reporting

on a survey of library resources carried out at Columbia in 1958, Tauber reported that 2,250 faculty members were questioned about collection scope and adequacy, and 644 responses were received "in fairly complete shape." Tauber dealt with faculty as the proper principal selectors but expressed concern with the inconsistency, lack of system, lack of agreement, and inadequate levels of interest and support such a faculty pattern provided. He suggested that the alternative for a good research library would be to obtain subject and field specialists who could anticipate future collection requirements as well as respond to faculty expression of need. He seems to have regarded this possibility as remote, however, for he appealed to universities to get their faculties, departments, and schools to assume a greater collection development responsibility, for libraries to involve them more, and even exhorted departments to hire faculty members with respect to their ability to strengthen library collections!<sup>7</sup>

During the next decade, academic libraries—especially large research libraries—found Tauber's idealism increasingly impractical and utopian. In addition to the fact that departments had come less and less frequently to hire faculty because of their capacity to develop systematic library collections, university enrollments, faculties, and physical plants began the nearly unrestrained growth and expansion pattern of the 1960s. At the same time, the output of library materials was growing rapidly, the "publish or perish" syndrome overtook U.S. universities everywhere, and the "new professionalism" in American scholarship resulted in greater demand on the materials resources of libraries and, at the same time, less interest in "doing the library's work for it," as the faculty said. The older pattern of library dependence on the faculty for collection development became increasingly impractical and unrealistic, and many libraries began to build staffs of subject area or language specialists in acquisitions or reference departments, as independent units, or a combination of these. In an important review article published in 1968, David Lane found that most librarians had changed to the view that collection development was a responsibility of librarians rather than faculty but that change in practice has followed rather more slowly.<sup>8</sup>

The value of having subject-educated specialist librarians responsible for collection development in research libraries had been demonstrated as early as 1936 by Waples and Lasswell in a study of holdings of 573 notable titles in 4 social science fields by 6 major American research libraries. Excluding data for the Library of Congress, they discovered that the New York Public Library (NYPL) held 92 percent of the titles compared to 68 percent at Harvard, the next highest ranking institution.<sup>9</sup> In commenting on the results of this study, Danton inquired into the reasons why this should be so; he found that the significant difference was that selection at NYPL, unlike the universities, was the ongoing responsibility of a corps of subject specialist librarians.<sup>10</sup> The other libraries depended at that time entirely—or al-

most entirely—on a less consistent and systematic pattern of faculty selection and recommendation.

Another point made by Waples and Lasswell that deserves repeating today is that while NYPL made a better showing in books and journals than any other library evaluated—American or European—it spent less money than any of the other five American libraries.<sup>11</sup> As Waples and Lasswell observed: "In terms of our previous discussion and of the data available, the New York Public Library appears more attentive to the future needs of American scholars in the social sciences. It pays greater deference to posterity."<sup>12</sup>

The move of primary collection development responsibility from the faculty to the library, the increase of attention devoted to collection development which resulted, and the common attempt to systematize, rationalize, and improve the planning and procedures of library collection development during the ensuing decade and a half (which in a sense has produced the present preconference) has been one of the most significant and original contributions to the growth of professional librarianship in the United States during the last generation.<sup>13</sup>

One of the first problems academic libraries have faced as a result of this quiet revolution has been a need to discover where they are with respect to the adequacy and quality of their collections and to develop plans and programs for more systematic improvements and growth of their library resources. This has produced between the late 1950s and the present "more work . . . on the evaluation of collections . . . than on any other facet of the library."<sup>14</sup>

#### *Literature and Methodology*

The two best and most comprehensive recent surveys of the literature of collection evaluation are those of George Bonn and F. W. Lancaster.<sup>15</sup> The American Library Association has prepared guidelines for the evaluation of collections that provide the most complete and helpful information on principles and methods of collection analysis, and the Association of Research Libraries has assembled a useful collection of documents from libraries and organizations actively involved in collection evaluation efforts.<sup>16</sup> Bonn provides a general literature survey with commentary, followed by some very useful pages of summary and criticism. Lancaster has as his primary aim the evaluation of collection use levels with the ultimate aim of "optimizing storage of the collection," but the section on evaluating the quality of collections and his up-to-date bibliography are highly valuable. The ARL Office of University Library Management Studies is also sponsoring a self-study collection analysis project (CAP) among research libraries, which is developing a set of reports and procedures concerning collection evaluations.<sup>17</sup>

Discussion of the purpose, goals, and methods of collection evaluation during the last fifteen years has been plagued by controversy and

debate over two issues. The major debate concerns the issue of objectivity versus subjectivity, simplistically transformed by some into an apparent conflict between quantity versus quality, which itself resolves into issues of methodology and philosophy: Should evaluation of collections be qualitative or quantitative in nature? Is quantitative study or data more objective and "scientific" than evaluative or judgmental data? The second issue revolves around matters of the purpose or goal of an evaluation. Often techniques are interpreted as more broadly useful or applicable than they were intended to be. A technique appropriate for an accreditation survey, for example, may convey little useful information to a librarian or faculty member concerned about a library's collection in classical studies, however easy the data may be to gather.

The value of using statistical measurements in evaluating library collections and services, and as a survey technique for purposes of accreditation or measuring gross collection adequacy, has been propounded by a number of librarians: perhaps the most important and influential among them has been Robert Downs. In "Resources of American Libraries: A Quantitative Picture," his first substantial statistical survey, Downs eschewed definitions of quality or of research materials, a term he said he was using "almost without definition." "No differentiation is being made," he said, "between items on any basis of quality or applicability to research needs."<sup>18</sup>

Downs followed this initial survey with statistical evaluations of many other libraries, library systems, and the research collection resources of Canada as well as the U.S. While Downs admitted that large collections do not themselves a great library make, he nevertheless concluded that the absolute size of collections had been demonstrated to be an important factor in judging collection adequacy. He also found strong correlation between the size of an academic library and the prestige of a parent university or college and listed other correlations between library—or even university—quality and rate of annual collection growth, comparison of size with other universities, the number of journal subscriptions, and the number of volumes in the reference collection.<sup>19</sup>

Quantitative methods of evaluating library collections have become increasingly popular in recent years for reasons other than Downs' influence. Quantification is currently in high fashion in the social sciences, and librarianship, calling itself "library science," sometimes identifies closely with this development of the larger group of disciplines of which many librarians feel themselves a part.

Many scholars of librarianship have devised formula approaches to library size, models to measure the effectiveness of collections, computer evaluation techniques, and a host of other statistical measures of collection adequacy and quality that are described and debated in the literature. George Bonn tends to favor "quantitative" as more realistic than "qualitative" ones but ultimately argues for a balance of methods

designed to match collections with user needs.<sup>20</sup> Lancaster devotes a good deal of effort to quantitative methods of evaluating collection quality; while he takes Downs, Williams, and Blau and Margulies to task, for example, for their conclusion that there is a positive correlation between absolute collection size and the academic excellence of an institution, he then goes on to argue that rate of growth of collection size is a better indicator of collection quality, but that the rate of increase in the absolute number of volumes added plus the absolute size of the library in number of volumes together correlate to the best indicator of "academic excellence of the university."<sup>21</sup> Without regard for the logical problems of trying to assess *university* excellence on the basis of *library* collection measures, one may question whether correlations may have been extended too far.

The tendency to rely more and more often on quantitative measures of collection quality and adequacy has come under attack from a number of directions. A group of Canadian librarians centered at Dalhousie University, coming from a tradition of conservative and scholarly library service, decry what they see as a growing reliance on "numerology," taking quantifiers like Downs to task for their failure to address questions of quality and for their *ex post facto hoc ergo propter hoc* mentality. John Ettlinger and John Miller are often witty and sometimes devastating in pointing out the logical and procedural fallacies of their "numerologist" opponents, but occasionally allow passion to dominate what others might call common sense.<sup>22</sup>

There have been second thoughts and words of caution even from within the ranks of Ettlinger's "numerologists." George Piternick has described basic fallacies in attempts to correlate the rate of growth of collections and the overall quality of an educational program, observing that

strong positive correlation between two sets of data does not in itself establish causal connection between them. The existence of a high correlation between absolute size of library holdings and the academic quality of the institution does not establish that the former causes the latter. . . . It does suggest strongly, however, that the two are not independent—very likely cause and effect are intertwined, as Clapp has concluded.<sup>23</sup>

James Krikelas has commented that statistics are useful to support administrative decision making and to describe various types of library activities. But he also points out that statistics can be meaningless, data can be inadequate, categories can be ambiguous, and published data may contain internal inconsistencies. "Mensuration is an important element in any scientific discipline, but it is not the primary objective; it is, rather, a means to an end."<sup>24</sup>

Krikelas put his finger directly on the knot of the dispute: the issue is not really one of "subjective" versus "objective," which even Bonn has suggested. Quantitative data is not necessarily "more objective" than qualitative or evaluative data. A statistic is susceptible of mislead-

ing presentation and interpretation, just as a judgment may be based on insufficient information. An accurate, soundly based judgment may be as objective a datum as a set of statistics.

There is a useful adage about collection development librarians which asserts: "In nature, it is the early bird which catches the worm; in collection development the bird that catches the worm is the one who recognizes a worm when he sees it!" The point is that librarianship is a profession that requires a capacity for discrimination and judgment. The acquisition, storage, and provision of access to thousands or millions of items requires that librarians be capable of making many judgments and discriminations each day. This is certainly true in collection development. These can be made *less* subjective by the use of appropriate criteria, policies, guidelines, and tools, but the exercise of choice and judgment cannot be completely eliminated from the practice of librarianship. These judgments can and should be as well informed as possible, based on evidence both quantitative and qualitative, on faculty and specialist opinion, upon statistics of size and growth rate, upon analysis of teaching and research programs, and the study of research patterns in various disciplines. The issue is not one of subjectivity versus objectivity or quantitative versus qualitative measures; it is an issue of choosing the right methodological tools to produce the kind and quality of result needed.

Many tools are available to the librarian who wishes to evaluate library collections, and it is the librarian's business to ascertain on the basis of the institution's nature, mission and goals, the collection development policy statement, and on the nature of the information needed, what evaluative techniques will be most helpful and meaningful.

I would like to endorse the view of the draft ALA "Guidelines" and of Harry Bach that in most cases the best evaluation will normally involve the application of more than one technique and a comparison or combination of the resulting data.<sup>25</sup> But the extent of one's means and the length of time available for evaluation will often provide external limits to the depth and length of an evaluation project, and these factors will have to contribute to the technique one chooses.

#### *Conduct of Evaluations*

One should begin any process of collection evaluation with the institution's missions and goals firmly in mind. Possession of a recent, accurate collection development policy statement is vital for determining areas and objectives for the assessment of collections. An evaluation planner should also be sure to remember any coordinated collection arrangements or assignments among branches, or with libraries at other institutions.

It is important to identify the questions to be answered by the evaluation or analysis. Is the collection of adequate size? Is it serving the needs of the patrons? What materials can be pruned for storage?

Does the library possess the most needed materials for graduate field exam reading in anthropology? Can data be obtained on collection adequacy as ammunition for the next budget go-around? What is the quality, for research purposes, of the Latin American collections? What is the adequacy of the undergraduate collections in history? An evaluator should always keep in mind the nature of the library and the nature of the need. For example, the Clapp-Jordan methodology will supply useful data on collection size *adequacy* if you want to know if your collection is adequate in numbers, but it will provide no help in ascertaining the *quality* of collections.<sup>26</sup>

With this material and these decisions in hand consult the ALA "Guidelines," and guides such as Bonn and Lancaster, and develop a plan or campaign of action. In consultation with an appropriate group of informed users—or administrators if budget justification is desired—seek to measure the utility and the practicality of the plan and tools; you will want to use those that will produce the most useful, appropriate, practical, and beneficial results.

Research or large academic libraries usually consist of a congeries of branches—research, special, undergraduate, high-use—with a central main core collection. Each branch or major collection of such an institution should normally be considered as a unique unit when planning an evaluation, and the methodology should be specially designed around its nature and goals.

Evaluation of the humanities and social sciences collections of the main library and branches remains a special problem of academic librarianship. Gross measures of collection size by subject breakdown, such as the comparative shelflist measurement project carried out under the aegis of the RTSD Chief Collection Development Officers of Large Research Libraries Discussion Group, have provided quantitative comparisons giving indicators of relative collection size, which may be interpreted as a crude measure of adequacy if one keeps in mind problems such as obsolescence, duplication, and erroneous or omitted data.<sup>27</sup> Such comparative data does provide information for some of the uses of collection evaluations mentioned at the beginning of this paper, but it does not adequately address the issue of quality in the subject collections of research libraries.

In 1958 Harry Bach addressed the issue of the difference in library needs between colleges and universities, and the problems in evaluating quality in large research library collections. He concluded that there was no single satisfactory method of collection evaluation and called for the devising of sampling techniques, using specialized rather than standard bibliographies, for evaluating research subject collections. He further suggested that some sort of "tiered" process might be useful, beginning with a rapid sampling process as a "strainer" to identify strength or weakness in areas of the collection.<sup>28</sup>

Since that time there have been a number of evaluations of research collections, and at least four such efforts have been reported in the

literature. R. P. Coale undertook a systematic evaluation of the research utility of the Newberry Library's holdings in the colonial history of Latin America. He used the "monographic bibliography" approach, checking bibliographies of 100 to 400 titles in classic books written by recognized scholars and, for current publications, subject samples from the *Handbook of Latin American Bibliography*. The purpose of the evaluation was to ascertain: (1) whether the scholars who had written the sampled books could have done so at Newberry, (2) whether the Newberry collection contained the significant printed primary source material and the important secondary accounts, and (3) whether the library had been keeping up with current scholarly output. To give the resulting data greater meaning, Coale took comparative samples at the University of California, Berkeley, the University of Texas at Austin, and the Library of the Hispanic Society of America.<sup>29</sup> The study was successful, and its only significant failure was the inability to develop a program of improvement based on the findings.

Robert Burns conducted an evaluation of holdings in science and technology at the University of Idaho in 1968. This evaluation updated and improved an earlier study and was designed both to measure the basic graduate research support adequacy of the collections and to chart directions for future growth in support of a new Ph.D. program. The survey was made against standard published lists and guides in the various fields and concentrated on serials and professional society publications. The evaluation summarized the strengths and weaknesses of the collection for each field and its capacity to support basic graduate instruction and research. Significantly, it developed a program for future acquisition, both current and retrospective, based upon the conclusions.<sup>30</sup>

Mary Cassata and Gene Dewey published some helpful and suggestive guidelines for the conduct of collection evaluations based upon the results of an evaluation carried out by subject bibliographers at SUNY-Buffalo in 1969. The Buffalo evaluation utilized a variety of techniques. The faculty contributed subjective evaluations of the research collections in their fields. The bibliographers developed goals statements for their areas of collection and examined the relationship between these goals and academic program scope by gathering and analyzing data on graduate and undergraduate course offerings and enrollments, graduate programs, faculty staffing plans, reading lists, etc. Undergraduate collections were measured against basic lists such as *Choice's Opening Day Collection* and *Books for College Libraries*. Research collections were tested by a number of means: checking the bibliography and footnotes of definitive books or dissertations, checking the percentage of authors in the Library of Congress classification schedule represented in the library's card catalog, and checking the library's holdings against the catalogs of special collections listed in Lee Ash's *Subject Collections*. The results of each specific subject evaluation

were stated in a report by the subject bibliographer who conducted or supervised the evaluation.<sup>31</sup>

The approach and methods suggested by Coale and Burns received further improvement and refinement in the work of William Webb on the process and results of a series of collection evaluations carried out at the University of Colorado libraries.<sup>32</sup> The approach used by Webb has served as a foundation for a series of research collection evaluations at Stanford University.

The purposes of the Colorado assessments were to provide a foundation of collection knowledge for the development of a collection development program, to develop a campaign of more thorough evaluations of and improvements in areas of weakness, to use the findings as support for the library's book budget requests, and to contribute to the professional growth of subject specialist, staff and their relationships with faculty.

Webb found that the "standard bibliographies" touted by most earlier articles on evaluation did not exist for most research collections in the humanities, so that specialized bibliographies had to be chosen by subject specialists and faculty working together. He also found that different categories and fields require different approaches; a sampling approach was not appropriate for evaluating reference collections, for example, because they really wanted to find out what they didn't have and develop a campaign to procure the material.

Webb's methodology was based upon a "tiered" approach, as is Stanford's. After bibliographies were chosen for each field, samplings were selected (for this purpose Webb argues that statistically correct or random samples are not necessary) on the following basis: 10 percent of bibliographies containing 100 to 1,000 items, 5 percent from 1,000 to 2,000 items, and 1 percent over 2,000 items. The data were tabulated and reported subfield by subfield, providing meaningful information without telling in detail what was or was not possessed. The data gave a good indication of strength and weakness and in some cases shocked the librarians who had very different perceptions of library holdings in certain areas. At this stage, graduate student assistants did the checking, and the average evaluation took sixty to seventy-five hours. When areas of weakness were thus identified, or when title-by-title checks of bibliographies were needed, trained support staff were used to guarantee greater accuracy, and they produced lists of *lacunae* and *desiderata* for focused collection building efforts.

The Colorado evaluations showed that economies of effort are appropriate to research collection evaluation: percentages of acceptable error are great enough in broad sampling evaluations to allow the use of graduate student assistance at relatively low cost with a high yield of useful data. I should mention, however, that both at Colorado and Stanford it has proved desirable to give graduate assistants some basic training and orientation in catalog, serial record, and government

document searching. Corporate entries, as might be expected, are especially problematic.

The results of the Colorado experience were heartening: a solid basis of information and direction for collection building was laid; a great deal of useful information about collection strengths and weaknesses was gained at reasonable cost; excellent data were obtained for budget presentations to university administrators and the legislature; and the administration was persuaded that something was being done about the "bottomless pit" of library acquisitions budgets.

#### *Collection Evaluation Program at Stanford*

A series of evaluations similar to the Webb model has been carried out at Stanford during the last year and a half and will be continued during the next few years in order to gain a fairly systematic picture of the quality—that is, the research and teaching adequacy—of the collections. Funded by a portion of special endowment revenues, the evaluations are intended: (1) to provide us with means for our bibliographers to learn and to know better their own portions of the multimillion volume research collections, (2) to establish useful bases for the revision of Stanford's collection development policy statement, (3) to provide data for us to use in attempting to coordinate some aspects of Stanford's collection development process with the University of California at Berkeley, (4) to provide information on which to build deliberate and shaped plans and specific recommendations for the improvement of Stanford's library collections, and to do so in the most rapid, useful, and efficient way, (5) to add support to our book budget requests, and, (6) to make possible the most productive and cost-efficient expenditure of library resources.

The program is tiered in a manner similar to that developed by Webb at Colorado. The need for evaluation is sometimes suggested by faculty responses to collection development staff interviews in which faculty are asked to evaluate the adequacy of the library collection in their own fields for teaching and research.<sup>33</sup> At other times, evaluations are undertaken as part of a systematic effort to sample collection effectiveness in large subject fields within the library's collections, especially where oral tradition or complaint suggest problems.

Normally a sampling is first searched, with a statistically significant number of titles in each of a number of subfields. For example, to evaluate the collections in early modern French history, we searched 165 titles from Robert Mandrou's *Introduction à la France Moderne*, divided into categories such as humanism, philosophy, history of science, religious life, art and artists, political and social theory, popular life, etc. When, by consulting the academic program needs and collection development policy statement levels, we have found that holdings are substantially lower than programs warrant, we mount more thorough searches and develop both current and retrospective acquisitions campaigns to rectify the inadequacies discovered.



In other instances, such as the sciences or "hard" social sciences, evaluative techniques are specially developed or borrowed from the literature to provide information about specific problems of collection availability or quality.

Evaluations have been completed or are underway in over twenty-two different subjects in the humanities and social sciences, and effort in the sciences is underway. Graduate students who have completed their comprehensive examinations in the field of study being evaluated are normally used to conduct the evaluations and, as indicated above, we have found it useful to provide some orientation and training for them. Graduate student bibliographers are currently paid at the upper levels of the hourly student range depending on training and experience. To date, Stanford has completed about fifty evaluations at an average cost of about \$400 which, to put it in scale, is about the same as the cost of acquiring and processing a dozen books. I would agree with Webb that the benefits have far outweighed the costs. Each graduate student is supervised by a library subject specialist in the area being evaluated. The subject specialist, the department faculty, and, in many cases, the graduate student, join in evaluating and selecting bibliographies for searching, keeping in mind the goals and scope of the academic program, the mission of the library collection, and the collection development policy statement parameters. The choice of bibliographies or other tools used for evaluation is a crucial one.

Four basic survey techniques for evaluation have thus far been used, each having a rather different goal:

1. A check of a sampling chosen from one or more accepted important subject bibliographies.
2. A check of one or more monographic or journal article bibliographies of works on the cutting edge of research in a discipline or subject.
3. A careful search of all titles in a significant subject bibliography in areas where weakness was revealed by methods one or two above.
4. For initial surveys in the social sciences, checks of basic lists of most-used and valuable titles to determine collection adequacy for graduate instruction and the availability of basic materials.

Methods one, two, or three often also involve checking holdings at Berkeley to provide comparable data and to take advantage of the Berkeley-Stanford Research Library Program cooperation. If Berkeley has an item we may well decide not to purchase it; if it is an expensive item, a journal, or a series, the two libraries may agree on a single location or share the purchase.

Each evaluation must result in a formal report submitted through the supervising subject specialist to the assistant director for collection development. The report is normally divided into two sections. The first consists of evaluative information and a description of the evalua-

tion process, and the second part contains the data. At Stanford this usually tells us—as a minimum—by subfield and total, (1) the number of titles checked, (2) the number and percentage at both Stanford and Berkeley, (3) the number and percentage at Stanford, (4) uniquely at Berkeley, and (5) at neither. Other types of data relevant to the goals of the particular evaluation may also be added. The introductory data includes: (1) the nature and goal of the evaluation, (2) the source, (3) the sampling used, (4) problems encountered (including unusual patterns of difficulty in finding the titles and materials found while searching which may suggest a different result than the data would imply), (5) general comments about the collections, (6) a list of specific areas of unusual strength or weakness, and finally—and the importance of this part of an evaluation cannot be overemphasized—(7) a plan or campaign of action to develop and build collections in areas of *undesirable* weakness (areas the collection development policy indicates should be stronger than the findings indicate that they are). Often a list of desiderata, *ranked in importance* is included, but this is usually an appendix rather than the substance of the plan for improvement. Normally the subject specialist, and perhaps faculty, will be involved in preparing this section of the report.

We always prepare a summary of each report. Gathered together and arranged, the summaries are used in reporting on the progress of the program to administrators as guides for collection planning, etc. The summaries normally include: (1) the evaluation title, (2) the departments or fields which the evaluation covers or touches, (3) a note about the sampling, (4) a summary of findings including areas of special strength or weakness, and (5) a conclusion summarizing action necessary to bring the collection to the desired level.

We have found that areas requiring particular care in planning collection evaluations are the selection of fields for evaluation (based upon faculty recommendation or suspicion of areas of weakness), the selection of bibliographies and other evaluative tools, the plan of the search, the selection of able, careful, and precise searchers, and the preparation of reports of quality and accuracy containing standardized, consistent, and program-related information and data. Reports presenting the perceptions and judgments of a trained graduate student with faculty and subject specialist contributions, as well as standardized statistical data representing the library's holdings, have proved as useful to us as data concerning specific items held or not held.

As a result of its collection evaluation program, Stanford shares in the benefits suggested at the beginning of this paper. We have certainly found the effort worthwhile.

An ongoing, consistent, and well-thought-out campaign of collection evaluation can be carried out over a number of years as need, time, and resources dictate and can assist a library to obtain and maintain the excellent and responsive collection that is its goal. A sound collec-



tion evaluation program interacts with the collection development policy, with the allocation of funds in support of collection development, and with the pruning of collections in forming the principal and most useful constituents of any library's collection development program.

## References

- Rutherford D. Rogers and David C. Weber, *University Library Administration* (New York, 1971), p.113.
- M. Llewellyn Raney, *The University Libraries* (University of Chicago Survey 7 [Chicago, 1933]); Alfred C. Potter, *The Library of Harvard University: Descriptive and Historical Notes* (Library of Harvard University Special Publication 6 [Cambridge, 1934]); Bibliographical Planning Committee of Philadelphia, *A Faculty Survey of the University of Pennsylvania Libraries* (Philadelphia Library Resources I [Philadelphia, 1940]); Donald Coney, *Report of a Survey of the Indiana University Library for the Indiana University, February-July 1940* (Chicago, 1940); Louis R. Wilson and Raynard C. Swank, *Report of a Survey of the Library of Stanford University, November 1946-March 1947* (Chicago, 1947); Stephen A. McCarthy, *Report of a Survey of the Library of the University of New Hampshire, January-February 1949* (University of New Hampshire Library 1949); Louis R. Wilson and Robert W. Orr, *Report of a Survey of the Libraries of the Alabama Polytechnic Institute, November, 1948-March, 1949* (Auburn, 1949); Louis R. Wilson, *Report of a Survey of the Libraries of Cornell University, October 1947-February 1948* (Ithaca, 1948); Robert W. Orr and William H. Carlson, *Report of a Survey of the Texas A and M College, October 1949-February 1950* (College Station, 1950).
- Robert B. Downs, *Union Catalogs in the United States* (Chicago, 1942), p.58-59, and *Resources of Canadian Academic and Research Libraries* (Ottawa, 1967).
- University of Michigan, Survey Research Center, *Faculty Appraisal of a University Library by the Survey Research Center of the University of Michigan* (Ann Arbor, 1961).
- Harry Bach, "Acquisition Policy in the American Library," *College & Research Libraries* 18:444 (Nov. 1957).
- J. Periam Danton, *Book Selection and Collections: A Comparison of German and American University Libraries* (Columbia University Studies in Library Science, 12 [New York, 1963]), p.71. Danton's book is also important as one of very few to focus on the subject of quality in research library selection.
- Maurice F. Tauber, "The Faculty and the Development of Library Collections," *Journal of Higher Education* 32:454-58 (Nov. 1961).
- David O. Lane, "The Selection of Academic Library Materials, A Literature Survey," *College & Research Libraries* 29:364-72 (Sept. 1968). Lane includes a useful study of the collection problems to which exclusive faculty selection had led but rightly observes that while primary responsibility for collections should belong to librarians, extensive faculty participation in the process is highly desirable and should be actively sought. A number of the library surveys cited above describe collection development and acquisitions problems which had developed over time through exclusive reliance and overdependence on faculty for long-term development of the collections. See, for example, Stephen McCarthy, *Report of a Survey* (see note 2).
- Harvard, Chicago, California, and Michigan scored, respectively, 68, 57, 45, and 36 percent. Douglas Waples and Harold D. Lasswell, *National Libraries and Foreign Scholarship. Notes on Recent Selections in Social Science* (Chicago, c.1936), p.71; chapter V, "American Libraries," p.69-82, was innovative in methodology for its time. The chapter remains interesting for its methodology and conclusions.
- Danton, *Book Selection and Collections*, p.75. It would be useful and informative to have a new study made on the Waples-Lasswell model but encompassing more fields to test the relative effectiveness of collection development programs in research libraries with and without subject specialist librarians.
- Waples and Lasswell, p.71.
- Ibid.*, p.74-75.
- George Bonn confirms this view: "Evaluation of the Collection," *Library Trends* 22:295 (Jan. 1974).
- F. W. Lancaster, *The Measurement and Evaluation of Library Service* (Washington, 1977), p.165. I would like to express my gratitude to Professor Lancaster for providing me with corrected galleys of his chapter "Evaluation of the Collection" for use in preparing this paper.
- George S. Bonn, "Evaluation of the Collection," p.265-304, and F. W. Lancaster, *The Measurement and Evaluation of Library Service*, p.165-206. A selective survey of literature relating to optimum library size with summaries of cited works may be found in Sam E. Ildon, "Qualitative/Quantitative Evaluation of Academic Library Collections: A Literature Survey," *International Library Review* 8:299-308 (June 1976). Signe Otterson has prepared a "Bibliography on Standards for Evaluating Libraries," *College & Research Libraries* 32:128-41 (March 1971), which contains some material for collection as well as library evaluation.
- American Library Association RTSD Resources Section, "Guidelines for the Evaluation of Library Collections," Final Draft, (June 1978). Prepared by the Collection Development Committee. After approval of the final draft these guidelines will be published by ALA. In the meantime, copies of the unapproved draft may be obtained from the office of the executive secretary of RTSD. An earlier version of the guideline statement is contained in the ARL SPEC kit: Association of Research Libraries, Office of University Library Management Studies, Systems and Procedures Exchange Center, *Collection Assessment*, Kit 41, February 1978.
- Association of Research Libraries, Office of University Library Management Studies, *Collection Analysis in Research Libraries; An Interim Report on a Self-Study Process* (Washington, 1978).
- Downs, *Union Catalogs in the United States*, p.58-59.
- Robert B. Downs, *Resources of Canadian Academic and Research Libraries* (Ottawa, 1967), p.207, 213.
- Bonn, "Evaluation," p.291, 294.
- Lancaster, *Measurement*, p.172-73. Concerning various quantitative measures of collection quality or adequacy, see in addition to Bonn and Lancaster, Maurice B. Line, *Library Surveys, an Introduction to their Use, Planning, Procedure and Presentation* (Hamden, 1967); R. Marvin McInnis, "The Formula Approach to Library Size: An Empirical Study of its Efficacy in Evaluating Research Libraries," *College & Research Libraries* 33:190-98 (May 1972); Peter Simmons, *Collection Development and the Computer: A Case Study in the Analysis of Machine Readable Loan Records and their Application to Selection* (Vancouver, 1971); and for model use the work of Philip M. Morse, such as "Measures of Library Effectiveness," *Library Quarterly* 42:15-30 (Jan. 1972); or "Library Models," in Alvin Dralee, et al., eds., *Analysis of Public Systems* (Cambridge, Mass., 1972), p.235-56.
- John R. T. Ettlinger, "Nation-wide Rationalization of Acquisition Policies in Canadian College and University Libraries." *Three Papers on Collections Delivered at the Canadian Association of College and University Libraries Workshop on Collection Development Held at Sackville, New Brunswick, June 17, 1973* (Halifax, 1975), p.11-27; Ettlinger, "Through a Glass Darkly—Academic Book Selection in Crisis," *APLA Bulletin* 32:32-40 (June 1968); John Miller, "Problem of Fall-Out from the Knowledge Explosion," *APLA Bulletin* 33:21-27 (Sept. 1969).
- George Piternick, "Library Growth and Academic Quality," *College & Research Libraries* 24:229 (May 1963).
- James Krikelas, "Library Statistics and the Measurement of Library Services," *ALA Bulletin* 60:496-97 (May 1966). Alan Singleton, in "Journal Ranking and Selection: A Review in Physics," *Journal of Documentation* 32:258-89 (Dec. 1976), concludes that most quantitative measures of the utility and quality of collections "do not prove satisfactory. Continued reliance on subjective judgment of librarian and user is called for."

25. Harry Bach, "Evaluation of the University Library Collection," *Library Resources & Technical Services* 2:24-29 (Winter 1958).
26. Verner W. Clapp and Robert T. Jordan, "Quantitative Criteria for Adequacy of Academic Library Collections," *College & Research Libraries* 26:371-80 (Sept. 1965).
27. *Titles Classified by the Library of Congress Classification: National Shelflist Count* (General Library, Univ. of California, Berkeley, 1975). An updated and expanded edition of this document nears completion under the editorship of Le Roy Ortopan of the Catalog Department of the General Library, University of California, Berkeley.
28. Bach, "Evaluation of the University Library Collection."
29. R. P. Coale. "Evaluation of a Research Library Collection: Latin American Colonial History at the Newberry," *Library Quarterly* 35:173-84 (July 1965).
30. Robert W. Burns, *Evaluations of the Holdings in Science/Technology in the University of Idaho Library* (University of Idaho Library Publication Number Two [Moscow, Idaho, 1968]).
31. Mary B. Cassata and Gene L. Dewey, "The Evaluation of a University Library Collection: Some Guidelines," *Library Resources & Technical Services* 13:450-57 (Fall 1969).
32. William Webb, "Project Co Ed: A University Library Collection Evaluation and Development Program," *Library Resources & Technical Services* 13:457-62 (Fall 1969). The evaluations included: medieval studies, art history, political science, physics, Slavic studies, and English and American literature and social history.
33. Mailed questionnaires have consistently proved less useful than personal contact; responses are usually too few to be significant. However, brief questionnaires filled out in the course of interviews have proved useful.