

# L'évaluation des programmes de main-d'oeuvre : Observations méthodologiques

## Evaluation of Manpower Programs : Some Methodological Notes

Jean Sexton

Volume 28, Number 3, 1973

URI: <https://id.erudit.org/iderudit/028421ar>

DOI: <https://doi.org/10.7202/028421ar>

[See table of contents](#)

Publisher(s)

Département des relations industrielles de l'Université Laval

ISSN

0034-379X (print)

1703-8138 (digital)

[Explore this journal](#)

Cite this article

Sexton, J. (1973). L'évaluation des programmes de main-d'oeuvre : Observations méthodologiques. *Relations industrielles / Industrial Relations*, 28(3), 583–601. <https://doi.org/10.7202/028421ar>

### Article abstract

The purpose of this article is to present a survey of the literature on the evaluation of manpower programs in French. To do so, it has been chosen to successively present the origin, the nature, the methodology and the limits of such evaluative efforts.

Dating back to 1897, the evaluation preoccupation has known a renewed popularity since the mid 1960's in the manpower field. The movement has originated in the U.S.A. and has been and still is spreading in other North American jurisdictions and in certain European countries.

Public dissatisfaction, limited public funds, competing needs, changes in the nature of social problems, an increased professionalization of the public sector, a more educated and more sophisticated public, and many obvious failures are some of the reasons that have led to this increased popularity of evaluating research. The purpose of evaluation is then to supply information leading to a better decision making.

Fundamentally, evaluation is a subjective and complex process. In fact evaluation is an appraisal of the extent to which a program has met its objectives, and why such a desired result was or was not fully achieved. The extent to which scientific methods are used in making that judgment will reduce but not eliminate the subjectivity involved in any evaluation effort. Evaluative research is therefore one from of applied research.

The first step of any evaluation is generally accepted to be the identification of the program objectives. While this may theoretically seem obvious and easy, many difficulties, especially in the manpower field, are met in such an exercise because few manpower programs have a simple, clearly defined single objective or even a dominant one. Then the evaluation process requires that the identified objectives are operationalized for success is rarely directly observable.

The methodology of evaluation requires that the type of evaluation done be chosen and that a control group be defined. As to the type of evaluation to be chosen, it must be noted that there are as many kinds of evaluation that there are criterias to judge the success of a program : input evaluation, output evaluation, adequacy of performance evaluation, cost-benefit evaluation, formative, summative, absolute, relative evaluation, etc.

The most accepted methodological requirement of evaluative research is undoubtedly the need for a control group. Indeed, in addition to searching if a program has met its objectives, it must be specified that such results were in fact due to the program evaluated. A control group is therefore used as a proxy for the absence of program. While there is general acceptance of the need for a control group, the consensus surely does not exist as to method of choosing the members of such a group. Ideally such members should be on all points similar to those exposed to the program with the sole exception of the program participation. Such an ideal principle is surely impossible to meet fully. Second best methods have been developed for the construction of control groups. Random assignment of individuals to an experimental and a control group before and after measures, comparison groups are some of them. The most popular however is to chose qualified and interested non-participants. The fact that perfect comparability between the two groups is impossible requires that an effort be made to establish the (statistical) degree of comparison. Moreover, the experimental and the control group must be independent. Again such a requirement may be difficult to fulfil.

This suggests that manpower program evaluation has many limitations. The presence of subjectivity, the dynamic character of the programs evaluated, the numerous criteria of success, objectives often badly defined, the problems to build operational definitions, the particular methodological difficulties of the control group, and the general methodological problems of any applied research are some of them. Moreover, serious limits may be due to the evaluator himself, and to the fact that many programs are evaluated at the same time. Finally, the way to really evaluate externalities or third party effects is yet to be found.

Experience suggests that manpower program evaluation is difficult, time consuming, and expensive. Moreover its results may be difficult to interpret and its methodology is often qualified as poor. Manpower program evaluation therefore has a bad reputation in many circles. This surely limits its use as a change agent. While further work is needed and will surely be done to ameliorate the quality of evaluation research methodology, there is also work to be done to favorize a better consumption of evaluation results.

Tous droits réservés © Département des relations industrielles de l'Université Laval, 1973

This document is protected by copyright law. Use of the services of Érudit (including reproduction) is subject to its terms and conditions, which can be viewed online.

<https://apropos.erudit.org/en/users/policy-on-use/>

# L'évaluation des programmes de main-d'oeuvre:

## Observations méthodologiques

Jean Sexton

*Dans cet article, l'auteur présente les origines, la nature, la méthodologie et les limites de l'évaluation des programmes de main-d'oeuvre.*

Dans le domaine de la politique et des programmes de main-d'oeuvre, la période qui s'étend de 1960 à 1970 est généralement désignée comme étant celle de la grande expansion. La décennie actuelle le sera probablement comme étant celle de l'évaluation tant en Amérique du Nord que dans plusieurs pays européens. L'évaluation en effet est devenue une préoccupation générale. Cependant, la notion même d'évaluation est aussi relative que celle de succès, autant pour les programmes de main-d'oeuvre que pour tout autre programme. On sait qu'un programme de main-d'oeuvre consiste dans une série d'actions, de mesures ou de projets visant le développement et le déploiement rationnel de *toutes* les ressources humaines d'une société pour et par un travail productif, librement choisi, gage de sécurité et de croissance individuelle et collective<sup>1</sup>.

Cet article poursuit un double but : d'abord présenter la nature, la méthodologie et certaines limites de

SEXTON, J., professeur-assistant, Département des relations industrielles, Université Laval, Québec.

\* Cet article est extrait d'une recherche plus vaste présentement en cours : *Les licenciements collectifs : une étude des activités de reclassement au Québec*. L'auteur complète cette étude avec M. Serge Laplante de l'université Harvard avec l'aide financière du ministère du Travail et de la Main-d'oeuvre du Québec et du Département des relations industrielles de l'université Laval. Nous tenons expressément à remercier le Centre de formation et de consultation (CFC) pour sa précieuse collaboration.

<sup>1</sup> Nous avons présenté cette conception plus en détail dans Jean SEXTON « La politique active de main-d'oeuvre et la fonction 'information' ; une opinion » Claude Laflamme et André Petit (éd.), *L'information scolaire et professionnelle dans l'orientation : approche multidisciplinaire*. Centre de documentation scolaire et professionnelle, Faculté des sciences de l'éducation, Sherbrooke, Université de Sherbrooke, 1973, pp. 245-264.

l'évaluation de programmes de main-d'œuvre au moyen d'une brève synthèse des études, surtout américaines, consacrées à ce sujet ; ensuite, combler un manque flagrant de documentation en langue française dans ce domaine. Nous n'avons pas la prétention d'être exhaustif, car, en plus d'être nombreux, les écrits portant sur l'évaluation ne sont pas exclusifs au domaine de la main-d'œuvre.

#### ORIGINE ET CAUSES DE L'ÉVALUATION

L'évaluation des programmes sociaux n'est pas une découverte récente en sciences sociales. La première évaluation connue de tels programmes a été effectuée en 1897 par J. M. Rice dans le domaine de l'éducation aux États-Unis. Entre 1920 et 1940, il se fit une série de tentatives pour utiliser les méthodes de recherches empiriques afin de mieux déterminer les effets des programmes sur le changement social à l'intérieur de divers contextes<sup>2</sup>. C'est au cours de cette période que Sydenstricker a établi les premiers principes d'évaluation<sup>3</sup>. Entre 1940 et 1960, on retient surtout le travail des psychologues sociaux en ce domaine<sup>4</sup>.

Durant la dernière décennie, un nouvel intérêt pour la pauvreté et les problèmes s'y rattachant entraîna une certaine renaissance de la recherche évaluative particulièrement aux États-Unis. On vit alors plusieurs lois américaines exiger d'une façon systématique que soit évalué l'impact de certains programmes tels ceux prévus par l'*Elementary and Secondary Education Act* de 1965 et ceux relatifs à la pauvreté et à la main-d'œuvre. Ce dernier point est important, car nulle part ailleurs, en politique publique américaine, la demande n'a été aussi forte pour l'évaluation. De plus, on notera que la réaction du législateur et de l'administrateur des programmes fut extrêmement rapide à la suite des pressions exercées<sup>5</sup>. Le Congrès, par exemple, chercha et cherche encore à connaître l'efficacité de la principale loi américaine en matière de main-d'œuvre, le M.D.T.A.<sup>6</sup>, et d'une foule d'autres programmes.

<sup>2</sup> Francis C. CARO (ed.), *Readings in Evaluation Research*, New York, Russell Sage Foundation, 1971, p. 4.

<sup>3</sup> Edward A. SUCHMAN, *Evaluative Research : Principles and Practice in Public Service and Social Action Programs*, New York, Russell Sage Foundation, 1967, p. 15.

<sup>4</sup> Francis C. CARO (ed.), *op. cit.*, p. 5.

<sup>5</sup> Par exemple, voir à ce sujet : Garth L. MANGUM, « Manpower Research and Manpower Policy », *A Review of Industrial Relations Research*, vol. II, I.R.R.A. (ed.), Madison, 1971, p. 95.

<sup>6</sup> Le Manpower Development and Training Act de 1962.

Cependant, cette attention pour l'évaluation a été longtemps exclusive aux États-Unis, comme le soulignait Somers en 1968 :

« Même si des pays comme la Suède, la Grande-Bretagne et la France sont capables de fournir des données... il n'y a aucune preuve d'un véritable rappel suivi dont le but est déterminer l'expérience subséquente sur le marché du travail des gens qui ont été formés ou reclassés géographiquement »<sup>7</sup>.

Si le Canada, comme le fait remarquer Somers<sup>8</sup>, n'a pas échappé à la règle, il n'en demeure pas moins que, depuis 1970, les gouvernements, tant au Canada que dans certains pays européens, et les chercheurs manifestent un intérêt croissant pour l'évaluation des programmes de main-d'œuvre. Par exemple, le Québec, d'une part a récemment établi une structure et une série de mécanismes prévoyant l'évaluation des programmes gouvernementaux de formation professionnelle et il subventionne une évaluation de son programme de reclassement<sup>9</sup>. D'autre part, le département des relations industrielles de l'université Laval avait pour thème l'évaluation de l'expérience québécoise dans le domaine des programmes de main-d'œuvre lors de son 27<sup>e</sup> Congrès annuel<sup>10</sup>.

La raison d'être de l'évaluation et les causes de l'intérêt croissant pour cette sorte de recherche semblent, à toute fin pratique, être les mêmes au Québec, au Canada et aux États-Unis. Au début du renouveau des politiques de main-d'œuvre durant les années '60 en Amérique du Nord, on insista surtout sur le développement de programmes et sur la façon de les offrir à ces populations choisies. Après cet enthousiasme initial, on en vint peu à peu à s'interroger sur l'efficacité de ces programmes, souvent définis indépendamment les uns des autres. Plusieurs facteurs ont motivé cet intérêt soudain pour l'évaluation des programmes de main-d'œuvre. L'insatisfaction du public, les limitations des fonds publics, la nécessité d'établir des priorités, les changements dans la nature des problèmes sociaux, la croissance de la professionnalisation dans le service

<sup>7</sup> Gerald G. SOMERS, « Evaluating Manpower Policies », U.S. Congress, Joint Economic Committee. *Federal Programs for the Development of Human Resources*, Washington, G.P.O., vol. 1, 1968, p. 159 (notre traduction).

<sup>8</sup> Gerald G. SOMERS « Federal Manpower Policies ». Richard U. Miller and Fraser Isbester (eds.), *Canadian Labour in Transition*, Scarborough, Prentice-Hall of Canada, 1971, p. 84.

<sup>9</sup> Jean SEXTON et Serge LAPLANTE, *Les licenciements collectifs : une étude des activités de reclassement au Québec (à paraître)*.

<sup>10</sup> Département des relations industrielles, *Politiques de main-d'œuvre : évaluation de l'expérience québécoise*, Québec, P.U.L., 1972, 188 pp.

public, la présence d'un public plus éduqué et plus exigeant, les modifications et innovations presque constantes dans les programmes publics et une multitude d'échecs évidents de certains programmes <sup>11</sup> sont quelques-unes des raisons qui en ont conduit plusieurs à croire que l'évaluation de tout programme, incluant ceux visant au développement des ressources humaines, était essentielle <sup>12</sup>. On est donc graduellement venu à exiger de tout programme de main-d'œuvre qu'il fasse la preuve de son efficacité afin d'obtenir l'appui du public et des hommes politiques.

Le but principal d'une telle évaluation est alors de fournir une information permettant une meilleure prise de décision à l'intérieur de ces programmes <sup>13</sup>.

#### NATURE DE L'ÉVALUATION

Suchman prétend que le terme « évaluation », malgré son usage répandu, est très mal défini et conséquemment mal utilisé <sup>14</sup>. Fondamentalement, l'évaluation est un jugement de valeur. De par sa nature même, ce processus est souvent subjectif et très complexe. La possibilité d'utiliser les principes et les méthodes de recherche empirique pour évaluer les programmes publics, comme nous le verrons plus loin, permet de contrôler dans une certaine mesure cette subjectivité intrinsèque au processus d'évaluation sans toutefois pouvoir l'éliminer complètement.

Les tentatives récentes de définition du concept d'évaluation reflètent des préoccupations tant au niveau de la cueillette de l'information sur les résultats des programmes qu'à celui des conclusions sur la valeur de ces mêmes programmes <sup>15</sup>. On peut alors en déduire qu'il est possible d'évaluer

<sup>11</sup> Par exemple, voir à ce sujet : Francis C. CARO, *op. cit.*, p. 1 ; Edward A. SUCHMAN, *op. cit.*, pp. 3-5 ; Carol H. WEISS, « Evaluating Educational and Social Action Programs : A Treeful of Owls ». Carol H. WEISS (ed.), *Evaluating Action Programs*, Boston, Allyn and Bacon, 1972, p. 4.

<sup>12</sup> Par exemple, voir à ce sujet : Gerhard, COLM, « On the Problems of Program Appraisal in the Field of Human Resource Development », U.S. Congress, Joint Economic Committee, *op. cit.*, p. 37.

<sup>13</sup> Voir par exemple : John W. SCANLON, Joe R. RAY and Joseph S. WHOLEY, « An Evaluation System To Support a Decentralized, Comprehensive Manpower Program », Michael E. BORUS (ed.), *Evaluating the Impact of Manpower Programs*, Lexington, Lexington Books, 1972, p. 29.

<sup>14</sup> Edward A. SUCHMAN, *op. cit.*, p. 27.

<sup>15</sup> Francis C. CARO, *op. cit.*, p. 2.

selon différents critères. Guba insiste sur ce point en rappelant que l'évaluation peut être définie en termes de mesure, en termes de conformité entre la performance et les objectifs et en termes de jugement professionnel<sup>16</sup>.

De façon générale, cependant, les auteurs s'accordent à dire que l'évaluation est un jugement dont le but est de déterminer dans quelle mesure un programme réalise ses objectifs<sup>17</sup>. L'acceptation de l'évaluation comme un processus visant surtout à recueillir de l'information sur les résultats des programmes constitue de loin la caractéristique propre à la recherche évaluative.

Cette première école de pensée fut graduellement enrichie de l'apport complémentaire d'une seconde école prétendant qu'il n'est pas suffisant de déterminer si un programme a rencontré ses objectifs ou non. Guba<sup>18</sup> et Scriven<sup>19</sup> soulignent le manque d'analyse dans l'évaluation. Aussi plusieurs se sont joints à eux, (v.g. Alkin<sup>20</sup>, Niland<sup>21</sup>, Riecken<sup>22</sup> et Suchman<sup>23</sup>) pour ajouter la notion fondamentale d'analyse dans l'évalua-

---

<sup>16</sup> Egon G. GUBA, « The Failure of Educational Evaluation ». Carol H. WEISS (ed.), *op. cit.*, p. 254.

<sup>17</sup> Par exemple, cf. Gerhard COLM, *op. cit.*, p. 29 ; Garth L. MANGUM, *op. cit.*, p. 68 ; John R. NILAND, *op. cit.*, p. 16 ; Henry M. RIECKEN, « Memorandum on Program Evaluation ». Carol H. WEISS (ed.), *op. cit.*, p. 86 ; Edward A. SUCHMAN, *Evaluative Research . . .*, *op. cit.*, pp. 31, 32, 37, 75 ; Edward A. SUCHMAN, « Action for What ? . . . », Carol H. WEISS (ed.) *op. cit.*, p. 53 ; Burton A. WEISBROD, « Benefits of Manpower Programs : Theoretical and Methodological Issues », G. G. SOMERS and W. D. WOOD (eds.), *Cost-Benefit Analysis of Manpower Policies*, Proceeding of a North American Conference, Kingston, Queen's University, 1969, pp. 3, 14 ; Burton A. WEISBROD, « Conceptual Issues in Evaluating Training Programs », *Monthly Labor Review*, Oct. 1966, p. 1091 ; Carol M. WEISS, « Evaluating . . . », Carol H. WEISS (ed.), *op. cit.*, pp. 3, 4, 21 ; Robert S. WEISS and Martin REIN, « The Evaluation of Broad-Aim Programs : Difficulties in Experimental Design and an Alternative » Carol H. WEISS (ed.), *op. cit.*, p. 247.

<sup>18</sup> Egon G. GUBA, *op. cit.*, p. 257.

<sup>19</sup> Michael SCRIVEN, *op. cit.*, p. 127.

<sup>20</sup> Marvin C. ALKIN, « Evaluation Theory Development », Carol H. WEISS (ed.), *op. cit.*, p. 107.

<sup>21</sup> John R. NILAND, « The Nature of the Manpower Function », John R. NILAND (ed.), *The Production of Manpower Specialists : A Volume of Selected Readings*, Ithaca, New York State School of Industrial and Labor Relations, Cornell University, 1971, p. 16.

<sup>22</sup> Henry W. RIECKEN, *op. cit.*, p. 86.

<sup>23</sup> Edward A. SUCHMAN, *op. cit.*, pp. 66, 109.

tion. Jusqu'à quel point les buts sont-ils atteints ? Pourquoi a-t-on atteint tel ou tel niveau de succès ? Est-ce que les programmes sont désirables ou non ? Etc. La première conception d'évaluation peut être qualifiée « d'évaluation factuelle » ou « d'évaluation descriptive » alors que la seconde est beaucoup plus de « l'évaluation qualitative » ou de « l'évaluation explicative ».

On a déjà dit de l'évaluation qu'elle était fondamentalement un jugement de valeur, un processus subjectif. Sans entrer dans les détails méthodologiques d'une telle évaluation, on peut immédiatement établir qu'un tel jugement de valeur peut être fait tant par des méthodes scientifiques que non scientifiques. Cette constatation amena Suchman à distinguer entre l'évaluation et la recherche évaluative<sup>24</sup>. Il définit alors l'évaluation comme un processus social d'établissement de jugement de valeur... processus qui ne requiert aucune procédure systématique pour rassembler ou présenter des données objectives supportant le jugement. Cette évaluation est de l'ordre de l'opinion. Par recherche évaluative, il réfère à l'utilisation de méthodes et de techniques de recherche scientifique dans l'évaluation. Donc, la recherche évaluative ne se contente pas seulement d'offrir un jugement de valeur sur une activité sociale, mais doit en plus le fonder sur des preuves. L'usage de cette dernière méthode est de loin la plus acceptée vu qu'elle est la seule susceptible d'éliminer certains des éléments de subjectivité inhérents au processus d'évaluation. La recherche évaluative est donc d'abord et avant tout de la recherche appliquée et plus spécifiquement, une forme de recherche appliquée puisqu'il est possible de faire de la recherche appliquée sans tenter d'évaluer quoi que ce soit.

#### LA SPÉCIFICATION DES OBJECTIFS

Les auteurs s'accordent généralement à dire qu'une des premières étapes, sinon la première de toute recherche évaluative est d'identifier les buts ou les objectifs d'un programme<sup>25</sup>. Il est tout aussi important de s'attarder à cette définition des objectifs d'un programme pour la recherche évaluative qu'il est fondamental de formuler d'abord une hypothèse en recherche de base. Cette identification d'objectif est essentielle puisqu'on cherche à répondre à la question suivante : l'objectif du programme a-t-il été atteint ?

La réponse positive à cette question exige fondamentalement qu'on prouve deux choses : premièrement, que le programme fournit les résultats

---

<sup>24</sup> *Ibid.*, pp. 7, 21, 31, 32, 75-81.

<sup>25</sup> *Ibid.*, p. 87, par exemple.

voulus et ensuite que ces résultats sont réellement dus au programme. Cette seconde exigence suggère l'importance de variables intervenantes et se trouve résumée dans l'expression d'« effet de Hawthorne »<sup>26</sup>.

Celui qui tentera de déterminer les objectifs d'un programme de main-d'œuvre pour fin d'évaluation se heurtera à plusieurs obstacles. Sans les expliquer en détails, en voici quelques-uns. Les objectifs ont-ils été clairement définis avant l'établissement du programme ? S'agit-il d'objectifs officiels ou officieux ? Quels étaient-ils au moment de la mise en oeuvre du programme ? Y a-t-il une différence fondamentale dans la façon dont l'administrateur et le législateur perçoivent les objectifs définis ? Y a-t-il plusieurs objectifs ? Si oui, existe-t-il un ordre de priorité ?

« Peu de programmes de main-d'œuvre ont un objectif simple, clairement défini et même un objectif dominant »<sup>27</sup>.

De plus, est-ce que l'objectif est stable ou a-t-il changé en cours de route ? Est-ce que les objectifs sont vagues ou précis ? Finalement, le programme à évaluer ne poursuit-il pas uniquement l'objectif (implicite) d'assurer sa pérennité ?

Les programmes de main-d'œuvre ont généralement tendance à définir des objectifs globaux ou à poursuivre plusieurs objectifs en même temps. Il devient alors souvent possible d'identifier une chaîne plus ou moins longue d'objectifs. L'évaluateur devra alors tenter de distinguer entre les objectifs immédiats, intermédiaires et ultimes, tels que définis par Suchman<sup>28</sup>. Un tel choix est très important vu qu'il influencera tout probablement les résultats de l'évaluation. Notons également qu'il n'y a pas de demi-succès. L'évaluation doit plutôt chercher rationnellement les causes de la non-atteinte des buts visés. Dans cette recherche, il pourra souvent savoir si les administrateurs du programme adoptent un compor-

---

<sup>26</sup> Glen G. CAIN et Robinson G. HOLLISTER « Evaluating Manpower Programs for the Disadvantaged », G. G. SOMERS and W. D. WOOD, *op. cit.*, p. 122 ; Francis C. CARO, *op. cit.*, pp. 3, 21 ; David K. COHEN, « Politics and Research : Evaluation of Social Action Programs in Education », Carol H. WEISS (ed.), *op. cit.*, p. 145 ; Henry W. RIECKEN, *op. cit.*, p. 92 ; Robert E. STOKE, « The Countenance of Educational Evaluation », Carol H. WEISS (ed.), *op. cit.*, p. 41 ; Edward A. SUCHMAN, *op. cit.*, pp. 32, 38 ; Burton A. WEISBROD, « Conceptual Issues... » *op. cit.*, p. 1091 ; Carol H. WEISS, « Evaluating... », *op. cit.*, p. 6.

<sup>27</sup> Glen G. CAIN et Robinson G. HOLLISTER, *op. cit.*, p. 122 (notre traduction).

<sup>28</sup> Edward A. SUCHMAN, *op. cit.*, pp. 68 sq.



tement de maximisation ou de satisfaction des avantages par rapport à l'atteinte des objectifs exprimés, pour utiliser la typologie de Simon<sup>29</sup>.

Les objectifs d'un programme ne sont pas toujours définis de façon opératoire. On doit alors préciser *a posteriori* des indicateurs qui constituent des manifestations empiriques du succès. Il y a sûrement des risques à procéder de la sorte. Cette identification d'indicateurs constitue pour plusieurs auteurs la deuxième étape nécessaire à toute recherche évaluative, après avoir déterminé les objectifs.

« La mesure des résultats d'un programme est la tâche principale de l'évaluateur. Il doit déterminer les objectifs de ce programme de façon opératoire au moyen d'indicateurs reflétant les résultats désirés »<sup>30</sup>.

Cette étape peut sembler évidente et même facile à réaliser. Cependant, les auteurs reconnaissent qu'une telle définition opératoire des objectifs constitue habituellement une des tâches les plus ardues d'une étude d'évaluation. Point n'est besoin d'insister sur le fait que c'est à partir de ces indicateurs que la collecte des données se fera.

## LA MÉTHODOLOGIE DE L'ÉVALUATION

On peut maintenant se demander comment s'effectuera cette évaluation ? Pour répondre à cette question, on doit s'arrêter à préciser le genre d'évaluation à faire et la méthode appropriée.

### Sortes d'évaluation

Fondamentalement, on peut identifier autant de sortes d'évaluation qu'il y a de critères pour juger du succès d'un programme quelconque. Un examen des ouvrages existant sur le sujet révèle une grande diversité de types d'évaluation. À toute fin pratique, il y a presque autant de typo-

<sup>29</sup> Herbert A. SIMON, *Administrative Behavior*, New York, Macmillan, 1957, p. XXV. Un comportement de maximisation des avantages (maximizing) est caractérisé par le choix du meilleur moyen parmi plusieurs pour atteindre une fin, tandis qu'un comportement de satisfaction des avantages (satisficing) se reconnaît par la loi du moindre effort. On peut illustrer cette différence de comportement par l'exemple suivant : un individu à la recherche d'un chapeau et adoptant un comportement de maximisateur, cherche un chapeau qui lui va très bien, qui sera durable, etc., alors que celui qui adopte un comportement de satisfaction des avantages choisira n'importe quel chapeau qui lui fait.

<sup>30</sup> Carol H. WEISS, « Evaluating... », *op. cit.*, p. 22 (notre traduction).

logies de l'évaluation qu'il y a d'auteurs sur le sujet. Par exemple, certaines de ces typologies sont basées sur l'étendue du programme<sup>31</sup>, quelques-unes le sont sur les phases du programme<sup>32</sup> et d'autres sont basées sur différents aspects du programme<sup>33</sup>, etc. Les typologies de Scriven<sup>34</sup>, Cain et Hollister<sup>35</sup>, Provus<sup>36</sup>, et Hemphill<sup>37</sup> utilisent fondamentalement un ou plusieurs des critères suivants : les activités pré-programme, durant le programme ou après le programme.

Suchman identifie cinq catégories d'évaluation : l'évaluation quant à l'effort a pour objet la quantité et la qualité des ressources utilisées ; l'évaluation quant au résultat de l'effort se préoccupe de l'effet produit et non de la qualité de cet effet ; l'évaluation quant à la pertinence de l'effort par rapport au besoin à satisfaire ; l'évaluation quant aux coûts et avantages des résultats atteints et l'évaluation dite analytique<sup>38</sup>. Avec ce dernier type d'évaluation, apparaissent les préoccupations plus récentes d'explications des résultats obtenus, quels que soient les critères de succès utilisés. Cette évaluation analytique peut se faire par rapport à quatre dimensions principales : les caractéristiques du programme lui-même ; la population exposée au programme ; le contexte dans lequel se situe le programme ; les différentes sortes d'effets produits par le programme. Ce dernier genre d'évaluation est évidemment un idéal à atteindre. La formulation des objectifs, les ressources de recherche disponibles et l'attitude de l'administrateur du programme peuvent constituer des contraintes sérieuses qui limiteront l'atteinte de cet idéal.

Suchman n'est pas le seul à avoir bâti une typologie de l'évaluation. Nous avons déjà mentionné qu'il y en avait presque autant qu'il y a d'auteurs traitant du sujet. En voici rapidement quelques autres afin d'identifier quelque peu l'univers des possibilités en ce domaine.

---

<sup>31</sup> Par exemple, voir à ce sujet : Joseph S. WHOLEY, *et al.*, « Proper Organizational Relationships », Carol H. WEISS (ed.), *op. cit.*, p. 116.

<sup>32</sup> Par exemple, voir à ce sujet : D. L. STUFFLEBEAM, *Evaluation As Enlightenment for Decision-Making*, Columbus, Ohio State University, 1969.

<sup>33</sup> Par exemple, voir à ce sujet : Henry W. RIECKEN, *op. cit.*, pp. 86-87. Edward A. SUCHMAN, *op. cit.*, pp. 61ff.

<sup>34</sup> Francis C. CARO, *op. cit.*, p. 4 ; Michael SCRIVEN, *op. cit.*, pp. 86-87.

<sup>35</sup> Glen G. CAIN et Robinson G. HOLLISTER, *op. cit.*, p. 120.

<sup>36</sup> PROVUS est cité dans Marvin C. ALKIN, *op. cit.*, p. 112.

<sup>37</sup> *Ibid.*

<sup>38</sup> Edward A. SUCHMAN, *op. cit.*, pp. 61 sq.

Scriven distingue entre l'évaluation « formative » et l'évaluation du changement<sup>39</sup>. Alors que l'évaluation « formative » réfère à celle qui peut servir à l'établissement d'un programme, l'évaluation dite du changement se concentre sur la valeur d'un programme après qu'il eut été en opération. On voit alors que l'évaluation du changement peut être dans une certaine mesure « formative » en ce sens qu'elle peut proposer des modifications au programme déjà existant. Stake, pour sa part, prétend qu'il existe deux critères principaux d'évaluation<sup>40</sup>. Comme son nom l'indique, l'évaluation absolue d'un programme se fait par rapport à des critères absolus tels que peut le refléter le jugement personnel. L'évaluation relative, comme son nom l'indique également, examine un programme par rapport à des standards relatifs comme peuvent le refléter les caractéristiques d'un programme alternatif. Il serait sûrement possible de continuer et d'ajouter une série de typologies encore différentes à celles déjà mentionnées. Cependant, de façon opératoire, il devient sûrement impossible de réellement compléter toutes ces catégories d'évaluation. C'est pourquoi, selon chaque circonstance, il faudra poser un choix.

Rappelons finalement que l'analyse coûts-avantages n'est pas nécessairement toujours la meilleure méthode en raison des critères de succès utilisés, et des difficultés quant à la définition des avantages et du taux d'escompte à retenir. Certains auteurs soutiennent qu'une telle évaluation est inutile si on n'a pas prouvé par d'autres travaux préalables que le programme atteignait ses objectifs<sup>41</sup>. En d'autres termes, l'analyse coûts-avantages représente le stade le plus avancé de l'évaluation. Cependant cette analyse n'est pas un substitut aux autres méthodes d'évaluation, mais bien l'extension logique de celles-ci.

#### **Quelques notes sur les groupes contrôle**

De même qu'il existe différentes approches pour effectuer une recherche appliquée, de même il y a différentes méthodologies possibles de l'évaluation. Le commun dénominateur de toutes ces approches semble être le principe généralement accepté que la recherche évaluative doit répondre aux mêmes exigences scientifiques que la recherche appliquée.

Plusieurs auteurs ont défini les étapes à franchir pour faire de l'évaluation. De façon générale, il semble qu'il y ait entente sur les points suivants : identification des objectifs du programme ; définition opératoire de ces objectifs ; collecte des données.

<sup>39</sup> Michael SCRIVEN, *op. cit.*, p. 128 ; Francis C. CARO, *op. cit.*, p. 4.

<sup>40</sup> Robert E. STAKE, *op. cit.*, p. 46.

<sup>41</sup> Edward A. SUCHMAN, « Action for What ? ... », *op. cit.*, p. 71.

En plus, tous ces auteurs sans exception insistent sur ce que nous appellerions la caractéristique principale de la méthodologie évaluative, c'est-à-dire la nécessité d'un groupe contrôle<sup>42</sup>.

La raison d'être de l'utilisation de groupes contrôle est de répondre à la question fondamentale et essentielle de toute évaluation, c'est-à-dire que serait-il arrivé sans le programme X ? Ou encore, quelle différence le programme a-t-il produite ? Il est admis qu'idéalement on devrait juger un programme de main-d'œuvre par rapport à des standards absolus de contribution maximum au déploiement rationel des ressources humaines<sup>43</sup>. La définition opératoire d'un tel critère est cependant impossible. Il est donc illusoire de vouloir mesurer après un événement ce qui serait arrivé si ce dernier ne s'était pas produit<sup>44</sup>. En l'absence d'une telle mesure, le meilleur substitut semble être de définir un groupe contrôle qui n'a pas été exposé au programme évalué mais dont l'expérience générale est *en tout point* similaire à celle des bénéficiaires de ce programme. On peut alors voir que le groupe contrôle idéal est un groupe absolument identique à celui des participants à l'exception de la participation au programme évalué. L'évaluation devient alors très relative.

Les groupes contrôle sont si importants pour évaluer un programme public que certains préfèrent utiliser des groupes absolument non identiques et non équivalents plutôt que de n'utiliser aucun contrôle<sup>45</sup>. Cependant, nous devons retenir, comme le mentionne Cohen, que toute la

---

<sup>42</sup> Michael E. BORUS et William R. TASH, *op. cit.*, pp. 4, 15-16; Michael E. BORUS (ed), *op. cit.*; Michael E. BORUS et Charles G. BUNTZ, « Problems and Issues in the Evaluation of Manpower Programs », *Industrial and Labor Relations Review*, vol. 25, no 2, Jan. 1972, pp. 234-245; Glen G. CAIN et Robinson G. HOLLISTER, « The Methodology of Evaluating Social Action Programs » *Public Private Manpower Policies*, Madison, I.R.R.A., 1969, p. 4; Glen G. CAIN et Robinson G. HOLLISTER, « Evaluating... », *op. cit.*, p. 125; Francis C. CARO, *op. cit.*, p. 23; David K. COHEN, *op. cit.*, p. 154; Conseil économique du Canada, *L'État et la prise de décision*, 8e exposé annuel, Ottawa, Information Canada, 1971, p. 132; Thomas G. GLENNON Jr., « Evaluating Federal Manpower Programs : Notes and Observations », Carol M. WEISS (ed), *op. cit.*, p. 182; Einar HARDIN, *op. cit.*, p. 106; John R. NILAND, *op. cit.*, pp. 16-17; Henry W. RIECKEN, *op. cit.*, p. 96; Peter H. ROSSI, *op. cit.*, p. 225; Michael SCRIVEN, *op. cit.*, p. 135; Gerald G. SOMERS, « Evaluation... », *op. cit.*, pp. 160-161; Edward A. SUCHMAN, *op. cit.*, pp. 93-95.

<sup>43</sup> John R. NILAND, *op. cit.*, p. 17.

<sup>44</sup> Michael E. BORUS et William R. TASH, *op. cit.*, pp. 15-16; Thomas K. GLENNON Jr., *op. cit.*, p. 181.

<sup>45</sup> Carol H. WEISS, « Evaluating... », *op. cit.*, p. 9.

stratégie de la recherche évaluative semble reposer sur la valeur de tels groupes contrôle <sup>46</sup>.

Alors que tous les auteurs s'entendent sur l'importance et le besoin théorique de définir un groupe contrôle pour compléter des évaluations relatives, la méthode de choix de ce groupe contrôle demeure un problème à résoudre et, pour utiliser l'expression de Borus, constitue un des domaines les plus controversés, débattus et incertains de l'évaluation <sup>47</sup>.

Pour plusieurs <sup>48</sup>, la méthode idéale est l'assignation au hasard d'une partie des sujets à un groupe expérimental qui est exposé au programme et d'une autre partie à un groupe contrôle qui n'est pas exposé au programme ou à qui on administre un placebo. Alors que cette méthode est très populaire et sûrement possible dans le domaine médical, il devient presque impossible de l'utiliser dans les programmes sociaux. Il est clair que des motifs d'ordre politique et administratif constituent des obstacles majeurs à de telles expériences contrôlées. Face à cette impossibilité d'atteindre un tel idéal, des mesures dites de second ordre ont été développées pour définir un groupe contrôle qui permette une comparaison avec le groupe original.

L'étude de cas, pour reprendre une constatation de Suchman, est paradoxalement la méthode de recherche évaluative la plus utilisée et en même temps la plus faible, surtout parce qu'elle ne fournit aucune comparaison <sup>49</sup>. Selon cette méthode, les observations ou les données recueillies sur les individus ou les groupes le sont seulement après exposition au programme évalué. C'est de l'évaluation absolue.

Une autre façon d'établir un groupe contrôle est d'utiliser la méthode dite « avant et après ». Par cette méthode, l'individu ou le groupe est son propre contrôle <sup>50</sup>. En effet, on introduit une mesure de base avant le début

---

<sup>46</sup> David K. COHEN, *op. cit.*, p. 154.

<sup>47</sup> Michael E. BORUS et Charles BUNTZ, *op. cit.*, pp. 235-327 ; Michael E. BORUS (ed.), *op. cit.*, p. XII ; Abraham STAHLER, « On the Choice of Control Group : Comment », Michael E. BORUS (ed.), *op. cit.*, p. 67.

<sup>48</sup> Francis C. CARO, *op. cit.*, p. 154 ; Thomas K. GLENNON Jr., *op. cit.*, p. 181 ; Einar HARDIN, « On the Choice of Control Groups », Michael E. BORUS (ed.), *op. cit.*, p. 57 ; Peter H. ROSSI, *op. cit.*, p. 231 ; Carol H. WEISS, « Evaluating . . . », *op. cit.*, pp. 18, 329.

<sup>49</sup> Edward A. SUCHMAN, *op. cit.*, p. 93.

<sup>50</sup> Glen G. CAIN et Robinson G. HOLLISTER, « Evaluating . . . », *op. cit.*, p. 125.

du programme et on réintroduit une autre mesure à la fin du programme. Borus prétend qu'une telle méthode d'établissement d'un groupe contrôle est de loin la moins désirable<sup>51</sup>. Fondamentalement, cette méthode, comme toutes les autres méthodes, est une approximation de l'exposition et de la non-exposition au programme<sup>52</sup>, avec, cependant, l'hypothèse très douteuse que tous les autres facteurs extérieurs et intérieurs au programme et au groupe n'ont pas changé. Mais, quelles sont précisément ces autres choses ou ces autres variables intervenantes pouvant agir entre ces deux moments? Les conditions du marché du travail peuvent changer, les facteurs climatiques peuvent être différents, la seule présence d'un programme peut en soi produire un effet, les services spéciaux aux participants, peuvent introduire un biais, etc. Hardin insiste sur ce point en avançant que des changements dans l'environnement peuvent sérieusement biaiser la mesure du succès d'un programme de main-d'œuvre selon cette méthode particulière<sup>53</sup>. Sur cette façon d'établir des groupes contrôle, il conclut :

« En somme, la simple comparaison avant et après est une source dangereuse de données sur les effets d'un programme. C'est surtout à cause des biais possibles qu'on ne peut pas réellement mesurer l'ampleur ou même la présence des effets d'un programme... C'est donc dire que les simples comparaisons avant et après créent des biais substantiels dans l'estimé de l'impact d'un programme. Il peut être pire d'utiliser de telles comparaisons que de ne pas faire d'évaluation du tout »<sup>54</sup>.

La méthode la plus populaire du choix d'un groupe contrôle serait d'examiner un autre groupe grandement similaire au groupe des participants avec cette différence que ces participants n'ont eu aucun contact avec le programme évalué. Dans l'état actuel des connaissances, comme le concluent Hardin et Borus, des non-participants qualifiés et intéressés semblent être le groupe contrôle le plus approprié plutôt qu'un simple échantillon de la population visée<sup>55</sup>. De façon opératoire, ceci s'est traduit

---

<sup>51</sup> Michael E. BORUS et William R. TASH, *op. cit.*, p. 19 ; Einar HARDIN et Michael E. BORUS, « An Economic Evaluation of Retraining Program in Michigan : Methodological Problems of Research », *Proceedings of the American Statistical Association*, 1966, pp. 133-137.

<sup>52</sup> Conseil économique du Canada, *op. cit.*, p. 132 ; Burton A. WEISBROD, « Conceptual... », *op. cit.*, p. 1092.

<sup>53</sup> Einar HARDIN, « On the Choice... », *op. cit.*, p. 42.

<sup>54</sup> *Ibid.*, pp. 44 et 57 (notre traduction).

<sup>55</sup> Einar HARDIN, « On the Choice... », *op. cit.*, p. 57 ; Michael E. BORUS et William R. TASH, *op. cit.*, p. 18.

par le choix d'individus du même âge et du même statut économique que les participants du programme<sup>56</sup>. Ensuite, un rappel suivi (follow-up) des groupes en question permet d'établir une comparaison des expériences d'emplois et de revenus des personnes exposées et non-exposées au programme<sup>57</sup>.

On peut idéalement compléter cette méthode populaire, évidemment lorsque c'est possible, par la proposition de Caro en construisant un groupe contrôle et un groupe de comparaison. Le groupe dit de comparaison est formé d'individus qui ont reçu un traitement autre que celui du groupe expérimental, alors que le groupe dit de contrôle est formé d'individus qui n'ont reçu aucun traitement<sup>58</sup>. Ceci permettrait de donner une image beaucoup plus fidèle de la réalité. Cependant, on doit se rendre compte que c'est un idéal à atteindre et que souvent les conditions particulières à un programme ou à un problème donné ne permettent pas un tel luxe.

Il y a évidemment plusieurs limites rattachées à l'utilisation de cette méthode, dite populaire, d'établir un groupe contrôle. L'impact relatif d'un programme de main-d'œuvre pouvant être défini comme la différence d'effet entre la présence et l'absence du programme<sup>59</sup>, nous avons établi que la meilleure approximation possible de cette réalité est obtenue par la constitution d'un groupe contrôle à peu près similaire au groupe expérimental en termes de caractéristiques générales. *Cependant, on se rendra rapidement compte que la parfaite comparaison entre deux groupes est impossible.* Niland insiste sur ce point de la façon suivante :

« Vu qu'une foule de facteurs autres que le fait d'avoir ou de ne pas avoir suivi un programme peut influencer le succès subséquent, on exigera du groupe contrôle qu'il soit identique, au moins statistiquement parlant, quant à l'éducation, l'âge, le sexe, l'arrière plan social, les attentes, et ainsi de suite pour une série de facteurs qui, en fait, rendent l'exercice impossible »<sup>60</sup>.

On devra alors recueillir des données sur le plus de variables possibles afin de bien établir le degré de comparaison entre les deux groupes utilisés

<sup>56</sup> Michael E. BORUS et Charles BUNTZ, *op. cit.*, p. 237 ; Einar, HARDIN, « Benefit Cost... », *op. cit.*, p. 106 ; Carol H. WEISS, « Evaluating... », *op. cit.*, p. 8.

<sup>57</sup> Conseil économique du Canada, *op. cit.*, p. 132.

<sup>58</sup> Francis C. CARO, *op. cit.*, p. 23.

<sup>59</sup> David A. MILLER, « On the Choice of Control Groups : Comment », Michael E. BORUS (ed.), *op. cit.*, p. 59.

<sup>60</sup> John R. NILAND, *op. cit.*, p. 16 (notre traduction).

dans l'évaluation. Comme dans la plupart des recherches évaluatives, on devra en plus faire l'hypothèse douteuse qu'il n'existe aucune différence dans les caractéristiques non-considérées, comme par exemple, les attitudes, la motivation, etc. En plus, même si les groupes étaient parfaitement comparables comme groupes, il n'y a aucune preuve que deux individus de groupes différents se comporteraient de la même façon sous l'effet du programme ou en son absence. Cependant cette méthode, dite populaire est sûrement préférable à la méthode dite « avant et après ». Nous sommes alors entièrement d'accord avec le Conseil économique du Canada lorsqu'il affirme dans son huitième exposé que :

« Quoiqu'il en soit, l'utilisation d'une forme quelconque de groupes contrôle, même si elle ne satisfait pas aux conditions rigoureuses de l'expérience contrôlée, est probablement encore préférable à l'approche « avant et après ». Du moins, elle fournit un moyen de contrôle supplémentaire quant à la validité de cette dernière approche »<sup>61</sup>.

En plus, le groupe expérimental et le groupe contrôle doivent être indépendants. Même si pour le groupe contrôle on choisit des individus qui n'ont jamais participé à un programme quelconque, il est clair que la véritable indépendance est très difficile à obtenir. La connaissance de l'existence d'un programme peut en elle-même avoir un effet et causer un comportement différent chez les membres du groupe contrôle. Cependant, ceci est à peu près impossible à mesurer et à contrôler.

Le groupe contrôle constitue une caractéristique propre et un pré-requis à toute recherche évaluative. Cependant, il est extrêmement difficile en pratique de satisfaire à cette condition. Plusieurs auteurs soutiennent que la méthodologie du groupe contrôle a été et continuera probablement d'être une des faiblesses les plus sérieuses de l'évaluation des programmes de main-d'œuvre<sup>62</sup>. Ceci suggère que l'évaluation des programmes de main-d'œuvre a plusieurs limites.

#### LIMITES DE L'ÉVALUATION

La recherche évaluative comporte plusieurs limitations. Un examen même rapide des écrits sur ce sujet fournira amplement de preuves à cet égard.

<sup>61</sup> Conseil économique du Canada *op. cit.*, p. 132.

<sup>62</sup> David A. MILLER, *op. cit.*, p. 590 ; R. S. GOLDFORD, *The Evaluation of Government Programs : The Case of New Haven Manpower Training Activities*, Yale Economic Essays, Yale University, Autumn 1969, pp. 59-104 ; Thomas K. GLENNON Jr., *op. cit.*, p. 181.



Rappelons quelques-unes de ces limites : la présence de la subjectivité, le caractère dynamique des programmes évalués, les nombreux critères de succès, les objectifs souvent mal définis, les difficultés des définitions opératoires et les problèmes méthodologiques. Sur ce dernier point, mentionnons que la recherche évaluative est un compromis entre les prérequis méthodologiques et les limites administratives <sup>63</sup>.

Plusieurs des limites de l'évaluation sont rattachées à l'évaluateur lui-même. En plus de la naïveté de l'évaluateur sans expérience, il existe le problème de choisir entre un évaluateur interne et un évaluateur externe. L'évaluateur de l'extérieur sera souvent considéré comme un « curieux ». On l'accusera alors de ne rien connaître au programme, de ne pas en connaître les rouages administratifs, et souvent même d'être incompetent. Quant à l'évaluateur de l'intérieur, les critiques l'accuseront souvent de manquer d'expérience de recherche, de n'avoir jamais fait d'évaluation, et plus spécialement de vouloir défendre son programme (surtout si c'est lui qui l'administre). Il existe alors autant d'avantages que d'inconvénients à utiliser un évaluateur de l'extérieur que de l'intérieur. La solution proposée vise évidemment le compromis.

Soulignons aussi qu'il existe une difficulté particulière pour l'évaluateur lorsqu'on utilise une combinaison de plusieurs programmes. Dans un tel cas, il risque d'attribuer tous les bénéfices à un même programme. Par exemple, devons-nous imputer tous les bénéfices à un programme de formation lorsqu'il y a eu des efforts spéciaux de placement pour trouver un emploi aux gens formés ? On devra donc tenir compte d'un tel facteur dans l'analyse. Également, quand un programme a plusieurs composantes, les résultats de l'évaluation vaudront pour l'ensemble des composantes, et non pas pour une ou certaines de ces composantes.

Finalement, une des critiques les plus fréquemment adressées à l'évaluation des programmes de main-d'œuvre insiste sur le fait qu'on ne tient jamais entièrement compte des « externalités » i.e. des effets affectant des tiers <sup>64</sup>. Ces « externalités », souvent appelées effets secondaires, sont surtout de trois ordres : les effets de déplacement, de substitution et de vide <sup>65</sup>. Un effet de déplacement est noté lorsque les participants à un pro-

<sup>63</sup> Edward A. SUCHMAN, *op. cit.*, pp. 16, 22, 82 et 91 (notre emphase) ; Burton A. WEISBROD, « Conceptual . . . », *op. cit.*, p. 1097 ; Carol H. WEISS, « Evaluating . . . », *op. cit.*, p. 6.

<sup>64</sup> Conseil économique du Canada, *op. cit.*, p. 130.

<sup>65</sup> *Ibid.* ; Michael E. BORUS et Charles BUNTZ, *op. cit.*, pp. 241-242.

gramme occupent après celui-ci des emplois que d'autres personnes auraient occupés en l'absence d'un tel programme. Il y a effet de substitution lorsque des travailleurs dans une industrie autre que celle visée par le programme sont touchés. Alors que ces deux premiers effets peuvent comporter surtout des aspects négatifs, l'effet de vide est plus avantageux. Il y aura effet de vide lorsque des personnes sans travail ou sous-employées pourront occuper des emplois ou de meilleurs emplois qu'elles n'auraient pas pu occuper si les participants du programme évalué n'avaient pas réussi et n'avaient pas accepté un tel emploi. En fait, il s'agit ici d'un effet multiplicateur d'emploi. Il vaudra surtout lorsque les programmes de main-d'œuvre viseront des postes clefs. Alors qu'on accepte théoriquement que de tels effets doivent être évalués, très peu d'auteurs se sont attardés à ce problème dans le passé. En effet, comme le notent Borus et Buntz, il n'existe ni théorie ni méthodologie opératoire pour identifier qui, à part les participants à un programme, est touché par les programmes de main-d'œuvre et pour mesurer ces effets. Il y a encore beaucoup de place pour des recherches plus poussées sur ces effets secondaires.

L'expérience suggère que l'évaluation des programmes de main-d'œuvre est difficile, qu'elle prend beaucoup de temps à être complétée et qu'elle est très dispendieuse. En plus, on prétend que les résultats de telles évaluations sont difficiles à interpréter. En résumé, on accuse souvent l'évaluation des programmes de main-d'œuvre d'avoir failli, d'être remarquablement pauvre, et de représenter un gros point d'interrogation au plan scientifique. L'évaluation a donc une mauvaise réputation dans plusieurs milieux et une telle réputation limite sérieusement son usage comme agent de changement de programme ou de politique. Cependant alors qu'il existe beaucoup d'améliorations à apporter à la production de recherches évaluatives, il existe également beaucoup de changements et de formation à effectuer chez le consommateur de ces évaluations.

## **Evaluation of Manpower Programs : Some Methodological Notes**

The purpose of this article is to present a survey of the literature on the evaluation of manpower programs in French. To do so, it has been chosen to successively present the origin, the nature, the methodology and the limits of such evaluative efforts.

Dating back to 1897, the evaluation preoccupation has known a renewed popularity since the mid 1960's in the manpower field. The movement has originated in the U.S.A. and has been and still is spreading in other North American jurisdictions and in certain European countries.

Public dissatisfaction, limited public funds, competing needs, changes in the nature of social problems, an increased professionalization of the public sector, a more educated and more sophisticated public, and many obvious failures are some of the reasons that have led to this increased popularity of evaluating research. The purpose of evaluation is then to supply information leading to a better decision making.

Fundamentally, evaluation is a subjective and complex process. In fact evaluation is an appraisal of the extent to which a program has met its objectives, and why such a desired result was or was not fully achieved. The extent to which scientific methods are used in making that judgment will reduce but not eliminate the subjectivity involved in any evaluation effort. Evaluative research is therefore one form of applied research.

The first step of any evaluation is generally accepted to be the identification of the program objectives. While this may theoretically seem obvious and easy, many difficulties, especially in the manpower field, are met in such an exercise because few manpower programs have a simple, clearly defined single objective or even a dominant one. Then the evaluation process requires that the identified objectives are operationalized for success is rarely directly observable.

The methodology of evaluation requires that the type of evaluation done be chosen and that a control group be defined. As to the type of evaluation to be chosen, it must be noted that there are as many kinds of evaluation that there are criterias to judge the success of a program : input evaluation, output evaluation, adequacy of performance evaluation, cost-benefit evaluation, formative, summative, absolute, relative evaluation, etc.

The most accepted methodological requirement of evaluative research is undoubtedly the need for a control group. Indeed, in addition to searching if a program has met its objectives, it must be specified that such results were in fact due to the program evaluated. A control group is therefore used as a proxy for the absence of program. While there is general acceptance of the need for a control group, the consensus surely does not exist as to method of choosing the members of such a group.

Ideally such members should be on all points similar to those exposed to the program with the sole exception of the program participation. Such an ideal principle is surely impossible to meet fully. Second best methods have been developed for the construction of control groups. Random assignment of individuals to an experimental and a control group, before and after measures, comparison groups are some of them. The most popular however is to chose qualified and interested non-participants. The fact that perfect comparability between the two groups is impossible requires that an effort be made to establish the (statistical) degree of comparison. Moreover, the experimental and the control group must be independent. Again such a requirement may be difficult to fulfil.

This suggests that manpower program evaluation has many limitations. The presence of subjectivity, the dynamic character of the programs evaluated, the numerous criteria of success, objectives often badly defined, the problems to build operational definitions, the particular methodological difficulties of the control group, and the general methodological problems of any applied research are some of them. Moreover, serious limits may be due to the evaluator himself, and to the fact that many programs are evaluated at the same time. Finally, the way to really evaluate externalities or third party effects is yet to be found.

Experience suggests that manpower program evaluation is difficult, time consuming, and expensive. Moreover its results may be difficult to interpret and its methodology is often qualified as poor. Manpower program evaluation therefore has a bad reputation in many circles. This surely limits its use as a change agent. While further work is needed and will surely be done to ameliorate the quality of evaluation research methodology, there is also work to be done to favorize a better consumption of evaluation results.

## LE TRAVAIL FÉMININ

### XXI<sup>e</sup> Congrès des Relations industrielles de l'université Laval

Évolution de la participation des femmes au monde du travail (JEAN-PIERRE DESPRÉS). La réglementation du travail féminin (RÉJEANNE COLAS). Aspects sociologiques du travail féminin (GÉRARD FORTIN). Aspects économiques de la participation des femmes au marché du travail (JACQUES ST-LAURENT). Le travail féminin — Témoignages (KATHLEEN FRANCOEUR, YOLANDE BOISSINOT, LOUISE MARCIL, DOROTHY PERTUISET, CHARLES LEBRUN, JACQUES VILLENEUVE, CLAUDE DUHAMEL, JEAN SIROIS). Participation des femmes aux mouvements syndicaux (YVETTE CHARPENTIER, LUCIE DAGENAIS). La femme dans le monde du travail d'aujourd'hui (PIERRETTE SARTIN).

*6 x 9, 178 pages, broché, \$4.00*

**LES PRESSES DE L'UNIVERSITÉ LAVAL**

**C.P. 2447, Québec 2.**

**Tél.: 656-2590**