Systems Analysis as a Decision-Making Tool for the Library Manager

A. GRAHAM MACKENZIE

DEFINITIONS

"SYSTEMS ANALYSIS," in this context, is not the process which is the necessary prelude to library automation; it is something much more basic. It involves seeking out the fundamentals of a situation and applying rigorous scientific methods to their study, with the aim of finding an optimal solution to the problems facing the manager. It is, in all but name, indistinguishable from operations research (usually referred to as OR), which has been defined as "(1) the application of scientific method (2) by interdisciplinary teams (3) to problems involving the control of organized (man-machine) systems so as to provide solutions which best serve the purpose of the organization as a whole."[1]

In this process it may be necessary to collect vast masses of data on the operations being considered, to establish unit costs or times, and to use the various other techniques of work-study; it will probably be necessary to devise some kind of model or simulation of various processes which are integral to the operation; but the one inescapable task in every OR study is to try to define, in quantifiable terms, the objectives of the organization as a whole. It may not be possible to light upon one final, all-embracing objective; but at the very least there must be intermediate objectives which, taken together, will meet a large part of this need.

Decisions may be short-term and tactical: If the book fund is increased by 20 percent, how many extra catalogers will be required to process the extra books? Or should catalog data be obtained from an external source? More important, however, are long-term or strategic decisions: What level of financial support should a university give to its

A. Graham Mackenzie is University Librarian and Director of the Library Research Unit, University of Lancaster, England.

library? What is the optimum collection size in a given situation? Traditional wisdom in this situation says "more means better;" but is this necessarily so if the library is seen as part of a larger corporate entity?

The library has advanced a long way from the time when it was merely a collection of books, with a relatively small number of users, each knowing his way about its shelves. With the expansion of higher education and advances in technology it has become much more a precision instrument for the transfer of information from the author to large numbers of users—often in a very short space of time, because of the numbers involved—who need a great deal of guidance and help. From a modest extension of the scholar's own personal collection of books it has become an essential segment of the educational process, as well as a tool of the advanced research worker.

The library manager is a relatively new concept. Traditionally the librarian has been a scholar who has acquired a veneer of professional skills to enable him to exercise bibliographical control over his collections, and a patina of the techniques of public relations and personnel management by which he related himself to his environment. In recent years, however, with the absolute growth of literature and increasing budgetary limitations, the emphasis has begun to change. The professional skills and techniques are still essential, but to them is being added a realization that a library, if it is to perform its full function, calls for the study of cost-effectiveness and similar concepts drawn from the world of industry. Without these concepts the manager only has experience, training and intuition to rely on, none of which is a satisfactory substitute for scientific method if limited resources are to be used to the best advantage.

METHODS

Their detractors have said that systems analysis and operations research are merely applied common sense. There is indeed a certain truth in this, but it is not the whole truth: the difference lies in the way in which the common sense is applied. The traditional OR study takes place in four discrete phases. The system under review is described (often by flow charting the processes which make it up), and particular attention is paid to delineating the key points in quantifiable terms. Once this has been done a series of mathematical models are built to define the various interrelationships within the system. The second stage is to measure the system as it is, by collecting objective data if possible, or otherwise by making assumptions (these must, however, be

quite explicit). With the input of this information we then have a working model, and by varying one or more of the inputs or parameters the performance of the new system can be described. This information is presented to the manager, who can then make operational decisions between different possible courses of action in the light of his new knowledge of the probable consequences of each. The last stage of an OR study is the achievement of operational control of the system by providing the manager with the means of achieving his objectives over a period of time, and a system of feedback of information so that he can monitor system performance.

The essential differences between OR and common sense thus appear to be quantitative thinking, model-building, and the mathematical manipulation of data. Unfortunately, few librarians tend to think in a quantitative manner, and even fewer have the training to construct or work with models. It is worth noting, however, that the same is probably true of many captains of industry, although this has not prevented the widespread adoption of such techniques there, to the extent that many firms have their own teams of OR experts to advise top management. The same could well become true of libraries in the near future; indeed it is possible, as well as highly desirable, for librarians to become an integral part of an OR team.

OBJECTIVES

There is, however, one major difference between libraries and industry which makes the application of OR more difficult. In industry, broadly speaking, there is one overall objective—financial profit—and hence there is one single significant measure of the output of the system. This means that the conceptual analysis involved can be more directly and obviously aimed towards the end-product, which is clearly measurable. (This does not, of course, imply that an industrial systems study must aim at maximizing profit, but it does place the decision-makers in the position of being able to evaluate the results of different courses of action against the objective yardstick of the likely financial returns.) In a library, however, "profit" is a meaningless term. The profits of an industrial special library may perhaps be inferred or measured indirectly as a contribution to the overall production of the parent company, but this is a special case not applicable to public or academic libraries. In their situation even modern economic theory cannot make more than a guess at the benefits conferred on the community served.

The central problem in any systems study of libraries therefore becomes not the techniques which should be used, but the precise specification of the objectives of the system. In this context it is useless to talk, for example, of serving "the reading, reference and research needs of its users," or of "library services up-to-date and commensurate with their needs . . . founded on adequate collections," as does the ACRL statement,[2] if for no other reason than these concepts have no quantitative meaning. They may be taken into consideration at a later stage as qualitative factors which affect the subjective decision of the manager about the relative merits of different solutions to a problem, and indeed some arbitrary values may be applied to such considerations, if only to insure that they are not forgotten; they cannot, however, be specified as integral parts of the model.

Libraries may perform many different and often conflicting functions: some acquire and preserve books for posterity, others cater more to the immediate needs of their users—be these undergraduates, senior research workers or someone looking for recreational material to while away an idle hour. Some libraries provide an information service at a sophisticated level, others little or none; some are centers of scholarly research, others concentrate on curricular books; some have elaborate catalogs, others little more than finding lists; one may exist almost entirely for lending, while another keeps all its books within the building. With this diversity of purpose there can be no single, all-embracing objective, valid for all libraries. However, if we take as an example the academic library, and place it in its context as a means of making documents (books, journals, abstracts, nonbook materials, etc.) available to users, we can perhaps measure its effectiveness by the amount of success with which it does this.

A number of objective measures of library performance have been suggested, based more or less directly on this concept. The most popular has been the probability that a reader will find on the shelf the document he is seeking, although researchers have used different techniques for measuring and defining this probability. Buckland[3] talks of satisfaction level and collection bias (the latter being a measure of the suitability of a library for browsing, as opposed to seeking a specific title), Urquhart[4] of reader failure at the shelf, and Orr[5] of a document delivery test which uses as its measure the delay in providing material. "Findability" alone, however, is not the whole story: postulate an (improbable) library which does not permit borrowing, and where readers are allowed into the stacks, but must leave the building again within

thirty minutes. Clearly the availability of its stock would be close to 100 percent, since the chance of any specific title being required by two people in the same thirty minutes is very small. In terms of people using documents, the library would be largely useless.

This consideration has led some investigators to consider as a measure of effectiveness the amount of time for which the reader uses a book: this is, after all, the end-product of the majority of library activities. Meier proposes "item use-days," thus equating ten loans for one day each with one loan for ten days.[6] This is an improvement on the simple amassing of circulation statistics—although both methods are to some extent dependent on the loan period allowed—but says nothing about the quality or intensity of use. Hamburg chooses "document exposure"—the total length of eyeball-to-page contact per circulation.[7] According to some tentative and unconfirmed findings quoted in Brophy, this time appears to be in a negative exponential relationship to the length of the loan period.[8] This measure therefore seems to carry more conviction than the others in an academic library situation, since a very short loan period—which is, subjectively, less valuable to the reader than one of medium length—will carry a lower weighting. Conversely a very long loan period, when a high proportion of the books on loan will be lying unused on the borrower's desk at any given time, weighs only little more than the loan period of medium length.

It may not be necessary for some purposes to define a single all-embracing objective for a library: certain aspects may be divorced from the main concept of book use and be treated in isolation. It is perhaps in this area that most systems work has so far been done, simply because it is easier to identify and analyze smaller problems. The economies of storage in stacks have been studied by the Purdue team under Leimkuhler[9] and, as book retirement, by Morse[10] and Raffel and Shishko.[11] Burkhalter[12] has produced a series of case studies on, for example, exit control and charge-out and accounting systems, although some of these are nearer to conventional work-study than to systems analysis. The Lancaster team has produced models of book processing and in-house binding.[13] There is, however, a grave danger of suboptimization in looking at only one aspect of a library in isolation. Efficiency may well be the enemy of effectiveness, as in the case of a circulation system designed for economy of operation which actually reduces the availability of books to readers, and therefore the amount of use that can be made of the library.

MODEL BUILDING

"The basis of all scientific work is experimentation. Experimentation is just what apparently cannot be done with administrative systems. . . . OR does not experiment with the system itself, it experiments with a model of the system."[14]

Whenever human beings are concerned, the only certain thing is that nothing is certain; therefore we cannot say that a library and its users will, on any given occasion, interact in one specific way. However, if a large number of observations are made and analyzed, we can say a great deal about the *probability* that a particular event will occur; it is on this basis that library models are constructed. Of course the predictions they generate will often be in error about individual readers or books; but in the aggregate they will be more right than wrong, and thus the library manager will be able to base his actions on statistical predictions rather than guesses.

This is not the place to examine in detail the technical aspects of model building; many varieties are possible, ranging from a simple description of the probability of some event occurring to a string of equations representing a series of complex interactions between the library and its users. It must, however, be emphasized that the essential character of any model is that it should represent all those parts of the real life situation which may significantly affect the outcome. If the analyst has been able to identify and describe each part of the system in a logical way, he may be able to eliminate some of them from the model by sensitivity analysis; e.g., in studying book availability in an academic library it may be found that in-library use (which poses difficulties for the data collector) can be ignored, since neither a very large nor a very small value for this will make a significant difference to the results which the model produces. Similarly, if data cannot be found for some part of the model (for example the effect on demand of differences in the pattern of teaching in various disciplines), it may be possible to treat this effect as a "black box," and earmark it for future study. In the meantime, the investigator can input a range of values and make a subjective judgment on the most likely ones.

DATA COLLECTION

Most libraries produce statistics annually; in many cases, however, these are merely the conventional figures of books acquired and cataloged, numbers of loans (perhaps analyzed by status of borrower or by

subject), and overall figures of expenditure. This kind of gross material is not what is required to quantify a model; needed instead is information on what the individual user does when he enters the building or borrows a book, on the number of times an individual title is requested by borrowers (or needed by borrowers—this is not the same thing), and similar detailed facts which are not normally recorded during the routine processes of most libraries.

This information must then be the object of special data-collection exercises. These may be direct, e.g., the analysis of issue records over a period to find out the distribution of borrowing among various categories of readers, the examination of date labels to discover the distribution of use among different categories of books or individual titles, or visual observation of readers' behavior to see how much the catalog is used. Or they may be indirect; the usual technique here is the questionnaire addressed to a sample of readers, although blanket coverage has on occasion been attempted. Both methods have their advantages and disadvantages: direct observation is more accurate and objective, but can be very tedious and time-consuming; the indirect method has all the normal pitfalls of survey techniques (the danger of loaded questions, the difficulty of insuring that the sample is representative, the risk of influencing behavior by the mere fact of asking questions about it). Also the indirect method is the only way of getting certain types of information, and is usually cheaper than direct observation because a well-chosen sample can often give as much information as complete coverage, with as much accuracy as the model needs.

It is worth noting that automation is often heralded as the answer to the data-collection problem. It can indeed provide more information more easily than can manual methods, but only if the designers of the system have foreseen from the start what will be needed for statistical and control purposes. This does not seem to be the case in most of the systems presently in use in libraries.[15]

Data on system performance and requirements are perhaps the most difficult to collect; but equally important is information on costs. It is comparatively easy, by applying standard work-study techniques to library operations, to establish the direct unit cost of any given operation —acquisition, cataloging, circulation. It is more complicated, but possible, to establish the total direct and indirect cost of, for example, the loan of a single book, made up of a proportion (because the book will probably be lent more than once) of the purchase price and processing, storage and circulation costs. This is only one side of the equation;

for every cost there ought to be a corresponding benefit. If a single circulation of an average title costs $3.50 in total, then either the benefit to the institution is greater than this, or the library is misusing money which could be spent more effectively on something else within the institution. One cannot measure benefits directly; Raffel and Shishko[11] at M.I.T. and Hawgood[16] and his team at Durham have used sophisticated PPBS and regression analysis techniques respectively, based to a large extent on the values which can be imputed to a library by its users. This approach, however, seems to be surrounded by dangers. Although it measures a consensus of opinion, it is nevertheless still subjective and can at times degenerate into a circular argument. One feels that there is a need for a quite different method of analysis, as yet undiscovered.

USE OF THE MODEL

Once the model has been constructed and tested, and quantities have been attached to its terms, it is ready for use. The objective of the manager is usually to optimize one specific aspect of his library; models have not yet been developed, and may never be, which are capable of describing all the complex interrelationships which constitute the library and its environment. He therefore has before him a choice of possible actions, some of which may have been suggested by the model itself. His need is first of all to discover what are the likely results of each, in terms of performance and cost, by simulating the performance of the system. Once this has been done he may have to assess the political implications of the (theoretically) optimum solution. If, for example, this can only be obtained at the cost of reducing the privileges traditionally accorded to faculty borrowers (who are often, indirectly, the policy-makers for the library through their participation in committees), this may be considered inexpedient or even impossible. Such intangibles are difficult to program into a model, and yet they are a real factor in decision-making; the model can therefore only be an aid, and qualitative judgment still has a major part to play. Perhaps it is at this stage, rather than at the earlier one of quantifying the model, that the assessment of benefits can best be taken into account, albeit in a less formal manner.

SYSTEMS AT WORK

The last few years have seen an enormous growth in the amount of effort directed towards systems analysis in libraries; but to see one of the

most elegant and far-reaching applications (simple though it is by comparison with much of today's work) one must go back to 1959, when plans were being made in Britain to improve the performance of the national interlending service for scientific journals. At that time there were, broadly speaking, four channels through which a library might borrow a title it did not possess: through one of the ten Regional Library Systems which cooperatively maintained union catalogs of the holdings of libraries in their respective regions; through the National Central Library in London, which held very little material itself, but had an incomplete national union catalog; by direct, semiformal, request to another library known to hold the required title; or by application to the Science Museum Library, which was in effect a national lending library for this type of material. The first two channels were relatively little used because of their slowness and uncertainty. The Science Museum Library, although the most important single source of loans, was becoming less and less effective as demands on it grew, since its lending role was only secondary.

The Department of Scientific and Industrial Research was given the task of providing a comprehensive and efficient service, and, to the disgust of the library profession, a comparative unknown—D. J. Urquhart —with no professional library qualifications was put in charge of what was then known as the Lending Library Unit. Urquhart, luckily for the future of the National Lending Library for Science and Technology, was uninhibited by traditional professional expertise. He was a scientist, and applied the techniques which have since become known as systems analysis.[17] His stated objective was to insure that there were, somewhere in the country, sufficient copies of frequently used periodicals to meet the total interlibrary loan demand, and at least one copy of infrequently used titles. By analysis of loan records in the Science Museum Library he demonstrated that the demand for its titles was a rough measure of the demand nationally, and that Bradford's Law of Scattering still held. He then used the Poisson distribution to predict the rate of arrival of demands, and his model showed that a single complete collection, with multiple copies is required, would perform much more efficiently and cheaply than the alternative of a decentralized collection spread over ten or more self-sufficient regions. The history of the National Lending Library has justified at least the general tenor of his calculations; at a more personal level Urquhart himself has succeeded in overcoming the hostility and suspicions of his more conven-

tional colleagues to the extent that in 1972 he was elected president of the Library Association.

The second example of a model at work is the variable loan and duplication policy developed at the University of Lancaster, and reported by Buckland.[3] In the course of a research project[13] it was discovered that a previous attempt to provide an efficient circulation system by using a long loan period in conjunction with an efficient recall procedure had in fact resulted in suboptimization, since the probability of a reader finding a specific book (satisfaction level) was only about 0.6. This was considered too low, and the research team was given the objective of raising the probability to 0.8. A model was constructed using a Monte Carlo simulation rather than the queuing theory favored by Morse,[10] and a number of alternative courses of action were proposed, consisting of duplication or one of several different sets of loan policies. The tangible and intangible benefits and costs were compared, and the solution chosen was a combination of duplication and a loan period of seven days for the most popular 10 percent of the books. The result of implementation was that satisfaction level rose by the desired amount; however, demands on the service also rose, and, in the three years since 1969, the satisfaction level appears to have gradually reverted to approximately its original figure of 60 percent. The overall performance of the system has nevertheless improved considerably, although in an unforeseen way: since demand has risen (probably because of users' reaction to the temporary improvement in satisfaction level) and has stayed high, book use, whether measured in terms of issues per head, or of document exposure, is about half as high again as it was prior to implementation. Further models are being developed to investigate the new situation, and more attention is being paid to the interaction between the system and the user than previously, when the library was thought of more as a physical than a biological system, and the user was considered as a "black box."

It is somewhat surprising to realize that, in spite of the activity on both sides of the Atlantic in recent years, very little appears to have been done in practical terms. M.I.T., the University of Michigan and Purdue University all have extensive graduate programs in OR oriented towards library service, and substantial numbers of papers and books have been published, notably those of Morse,[10] Raffel and Shishko,[11] and Burkhalter[12] (although perhaps the latter leans toward work-study). In Great Britain there have been major studies at Durham,[16]

Lancaster[13] and Cambridge,[18] all with the avowed intention of examining library systems, and all with support from the Office for Scientific and Technical Information. From the published literature it appears, however, that only at Lancaster and Michigan have there been any major changes in an operational library system as a result of OR studies. (This may be an unjust comment; Newhouse and Alexander[19] imply, although they do not specifically describe, changes in the Beverly Hills Public Libraries, and Houghton[20] describes a new policy for special library journal subscriptions which may have been put into practice. However, if librarians are actually basing operating decisions on such techniques, their light appears to be small and their bushel exceedingly large: one would expect full and explicit publication of a technique as new as OR.)

It may be that at this stage there is more need for enabling and theoretical studies than for practical operations research in libraries. It is more likely, however, that the profession as a whole is either not convinced, or ignorant, of the potential value of these attitudes. There is a need for librarians to be more intimately concerned with the day-to-day realities of OR as members of the team. If OR is left to its own professional exponents there is a risk that elegant theoretical model building will gain preference over more rough-and-ready, but more practical, techniques which can benefit libraries. There is also a need for library schools to teach management in a more quantitative way; it is true that the average student may not have much opportunity to put these theories into practice during the early part of his career, but effective management depends on attitude as well as on techniques, and if the student can be brought to appreciate this at an early stage in his career, he will be more receptive when he is later placed in a position where such skills can be of value to him.

References

1. Ackoff, Russell L., and Sasieni, Maurice W. *Fundamentals of Operation Research.* New York, Wiley, 1968, p. 6.

2. "Statement of Service to Library Users," *ACRL News,* 27:21, April 1966.

3. Buckland, Michael K. "An Operations Research Study of a Variable Loan and Duplication Policy at the University of Lancaster," *Library Quarterly,* 42:97-106, Jan. 1972.

4. Urquhart, John A., and Schofield, J. L. "Measuring Readers' Failure at the Shelf," *Journal of Documentation,* 27:273-86, Dec. 1971, and 28:233-41, Sept. 1972.

5. Orr, Richard H., *et al.* "Development of Methodologic Tools for Planning

and Managing Library Services: II. Measuring a Library's Capability for Providing Documents," *Bulletin of the Medical Library Association,* 56:241-67, July 1968.

6. Meier, Richard L. "Efficiency Criteria for the Operation of Large Libraries," *Library Quarterly,* 31:215-34, July 1961.

7. Hamburg, Morris, *et al. A Systems Analysis of the Library and Information Science Statistical Data System. . . . Interim Report.* Philadelphia, University of Pennsylvania, 1970; and ———. "Library Objectives and Performance Measures and Their Use in Decision Making," *Library Quarterly,* 42:107-28, Jan. 1972.

8. Brophy, P., *et al. A Library Management Game* (University of Lancaster Library Occasional Papers, no. 7). Lancaster, University of Lancaster Library, 1972.

9. Leimkuhler, Ferdinand F. "Library Operations Research: A Process of Discovery and Justification," *Library Quarterly,* 42:84-96, Jan. 1972.

10. Morse, Philip M. *Library Effectiveness: A Systems Approach.* Cambridge, Mass., M.I.T. Press, 1968; and ———. "Measures of Library Effectiveness," *Library Quarterly,* 42:15-30, Jan. 1972.

11. Raffel, Jeffery A., and Shishko, Robert. *Systematic Analysis of University Libraries: An Application of Cost-Benefit Analysis to the M.I.T. Libraries.* Cambridge, Mass., M.I.T. Press, 1969.

12. Burkhalter, Barton R., ed. *Case Studies in Systems Analysis in a University Library.* Metuchen, N.J., Scarecrow Press, 1968.

13. Buckland, Michael K., *et al. Systems Analysis of a University Library* (University of Lancaster Library Occasional Papers, no. 4). Lancaster, University of Lancaster Library, 1970.

14. Duckworth, W. E. *A Guide to Operational Research.* 2nd ed. London, Methuen, 1965, p. 19.

15. Maidment, William R. "Management Information from Housekeeping Routines," *Journal of Documentation,* 27:37-42, March 1971.

16. Hawgood, John, *et al. Project for Evaluating the Benefits of University Libraries: Final Report.* Durham, University of Durham, 1969.

17. Urquhart, D. J., and Bunn, R. M. "A National Loan Policy for Scientific Serials," *Journal of Documentation,* 15:21-37, March 1959.

18. Plumb, Philip W. "The Cambridge University Library Management Research Unit," *Library Association Record,* 73:187-88, Oct. 1971.

19. Newhouse, Joseph P., and Alexander, Arthur J. *An Economic Analysis of Public Library Services.* Santa Monica, Rand Corporation, 1972.

20. Houghton, Bernard. "Zipf!" *New Library World,* 73:130, Nov. 1971.