

ÉTUDE

L'évaluation de programme : modèles multiples et caractéristiques souhaitables pour une application en milieu archivistique

Jacques Grimard

INTRODUCTION

S'ils ne pratiquent pas systématiquement l'évaluation de programme, les archivistes, y compris les gestionnaires de documents, n'en ignorent pas pour autant la nature et les applications. Dans un article paru à l'automne 2004 (Grimard et Pagé 2004), dans la revue *Archival Science*, Lucie Pagé et moi avons montré que les unités administratives ou les programmes à l'intérieur desquels œuvrent les archivistes font l'objet d'évaluation de la part des vérificateurs ou autres auditeurs de leurs organisations tutélares. Nous avons également mis en lumière le rôle « expert » de certaines institutions publiques d'archives dans l'évaluation des pratiques de gestion documentaire de ministères, agences ou autres organisations gouvernementales.

Par ailleurs, les travaux qui, au cours des années 1980 et 1990, ont mené à la publication, en 1994, des *Guidelines for the Evaluation of Archival Institutions* de la Society of American Archivists (SAA 1994) indiquent que la question fait l'objet, depuis un bon moment, de réflexions et de discussions dans la communauté archivistique. Moins présente dans la littérature archivistique des dernières années, l'évaluation de programme semble cependant effectuer un retour. Ainsi, la formulation de la thématique du congrès 2005 de l'Association des archivistes du Québec – « Avons-nous les moyens de nos ambitions? La réponse est-elle d'ordre législatif, stratégique ou professionnel? » – semble exprimer des préoccupations évaluatives. Autre signe encore plus explicite : l'Association of Canadian Archivists a inscrit à son programme de la Conférence 2005 une session portant spécifiquement sur l'Évaluation de programme : « Are we good enough? »

Do Archives make a difference? Evaluating Performance and the Impact of Archives ». Faut-il y voir l'expression d'un intérêt passager ou, au contraire, l'amorce d'un chantier de réflexion ou d'un programme de travail? S'il faut en juger par des travaux comme ceux de l'International Records Management Trust (IRMT) sur le Records Management Capacity Assessment System (RMCAS)¹ ou encore par la diffusion de nombreux outils d'évaluation de la capacité de gestion de l'information², il est permis de penser que nous assistons à un regain d'intérêt, de la part des professionnels de l'information, particulièrement des « *records managers* »/gestionnaires de documents ainsi que des archivistes, pour l'évaluation de leurs activités.

L'implantation d'une culture et de pratiques d'évaluation de programme au sein d'une communauté professionnelle suppose que cette dernière comprenne bien de quoi il s'agit et en mesure bien l'intérêt pour ses propres activités, voire pour son propre développement. Je me propose dans le cadre de cet article d'explorer plus avant cette discipline, dont les premiers balbutiements remontent à près de deux siècles et d'en présenter brièvement les approches les plus utilisées par les spécialistes du domaine. Je compte également, à défaut de pouvoir proposer une ou des approches directement applicables en milieu archivistique, suggérer des caractéristiques à rechercher dans les divers modèles existants de manière à ce que les archivistes puissent éventuellement tirer le meilleur parti possible de l'évaluation de programme, à la fois outil de gestion des organisations et véhicule d'apprentissage sur leur fonctionnement et sur leur développement.

Avant d'aller plus loin, le lecteur doit aussi être averti qu'il ne trouvera pas dans le texte qui suit une proposition de cadre méthodologique défini sous tous ses aspects ou encore une grille d'analyse immédiatement applicable dans les services d'archives. L'objectif est ici beaucoup plus modeste. S'inscrivant dans une démarche lentement amorcée il y a quelques années, ce texte fait en quelque sorte suite à deux articles déjà parus dans les revues *Archivaria* et *Archival Science* (Grimard 2004, 69-87; Grimard et Pagé 2004, 99-126) et vise simplement à présenter à la communauté archivistique l'évaluation de programme dans toute la multiplicité et dans toute la complexité de ses approches. C'est que cet outil de gestion est bien davantage qu'une simple boîte à outils remplie d'instruments de mesure applicables à tous égards. C'est d'abord une approche avant d'être une méthode assortie d'instruments de collecte, d'analyse et de mesure. Il m'est apparu utile de tracer ici à l'attention des lecteurs de la revue *Archives*, un portrait de l'évaluation de programme qui fasse état des multiples approches qui la caractérisent. C'est, me semble-t-il, un préalable incontournable pour qui veut par la suite arrêter des choix quant aux méthodes les plus prometteuses et quant aux outils les plus efficaces.

BREF RETOUR ... EN ARRIÈRE

Mon intérêt pour l'évaluation de programme en milieu archivistique remonte au temps où, travaillant dans des institutions publiques d'archives canadiennes, j'étais occasionnellement amené à participer à des études d'évaluation « formative » ou « sommative »³, soit à titre de membre de groupes de travail mandatés pour recevoir les résultats d'études évaluatives, soit à titre de fournisseur de renseignements ou soit encore à titre de gestionnaire responsable de secteurs ou d'opérations archivistiques.

Par ailleurs, en dehors de ces activités proprement évaluatives, j'étais régulièrement appelé à collaborer à des exercices de planification stratégique ou opérationnelle au terme desquels il fallait arriver à déterminer les facteurs de succès des activités ou projets envisagés et à préciser des critères aussi bien que des instruments sur lesquels nous entendions prendre appui pour juger des résultats atteints. Je retiens de ces expériences un certain nombre d'observations quant aux difficultés qui se posent lorsqu'il s'agit de porter un jugement sur le déroulement, sur les résultats ou encore sur les retombées des activités archivistiques :

- Les exercices d'identification d'indicateurs de succès donnent souvent lieu à des choix dont on se rend compte à l'usage qu'ils sont moins révélateurs qu'on ne l'avait espéré au départ et qu'ils sont parfois, voire souvent, difficilement utilisables ;
- Les données et informations sur lesquelles prennent appui les observations aussi bien que les conclusions des études évaluatives ne sont pas toujours des plus appropriées, soit qu'elles soient fragmentaires et se prêtent difficilement à toute tentative de généralisation, soit qu'elles ne révèlent qu'un aspect de la réalité et ne permettent pas de juger adéquatement de l'ensemble de la problématique étudiée ;
- Les études d'évaluation suscitent généralement de l'inquiétude et de la méfiance chez ceux et celles dont les activités font l'objet d'un regard scrutateur, ce qui ne va pas sans affecter la qualité même des résultats de l'exercice d'évaluation, et sans compromettre les leçons qu'on en tirera ou l'usage qu'on en fera – usage d'ailleurs plutôt mal connu qu'on n'a habituellement pas précisé d'entrée de jeu.

C'est dans la foulée de ces observations tirées de l'expérience que j'ai choisie, en joignant l'École de bibliothéconomie et des sciences de l'information de l'Université de Montréal, de pousser plus loin la réflexion sur l'évaluation de programme. Des recherches préliminaires dans la littérature scientifique et professionnelle m'ont fait voir que si cette discipline était bien connue et pratiquée systématiquement par les bibliothécaires et les muséologues, elle l'était vraisemblablement moins chez les archivistes. Il m'est alors apparu que la communauté archivistique gagnerait à maîtriser les fondements théoriques et les méthodes d'évaluation de programme et à se doter de méthodes d'évaluation qui soient adaptées à ses besoins et à ses attentes. C'est dans cet esprit que je me suis adressé au Conseil de recherche en sciences humaines du Canada qui a accepté de financer ma recherche sur *l'Évaluation de programme en milieu archivistique : principes et pratiques*. Le projet a donc pour objet les fondements théoriques et les méthodes d'évaluation de l'efficacité (processus), de l'efficacités (résultats) et de l'impact (portée ou retombées) des activités, des produits et des services offerts aux clientèles, dans les organisations archivistiques. En fait, il s'agit de concevoir une approche d'évaluation de programme plurielle et modulaire 1) qui permette aux institutions et aux services d'archives d'apprécier l'efficacité, l'efficacités et la portée de leurs activités et de leurs réalisations ; et 2) qui tienne compte à la fois des objectifs de l'évaluation, des types d'organisations et de la nature de leurs activités et de leurs réalisations. Plus spécifiquement, le projet a pour objectif :

- de contribuer à l'établissement d'assises théoriques pour une démarche d'évaluation de programme en milieu archivistique ;

- d'identifier, selon différentes perspectives évaluatives, des critères, des instruments de mesure et des indicateurs adaptés aux objets et aux finalités de l'évaluation ;
- de construire un cadre méthodologique qui permette d'évaluer l'efficacité, l'efficacéité et l'impact des activités et des réalisations destinées aux clientèles internes et externes des services d'archives ;
- de proposer des instruments spécifiques d'évaluation de programme adaptés aux besoins et aux pratiques en milieu archivistique ;
- et ainsi, de lancer un chantier de réflexion, de susciter l'intérêt de la communauté archivistique pour les questions relatives à l'évaluation de programme et conséquemment de paver la voie à d'autres chantiers de recherche sur d'autres aspects de la problématique d'évaluation de programme en milieu archivistique.

On aura compris que cette démarche de recherche – ambitieuse et complexe – ne saurait donner de résultats satisfaisants si elle n'est pas d'abord assise sur une bonne connaissance, voire sur une solide compréhension, des modèles ou des approches évaluatives les plus susceptibles d'être utiles aux archivistes.

DE L'ÉVALUATION DE PROGRAMME

Ma collègue Lucie Pagé et moi avons eu l'occasion, dans des articles parus récemment (Grimard 2004, 70-73 ; Grimard et Pagé 2004, 101-102), de suggérer quelques-unes des définitions les plus englobantes de l'évaluation de programme et de mettre en lumière le fait que l'exercice va bien au-delà de la simple appréciation de l'atteinte d'objectifs et qu'il porte davantage sur les caractéristiques inhérentes des programmes, sur les activités qui s'y déroulent, sur les résultats obtenus, sur les retombées, voire sur la pertinence même du ou des programmes étudiés. Depuis au moins une décennie, les spécialistes du domaine mettent aussi l'accent sur le fait que l'évaluation de programme, quel que soit le qualificatif qu'on lui attribue, suppose aujourd'hui l'implication – à des degrés divers et selon des formules variées ainsi qu'en fonction de l'usage qu'on entend faire des résultats de l'étude évaluative – de tous ceux et celles qui participent au programme : depuis les allocateurs de ressources et autres décideurs jusqu'aux utilisateurs du ou des groupes visés par le ou les programmes évalués, en passant par les divers acteurs chargés de la réalisation même des activités. Les travaux de Patton sur le « Developmental Evaluation » (Patton 1994, 312-313) et ceux de Fetterman sur l'« Empowerment Evaluation » (Fetterman 1996, 40) confirment, parmi d'autres approches à caractère démocratique, l'importance du rôle des participants dans l'établissement des objectifs et dans la conduite même de l'évaluation.

Un peu d'histoire

Cela dit, l'évaluation de programme n'est pas nécessairement une discipline nouvelle. Les spécialistes du domaine en font remonter les premières applications en Grande-Bretagne, au début du XIX^e siècle, alors que la Révolution industrielle imposait des ajustements et des modifications en profondeur aux législations anti-pauvreté, aux programmes d'éducation et aux services hospitaliers et sociaux. Il s'agissait alors d'évaluer dans quelle mesure ces lois et ces institutions répondaient aux besoins et aux exigences

des nouvelles réalités du monde industriel. Il va sans dire qu'au cours des décennies suivantes les instruments se sont raffinés et que les approches se sont diversifiées : à un premier « Âge de la Réforme », qui recouvre tout le XIX^e siècle⁴ et où les pratiques évaluatives nous apparaissent, à l'aune de nos standards actuels, plutôt informelles et impressionnistes – encore que ces premières tentatives aient jeté les bases des approches empiriques qui marquent toujours certains modèles encore en application –, succède un « Age of Efficiency and Testing », qui marque le premier tiers du XX^e siècle.

Normalisation, systématisation et efficience deviennent les maîtres mots de l'évaluation particulièrement dans le domaine de l'éducation, alors qu'on s'intéresse à la performance des enseignants dorénavant appréciée selon des critères déterminés et qu'apparaissent les premiers tests objectifs destinés à juger de la qualité des enseignements. Au cours des années 1930, Ralph Tyler influencera considérablement les méthodes d'évaluation dans le monde de l'éducation aux États-Unis – on a même parlé d'un « Tylerian Age » pour la période 1930-1945 – par son approche comparative visant à mesurer les résultats obtenus en regard des objectifs initialement définis. La décennie suivante, qualifiée d'« âge de l'innocence » – voire d'« âge de l'ignorance » en raison du peu d'intérêt soulevé, toujours aux États-Unis, par la pauvreté et par le gaspillage des ressources – voit les méthodes et approches développées durant les périodes subséquentes se raffiner. Dans le monde de l'éducation, l'usage des tests normalisés se généralise et des efforts sont consentis à l'apprentissage de l'élaboration des objectifs. Au cours des quinze années qui suivent – « The Age of Expansion 1958-1972 » –, l'évaluation de programme connaît un développement fulgurant à la faveur, entre autres et surtout, du lancement de multiples programmes fédéraux américains de lutte à la pauvreté. L'évaluation devient alors un instrument essentiel de l'appréciation de l'efficacité de tous ces programmes. Elle donne lieu à un gigantesque effort de réflexion dont témoignent les multiples écrits sur de nouvelles théories et sur de nouveaux modèles ou approches évaluatives qui « recognized the need to evaluate goals, look at inputs, examine implementation and delivery of services, as well as measure intended and unintended outcomes of the program. They also emphasized the need to make judgements about the merit or worth of the object being evaluated » (Madaus, Stufflebeam et Scriven 1983, 14). Puis vient l'« âge de la professionnalisation » (1973-1982) marqué par un questionnement identitaire chez les évaluateurs eux-mêmes qui prennent de plus en plus conscience que « the techniques of evaluation must achieve results previously seen as peripheral to serious research; serve the information needs of the clients of evaluation; address the central value issues; deal with situational realities; meet the requirements of probity; and satisfy needs for veracity »⁵.

Qu'en est-il depuis les vingt dernières années ? Michael Quinn Patton voit dans la quinzaine d'années de la fin du XX^e siècle quatre faits marquants : 1) l'importance prise dans l'ensemble du processus évaluatif par l'utilisation des résultats de l'évaluation ; 2) la multiplication des approches et méthodes évaluatives davantage adaptées aux situations particulières rencontrées par les évaluateurs ; 3) la fin du débat sur le quantitatif et le qualitatif, dorénavant vus comme deux méthodologies utiles, voire complémentaires, et répondant à des besoins d'analyse différents ; 4) la reconnaissance de la nécessité d'adapter l'évaluation aux réalités culturelles multiples et plurielles qui caractérisent notre monde globalisant (Patton 2002, 1-2). Ray Pawson et Nick Tilley voient, quant

à eux, dans les quatre dernières décennies du XX^e siècle, s'affirmer quatre perspectives différentes : « *the experimental, the pragmatic, the naturalistic and the pluralist* ». ⁶ Dans le premier cas, l'accent est mis sur la recherche des causes observables et reconnaissables par une démarche expérimentale dans le cadre de laquelle un groupe est soumis à l'influence d'un ensemble de facteurs alors qu'un autre ne l'est pas. Pensons ici aux observations qui pourraient résulter d'une expérience au cours de laquelle un groupe de résidents d'un quartier donné bénéficierait de mesures de protection particulières par rapport à un autre groupe qui n'en recevrait aucune. Jugée d'application simple, la démarche n'en est pas moins vue par les experts comme comportant de sérieuses limites et comme étant carrément inappropriée lorsqu'il s'agit de porter un jugement sur les problématiques complexes qui se posent dans l'opérationnalisation des programmes sociaux, humanitaires ou éducatifs. Dans le second cas – la démarche « pragmatique » qu'on pourrait aussi bien qualifier de « réaliste » –, l'emphase est mise sur l'importance d'en arriver à des évaluations qui soient utiles et utilisables par ceux à qui elles s'adressent : qui tiennent compte du contexte organisationnel, de l'environnement « politique » des programmes soumis à l'évaluation et du degré de recevabilité des constats et des conclusions de l'étude évaluative. Ceux et celles qui placent la véracité, voire la vérité, au-dessus de tout, jettent généralement un regard critique sur ces approches pragmatiques et les jugent propres à amener les évaluateurs à des compromis inacceptables, voire à des compromissions. Troisième courant : l'approche naturaliste ou « constructiviste » qui place au centre de la réflexion et de la recherche évaluative les multiples perspectives, perceptions et « constructions » des différents acteurs impliqués et qui fait de l'évaluateur un « orchestrateur » ou un négociateur au service de participants ayant des points de vue différents plutôt qu'un spécialiste apte à juger lui-même d'une situation ⁷. « Thus [écrivent Pawson and Tilley] whilst the experimentalist might only have eyes for experimental and control subjects, and the pragmatists would look in the main towards policy makers, constructivism attempts to accommodate « stakeholders » by the dozen » (Pawson et Tilley 1997, 18). On aura compris que ces deux auteurs sont plutôt sceptiques face à cette approche par trop relativiste et inapte à saisir les traits objectifs d'une réalité qui existe en elle-même ou en dehors de la perception ou de la compréhension de ceux et celles qui la vivent. Quatrième et dernier courant : la démarche pluraliste ou « comprehensive », c'est-à-dire globale, qui retient des caractéristiques des trois précédentes et qui s'intéresse aux intentions initiales telles que définies en phase de conceptualisation du programme, à son implantation, à son opérationnalisation et à son utilité, y compris à ses résultats et à ses retombées. Malgré le « gros bon sens » qui semble la caractériser, cette démarche n'est pas non plus sans faiblesses et peut apparaître ambitieuse et idéaliste. ⁸

En somme, au cours des deux derniers siècles, les fondements théoriques de l'évaluation se sont enrichis et modifiés, les approches ou les modèles se sont multipliés et les méthodologies et les processus se sont raffinés. L'évaluation elle-même est devenue une discipline scientifique et une profession reposant toutes deux sur des assises conceptuelles aux contours définis. Toutes deux dotées de méthodes et procédés propres à soutenir les démarches analytiques de plus en plus complexes et adaptées aux cultures et aux caractéristiques des organisations de notre temps : plus imputables et plus rentables, tant devant leurs allocateurs de ressources et autres dirigeants que devant leurs utilisateurs – clients ou citoyens –; plus participatives et démocratiques

dans leur fonctionnement ; et davantage centrées sur la qualité des résultats, voire des retombées, en nos sociétés de consommation et de droit exigeantes quant à la qualité des produits et services offerts et hautement soucieuses d'affirmation et de protection des droits des individus. Peu à peu la discipline évaluative a gagné du terrain et ses théories et ses méthodes ont trouvé de multiples applications en éducation, dans le domaine social, au sein des grandes bureaucraties publiques, dans les entreprises d'affaires, dans le monde de la recherche scientifique, dans les multiples organisations – humanitaires, culturelles ou autres – de la société civile, dont les bibliothèques et les musées, où la pratique de l'évaluation de programme remonte respectivement au milieu et au début du XX^e siècle.

De la multiplicité des modèles ou des approches

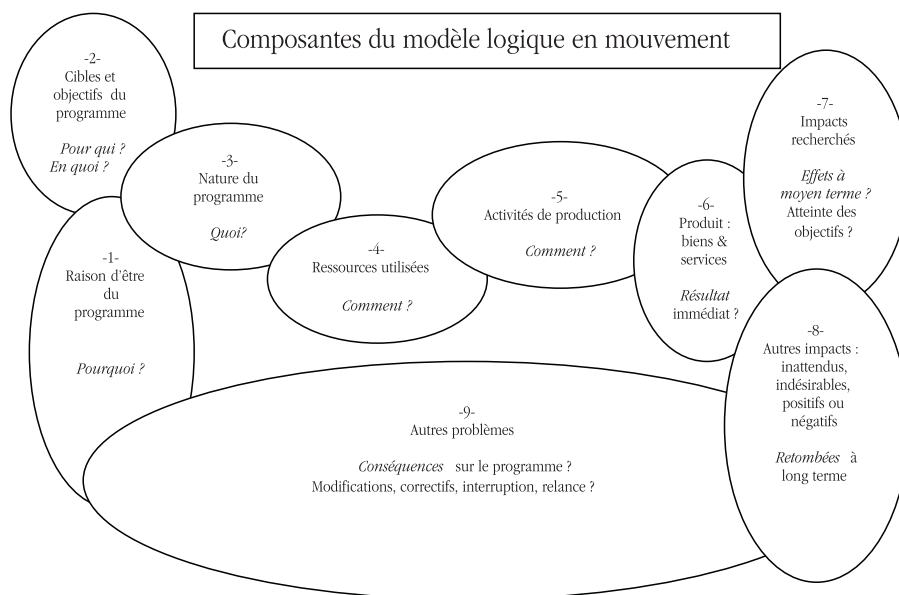
Quiconque suit, même de loin, ce qui se passe ces années-ci du côté de l'évaluation de programme, du moins dans les bureaucraties publiques canadiennes, aura noté l'usage de termes particulièrement révélateurs de l'importance qu'on accorde à la nécessité de se donner des instruments de plus en plus raffinés pour apprécier ou juger des résultats obtenus dans les programmes publics accessibles aux citoyens. On pense, pour ne citer que des exemples parmi les plus fréquemment notés, à : « gestion axée sur les résultats »⁹, « mesure de rendement », « Performance Management », « évaluation de programme », « théorie du programme », « modèle logique », « Balanced Scorecard », « Appreciative Inquiry » « Capacity Assessment ». Par ailleurs, la multiplicité – certains diront : la surabondance – des qualificatifs utilisés pour désigner les nombreux modèles évaluatifs est indicative d'un foisonnement¹⁰ qui ne facilite pas nécessairement la tâche lorsqu'il s'agit de choisir le modèle et les méthodes les plus appropriés à une situation donnée. D'autant plus qu'en matière d'évaluation tout comme en d'autres domaines, il est des modèles dominants, parfois proposés comme solutions universelles, au sujet desquels il est souvent difficile de se faire un jugement et d'apprécier en quoi ils sont susceptibles de répondre adéquatement aux besoins.

Ainsi en cours de recherche, j'ai été confronté à cette difficulté lorsque dans le cadre de discussions avec des spécialistes de l'évaluation j'ai été invité à utiliser le « modèle logique » jugé par des interlocuteurs, forts de multiples expériences évaluatives, comme étant l'un des modèles les plus performants parce qu'il prenait appui sur une solide compréhension théorique d'un programme donné et de sa logique interne et parce qu'il s'intéressait non seulement aux résultats immédiats mais également aux effets (impacts) à moyen et long termes des programmes évalués. Fallait-il dès lors immédiatement mettre de côté les autres modèles existants et s'en tenir à une application rigoureuse du « modèle logique » pour éventuellement mettre au point une méthodologie d'évaluation adaptée au milieu archivistique ? Comment par ailleurs justifier ce choix sans avoir vu les possibilités que pouvaient offrir d'autres modèles non moins bien documentés dans la littérature ? Il m'est apparu qu'avant d'arrêter un choix, je ne pouvais pas faire l'économie d'une étude, sinon approfondie, à tout le moins suffisamment éclairante pour me permettre de comprendre ces divers modèles, y compris leurs méthodes, d'en mesurer l'intérêt du point de vue de l'objectif de ma recherche et d'en tirer éventuellement le meilleur parti possible dans l'élaboration du ou des modèles, des méthodes et des instruments destinés à la communauté archivistique.

Le modèle logique

Qu'en est-il au juste de ce « modèle logique » (*Logic Model*) ? Essentiellement le modèle logique consiste en un instrument de représentation destiné à expliquer le flux de ce que nous pourrions appeler le système de production d'un programme donné. Il permet d'établir les liens de cause à effet et d'expliquer les articulations ou les rapports entre les diverses composantes de ce système, depuis les intrants jusqu'aux retombées à long terme (impacts), en passant par les activités, les processus et les résultats à court et moyen termes (Montagne 2000, 141). En cela, il apparaît comme un cadre d'analyse propre à faciliter une évaluation intégrée d'un programme. Coosky, Gill et Kelly parlent d'ailleurs d'un « integrative framework for analysis » qui porte à la fois sur les ressources, sur les activités, sur les résultats (outputs) et sur les retombées (outcomes) et proposent la définition suivante des modèles logiques : « [...] flow charts that display a sequence of logical steps in program implementation and achievement of desired outcomes. » (Coosky, Gill et Kelly s.d.) La littérature scientifique et administrative offre plusieurs représentations graphiques de la démarche. Toutes mettent en lumière les liens entre les différentes étapes du modèle qui, pour emprunter aux propos de David Julian (Julian 1997, 251), fournit un outil de conceptualisation des rapports entre les résultats d'un programme, ses effets sur ceux à qui il s'adresse et ses retombées à plus long terme et à plus large portée. Ou, autrement, ainsi que l'écrivent des chercheurs du RUSH Project : « A tool depicting the « chain of events » that link inputs and activities to outcomes ». ¹¹ Modèle qui, en plus de prendre appui sur des informations relatives à la description du programme, aux ressources investies, aux résultats escomptés et obtenus et aux impacts à moyen et long termes, inclut également une appréciation de la participation des acteurs ou des intervenants aux diverses étapes de conception, livraison et utilisation d'un programme ¹².

Figure 1 : Le modèle logique et ses composantes en interaction¹³



Un regard synthèse sur ces diverses composantes laisse en fait entrevoir quatre ensembles d'éléments reliés les uns aux autres et interagissant de manière à permettre de comprendre l'ensemble d'un programme à partir de plusieurs perspectives. Ce sont : 1) les *conditions* qui ont donné lieu à l'existence du programme – il peut s'agir soit d'une problématique particulière à laquelle on a voulu apporter une solution, soit d'une réponse à un besoin qu'une organisation entend satisfaire ; 2) la *nature* du programme – ou ce qu'on réalise et investit pour solutionner le problème ou satisfaire le besoin qui a amené la création du programme ; 3) les *conséquences* – la littérature parle ici d'impacts – de l'existence du programme exprimées en termes a) de *résultats immédiats* attendus ou obtenus par ceux et celles chargés de livrer les produits et services du programme, b) d'*effets à moyen terme* des produits et services sur la situation qu'on entendait corriger, sur le besoin qu'on entendait satisfaire ou sur les destinataires des interventions du programme et c) de *retombées à long terme* des interventions du programme¹⁴ ; et 4) d'*acteurs*, d'intervenants ou de participants qu'il s'agisse des concepteurs du programme, des allocateurs de ressources, des gestionnaires, des préposés au programme, des bénéficiaires du programme, voire de la population affectée par le programme – ensemble non explicitement illustré dans les premiers diagrammes du modèle logique mais occupant une place visible dans les représentations récentes¹⁵.

Deux caractéristiques principales du modèle semblent ressortir jusqu'ici : d'abord l'importance accordée au rapport de causalité entre les composantes du programme qui permet d'établir puis de juger des liens (*linkages*) qui les relient les uns aux autres ; puis, la hiérarchisation des résultats recherchés et/ou obtenus qui permet de raffiner l'analyse des rapports de causes à effets entre les objectifs visés par un programme et ce qui en résulte – hiérarchisation dans le temps, mais aussi dans la portée, étant entendu que les résultats immédiats atteignent des cibles plus étroites que ne le font les conséquences à long terme à portée généralement plus large dans le temps comme dans l'espace (Gilbert et al. 1995, 334-335 ; Rush et Osborne 1991, 95 ; Hirose et al. 1997, 252-253 ; Cooksie, Gill et Kelly s.d., 1-3).

Il est par ailleurs une autre caractéristique dont fait état la littérature et qui appelle quelques explications : le modèle logique est en effet associé à une démarche de théorisation du programme – les auteurs renvoient à des expressions comme : modèle « theory-based » ou modèle « theory-driven » ou plus simplement « program theory ». Cooksy et ses collaborateurs parlent de « program theory, in the form of a program logic model, as an integrative framework » et opposent les approches fondées sur une démarche de théorisation (*theory driven evaluations*) à celles axées davantage sur les dimensions méthodologiques de l'évaluation (*methods-driven evaluations*) (Cooksy, Gill et Kelly s.d., 2). Les chercheurs du RUSH Project présentent la démarche théorique comme une parmi quatre types de modèles logiques : « Theory Approach Logic Model », « Outcomes Approach Logic Model », « Activities Approach Logic Model » et « Research Performance Logic Models » (RUSH 2005, 15-18). Pour Carol Weiss,

« The root idea of TBE [theory-based evaluation] is that the beliefs and assumptions underlying an intervention can be expressed in terms of a phased sequence of causes and effects (i.e. a program theory). (...). The theory in question is the set of beliefs and assumptions that undergird program activities. (...). They are the hypotheses on which people, consciously or unconsciously, build their program plans and actions » (Weiss 1997, 501,503).

Ces propos, et plus particulièrement la première phrase de la citation qui précède, nous ramènent en quelque sorte à la première des deux caractéristiques du modèle logique notées plus haut. Ils font écho aussi à ceux de Chen qui, après avoir discuté de théorie descriptive – vraisemblablement la plus pratiquée par les évaluateurs – par opposition à la théorie prescriptive – qui met l’emphase sur ce qui doit être fait –, propose une définition de la notion de « program theory » qui reflète ces deux aspects et suggère un autre regard qui partage l’approche en deux sous-théories : « normative » et « causative », la première fournissant des indications sur ce que devrait être le programme et la seconde spécifiant comment fonctionne le programme et étant fondée « on empirical examination of the causal relationship between treatment and outcome » (Chen 1990, 8-10).

L’étude des rapports de cause à effet entre les composantes d’un programme¹⁶ et l’appréciation hiérarchisée des résultats atteints, deux démarches appuyées sur une solide compréhension de la logique, de la théorie du programme, constituent donc les deux piliers du modèle logique. Mais si ce dernier offre l’avantage d’être ouvert à l’usage de méthodes – multiples – les plus propres à garantir une évaluation crédible, il n’en présente pas moins certains désavantages. Ainsi, la compréhension et l’analyse de la logique d’un programme exigent du temps et des ressources, deux facteurs non négligeables particulièrement en situation de contrainte budgétaire. Par ailleurs, malgré la souplesse de la démarche elle-même, le danger subsiste que la construction théorique du programme soit utilisée de façon rigide : soit qu’il devienne difficile d’y ajouter d’autres éléments qui en compromettraient l’architecture ou l’équilibre ; soit – et c’est particulièrement vrai dans le cas où les acteurs ne sont pas tous mis à contribution dans l’élaboration du modèle théorique – qu’elle soit biaisée et jugée ne refléter qu’une dimension de la réalité¹⁷, soit enfin, qu’elle devienne l’unique étalon de mesure pour juger de la valeur du programme (Cooksy, Gill et Kelly s.d., 4). Pour Daniel Stufflebeam, l’évaluateur – à nos yeux, cela vaut aussi pour tous les participants à l’exercice évaluatif – est certes justifié de se donner une représentation théorique du fonctionnement d’un programme, de vérifier si les plans et opérations de celui-ci prennent appui sur des approches théoriques éprouvées et de reconnaître les variables clés d’un programme et les rapports d’interaction entre ces dernières. Cependant, créer une théorie du programme présente le risque majeur que celle-ci ne représente, au mieux, qu’une partie de la réalité ou, au pire, qu’elle soit tout-à-fait inappropriée, et même fautive. « Overall [conclut Stufflebeam], there really is not much to recommend theory-based program evaluation, since doing it right is usually not feasible and since failed or misrepresented attempts can be highly counterproductive. » (Stufflebeam 2001, 39) Sans rejeter le modèle logique, y compris sa démarche de théorisation, il est probablement prudent de garder ces réserves et mises en garde à l’esprit.

Autres modèles ou approches d’évaluation

Or, existe-t-il d’autres modèles qui, au premier regard, apparaissent susceptibles de fournir des assises conceptuelles et méthodologiques à la construction d’une démarche évaluative adaptée au milieu archivistique ? Le Tableau 1 fait état d’une cinquantaine d’approches présentées par des auteurs parmi les plus cités dans le domaine et dont la liste ne couvre pas tout le territoire évaluatif d’ailleurs. Ainsi, on ne semble pas y reconnaître de mention explicite de l’approche naturaliste (*Naturalistic Evaluation*)

Tableau 1 : Évaluation de programme : modèles ou approches

Glass & Ellett, 1980 (Glass et Ellet 1980, 213-217)	House, 1980 (House 1980, 21-43)	Stufflebeam & Webster, 1983 (Stufflebeam et Webster 1983, 23-40)	Stufflebeam, 2001 (Stufflebeam 2001, 13-80)
<ul style="list-style-type: none"> • Evaluation as Applied Science • Evaluation as System Management • Evaluation as Decision Theory • Evaluation as Assessment of Progress Towards Goals • Evaluation as Jurisprudence • Evaluation as Description or Portrayal • Evaluation as Rational Empiricism 	<ul style="list-style-type: none"> • The Systems Analysis Approach • The Behavioral Objectives (or Goal-Based) Approach • The Decision-Making Approach • The Goal-Free Approach • The Art Criticism Approach • The Professional Review (Accreditation) Approach • The Quasi-Legal (Adversary) Approach • The Case Study (or Transaction) Approach 	<p>Politically Oriented (or Pseudo-Evaluations) : Studies that promote a positive or negative view of an object irrespective of its worth</p> <ul style="list-style-type: none"> • Politically controlled Study • Public-Relations-inspired Studies <p>Questions-oriented (or Quasi-evaluation Studies) : Studies that address specified questions whose answers may or may not assess an object's worth</p> <ul style="list-style-type: none"> • Objective-based Studies • Accountability Studies • Experimental Research Studies • Testing Programs • Management Information Systems <p>Values-oriented Studies (or True Evaluation) : Studies that are designed primarily to assess some object's worth</p> <ul style="list-style-type: none"> • Accreditation / Certification Studies • Policy Studies • Decision-oriented Studies • Consumer-oriented Studies • Client-centered Studies • Connoisseur-based Studies 	<p>Pseudo-Evaluations</p> <ul style="list-style-type: none"> • Public Relations-Inspired Studies • Politically Controlled Studies <p>Questions-and-Methods-Oriented Evaluation Approaches (Quasi-Evaluation Studies)</p> <ul style="list-style-type: none"> • Objectives-Based Studies • Accountability, Particularly Payment by Results Studies • Objective Testing Program • Outcome Evaluation as Value-Added Assessment • Performance Testing • Experimental Studies • Management Information Systems • Benefit-Cost Analysis Approach • Clarification Hearing • Case Study Evaluations • Criticism and Connoisseurship • Program Theory-Based Evaluation • Mixed-Methods Studies <p>Improvement/Accountability-Oriented Evaluation Approaches</p> <ul style="list-style-type: none"> • Decision/Accountability-Oriented Studies • Consumer-Oriented Studies • Accreditation/Certification Approach <p>Social Agenda / Advocacy Approaches</p> <ul style="list-style-type: none"> • Client-Centered Studies or Responsive-Evaluation • Constructivist Evaluation • Deliberative Democratic Evaluation • Utilization-Focused Evaluation

qui met l'emphase sur la compréhension de la réalité à partir de l'expérience de ceux et celles qui la vivent (House 1980, 279-282; Guba et Lincoln 1983, 311-333) la collaboration active de ces derniers au processus évaluatif – participation ou collaboration qui renvoient à deux autres formes d'évaluation relativement récentes, la *Participatory Evaluation*¹⁸ et la *Collaborative Evaluation* (O'Sullivan 2004, 23-38) qu'on ne retrouve pas également dans les catégories énumérées au Tableau 1. Pas plus d'ailleurs qu'on y trouve de références précises à cette variété d'approches mentionnée par Patton dans sa présentation de l'« Appreciative evaluation » – qui met l'accent sur l'identification de ce qui fonctionne bien dans un programme ou une organisation et sur la construction de solutions à partir de ces forces (Coghlan, Preskill et Tzavaras Catsambas 2003, 5-7) – qui, à ses yeux :

« [...] offers another option for understanding learning-and-developmental diversity within evaluation, adding to the increasingly rich array of developmental diversity within evaluation that includes participatory evaluation, collaborative evaluation, empowerment evaluation, inclusive evaluation, developmental evaluation, multivocal evaluation, learning-oriented evaluation, democratic evaluation, feminist evaluation, and now Appreciative Inquiry evaluation or simply appreciative evaluation. » (Patton 2003, 88)

Approches dont plusieurs supposent une participation active de tous les acteurs d'un programme, depuis les concepteurs jusqu'aux bénéficiaires en passant par les agents d'implantation et d'exécution. Participation qui, selon la définition que donne Patton de la « Developmental Evaluation », engage même l'évaluateur dans une relation à long terme avec le client : « [...] in which the evaluator becomes the part of a design team helping to monitor what's happening, both processes and outcomes, in an evolving, rapidly changing environment of constant feedback and change. » (Patton 1994, 313, 317) Pour Robert Stake, tenant du modèle « Responsive Evaluation », malgré qu'il doive s'engager dans le processus d'évaluation sans cependant forcer les consensus au point d'en créer qui n'ont aucune prise sur le réel, l'évaluateur ne doit pas abdiquer sa responsabilité et son rôle d'expert : « To be responsive does not automatically yield design authority to stakeholders. It means coming to know the circumstances and problems and values well, then using professional talent and discipline to carry out the inquiry [which] belongs to the evaluator. » (Abma et Stake 2001, 8-9) D'autant plus, diront les tenants du modèle réaliste (*Realistic Evaluation*), que l'évaluateur ne peut s'en tenir à la surface des choses, qu'il doit, entre autres tâches, comprendre en quoi un programme donné est porteur de changement, démontrer la mécanique causale à la source du ou des problèmes observés, saisir les facteurs contextuels de cette problématique, expliquer le processus de production des effets (*outcomes*) obtenus et fournir aux intervenants impliqués des explications articulées et éprouvées quant au rapport entre le contexte et les mécanismes qui affectent le fonctionnement et les résultats du programme (Pawson et Tilley 1997, 214-220).

L'étude de chacun de ces modèles déborde largement le cadre et l'objectif du présent article. Au passage, cependant, on pourra noter que d'une catégorisation à l'autre – pour ne s'en tenir qu'aux données réunies au Tableau 1 –, il y a des similitudes, certaines plus explicites que d'autres cependant. Ainsi, on ne cherchera pas longtemps la différence entre l'« Accreditation / certification Studies » et l'« Accreditation / certification

Approach ». On aura également compris que la « Professional Review (Accreditation) Approach » nommée par House renvoie à la même démarche que les deux précédentes signalées par Stufflebeam et Webster. Malgré des dénominations différentes, il y a par ailleurs des ressemblances qu'on peut aisément apercevoir entre l'« Evaluation as Jurisprudence », « The Quasi-Legal (Adversary) Approach » et l'approche dite de « Clarification Hearing ». On devine aussi des parentés proches entre l'« Evaluation as Assessment of Progress Towards Goals », « The Behavioral Objectives (or Goal-Based) Approach » et les « Objectives-Based Studies » ; ou entre le « Connoisseur-based Studies » et les modèles « Art Criticism » et « Criticism and Connoisseurship ». Mais au-delà de la nature même de ces différentes approches, il y a plus important, leur portée ou mieux leur intérêt pour une recherche visant, parmi les objectifs exprimés plus haut, à construire un cadre méthodologique qui permette d'évaluer l'efficacité, l'efficacéité et l'impact des activités et des réalisations destinées aux clientèles internes et externes des services d'archives.

Regard critique et appréciatif

D'entrée de jeu, l'approche inspirée des relations publiques apparaît ici de moindre utilité. Qualifiée de « pseudo-évaluation », elle vise essentiellement à faire valoir le beau côté des choses auprès des auditoires auxquels elle est destinée ; elle ne tend pas à faire ressortir la véritable valeur d'un programme ou à mettre en lumière ce qui doit être fait pour l'améliorer. A vrai dire, on peut même s'interroger sur le caractère évaluatif de ce modèle qui s'apparente davantage à une démarche de communication – au sens le moins reluisant du terme d'ailleurs – et qui apporte fort peu lorsqu'il s'agit de déterminer si un programme fonctionne tel que prévu, s'il est pertinent ou si des ajustements sont requis pour en améliorer la performance. Il en va de même pour les études politiquement contrôlées. En effet, s'il est légitime pour un client de disposer à sa guise de l'information obtenue par le biais d'une évaluation – pourvu que ce soit fait en conformité avec la loi et dans le respect des engagements convenus au départ par les parties prenantes à l'évaluation –, il l'est certainement moins d'orienter, voire de diriger, une démarche évaluative de manière à ce qu'elle produise une vision tordue ou erronée de l'intérêt ou de la valeur d'un programme. Encore là, un exercice d'appréciation conçu et réalisé pour soutenir des conclusions – partielles et partiales – établies à l'avance peut difficilement tenir lieu d'évaluation de programme (Stufflebeam 2001, 13-16 ; Stufflebeam et Webster 1983, 25-26).

Par ailleurs, d'autres modèles, qui ont fait leurs preuves dans des domaines bien particuliers, semblent peu propres à soutenir une recherche du type de celle décrite plus haut. Je pense particulièrement ici aux évaluations faites dans le cadre d'examens destinés à vérifier, de manière objective ou non, les capacités et les performances d'individus. Ces formes d'évaluation, qu'il s'agisse, pour reprendre les appellations citées au Tableau 1, de « Testing Programs », d'« Objective Testing Programs » ou encore de « Performance Testing », sont largement utilisées dans les milieux éducatifs. Mais parce qu'elles ont pour objet les capacités de l'individu et qu'elles visent essentiellement à mesurer leurs aptitudes ou l'ampleur de leurs succès ou de leurs échecs, elles ne présentent aucun intérêt, du double point de vue de la portée et de la méthode, pour une recherche ayant pour objet des programmes dans leurs multiples dimensions et visant précisément à proposer des assises méthodologiques (Stufflebeam 2001, 21-22, 25-26 ;

Stufflebeam et Webster 1983, 29-30). Moins utiles également apparaissent les démarches de confrontation des points de vue en usage dans les tribunaux – « Jurisprudence as Evaluation », « Quasi-Legal (Adversory) Approach » et « Clarification Hearing » – qu'un auteur n'hésite d'ailleurs pas à qualifier d'« only marginally relevant to Program Evaluation ». L'approche, même si elle force à l'expression claire des facteurs à prendre en considération et si elle conduit nécessairement à une conclusion, à une décision ou à une action, demeure coûteuse. De plus, elle ne favorise pas nécessairement l'expression d'un large éventail de perspectives, et la formulation des points de vue y est soumise à la dynamique gagnant-perdant en plus d'être forcément partielle (Glass et Ellet 1980, 215; House 1980, 241-244; Stufflebeam 2001, 33). Quant aux modèles fondés sur l'expertise de spécialistes – « Art Criticism », « Connoisseur-based Studies » ou « Criticism and Connoisseurship », exploités notamment par les critiques d'art ou encore par les œnologues –, ils conduisent à des évaluations particulièrement dépendantes des qualifications, des connaissances et de la subjectivité de l'expert-évaluateur. Le recours au spécialiste ne comporte pas que des inconvénients et peut, au contraire, fournir à ceux et celles à qui sont destinés les résultats de l'évaluation des informations fort riches à partir desquelles des jugements éclairés pourront être posés (House 1980, 235-237; Stufflebeam 2001, 36-37; Stufflebeam 1983, 35). Mais dans un contexte d'évaluation de programmes sociaux, culturels ou humanitaires, où un large éventail de facteurs doit être pris en considération, il peut être difficile de recruter un expert disposant de toutes les connaissances requises; en outre, l'approche laisse peu de place à la participation de tous les intervenants au programme visé, ce qui, on le verra, peut compromettre l'utilisation et l'application des résultats de l'évaluation.

Les études évaluatives de type expérimental – « Experimental Research Studies » ou « Experimental Studies » auxquelles on peut associer, pour les fins de la présente discussion, l'« Evaluation as Applied Science » – s'appuient quant à elles sur une longue tradition et jouissent d'une haute crédibilité dans les milieux scientifiques. Centrées sur l'appréciation des résultats et fondées sur des méthodes éprouvées, elles permettent généralement d'établir la relation causale entre un traitement donné et les résultats obtenus. Dans le domaine des archives, elles sont déjà largement utilisées par les spécialistes de la préservation qui peuvent ainsi déterminer *a priori* ou *a fortiori* les effets d'une intervention de prévention, de conservation ou de restauration. Mais elles apparaissent d'application plus difficile dans les cas d'évaluation d'autres produits et services archivistiques. Elles fournissent indéniablement des informations précises sur un sujet donné. Elles sont cependant jugées offrir moins de possibilités lorsqu'il s'agit de se donner des instruments pour améliorer ou pour guider le développement d'un programme (Glass et Ellet 1980, 213; Stufflebeam et Webster 1983, 29) ou encore pour en apprécier la valeur intrinsèque (*merit*) ou la valeur extrinsèque (*worth*).¹⁹ Ces approches présentent donc, tant dans leur esprit que dans leur lettre, un faible intérêt du point de vue de la recherche en cours. En outre, à l'exception peut-être des études expérimentales appliquées aux activités de préservation, elles semblent peu appropriées à aider le milieu archivistique en général, et à évaluer l'efficacité, l'efficience ou la portée de ses interventions.

Par ailleurs, il est d'autres modèles qui, au premier regard, suscitent davantage d'intérêt pour le monde archivistique. Pensons, par exemple, à la démarche accréditive – « Accreditation / certification Approach » ou « Professional Review (Accreditation) »

Approach » – adoptée par le groupe de travail de la Society of American Archivists sur l'évaluation institutionnelle (*SAA Task Force on Institutional Evaluation*) dans le cadre de ses travaux sur l'évaluation des institutions d'archives qui ont conduit à l'adoption, en 1994, des *Guidelines for the Evaluation of Archival Institutions* (SAA 1994, 6-7). Alors que citoyens et consommateurs s'attendent à une plus grande imputabilité de la part de leurs gouvernements, des entreprises et des organisations de la société civile, ce type d'approche est de plus en plus utilisé. Parce qu'il prend appui sur l'analyse d'experts et spécialistes, il est jugé propre à permettre un jugement éclairé sur la qualité des institutions et des programmes visés. Mais son application ne va pas sans poser certains problèmes dont le fait 1) que la certification est basée sur la présomption que seuls les membres d'un champ d'expertise donné sont qualifiés pour juger d'une activité exercée par leurs pairs et 2) qu'elle met davantage l'accent sur les intrants et les processus que sur les résultats et les retombées (House 1980, 238-241 ; Stufflebeam et Webster 1983, 31-32 ; Stufflebeam 2001, 61-62). Autre exemple : les analyses des politiques – ou « Policy Studies » classées dans la catégorie des « Values-Oriented Studies » et qualifiées de « True-Evaluation ». Essentiellement comparatives et destinées à évaluer des politiques publiques – économiques, sociales, culturelles, humanitaires –, ces études s'attachent à mettre en lumière le rapport entre les coûts financiers et autres : les résultats obtenus permettent de mieux identifier et apprécier la politique la plus avantageuse (Stufflebeam et Webster 1983, 32). Enfin, on peut difficilement justifier que le modèle qui fait appel à une combinaison des méthodes quantitative et qualitative – « Mixed-Methods Studies » – soit sans intérêt. De telles études offrent l'insigne avantage d'être souples et ouvertes à l'exploitation des méthodes les plus aptes à fournir des réponses utiles et éclairantes aux questions évaluatives soulevées (Stufflebeam 2001, 41).

Cette possibilité de recours aux méthodes les plus appropriées par rapport à un sujet donné constitue un avantage marquant et implicitement souligné, par ailleurs, dans les Normes d'évaluation de programme adoptées par l'American Evaluation Association et supportées par la Société canadienne d'évaluation. Ces normes – regroupées en quatre catégories : utilité (*Utility Standards*), faisabilité (*Feasibility Standards*), appartenance (*Propriety Standards*) et exactitude (*Accuracy Standards*) (AEA 1994) – insistent, entre autres, sur la nécessité d'appuyer l'évaluation sur des informations justes et sur des processus rigoureux et adaptés aux besoins ou aux attentes, aspects que seules de multiples méthodes de cueillette, de validation et d'analyse de l'information provenant d'un large éventail de sources permettent de couvrir. Dans son étude synthèse d'une vingtaine de modèles évaluatifs parmi les plus utilisés, Daniel Stufflebeam note :

« Apart from pseudo-evaluations, there is among the approaches an increasingly balanced quest of rigor, relevance, and justice. Clearly, the approaches are showing a strong orientation to stakeholder involvement and the use of multiple methods. » (Stufflebeam 2001, 89)

Même l'approche classique d'évaluation en fonction de l'atteinte des objectifs – jugée généralement trop étroite dans son questionnement et peu propre à suggérer des pistes en vue de l'amélioration des programmes – prend appui sur plusieurs méthodes, contrairement, par exemple, à l'évaluation de type purement expérimental (House 1980, 228-229 ; Stufflebeam et Webster 1983, 27 ; Stufflebeam 2001, 17-18 ; Glass et Ellet 1980, 214-215).

Autre caractéristique d'une approche évaluative efficace : sa capacité à fournir des indications sur les améliorations possibles à apporter aux programmes étudiés. À cet égard, les modèles décisionnels et les approches centrées sur l'imputabilité – « Decision Making », « Evaluation as Decision Theory », « Decision Oriented Studies », « Decision Accountability Studies », « Accountability Studies » et « Accountability, Particularly Payment by Results Studies » – offrent d'intéressantes perspectives. On ne s'étonnera pas que ces types d'évaluation soient particulièrement bien vus dans les milieux politiques, chez les décideurs et chez les destinataires des programmes étudiés. Parce qu'ils contribuent à clarifier les attentes et parce qu'ils fournissent un outillage facilitant l'exercice de la responsabilité, ils gagnent même en popularité auprès des acteurs chargés de la production des biens et services résultant d'un programme. C'est qu'en plus de faire ressortir l'importance de l'utilité de l'information à collecter à des fins évaluatives, ces modèles tendent à mettre en lumière les résultats, les effets et les retombées des programmes, permettent de juger de leurs mérites et de leurs valeurs, favorisent la mise en place de processus d'évaluation systématique et continue et sont résolument tournés vers les décisions à prendre pour éventuellement améliorer les choses. En outre, ils supposent des analyses qui couvrent large : qui s'intéressent au contexte, aux intrants, aux processus et aux conséquences, qui requièrent un usage équilibré de méthodes quantitatives et qualitatives et une participation active de tous les intervenants à un programme donné. Pour peu qu'il n'y ait pas collusion entre évaluateurs et décideurs et que leur application se fasse sur une base de confiance et de franchise, ces modèles se présentent donc comme des sources d'inspiration obligées pour le milieu archivistique (House 1980, 230-232; Glass et Ellet 1980, 214; Stufflebeam et Webster 1983, 28,33; Stufflebeam 2001, 18-20, 58).

Par ailleurs, il est important que le modèle et les méthodes d'évaluation retenus permettent de couvrir sous tous ses aspects la problématique sous examen. À cet égard l'étude de cas – « Case Study (or Transaction) Approach », « Evaluation as Description or Portrayal » et « Case Study Evaluation » – offre d'intéressantes possibilités. Jugée parfois plus descriptive que proprement évaluative, l'approche n'en est pas moins largement utilisée; précisément parce qu'elle vise à une analyse approfondie de tous les aspects d'une question et qu'elle permet de porter un jugement complet et éclairé sur la réalité d'un programme, c'est-à-dire : appuyée sur une information riche, diversifiée, couvrant l'ensemble d'une problématique, fondée sur l'expression d'une multiplicité de points de vue formulés par tous les acteurs et compilée puis analysée à l'aide des méthodes qualitatives et quantitatives les plus appropriées (House 1980, 243-247; Glass et Ellet 1980, 215-216; Stufflebeam 2001, 34-36). Fondé sur une démarche empirique, c'est-à-dire sur une application de la raison et de l'observation de tous les aspects déterminants dans l'établissement de la valeur d'un programme (Glass et Ellet 1980, 216-217), le modèle tend vers une représentation la plus fidèle possible de la réalité et se rapproche en cela des approches naturaliste et « responsive » – sans cadre évaluatif prédéterminé donc (Scriven 1991, 240-243, 315-316). À cet égard, on notera que la *Méthode d'évaluation des pratiques de conservation préventive*, réalisée par le Centre interrégional de conservation du livre (Arles), – avec ses cent vingt-quatre indicateurs, ses grilles d'évaluation, ses fiches diagnostic et ses tableaux de synthèse (Ipert et al. 2002) – s'inscrit vraisemblablement dans cette catégorie d'approches descriptives.

L'étude de cas présente donc des avantages non négligeables, particulièrement si on la compare à des approches de portée plus étroite. Qu'on pense, par exemple, aux systèmes d'information de gestion – « Management Information Systems » – qui permettent de recueillir l'information nécessaire au gestionnaire pour diriger un programme, c'est-à-dire pour le planifier, le suivre et le contrôler et pour produire les rapports utiles sur son déroulement. Ces systèmes sont assurément d'une impérative nécessité et d'une indéniable utilité dans toutes organisations mais, règle générale, ils sont plus aptes à produire des informations sur les résultats immédiats obtenus que sur les effets ou sur les retombées à moyen et à long termes (Stufflebeam et Webster 1983, 30-31 ; Stufflebeam 2001, 29-30). Qu'on pense aussi aux analyses de coût-bénéfice – « Benefit-Cost Analysis Approach » – qui mettent en lumière le rapport entre un investissement dans un programme donné et ses retombées négatives ou positives. Encore là, les résultats qu'on en obtient aident à prendre les décisions appropriées pour corriger la démarche d'implantation d'un programme ou pour décider de son devenir. Mais parce qu'ils sont généralement exclusivement d'ordre financier, ces renseignements ne permettent pas d'entrevoir tous les aspects – contexte, intrants, processus et résultats – d'un programme et, ainsi, de procéder à une appréciation globale de l'efficacité et de l'efficacéité d'un programme (Stufflebeam 2001, 32). Qu'on pense enfin à l'évaluation des résultats d'un programme dans une perspective d'appréciation de sa valeur ajoutée – « Outcome Evaluation as Value-added Assessment ». Là encore, l'approche se révèle fort utile puisqu'elle conduit à la mise en place de systèmes d'information favorisant l'émergence d'instruments systématiques et normalisés pour suivre l'évolution d'un programme. On lui reproche cependant de faire peu de place aux informations qualitatives et d'accorder trop peu d'attention aux intrants et aux processus (Stufflebeam 2001, 21-24). Cela dit, il faut bien voir que, bien qu'elle tende vers une démarche holistique et vers des résultats qui traduisent davantage la réalité que l'application d'un modèle ou d'une grille d'analyse, l'étude de cas ne constitue pas une panacée car elle pose, entre autres, une difficulté majeure, inhérente en quelque sorte à sa nature. En effet, son application suppose qu'une attention particulière soit apportée à la nature, à la qualité et à la portée de l'information sur laquelle s'appuieront l'analyse et le jugement évaluatif. Autrement dit, elle suppose que les méthodes mises à contribution permettent de bien cerner tous les aspects de la réalité, y compris la nature des faits eux-mêmes et les perceptions et jugements des différents acteurs. À défaut d'y parvenir, le portrait de cette réalité sera incomplet et ne permettra pas de juger sûrement et équitablement des mérites et de la valeur d'un programme (House 1980, 243-247 ; Stufflebeam 2001, 34-36).

Autre trait dominant de l'évaluation telle qu'on la conçoit et telle qu'on la pratique ces années-ci : son ouverture à l'expression et à l'intégration des points de vue de tous les acteurs. Alors que nos organisations sont de moins en moins hiérarchiques, que le pouvoir de décision y est de plus en plus partagé, que les responsables de la réalisation des programmes sont impliqués dans leur conception aussi bien que dans leur implantation, que ces programmes, souvent mis en œuvre en partenariat, s'adressent à des groupes de plus en plus préoccupés de qualité et d'imputabilité, les modèles évaluatifs de type démocratique – « Deliberative Democratic Evaluation », « Constructivist Approach » et « Utilization-Focused Evaluation » – et les approches centrées sur les destinataires des programmes – « Client-Centered or Responsive Studies » et « Consumer-Oriented Studies » ont définitivement la cote. Bien qu'on les voit parfois

comme un idéal à atteindre, ces démarches évaluatives sont aussi jugées bénéfiques principalement parce qu'elles favorisent la participation active de toutes les parties intéressées à un programme donné. Cette participation peut d'ailleurs aller bien au delà de la seule expression de points de vue sur les résultats ou les effets d'un programme. Elle peut en effet comprendre l'implication des acteurs dans la formulation des questions évaluatives, dans le choix des méthodes, voire dans l'établissement des conclusions²⁰. Cet engagement des intervenants, toutes catégories confondues, permet par ailleurs de tirer le maximum de bénéfices de leurs connaissances et de leur expérience du sujet, favorise, par l'expression de multiples points de vue sur toutes les dimensions d'un programme²¹, une analyse complète de la réalité et facilite la réception des conclusions de l'évaluation. On conçoit alors que le rôle de l'évaluateur en est davantage un de guide ou de sherpa aidant les intervenants à produire leur propre évaluation, que celui d'un expert de l'extérieur livrant ses propres observations et apportant des réponses en forme de recommandations (Stufflebeam 2001, 70-80; Stufflebeam et Webster 1983, 34-35). Mais encore faut-il, pour que la démarche démocratique porte fruit, que les participants fassent confiance au spécialiste de l'évaluation – à ses connaissances et à son expertise – qu'ils évitent de se perdre dans des analyses détaillées du « comment », qu'ils gardent le cap sur les mérites et la valeur du programme et qu'ils conviennent à l'avance d'accepter les conclusions puis d'agir par la suite en conséquence. Mais c'est vraisemblablement là une autre histoire!

QUEL(S) MODÈLE(S) POUR LE MILIEU ARCHIVISTIQUE ?

Cela dit, peut-on, à la lumière du survol qui précède, déterminer le ou les modèles les plus susceptibles de rejoindre les préoccupations du milieu archivistique? D'entrée de jeu, il faut reconnaître que le choix d'une ou de plusieurs approches *ex abstracto* est en pratique impossible et inutile. Compte tenu de la diversité qui caractérise le milieu archivistique, tant dans la définition que dans l'aménagement des programmes d'archives, on peut raisonnablement penser que, là comme ailleurs, pour reprendre une expression populaire : « One size doesn't fit all ». Par ailleurs, la finalité de l'évaluation, les questions évaluatives et la nature des problématiques examinées sont déterminantes dans le choix de la démarche. Décider *à priori* d'un type de démarche, fut-elle fondée sur la mise à contribution de plusieurs modèles, sans avoir une vision claire des trois facteurs précités, présente un risque élevé d'échec : au mieux, on en tirera des renseignements à peu près inutilisables dont on ne saura que faire; au pire, on se retrouvera avec des informations inadéquates sur lesquelles on basera des décisions et des stratégies de redressement forcément inappropriées. Dans les circonstances, peut-être vaut-il mieux chercher du côté des caractéristiques de l'approche souhaitable et tenter ainsi d'en établir le profil.

Dans un rapport soumis en 2002 à la Société canadienne d'évaluation, les analystes de Zorzi & Associates reviennent sur la définition de l'évaluation de programme et, tout en reconnaissant qu'il n'en existe pas vraiment qui soit d'application universelle, suggèrent que le milieu des évaluateurs semble s'entendre sur les caractéristiques de base suivantes :

« Evaluation applies research design principles to answer practical questions about programs. Data is collected and processed systematically to provide evidence about what is happening in a program (processes and outcomes), why it is happening, and

how the program can be improved. This evidence can be used to make judgements about the program's merit or worth.

Evaluation is cross-disciplinary and draws methods from many different fields of study. » (Zorci, MCGuire et Perrin 2002)

Ces caractéristiques, dégagées à partir d'une large consultation auprès de la communauté des évaluateurs rejoignent, on ne s'en étonnera pas, les normes d'évaluation de programme évoquées plus haut²². Ainsi, il y a un rapport évident entre, d'une part, les normes d'utilité (*Utility Standards*), qui marquent la nécessité pour toute évaluation de satisfaire au besoin informationnel des destinataires, les normes d'exactitude (*Accuracy Standards*), qui garantissent que l'évaluation produira des informations adéquates sur les facteurs déterminant les mérites et la valeur d'un programme, et, d'autre part, l'énoncé selon lequel les données doivent être recueillies et traitées systématiquement de manière à fournir toutes les évidences utiles sur tous les aspects d'un programme et d'ainsi permettre d'apprécier les mérites et la valeur d'un programme. Par ailleurs, ces normes parlent aussi de faisabilité (*Feasibility Standards*), y compris de réalisme, de pragmatisme, de diplomatie et de frugalité dans l'exercice d'évaluation. Elles évoquent également l'importance de l'appartenance (*Propriety Standards*) et la nécessité d'une intervention évaluative qui soit juste, qui soit aussi respectueuse des personnes, de leurs attentes et de leurs droits et qui tienne compte des effets de l'évaluation sur les destinataires (AEA 1994). S'appuyant sur ces normes pour évaluer la vingtaine de modèles qu'il avait étudiés, Daniel Stufflebeam en a primé neuf qu'il n'hésite pas à qualifier de « Best Approaches for 21st Century Evaluation ». Son pointage donne le classement suivant : *Decision/Accountability* : 92 ; *Utilization-Focused* : 87 ; *Client-Centered/Responsive* : 87 ; *Deliberative Democratic* : 83 ; *Consumer-Orientation* : 81 ; *Constructivist* : 80 ; *Case Study* : 80 ; *Outcome Monitoring / Value-Added* : 72 ; *Accreditation* : 60 (Stufflebeam 2001, 90). Or, toutes ces approches ont en commun – à des degrés divers, s'entend – les caractéristiques suivantes : elles sont utilisées à des fins d'amélioration des programmes ; elles contribuent à la reddition de compte ; elles capitalisent sur la participation active et approfondie de tous les intervenants ; elles gardent le cap sur l'utilisation des résultats de l'évaluation ; elles tendent vers la production de renseignements adéquats et crédibles ; et, pour la plupart, elles s'appuient sur des méthodologies plurielles.

Qu'est-ce à dire alors ? Essentiellement ceci : qu'il n'y pas qu'un seul modèle qui soit applicable et que le choix d'une approche et des méthodes de cueillette et d'analyse de l'information dépend d'abord et avant tout de l'objet de l'évaluation – l'« évaluand » –, soit de la question évaluative elle-même. Ainsi, pour s'en tenir à des exemples simples mais évocateurs de mon propos, s'il s'agit de vérifier l'efficacité d'une intervention de protection d'une collection de photos anciennes fondée sur l'usage de produits chimiques particuliers, l'approche expérimentale et les méthodes comparatives afférentes peuvent fort bien convenir.

Par contre, si l'intention est plutôt de voir en quoi un programme de formation sur l'utilisation d'un nouveau logiciel d'indexation améliore effectivement la performance des usagers d'une organisation, il faudra probablement faire appel à des méthodes diversifiées. Notamment, il faudra considérer les intentions initiales et procéder à une évaluation en fonction des objectifs ; juger de la qualité des résultats, voire des effets et des retombées, en termes de valeur ajoutée. Il faudra aussi voir en quoi le nouveau

produit répond aux attentes des utilisateurs et adopter une démarche évaluative centrée sur le client.

Si le survol des diverses approches ne permet pas de déterminer LE modèle et LES méthodes applicables au milieu des archives, il met néanmoins en lumière les qualités premières à rechercher lorsqu'il s'agit de définir une démarche évaluative. Ainsi, le modèle et les méthodes applicables doivent :

- permettre de faire le tour de la question évaluative posée, de voir l'objet de l'évaluation dans toutes ses perspectives, sous tous ses angles ... ou, autrement, de couvrir la totalité du sujet ;
- offrir la possibilité de recourir à toutes les méthodes et techniques – qualitatives et (ou) quantitatives – utiles et nécessaires à la compréhension, à l'analyse et à l'interprétation des faits ;
- prendre appui sur une connaissance adéquate et réaliste, c'est-à-dire éloignée de toute application de grilles d'analyse prédéterminées, des intrants, des processus, des résultats, des effets, voire des retombées de l'objet soumis à l'évaluation, de manière à produire une représentation théorique fidèle d'un programme ou d'un de ses éléments ;
- s'appuyer également sur des données significatives par rapport à l'objet évalué et éviter la collecte aveugle de données inutilisables, de manière à ce que les informations qu'on en tirera soient vérifiables et crédibles et à ce que le jugement qui découlera de l'analyse et de l'interprétation soit valide et puisse résister à la critique ;
- impliquer, sans tomber dans les travers de la démocratie directe et sans dépouiller l'évaluateur de son rôle expert, une participation active de tous les intervenants : ceux qui déterminent les orientations d'un programme, ceux qui en assurent la réalisation et ceux enfin à qui il s'adresse de façon 1) à pouvoir tirer le maximum de leur compréhension, de leur connaissance et de leur expertise au moment de l'analyse et de l'interprétation des faits et 2) à faciliter la réception des conclusions par les milieux visés ;
- tout en tenant compte de l'utilisation qu'on entend faire des résultats de l'évaluation, permettre de dégager des conclusions qui serviront à améliorer les choses non seulement du double point de vue de l'efficacité et de l'efficacé mais aussi de celui de la pertinence ;
- être abordables aux plans de l'effort à consentir et des ressources à investir, c'est-à-dire réalisables de manière à ce que les coûts soient jugés raisonnables à l'aune des bénéfices qu'on en tire.

En revenant au point de départ, on aura noté que le modèle logique, largement utilisé par les praticiens de l'évaluation, comporte plusieurs de ces caractéristiques. En ce sens, il offre d'intéressantes possibilités et devrait se révéler plus qu'utile dans l'évaluation d'activités, de produits ou de services archivistiques. Mais, il n'est pas le seul et qui plus est, il affiche certaines lacunes. D'une part, il suppose une démarche de théorisation qu'il faut mener avec rigueur en évitant de tomber dans le piège des grilles préconçues. D'autre part, on le voit souvent comme étant davantage centré sur les préoccupations des décideurs et gestionnaires et comme faisant peu de place aux points de vue et aux intérêts des autres intervenants. Mais, on conviendra que ce sont

là des aspects qui peuvent être corrigés pour peu qu'on s'assure que les représentations théoriques s'appuient sur une réflexion empirique, voire naturaliste et constructiviste, et qu'on mette à contribution les points de vue de tous les intervenants impliqués. Cependant, l'évaluation faite à partir d'une étude de cas ou d'une étude de la valeur ajoutée et les démarches de type participatif offrent aussi des possibilités qu'en milieu archivistique nous aurions tort de ne pas exploiter.

CONCLUSION

Le survol qui précède visait à réunir des informations sur les modèles d'évaluation de programme pratiqués ces années-ci et à les examiner sous l'angle de l'intérêt qu'ils présentent pour l'élaboration d'approches applicables au milieu archivistique. Il en ressort que l'approche unique et idéale n'existe pas et qu'il faut, au contraire, s'en remettre à des démarches souples et ouvertes aptes à apporter réponses aux questions évaluatives posées par le milieu concerné. Il en ressort de fait que, quels que soient le modèle, les méthodes et les techniques choisies, ils se doivent d'être holistiques, pluriels, réalistes, démocratiques, abordables et fondés sur des données et des informations sûres; ils se doivent aussi d'aller au delà du regard appréciatif à court terme et de la seule critique du « comment » et de conduire à l'amélioration de la réalité.

Par ailleurs, le choix d'un modèle ne saurait être établi *ex nihilo*. La décision d'adopter un modèle plutôt qu'un autre doit prendre appui sur l'objet même de l'évaluation et sur l'usage qu'on veut en faire : les choix quant à l'approche, quant aux méthodes et quant à la nature de l'information à recueillir en dépendent. La démarche retenue doit aussi prendre appui sur une solide compréhension du programme visé, ce qui ne va pas nécessairement de soi pour une communauté professionnelle habituée à voir sa réalité – pratique et discipline – en termes de fonctions et activités et qui vraisemblablement est beaucoup moins familière par rapport à la notion de programme. Enfin, et par-dessus tout, les archivistes doivent vouloir se questionner non seulement sur la qualité de leurs activités et des résultats en découlant, mais aussi sur leurs effets et leurs retombées. Est-ce que le programme visé fonctionne? Est-ce qu'il est porteur de changement? Améliore-t-il les choses? Qu'advierait-il si le programme n'existait pas? Voilà autant de questions, parmi d'autres, auxquelles l'évaluation de programme, dans le domaine des archives comme ailleurs, peut et doit apporter réponse.

Prenant en compte ce qui précède et le soin qu'il faut mettre pour choisir l'approche évaluative, on ne s'étonnera pas que récemment encore la communauté archivistique canadienne ait vivement réagi à l'initiative du Conseil canadien des archives qui désirait évaluer ses programmes d'aide vieux d'une vingtaine d'années. On en retiendra probablement qu'il sera dorénavant important d'associer les intervenants de plus près à la formulation des questions évaluatives aussi bien qu'au choix du modèle et des méthodes. On en tirera probablement d'autres leçons dont la suivante : qu'il n'y a pas lieu de laisser aller des programmes sans les évaluer régulièrement, soit pour les suivre et ajuster le tir au fur et à mesure, soit pour se donner toute l'information adéquate pour permettre aux acteurs concernés de juger en tout temps de leur pertinence, et ce, en toute connaissance de cause.

Jacques Grimard Professeur agrégé. École de bibliothéconomie et des sciences de l'information. Université de Montréal

NOTES

1. Pour une version en ligne de l'instrument d'évaluation RMCAS, voir sur le site web des Archives nationales britanniques <http://www.nationalarchives.gov.uk/RMCAS/default.htm>. Page consultée le 3 juin 2005.
2. Voir entre autres : Évaluation des capacités de la gestion de l'information. Outil et méthodologie. Juillet 2003. Information Management Capacity Check. Tool and Methodology. July 2003, http://www.collectionscanada.ca/information-management/0603/06030105_e.html, page consultée le 6 juin 2005. Je remercie Andrew Lipchak de la firme Infotegrity Consulting qui a mis à ma disposition les documents utilisés lors de sa présentation : Evaluating Records and Information Management devant le chapitre torontois de l'ARMA, le 4 mai 2005. L'auteur y signale au moins une douzaine d'autres outils semblables à celui diffusé par Bibliothèque et Archives Canada.
3. Les études « formatives » mettent l'accent sur l'appréciation d'un programme en cours d'implantation et secondent et les travaux « sommatifs » intéressent davantage les résultats et les retombées des activités. Voir : Bingham, R.D., Felbinger, C.L. 2002. Evaluation in Practice. A Methodological Approach, 2nd Ed, New York, London, Chatham House Publishers / Seven Bridges Press, LLC, p. 4-7.
4. La périodisation qui suit est empruntée à : Madaus, G.F., Stufflebeam, D.L. et Scriven, M.S. 1983. *Program Evaluation : A Historical Overview*. In Madaus, G.F., Scriven, M.S., and Stufflebeam, D.L., *Evaluation Models. Viewpoints on Educational and Human Services Evaluation*, Boston, The Hague, Dordrecht, Lancaster, Kluwer-Nijhoff Publishing, p. 3-21.
5. Nos auteurs qualifient en effet la décennie précédant la parution de leur ouvrage de *Age of Professionalization*. Madaus, G.F., Stufflebeam, D.L., and Scriven, M.S. 1983, *Program Evaluation : A Historical Overview*. In Madaus, G.F., Scriven, M.S., and Stufflebeam, D.L., *Evaluation Models. Viewpoints on Educational and Human Services Evaluation*, Boston, The Hague, Dordrecht, Lancaster, Kluwer-Nijhoff Publishing, p. 16.
6. En italique dans le texte. Pawson, R., Tilley, N. 1997. *Realistic Evaluation*. London, Thousand Oaks, New Delhi, Sage Publications, p. 4.
7. Selon Pawson et Tilley, *The engine of the method is thus an exchange of meaning between the researcher and all program participants*. Pawson, R., Tilley, N. 1997. *Realistic Evaluation*, London, Thousand Oaks, New Delhi, Sage Publications, p. 18.
8. On lira avec intérêt les observations critiques – teintées d'un humour de bon aloi – de chacun de ces courants formulées par : Pawson, R., Tilley, N. (1997) : *A history of Evaluation in 28½ pages*. In *Realistic Evaluation*, London, Thousand Oaks, New Delhi, Sage Publications, p. 1-29.
9. L'expression a d'ailleurs tenu lieu de thématique au 9^e symposium annuel du Forum sur le rendement et la planification, tenu à Ottawa les 25 et 26 mai 2005 et portant sur : La gestion axée sur les résultats – Tirer des leçons de l'expérience des autres. Le grand échange.
10. Michael Quinn Patton parle d'une centaine d'approches évaluatives possibles : [Patton, M.Q.]. 2002. *A Conversation with Michael Quinn Patton*. Interview by Julia Coffman, in Section « Questions and Answers » of The Evaluation Exchange Newsletter. Family Support Issue, 8, 1 : 2. <http://www.gse.harvard.edu/hfrp/eval/issue18/qanda.html>. Page consultée le 13 juin 2005.
11. Je remercie M. Mark Linder, agent de communication au Research Utilization Support and Help Project d'avoir mis à ma disposition une version- non-définitive - augmentée et remaniée d'un texte connu jusqu'ici sous le titre de : Anatomy of a Logic Model et préparé par des chercheurs rattachés aux : Research Utilization Support and Help (RUSH) Project, National Center for the Dissemination of Sensibility Research (NCDDR) (Mars 2005), Effective Outcomes Planning - DRAFT, Austin (TX), Southwest Educational Development Laboratory (SEDL), p. 28.
12. Steve Montague rappelle en effet que : « Recent iterations of the logic model have included the concept of “reach”, the users and stakeholders involved in the delivery of a program, service, or initiative ». Voir

- Montague, S. (2000). *Focusing on inputs, outputs, and outcomes : are international approaches to performance management really so different ?* In The Canadian Journal of Program Evaluation, 15, 1 : 142.
13. Représentation largement inspirée des représentations sur les objets du programme et sur les questions d'évaluation proposées dans Évaluation des programmes publics. ENAP, programme 465. Québec, École nationale d'administration publique, p. 4, 13 et 16.
 14. Voir entre autres : Gilbert, R.A., Julian, D.A., Jones, A, Deyo, D. 1995. *Open systems evaluation and the logic model : program planning and evaluation tool*. In Evaluation and Program Planning, 18, 4 : 335, 336 ; Rush, B., Ogborne, A. 1991. *Program Logic Models : Expanding Their Role and Structure for Program Planning and Evaluation*. In Canadian Journal of Program Evaluation, 6, 2 : 96-100 ; ÉNAP, Évaluation des programmes publics. ÉNAP, programme 465, p. 4.
 15. Voir à ce sujet et parmi d'autres : Patton, M.Q. 1997. *Utilization-Focused Evaluation*. The New Century Text, 3rd ed., Thousand Oaks, London, New Delhi, Sage Publications, p. 235-236.
 16. Pour Gilbert et ses collaborateurs, cet aspect compte parmi les points forts de l'approche : *The logic model provides the opportunity to consider linkages between problem/conditions, activities, outcomes and impacts. This is one of the major strengths of the logic model as a planning and evaluation tool*. Gilbert, R.A., Julian, D.A., Jones, A, Deyo, D. 1995. *Open systems evaluation and the logic model : program planning and evaluation tool*. In Evaluation and Program Planning, 18, 4 : 335.
 17. D'autant plus qu'en matière de construction théorique, les consensus ne sont pas faciles à atteindre. E. Herose et ses collègues écrivent à ce sujet : « The development of system impacts is dependent on the ability to achieve consensus regarding a few critical community issue ». Hirose, E., Taneda, Y., Ishii, T., Julian, D.A. (1997), « The utilization of the logic model as a system level planning and evaluation device », in Evaluation and Program Planning, 20, 3 : 256.
 18. « Participatory evaluation is people-centered : project stakeholders and beneficiaries are the key actors of the evaluation process and not the mere object of the evaluation ». United Nations Development Program, Evaluation Office, Participatory Evaluation, p. 1. Consulté le 27 juin 2005.
 19. « Merit refers to the intrinsic value of a program, for example, how effective it is in meeting the needs of those it is intended to help. Worth refers to extrinsic value to those outside the program, for example, to the larger community or society » Patton, M.Q. 1997. *Utilization-Focused Evaluation*. The New Century Text, 3rd éd., Thousand Oaks, London, New Delhi, Sage Publications, p. 65.
 20. Elle est en cela éloignée de l'approche d'analyse de système – Systems Analysis Approach et Evaluation as Systems Management – particulièrement critiquée parce qu'elle tend à privilégier le point de vue des gestionnaires du programme et de leurs associés et à exclure celui des autres acteurs, parce qu'elle ne se détache pas de la logique systémique du programme et parce qu'elle s'intéresse davantage à l'efficacité qu'au mérite et à la valeur du programme. Voir : Glass, G.V., Ellett, Jr., F.S. 1980. *Evaluation Research*. In *Annual Review of Psychology*, 31 : 213-214. House, E.R. 1980. *Evaluating with Validity*, London, Beverly Hills, Sage Publications, p. 226-227.
 21. Encore que ce ne soit pas nécessairement le cas pour les évaluations mettant en lumière le point de vue des consommateurs qu'on qualifie également de « Goal-Free ». Ce type d'évaluation, particulièrement prisée chez les groupes de consommateurs, fait en effet abstraction des objectifs du programme et n'en prend en compte que les effets. House, E.R. 1980. *Evaluating with Validity*, London, Beverly Hills, Sage Publications, p. 232-234. Stufflebeam, D.L. 2001. *Evaluation Models*. In Henry, G.T, Greene J.C. (ed), *New Directions for Evaluation*, no. 89, p. 59-60. Stufflebeam, D.L., Webster, W.J. 1983. *An Analysis of Alternative Approaches to Evaluation*. In Madaus, G.F., Scriven, M.S., Stufflebeam, D.L., *Evaluation Models. Viewpoints of Educational and Human Services Evaluation*, Boston, The Hague, Dordrecht, Lancaster, Kluwer-Nijhoff Publishing, p. 34.
 22. Voir p. 91.

BIBLIOGRAPHIE

- ABMA, T.A., STAKE, R.E. 2001. Stake's Responsive Evaluation : Core Ideas and Evolution. In *Responsive Evaluation*, publié sous la dir. de J.C. Greene et T.A. Abma. San Francisco, Jossey-Bass, Collection New Directions in Evaluation, no 92.
- AMERICAN EVALUATION ASSOCIATION. 1994. *The Program Evaluation Standards. Summary of the Standards.* [En ligne]. <http://www.eval.org/EvaluationDocuments/progeval.html> (Page consultée le 14 juillet 2005).
- CHEN, H. 1990. Issues in Constructing Program Theory. *New Directions for Program Evaluation* 47 : 7-18.
- COGHLAN, A.T., PRESKILL, H. and T. TZAVARAS CATSAMBAS, T. 2003. An Overview of Appreciative Inquiry. In *Using Appreciative Inquiry in Evaluation*, publié sous la dir. de H. Preskill et A.T. Coghlan. San Francisco, Jossey-Bass, Collection New Directions for Evaluation, no 100.
- COOKSIE, L.J., GILL, P. et P.A. KELLY. (s.d.). *The Program Logic Model as an Integrative Framework for a Multimethod Evaluation.* [En ligne]. <http://www.hsrh.houston.med.va.gov/AdamKelly/Logic.html> (Page consultée le 15 juin 2005).
- FETTERMAN, D.M. 1996. Empowerment Evaluation : An Introduction to Theory and Practice. In *Empowerment Evaluation : Knowledge and Tools for Self-Assessment and Accountability*, Thousand Oaks, Fetterman, D.M., Kaftarian, S.J., and Wandersman, A. (eds), Sage.
- GILBERT, R.A., JULIAN, D.A., JONES et A, DEYO, D. 1995. Open systems evaluation and the logic model : program planning and evaluation tool. *Evaluation and Program Planning* 18, 4 : 333-341.
- GLASS, G.V. et F.S. ELLETT, Jr. 1980. Evaluation Research. *Annual Review of Psychology* 31 : 211-228.
- GRIMARD, Jacques et Lucie PAGÉ. 2004. Towards Program Evaluation in Archives. *Archival Science* 4, 1-2 : 99-126.
- GRIMARD, Jacques. 2004. Program Evaluation and Archives : « Appraising » Archival Work and Achievements *Archivaria* 57 : 69-87.
- GUBA, E.G. et Y.S. LINCOLN. 1983. Epistemological and Methodological Bases of Naturalistic Inquiry. In *Evaluation Models. Viewpoints on Educational and Human Services Evaluation*, publié sous la dir. de Madaus, G.F., Scriven, M.S., Stufflebeam, D.L. Boston, The Hague, Dordrecht, Lancaster, Kluwer-Nijhoff Publishing.
- HIROSE, E., TANEDA, Y., ISHII, T. et D.A. JULIAN. 1997. The Utilization of the logic model as a system level planning and evaluation device. *Evaluation and Program Planning* 20, 3 : 251-257.
- HOUSE, E.R. 1980. Naturalistic Evaluation. *Evaluating with Validity*. Beverly Hills, London, Sage publications.

- IPERT, Stéphane, LE GUEN, Gilbert, MÉRIC, Laure et Benoît de TAPOL. 2002. *Une méthode d'évaluation des pratiques de conservation préventive dans un service d'archives*. Paris, Archives de France.
- JULIAN, D.A. 1997. The Utilization of the Logic Model as a System Level Planning and Evaluation Device. *Evaluation and Program Planning* 20, 3 : 251-257.
- MADAUS, G.F., STUFFLEBEAM, D.L. et M.S. SCRIVEN. 1983. Program Evaluation : A Historical Overview. In *Evaluation Models. Viewpoints on Educational and Human Services Evaluation*, Madaus, G.F., Scriven, M.S., and Stufflebeam, D.L. Boston, The Hague, Dordrecht, Lancaster, Kluwer-Nijhoff Publishing.
- MONTAGUE, S. 2000. Focusing on inputs, outputs, and outcomes : are international approaches to performance management really so different? *The Canadian Journal of Program Evaluation* 15, 1 : 139-148.
- O'SULLIVAN, R.G. 2004. *Practicing Evaluation : A Collaborative Approach*, Thousand Oaks, Sage.
- PATTON, M. Q. 1994. Developmental Evaluation. *Evaluation Practice* 15, 3 : 311-320.
- PATTON, M. Q. 2002. A Conversation with Michael Quinn Patton. Interview by Julia Coffman. In Section « Questions & Answers » of *The Evaluation Exchange Newsletter. Family Support Issue* 8, 1 : 1-2. [En ligne]. <http://www.gse.harvard.edu/hfrp/eval/issue18/qanda.html> (Page consultée le 13 juin 2005).
- PATTON, M.Q. 2003. Inquiry into Appreciative Evaluation. In *Using Appreciative Inquiry in Evaluation*. Preskill, H., Coghlan, A.T., San Francisco, Jossey-Bass, Coll. « New Directions for Evaluation », no 100.
- PAWSON, R. et N. TILLEY. 1997. The New Rules of Realistic Evaluation. In *Realistic Evaluation*. London, Thousand Oaks, New Delhi, Sage Publications.
- RESEARCH UTILIZATION SUPPORT AND HELP (RUSH) Project, National Center for the Dissemination of Sustainability Research (NCDDR). Mars 2005. *Effective Outcomes Planning – DRAFT*, Austin (TX), Southwest Educational Development Laboratory (SEDL).
- RUSH, B. et A. OGBORNE. 1991. Program Logic Models : Expanding Their Role and Structure for Program Planning and Evaluation. *Canadian Journal of Program Evaluation* 6, 2 : 95-106.
- SCRIVEN, M. 1991. *Evaluation Thesaurus*. 4^e éd., Newbury Park, London, New Delhi, Sage Publications.
- SOCIETY OF AMERICAN ARCHIVISTS (SAA). 1994. Guidelines for the Evaluation of Archival Institutions. *Archival Outlook* (Mars 1994) : 6-7.
- STUFFLEBEAM, D.L. et W.J. WEBSTER. 1983. An Analysis of Alternative Approaches to Evaluation. In *Evaluation Models. Viewpoints of Educational and Human Services Evaluation*, Madaus, G.F., Scriven, M.S., Stufflebeam, D.L. Boston, The Hague, Dordrecht, Lancaster, Kluwer-Nijhoff Publishing.
- STUFFLEBEAM, D.L. 2001. Evaluation Models. In Henry, G.T, Greene J.C. (ed), *New Directions for Evaluation* 89 : 8-98.

- WEISS, C.H. 1997. How Can Theory-Based Evaluation make a Greater Headway?
Evaluation Review 21, 4 : 501-524.
- ZORZI, R., MCGUIRE, M. et B. PERRIN. 2002. *Canadian Evaluation Society Project
in Support of Advocacy and Professional Development. Executive Summar.*
Non-published Report, Toronto, Zorzi & Associates, p. ii.