

Counterpoint

Program Evaluation and Archives: “Appraising” Archival Work and Achievements

JACQUES GRIMARD

RÉSUMÉ L'objectif de cet article est d'illustrer le potentiel de l'évaluation de programme pour les institutions d'archives et de poser les fondements du développement d'un projet de recherche sur le sujet. L'auteur explore deux questions essentielles: « Qu'est-ce que l'évaluation de programme? » et « Est-ce que ce concept est applicable pour les institutions d'archives? ». En se basant sur un examen approfondi de la littérature, il démontre que dans les institutions d'archives, comme ailleurs, ce genre de programme doit se fonder sur un mélange judicieux d'approches. Il offre également des réflexions sur les critères, les mesures et les indicateurs des activités, produits et services archivistiques. L'article conclut que les archivistes trouveraient opportun et stratégique de développer leurs connaissances et leur expertise sur l'évaluation de programme et d'en explorer plus avant la théorie et les applications potentielles au niveau méthodologique. Il suggère le développement d'un projet de recherche sur l'évaluation de programme applicable aux institutions d'archives.

ABSTRACT The purpose of this article is to illustrate the potential of program evaluation for archives and to lay the foundation for the development of a research initiative on this subject. It explores two basic questions: “What is program evaluation?” and “Is the concept applicable to archives?” Based on an extensive review of the literature, it demonstrates that in archives, as elsewhere, program evaluation has to rely on a judicious mix of approaches. It also offers considerations on criteria, measurements, and indicators for archival activities, products, and services. This paper concludes that archivists would find it strategically opportune to develop their knowledge and expertise on program evaluation and to explore further the theory and the potential applications of program evaluation methodologies. It suggests the development of a research initiative on program evaluation applicable to archives.

Introduction

A review of the professional and academic literature in archival science indicates that despite a long tradition and expertise in appraising and in evaluating information and in collecting data on their activities,¹ archiv-

¹ The annual publication of *L'activité de la direction des Archives de France et des services*

ists² have not significantly addressed evaluation of their programs, either from a theoretical or from a methodological perspective.³

This paper proposes that archivists would find it strategically opportune to develop further their existing knowledge and expertise by exploring the theory and the potential applications of program evaluation methodologies. This is particularly important given the varied challenges of the current environment such as: limited resources, professional and technological convergence, and fierce competition among cultural institutions and information management services. High expectations with regards to availability of and access to authentic and reliable information and growing concerns regarding professional and social accountability add to these challenges. By taking an evaluative approach, the position of archives in such an environment would be strengthened. Archivists would be less vulnerable when demonstrating the value-added of their contribution to society, participating in generic program evaluation initiatives, reviewing their own programs and budgets, or exploring ideas to improve programs and services.

In order to raise interest in and build understanding of program evaluation, this article will answer the following two basic questions: “What is program evaluation?” and “Is the concept applicable to archives?” It will also suggest considerations for the development of a research initiative on program evaluation applicable to archives.

What is Program Evaluation?

From the perspective of its fundamental *raison d'être*, “... evaluation is the process of determining the extent to which the goals and objectives of a program are being attained.”⁴ For Emil J. Posavac and Raymond G. Carey, who look more at its nature and purpose,

Program Evaluation is a collection of methods, skills, and sensitivities necessary to determine whether a human service is needed and likely to be used, whether the service is sufficiently intensive to meet the unmet needs identified, whether the service is

d'Archives publiques (Paris, 1998), represents an illustration of this tradition (available at ministère de la Culture et de la communication).

2 In this paper, the term “archivists” refers to professionals involved in the management of records throughout the entire life cycle or the whole life continuum of recorded information.

3 See Lucie Pagé, « Méthodes et techniques d'évaluation des services d'archives. Des pistes à développer pour les documents historiques », *Archives*, vol. 23, no. 4 (1992), p. 57–78; Lucie Pagé, « Évaluer pour évoluer. Méthodes et techniques d'évaluation des centres d'archives : des pistes à explorer », dans Louise Gagnon-Arguin et Jacques Grimard, dir., *La gestion d'un centre d'archives. Mélanges en l'honneur de Robert Garon* (Québec, 2003), p. 135–166.

4 Michael Quinn Patton, *Practical Evaluation* (Beverly Hills, London, New Delhi, 1982), p. 34.

offered as planned, and whether the service does help people in need at a reasonable cost without unacceptable side effects.⁵

In their view, the sole purpose for program evaluation activities is the contribution it makes to the provision of quality services by providing feedback.⁶ Michael Quinn Patton proposes the following definition:

Program Evaluation is the systematic collection of information about the activities, characteristics, and outcomes of programs to make judgements about the program, improve program effectiveness, and/or inform decisions about future programming.⁷

He further observes that “decision makers can use evaluation to reduce uncertainty, enlarge their option, increase control over problem activities, speed things up, and increase their sophistication about program process.”⁸ Similarly, Carole Weiss believes that the primary purpose of program evaluation is “the contribution to subsequent decision making and the improvement of future programming ...”⁹ Douglas Zweizig et al. add to the thinking by stating that the primary purpose for improving services in the future is not to prove, but to improve and he makes the link between evaluation and learning.¹⁰

In their book, *Fourth Generation Evaluation*, Egon G. Guba and Yvonna S. Lincoln introduce a new concept: fourth generation evaluation is a “form of evaluation in which the claims, concerns, and issues of stakeholders serve as organizational foci (the basis for determining what information is needed), that is implemented within the methodological precepts of the constructivist inquiry paradigm.”¹¹ The challenge then, is not to go beyond the traditional systematic collection of information and the positivist analytical approach, but rather, to look at the values and expectations of participants and understand their significance in relation to the object evaluated. This is somewhat in line with Posavac and Carey’s definition quoted above which also refers to a more “skills” and “sensitivities” based approach.

5 Emil J. Posavac and Raymond G. Carey, *Program Evaluation. Methods and Case Studies*, 5th ed. (Upper Saddle River, 1997), p. 2.

6 Ibid., p. 14.

7 Michael Quinn Patton, *Qualitative Research & Evaluation Method*, 3rd ed. (Thousand Oaks, London, New Delhi, 2002), p. 10.

8 Quinn Patton, *Practical Evaluation*, p. 61.

9 Carol H. Weiss, *Evaluation Research. Methods for Assessing Program Effectiveness* (Englewood Cliff, 1972), p. 4.

10 Douglas Zweizig et al., *The TELL IT! Manual. The Complete Program for Evaluation Library Performance* (Chicago, London, 1996), p. 5.

11 Egon G. Guba and Yvonna S. Lincoln, *Fourth Generation Evaluation* (Newbury Park, London, New Delhi, 1989), p. 50.

In a discussion of program evaluation, it is necessary to also discuss the concept of performance measurement, a concept well understood and used by librarians and other specialists of information sciences. As can be seen by the following definitions, performance measurement is similar to program evaluation. For Charles R. McClure and Betsy Reifsnyder, it is “a broad concept that includes both input and output measures, stresses assessment of organizational effectiveness, and provides a basis for planning.”¹² For Nancy A. Van House

Performance measures refer collectively to several kinds of measures that reflect the performance of the organization. These include: inputs or resources used; process or measures reflecting internal operations; productivity or the ratio of outputs to inputs; outputs, that is the extensiveness and effectiveness of services delivered; and outcomes, the most difficult to measure, the effects of the services provided on clients and society.¹³

Baker and Lancaster see “evaluation as a management tool, the main purpose of which is to identify current strengths, limitations, and failures, and to suggest ways to improve it,”¹⁴ while McClure and Samuels talk about a “measurement [instrument] of effectiveness in reaching some predetermined goal.”¹⁵ It is worth noting that these authors also refer to “judging”¹⁶ the performance of an organization (namely a corporate information centre and a library), the quality and the value of its activity and/or service,¹⁷ to assessing its effectiveness,¹⁸ its efficiency, and its ability to achieve its organizational and societal goals.¹⁹

It can be deduced that program evaluation implies assessment of organizational performance, that it offers methodologies and methods, including

12 Charles R. McClure and Betsy Reifsnyder, “Performance Measures for Corporate Information Centers,” *Special Libraries* 75, no. 3 (1984), pp. 193–94.

13 Nancy Van House, “Output Measures in Libraries,” *Library Trends* 38, no. 2 (1989), p. 268.

14 Sharon L. Baker and F.W. Lancaster, *The Measurement and Evaluation of Library Services* (Arlington, 1991), p. 3.

15 Charles R. McClure and A. Samuels, *Strategies for Library Administration: Concepts and Approaches* (Littleton, 1982), p. 233.

16 In an article on principles and practice for performance measurement, Geoffrey Ford wonders “How then do we judge the performance of a library?”. Geoffrey Ford, “Approaches to Performance Measurement: Observations on Principles and Practice,” *British Journal of Academic Librarianship* 4, no. 2 (1989), p. 76.

17 Geoffrey Ford, “Performance Measurements: Principles and Practice,” *IFLA Journal* 15, no. 1 (1989), p. 13.

18 Thomas W. Shaughnessy, “Assessing Library Effectiveness,” *Journal of Library Administration* 12, no. 1 (1990), p. 2.

19 Nick Moore, “Measuring the Performance of Public Libraries,” *IFLA Journal* 15, no. 1 (1989), p. 13.

related instruments, to examine the results achieved against stated objectives and assess to the quality, the value, and the pertinence of these achievements. As a corollary, program evaluation is a process involving the methodical collection of quantitative and qualitative data on performance of activities, use of resources, and results achieved in order to assess the quality and the impact of archival products and services.

Program Evaluation: Framework

Many authors refer to two broad categories of approaches to evaluation²⁰: the “formative” evaluation, which determines if a program is carried out and/or has been implemented as planned; and the “summative” evaluation, which measures the success of a program in achieving its objectives and assesses its impact.²¹ Mostly performed during the course of activities, formative evaluation is sometimes referred to as “in-term” evaluation.²² Essentially, the formative approach primarily looks at efficiency and responds to questions such as: “Is the preservation program operating according to plans, within budget and/or in line with conservation standards?” Summative evaluation is also known as “ex post” or “post hoc”²³ because it primarily looks at completed activities. Summative evaluation focuses on effectiveness and is appropriate for example, to verify if an acquisition program has reached its objectives, if a digitization initiative has effectively contributed to the diffusion of archival information, or to identify the impacts of a retention scheduling training pro-

20 In the last revision of his book, *Utilization-Focused Evaluation (Utilization-Focused Evaluation. The New Century Text*, 3rd ed. [Thousand Oaks, London, New Delhi, 1996]), Michael Quinn Patton identified over one hundred possible evaluation approaches. Talking about “breakthrough ideas” that have influenced evaluation practice over the last decade, Patton explained that “we have moved well beyond the idea that there are only formative and summative evaluations [and that] this diversity means that both the evaluator and the user have the opportunity to figure out the kind of evaluation that is most appropriate for their particular situation and information needs.” Michael Quinn Patton, “A Conversation with Michael Quinn Patton,” *The Evaluation Exchange Newsletter. Family Support Issue* (Spring 2002), available at <<http://www.gse.harvard.edu/hfrp/eval/issue18/quanda.html>> (accessed November 2003).

21 Government of Canada, Human Resources and Skills Development Canada (HRSD), *Quasi-Experimental Evaluation. Evaluation and Data Development. Strategic Policy* (1998), p. 6. <http://www11.hrdc-drhc.gc.ca/pls/edd/QEE_78000.htm> (accessed October 2003).

22 Evaluation Associates Ltd., *A Guide to Programme Evaluation* (1997). <<http://www.evaluation.co.uk/evaluation>> (accessed October 2003).

23 Michèle Antoine is also suggesting an *ante facto* approach, “l’évaluation préalable” to evaluate expectations of museum visitors. Michèle Antoine, « L’évaluation ... une pratique à l’usage des riches et bien nantis? », *La recherche universitaire en muséologie. Actes du 2^e colloque organisé par le programme conjoint de Maîtrise en muséologie de l’Université du Québec à Montréal / Université de Montréal, Montréal, 18 mai 1991* (Montréal, 1991), p. 113, 116–117.

gram on the activities of operational units or on the quality of records management operations.

Another important consideration is the need to determine whether a qualitative or a quantitative approach is to be taken. A quantitative approach facilitates comparisons and offers more potential in terms of generalization. Both the quantitative and qualitative approaches provide a reliable and valid picture of a given situation; however, a qualitative approach increases the depth of understanding. Patton adds that: "qualitative and quantitative methods involve different strengths and weaknesses and constitute alternative, but not mutually exclusive, strategies for research."²⁴ For Jean-Pierre Clavel, evaluation is based on both the quantitative and qualitative approach because it is fundamentally about researching objective (numbers) and subjective (qualities) values of an object.²⁵

It is also important to consider whether the evaluation uses a comparative perspective (assessing performance against outside organizations) or an autistic/self-diagnostic approach (assessing performance against itself). According to Charles Curran and Philip M. Clark, the comparative approach in the performance measurement of libraries is misleading because too many different and hardly comparable factors influence the activities and the results achieved from one library to the other. In their view: "Accountability is best determined by measuring the present self against the past self and by planning to target performance for a future self."²⁶ For Douglas Zweizig, however, results achieved in other libraries can be effectively used as a standard for comparison, providing the focus is on the ends and not on the means.²⁷

We can see that there is no such a thing as a one-dimensional approach to evaluation. In archives, as elsewhere, program evaluation has to rely on a variety and a judicious mix of approaches.

Program Evaluation: Methodology

Regardless of the approach, it is imperative to determine at the outset what is to be accomplished with the results of an evaluative inquiry. The following seven "responsibilities," identified by Posavac and Carey as a requirement for provision of effective services, can be useful in developing these accomplishment/results/outcome statements: "devote resources to meeting unmet need";

24 Michael Quinn Patton, *Qualitative Research*, p. 14.

25 Jean-Pierre Clavel et al., dir., *L'évaluation des bibliothèques universitaires* (Montréal, 1984), p. 7.

26 Charles Curran and Philip M. Clark, "Implications of Tying State Aid to Performance Measures," *Public Libraries* (November-December 1989), p. 350.

27 Douglas L. Zweizig, "So Go Figure: Measuring Library Effectiveness," *Public Libraries* (Spring 1987), pp. 22-23.

“verify that planned programs do provide services”; “examine results”; “determine which services produce the best results”; “select the type of programs that offer the most needed services”; “provide information needed to maintain and improve quality”; and, “watch for unplanned side effect.”²⁸

These seven responsibilities can be used to formulate expected results from an evaluation inquiry. Examples of expected results might include: whether budgetary resources are sufficient to support objectives; whether information is available to make decisions on reallocation; whether products and services are or are not produced and delivered in accordance with sets of predetermined standards; whether programs and activities support or do not support the strategic orientations or the priorities of the organization; whether impact and/or the side effects of a program on a targeted population are or are not monitored to ensure responsiveness.

To further illustrate, to assess to what extent a moving image copying program is performing in accordance with a commonly accepted set of technological standards, agreement is required on the validity of the benchmarks, the measurement techniques, and the scope of the samples. However, in order to measure the long-term cost-benefit of a similar program from a preservation perspective and/or to determine the value of its contribution to improved use of the audio-visual holdings at a national level, requires a different set of data, e.g., dollars invested, amortization of costs, quality of images, expectations of users, populations served, speed of delivery process, etc. Another illustration may be: the assessment of the impact of diffusion through the Web, of an unknown body of archival records on the development of knowledge in a targeted segment of a population, or on the development of historical research. To make decisions on the future development of a digitization program, one needs to go beyond the simple counting of “hits,” to capture significant data on the effective use of the information contained in these records (whether in areas of publications, cinema production, or television programming). We can therefore see that determining the use of the results of an evaluation initiative represents an important step. It has a direct effect on the choice of the approaches, on the determination of the methodological parameters, on the types of skills and sensitivities required from the evaluators and the participants in the evaluation process, and on the instruments and data required to substantiate the decisions to be made.

Having determined what is to be accomplished, it is then necessary to identify the criteria, the measures, the indicators,²⁹ including the data and informa-

28 Emil J. Posavac and Raymond G. Carey, *Program Evaluation*, pp. 3–7.

29 The OED is defining a **crit**erion as being “A canon ... by which anything is judged or estimated” or “a characteristic attaching to a thing, by which it can be judged or estimated.” It is essentially the prism or the angle through which an object is evaluated. The **meas**ure is the set

tion needed to substantiate the evaluation study. Experience has shown that this represents a challenging part of the evaluation. Archivists are often asked, when planning activities or developing projects, to determine at the outset the indicators of success, both in terms of results and of impacts. Moreover, the collection of data can be time-consuming and costly. As noted by the British authors of *Measuring Performance*: “Any collection and evaluation of information is time consuming and it is important to ensure that facts and figures collected are capable of being used to answer the question posed or genuinely shed light on a problem under consideration.”³⁰ It is therefore imperative to make sure the criteria, the measurements, and the indicators selected are relevant, that they are sufficient in number to provide a complete and valid picture, that they are reliable, based on sound data, not subject to error, and that the data collection is realistic and affordable in terms of cost.

Criteria

Looking at criteria, one has to test the pertinence and ability to sustain the development of a valid picture. In the “Age of the Internet,” no one will challenge the fact that the number of visitors in a reading room will provide an incomplete and misleading view of the use of archives. In the same vein, one would be justified to question the scope of a criterion such as, the volume of photography reproduced in a given period, to appreciate the impact of a diffusion or communication program. In such a context, other criteria would be needed such as: the purpose of the request; the profession of the client/user; the effective use of the copies produced by an archival service; and the population who has access to the information contained in the copy. In this regard, museologists and librarians have developed expertise and knowledge in the analysis of the characteristics, the needs, and the behaviours of their users, visitors, and populations served.³¹ Archivists should also look at using this type of data to develop their own approach.

of instruments – quantitative or qualitative – used to assess the criterion or characteristics. Gilles Deschâtelets refers to “measures” as « des qualificatifs, des aspects [...] bien définis, bien identifiables des paramètres [criteria] à évaluer » (Gilles Deschâtelets, « Évaluation d’un service de télé-référence », dans Jean-Pierre Clavel et al., dir., *L’évaluation des bibliothèques universitaires*, p. 89). The **indicator** derives from or is the result of the measurement for a given criterion. It is the information on which the judgement and the decision will be based.

30 Patricia Methven et al., *Measuring Performance. A Practical Guide for Use in the Setting of Standards and Measurement of Collection Policy, Document Production, Repository Facilities, Cataloguing, Indexing, Outreach and User Satisfaction* (London, 1993), p. 6.

31 See among many other studies: Minda Borun and Randi Korn, eds., *Introduction to Museum Evaluation* (Washington, DC, 1999); Geoffrey Ford, “Approaches to Performance Measurement,” pp. 73–87; F.W. Lancaster, *If you Want to Evaluate your Library*, 2d ed. (Champaign,

Measurements

The literature in information science provides abundant information on measurement methods³² and again archivists might find some inspiration for potential application in their milieu. Let’s look first at the comparative measures. Discussing the nature of the data to be used to substantiate a comparative analysis, Palmour and his collaborators state that

The comparative evaluation of one library with others depends upon the availability of comparable data [and that] where such statistics are available, the data coordinator should choose a number of libraries or systems which have environments similar to that of his or hers, or which have comparable sizes of population, and compare those resources and services for which data are available.³³

It is certainly easier to proceed with a comparative evaluation when standards can be used as predictors, as in the American public libraries system.³⁴ Since such commonly accepted factors have not been defined in the archival world, the population served might represent an acceptable basis for comparison, providing it is properly qualified. This is necessary data since, as has been observed in the public library environment, “use varies with population characteristics such as urban/rural populations, levels of education, per capita income, and medium age of population.”³⁵ Provision of a picture of the demographic environment, including socio-professional characteristics and population size, would help compare performance of archival programs.

From a comparative perspective, budget is another key factor to be considered. As suggested by Charles and Ruth Rockwood, the size of the population

1993); Ross J. Loomis, *Museum Visitor Evaluation: New Tool for Museum Management* (Nashville, 1987); Charles R. McClure and A. Samuels, *Strategies for Library*; Denis Samson, « La notion d’évaluation dans le champ muséal », Groupe interdisciplinaire de recherche en archivistique (GIRA), *L’évaluation des archives : des nécessités de la gestion aux exigences du témoignage* (Actes du 3^e Symposium en archivistique, Montréal, 1998); Douglas Zweizig et al., *The TELL IT! Manual*.

32 See among many others, the two complementary works of F. Wilfrid Lancaster, *The Measurement and If you Want*, frequently referred to by academics and information professionals. Museologists also have a long-standing tradition in program evaluation. They have developed numerous methods – well documented in the literature – « [visant] toutes à cerner un aspect ou un autre de la relation entre l’institution, ses programmes et ses expositions avec les usagers ». Denis Samson, « La notion d’évaluation », p. 47.

33 Vernon E. Palmour et al., *A Planning Process for Public Libraries* (Chicago, 1980), p. 110.

34 See Charles E. Rockwood and Ruth H. Rockwood, “Quantitative Guides to Public Library Operations,” *University of Illinois Graduate School of Library Science, Occasional Papers* 89 (Urbana, 1967), p. 4.

35 Palmour, *A Planning Process*, p. 110.

served, is a chief factor in the determination of budget requirements.³⁶ After testing different factors – population served, library staff, number of volumes,³⁷ circulation, and total budget for which they have established coefficients of determination – Charles and Ruth Rockwood concluded that: “a comparison between public libraries grouped according to their budgets for most purposes, will be much more meaningful than comparisons between public libraries according to any of the other plausible guides tested.”³⁸

In developing their profile of the Canadian archival system, the Consultative Group on Canadian Archives used, twenty-five years ago, the budget factor as a basis for comparison.³⁹ Actualized financial data, provided by a much larger group of respondents⁴⁰ and co-related with other factors, such as the ones studied by the Rockwoods, would provide a picture expressing different aspects of the Canadian archival system’s complexity and could be particularly useful in the review of funding strategies, be it at the local, regional or national levels. Comparative measures of financial data coming from various archives services are also of interest for establishing benchmarks for allocations of budgetary resources.

In addition to these external measures, internal measures are necessary to answer fundamental questions relating both to performance and to pertinence of programs. One internal measure is based on roles and mission of an organization. Charles McClure and his collaborators define mission as a concise expression of an organization’s purpose and the fundamental reasons for its

36 Rockwood and Rockwood, “Quantitative Guides,” p. 2.

37 By comparison, since long-term preservation of records with heritage value is a key element of an archives’ mission, and since, in many instances, archives store and circulate semi-active records, the size of the archival holdings and the volume of records stored in records centres cannot be ignored and might represent useful guides to determine the level of budget. From a public service perspective, the volume factor would probably need to be qualified before being co-related to the use factor.

38 Rockwood and Rockwood, pp. 10–11. Co-relating budget with other factors, they were able to suggest the added value of every additional slice of USD \$10,000 to the budget of public libraries in terms of persons served, staff required, purchase of books, and total circulation.

39 *Canadian Archives*, Report to the Social Sciences and Humanities Research Council of Canada by the Consultative Group on Canadian Archives (Ottawa, 1980), pp. 50–52. We are not aware of any other comparative study of the Canadian archival system. The NPT (National, Provincial, and Territorial Archivists) are compiling, in preparation for their annual conferences, various data on distribution of resources, volume of holdings, services offered, and use of archives in their respective institutions. However, these data are generated for internal consumption and are not accessible through publications.

40 The most recent edition of the *CCA Directory of Archives* provides a list of 800 Canadian archival repositories (<<http://cdncouncilarchives.ca/directory.html>>). In August 1978, 321 questionnaires were distributed in an attempt to contact all institutions involved in archival activity and 185 were returned “with explanatory letters from a further 31 archives” (*Canadian Archives*, p. 30). Assuming that a similar study is done in 2003, that the same attempt “to contact all institutions” is made, and that the response rate (*ca* 60%) remains the same, the number of respondents today could be close to 500.

existence. They recognize the mission of public libraries as eight distinct roles. One of them, that of “research centre,” corresponds to a key traditional role of archives: “... to assist scholars and researchers to conduct in depth studies, investigate specific areas of knowledge, and create new knowledge.”⁴¹

In order to assess the degree of success achieved, these authors suggest output measures that may also be useful for archival programs: “title fill rate,” which “reports the percentage of specific titles desired by ... users that are available in the library at the time of the request”; “reference completion rate,” which “is a measure of the number of reference transactions completed in proportion to the total number of reference transactions”; “in-library materials use per capita,” which is based on the number of items removed from storage locations and used within the institution and “reports the estimated annual number of materials used by patrons ... divided by the jurisdiction population”; and “document delivery,” which refers to “the delivery of material coming from outside the library.”⁴²

Although these measures would need to be adapted and refined for the archival community, co-relating documents requested to documents effectively available (in an archival environment, the concept of availability would have to be qualified), and establishing the ratio between reference transactions and reference completed, would provide significant information. They offer potential to assess the “fit” between roles and expectations and to determine the degree of success or the level of excellence achieved. Further, these three measures are useful to measure the outputs, i.e., the quality of reference and access services provided to users and to the population served.

Indicators

As discussed earlier, the primary purpose of performance measurement is to provide significant indicators from which judgements can be developed, accurate and sound decisions can be made, and appropriate actions can be undertaken to improve the quality of products or services and/or to increase the impact of on-going or specific activities. The most commonly used indicators in the literature refer to “inputs,” “throughputs,” “outputs,” and “outcomes/impact.”⁴³

Input indicators relate to the resources: human, informational, material and

41 Charles R. McClure et al., *Planning and Role Setting for Public Libraries: A Manual of Options and Procedures* (Chicago, 1987), p. 28.

42 Douglas L. Zweizig and Eleanor Jo Rodger, *Output Measures for Public Libraries: A Manual of Standardized Procedures* (Chicago, 1982), pp. 17, 39, 45, 53, and 87.

43 See Lucie Pagé, « Évaluer pour évoluer », pp. 143–44. The British Office of Arts and Libraries refers to a different categorization: “operational performance,” “effectiveness,” “cost-effectiveness,” “impact,” and “other derived indicators.” Office of Arts and Libraries, *Key to Success: Performance Indicators for Public Libraries. A Manual of Performance Measures and Indicators Developed by King Research Ltd.* (London, 1990), pp. 78–79.

financial, used for the production and the delivery of products or services. The holdings of archives or records can also be considered part of the resources available to an archival organization to generate products and to provide services. These indicators also include the mission, mandate, and strategic orientations as well as the numerous standards, directives, and instructions which all represent elements of an organization's framework for action.

Throughput is a second category of indicators documenting the energy in motion or in action or the "resources applied." Throughputs inform on the performance of activities; they are indicative of the work done with the resources invested and other input and aim at showing the magnitude of effort made to perform activities and to achieve results. In an archival environment for example, throughput tells evaluators what and how things are done with regards to assistance to records producers, development and implementation of methods, processes, and tools needed to manage the recorded information continuum, appraisal, disposition, acquisition, accessioning and transfer of archival records, description of records throughout their life cycle, preservation, circulation, availability, access, communication, and other uses.

Outputs are indicators that reflect the results achieved. Outputs document the products generated or the services delivered with the available resources and through their application. In archives, they provide indications, for example, on: the quality and the quantity of advisory services offered; retention schedules effectively developed and implemented; records appraised, acquisition, transfer, or elimination of archives; the nature and scope of instruments resulting from descriptive activities; and the different ways people use recorded information. These indicators shed light on what is being effectively produced and/or delivered.

Impact indicators document the effect of achieved results and represent the fourth and last category of indicators. Impact indicators are intended, for example, to measure the extent to which the production of a given set of finding aids facilitates, improves, or enlarges access to archival information and effectively contributes to the development of a specific field of knowledge. Impact indicators are also of interest to those who want to appreciate the impact of a change to a program or to evaluate how it affects a targeted audience in the long term. Authors and practitioners agree that assessment of outcomes of a program is extremely difficult because of intangible objectives. Peter Drucker even suggests that: "one would do well to abandon the idea of using desired outcomes as direct criteria for the evaluation of ... information services."⁴⁴ Baker and Lancaster, however, do believe that it is possible to analyze the impact of a service by looking at factors such as: existence, awareness, trial, adoption, referral, and true impact or benefit.⁴⁵

44 Quoted by Baker and Lancaster, *The Measurement*, p. 3.

45 *Ibid.*, p. 154.

If we were therefore to look at the impact of the “existence” of a new regional archival service,⁴⁶ one could argue that the archives acquired, preserved, and made accessible by this service could have been lost, that major aspects of the regional documentary heritage could not have been explored, and that knowledge growth of regional and even national history would have been impaired. It might be easier though to illustrate increases in awareness or adoption of a product or a service through indications of behaviours of visitors and users or on references in publications, in the press, and in the media or in *curricula*. Demonstration of monetary gains or impacts is also feasible through calculation of the differences between what one was ready to pay and what has been effectively paid to obtain information; or between the time and other resources, expressed in monetary terms, one would have spent to retrieve the information and the cost actually paid to have it delivered; or between the cost of an operation or of a decision with and without the information coming from the archives.⁴⁷

Program Evaluation and Archives

Is this reservoir of reflections, analysis, and studies, developed and applied in related fields of knowledge and practice, of interest for archivists? Does program evaluation offer to archival institutions and services a useful tool to assess performance and measure the quality and the impact of archival products and services? The answer to these questions is yes; but two conditions need to be considered. First, the archival community must include program evaluation in the list of its fundamental questions on theory and practice and on the development of the field.⁴⁸ Second, the knowledge and the various

46 Such as the recently created Centres agréés d’archives in Quebec.

47 Baker and Lancaster, *The Measurement*, pp. 154–61. To illustrate this last point, one could think of the Swiss “Fichier systématique des compétences” developed between 1848 and 1980 which contains relational indications on competencies and related responsibilities in the Swiss federal bureaucracy (Primus Monn, « Les compétences légales au service des Archives fédérales suisses », *Revue des Archives fédérales suisses* 26 [2000]), p. 337, et *passim*). Should there be a need today to collect all this information, the cost of such an activity would have to be calculated in millions of CH francs. The very existence of this source of information represents major savings which can be estimated through comparisons, on one hand, of the cost of its production, maintenance, and use, and, on the other hand, the eventual price which would have to be paid to generate this information or to obtain information of similar interest from elsewhere.

48 This may seem simplistic, but judging from the score of archival textbooks or manuals currently referred to or in use in archival education, management, including program evaluation, is not raising particular interest. Apart from few exceptions – Thomas Wilsted and William Nolte, *Managing Archival and Manuscript Repositories* (Chicago, 1991); Archives Association of British Columbia, *A Manual for Small Archives* (Vancouver, 1999), (<<http://www.aabc.bc.ca/aanc/msa>>, accessed May 2004); Anne-Marie Schwirtlich and Gunnell Bell-

applications of program evaluation developed by scientists and practitioners in other domains have to be looked at from an archival perspective and adapted to the reality of archives. To meet these conditions, the community may want to involve academics and practitioners from all horizons of the profession in a collective initiative, the purpose of which would be to develop a suitable program evaluation framework before the end of the current decade. As a corollary, the archival community should consider the development of a research program as a strategic instrument to facilitate the achievement of this goal. Though it is early in such an exercise, the following four dimensions may be taken into considerations at the conceptualization stage.

Knowledge Base

In her recently published text on program evaluation and archives, Lucie Pagé addresses key issues on evaluation methods and techniques.⁴⁹ She also introduces the half-dozen publications of the Society of American Archivists (SAA) on program evaluation in archives and a few other guides or studies from professional associations and from Library and Archives Canada.⁵⁰ For Pagé, these contributions “mettent davantage l’accent sur l’évaluation des intrants, les ressources, et négligent bien souvent les extrants, les services [et] s’intéressent peu aux relations entre les centres d’archives et la communauté.”⁵¹ Looking back at the SAA’s initiative on institutional evaluation, the approach was in line with the spirit of program evaluation in that its initial goal was to help repositories improve their services, assist archivists in their efforts to better serve the public by improving archival institutions, and allow archives to be evaluated in terms of the goals they set for themselves.⁵²

But judging from the set of ten principles adopted in 1994,⁵³ the orientations seem to have somewhat changed. Based on their substance and on their wording – the use of the conditional tense is particularly revealing – and since they were

viken, “Getting Organised: The Basics,” in Ann Pederson, ed., *Keeping Archives* (Sydney, 1987), pp. 21–72; Elizabeth Shepherd and Geoffrey Yeo, *Managing Records. A Handbook of Principles and Practice* (London, 2003) – authors are barely referring to this key function.

49 In which (« Évaluer pour évoluer », pp. 135–166), she is revisiting a text published ten years earlier (« Méthodes et techniques », pp. 57–78).

50 Lucie Pagé, « Évaluer pour évoluer », pp. 135–166.

51 Ibid., pp. 137–138.

52 “Report of the Task Force on Institutional Evaluation,” *SAA Newsletter* (January 1980), p. 7 and *Evaluation of Archival Institutions. Services, Principles, and Guide to Self-Study* (Chicago, 1982), p. 3.

53 Derived from the accreditation process of the American Museum Association and stated in the 1982 SAA *Guide to Self-Study*, these principles were adopted by the SAA Council as “Guidelines for the Evaluation of Archival Institutions,” *Archival Outlook* (March 1994), pp. 6–7.

conceived to serve both self-assessment studies and an accreditation processes, these principles or guidelines look more like conditions to be met in order to be recognized as a state-of-the-art archives. Nevertheless, the tools developed to support the process and designed to assist archivists in organizing information about their service and in “systematically” evaluating it⁵⁴ remain valid and provide helpful indications from a methodological perspective.

It would be useful to add to Pagé’s list, the guide developed by British archivists and published in 1993 by the Society of Archivists: *Measuring Performance* which is based on a wide analysis of practice and pilots tested nationwide and presented by its authors as an authoritative reflection of professional practice.⁵⁵ Similarly, but from a different perspective, the reflections of records management specialists on the value of their programs might also represent useful learning material for archivists. Contributions such as Robert L. Sanders’ article on *RIM Self-Evaluation*,⁵⁶ Susan K. Goodman’s exploration of “value added concepts for information processes in organizations,”⁵⁷ and William Saffady’s benchmarking study⁵⁸ are worth mentioning. Stating the fundamental goals of RIM programs – security, accessibility, and elimination – Sanders poses questions that are at the core of any program evaluation initiative:

What are the inherent program strengths and weaknesses? Is the program appropriate to internal and external environment? To what extent is our program integrated into the real business of our company? Does our program fit the company’s particular needs? Have we objectively evaluated our role in the organization?⁵⁹

Referring to authors who wrote about “value-added processes,” but discussing mostly the model proposed by Robert S. Taylor in *Value-Added Processes in Information Systems*,⁶⁰ Goodman introduces this theoretical framework which, in her view, represents a valid instrument “to describe, define, analyze and synthesize data about the operation and effects of information systems” and to “support decision, judgemental, analyzing and organizing processes.”⁶¹

54 Paul H. McCarthy, *Archives Assessment and Planning Workbook* (Chicago, 1989), p. 5.

55 Methven, *Measuring Performance*, p. 1.

56 Robert L. Sanders, “RIM Self-Evaluation. The Case for Truth and Realism,” *Records Management Quarterly* (July 1997), pp. 47–53.

57 Susan K. Goodman, “Measuring the Value-Added by Records and Information Management Programs,” *Records Management Quarterly* (April 1994), pp. 3–13.

58 William Saffady, *Records and Information Management: A Benchmarking Study of Large U.S. Industrial Companies* (Lenexa, 2002).

59 Sanders, “RIM Self-Evaluation,” p. 48.

60 Robert S. Taylor, *Value-Added Processes in Information Systems* (Norwood, 1986).

61 Goodman, “Measuring the Value-Added by Records and Information Management Programs,” p. 5.

If these authors are providing reflections potentially useful to set the foundations for a program evaluation methodology for archives, both Saunders' questions and Taylor's model, including Goodman's analysis, need to be substantiated with additional information on criteria, measures, and indicators. Saffady's examination of the status of RIM programs in forty-two American industrial companies, intended for specialists interested in RIM policies and practices of benchmarking and program evaluation, includes indications on questions to be addressed and will certainly help in developing evaluation frameworks.⁶² However, as illustrated throughout this article, in addition to this relatively small *corpus* of archival literature, further exploration in the numerous program evaluation studies of specialists in management and information sciences, in museology, or in education needs to be performed.⁶³

Current Program Evaluation Practices in Archives

The investigation of the literature can only be profitable if realized in parallel with an analysis of the evaluation requirements and practice in the archival system, including the quality and the evaluative potential of administrative data currently produced and available. A preliminary examination of administrative records from three archives services⁶⁴ indicates that archivists have developed experience and skills in setting goals and objectives, in monitoring activities, in identifying results expected, and in measuring the degree of achievement and the cost of archival products, services, and activities. A quick analysis though, of records such as strategic and operational plans and related status reports, annual or monitoring reports, or statistical "bilan" or "health of the organization" on various types of use of archives products or services, shows that in most cases these documents are generated for control purposes and that their content is not necessarily used in an evaluation perspective. Nevertheless, this data could be used in a program evaluation process, either to

62 Saffady's study is "designed to address ... the current status of records and information management programs in American business [and to provide] records managers [with a] method of comparing their companies' programs, policies, and practices with those of other companies."

63 We would like to thank Véronique Dupuis and Olga Sigal who identified – in pre-selected publications in archival, library, management sciences, and in museology – numerous criteria, measures, and indicators studied or in use in these milieus. The purpose of these "pre-research" projects – performed in 2002 and 2003 and funded through the Petites subventions de recherche du CRSH, administered by the Université de Montréal – was to verify the availability and the exploitability of the academic and professional literature, and to assess its potential for a larger research on methodology for program evaluation in archives.

64 Managers of these services have requested anonymity both for themselves and for their respective administrative units. As a matter of fact, confidentiality will be an easily understandable pre-requisite for any research in program evaluation.

document criteria, to support measurement, or to assist in the formulation of indicators.

Among the initial things to be looked at in the archival milieu should be the current perception, understanding, and use of program evaluation in the archival system. It is essential to learn about the expectations of archivists and other stakeholders with regard to program evaluation and to clarify what they would want to achieve through evaluation of their programs and/or activities. Moreover, there is an imperative need to understand what archivists currently do and what they want or require to know in order, for example, to maintain or improve the quality of their products and services or to ameliorate the knowledge of their clientele, including their profiles and their expectations. What would they want to achieve through an evaluation program? Answers to these questions will lead to an articulate categorization of the multiple facets of expectations and requirements and will set the context and assist in the development of the methodology and of the instruments of the research.

Content and Methodological Framework

The organization of the research work requires careful thought and attention. The number of services and/or institutions composing the Canadian archival system is close to a thousand and their status and magnitude varies considerably. At one end of the spectrum, there is a national institution with millions of dollars in budget and hundreds of employees; at the other end, there are scores of local heritage and/or historical societies with a few hundred dollars allocated to archival activities, most of the time performed over a few hours a week through volunteers. From a methodological perspective, these entities can hardly be put in the same basket. An exercise of categorization, based on multiple factors – e.g., populations and users served, jurisdictional status, level and quality of resources available, nature of activities, mandates, and roles, etc., will have to be performed. This should generate a multi-facet grid of archival institutions/services and determine the level of customization required in the production of the instrumentation for data collection and information analysis.

Moreover, it would be ambitious to try to do everything at once. A number of considerations will be required such as: sequences in the research program; categories of institutions and services to be studied; participants to be selected; and activities to be investigated. These will set the framework to determine the nature of the questions to be raised, the data required, and the methods for gathering information – parameters of inquiries, format of questionnaires, type of interviews, profiles of interviewees, grids for selection and analysis of administrative records, etc. – and will influence the definition of the methodology – qualitative, quantitative, mix of both, etc. One should reasonably expect that an approach “à géométrie variable” will be required and

that here, probably more than elsewhere, “one size does not fit all.” A clear definition of methodology is a key issue. In order to generate valid and credible results, the methodology needs to be based on sound theoretical foundations, on thorough knowledge and understanding of validated tools, and on fair appreciation of current practices.

Interaction and Communication

Program evaluation will be accepted by the archival community, including managers, professionals and technicians, if, and only if, it is deemed to be useful and usable to facilitate and sustain the implementation and the delivery of archival programs. This means careful attention will need to be given to this demonstration. One obvious means to achieve this is to involve the community in the development of the methodological framework and in the creation of the evaluation instruments. It is particularly important to have archivists on board to formulate the evaluative questions, i.e., those focussed interrogations from which will derivate all the methodological instrumentation. It will be no less essential to test regularly, in the milieu, the tools – criteria, measurements, indicators, calculation formula, data required – before applying them more largely, be it on a small sample of participants or on a larger group of institutions or services. This will permit appropriate adjustments, alleviate criticisms, and facilitate the performance of inquiries on larger grounds. In certain instances, it may be wise to plan the performance of pilot projects after the inquiry to confirm the validity of the findings and the applicability of the results. Finally, careful considerations will need to be given to communications. The community will need to be kept regularly informed of the progress achieved through regular and frequent postings of results obtained.

Information communication technologies offer a large variety of options in this regard, as well as for collecting information and for conducting the inquiries or the tests: an electronic newsletter, a Web site, a discussion forum, etc. The community should also have the opportunity to comment and to discuss this information at annual conferences, professional colloquia, or specific workshops. In other words, it will be important to develop an articulated and pro-active communication strategy, the objective of which should be, not only to keep the actors abreast of progress, findings, and results achieved, but also to allow them to actively participate in the initiative.

Conclusion

The underlying purpose of this article was to illustrate the potential of program evaluation for archives, and as well, to start the conceptualization process for the development a research initiative on the subject.

The academic and professional literature indicates that the archival commu-

nity has not reflected significantly on program evaluation, though, much has been written on this subject in other fields of endeavour. This article has attempted to provide a comprehensive view of program evaluation and has illustrated, somewhat, its applicability to archives. The suggestion is not to simply adopt the theory and methods used by others, but to carefully extract and apply the elements of knowledge and practice potentially suitable for archives and to the development our own theory, methodology, and practice.

The development of a research initiative has the potential to provide the archival profession with the methods, techniques, and tools demonstrated to be effective in related professional fields, to better assess the quality and impact of their programs, activities, products, and/or services. Such an initiative represents a complex challenge that will require the involvement of all the players in the archival community. For example, the academics' role would be key, through various research projects, in investigating and developing theoretical and methodological frameworks, in performing the required inquiries, and in analyzing and synthesizing their findings. It will be important that practitioners and stakeholders – including managers, professionals, technicians, and users who, on a daily basis, ensure the delivery of or utilize archival products and services – contribute and participate in the setting of the methodological framework, provide the information required, and validate the results and their applicability.

An initiative of this magnitude is not going to be achieved through a gigantic mega-project. It is probably going to be composed of multiple small projects investigating specific objects of program evaluation and iteratively contribute to the development and evolution of a methodology suitable for archives. Such an endeavour will be worthwhile to ensure the position of archives among its competitors in both the cultural and the information worlds and to provide significant, telling, and sound information that demonstrates its pertinence in our global world.

