

Évaluation de programme et recherche évaluative : des activités distinctes

Marthe Hurteau and Sylvain Houle

Volume 28, Number 3, 2005

URI: <https://id.erudit.org/iderudit/1087032ar>

DOI: <https://doi.org/10.7202/1087032ar>

[See table of contents](#)

Publisher(s)

ADMEE-Canada - Université Laval

ISSN

0823-3993 (print)

2368-2000 (digital)

[Explore this journal](#)

Cite this article

Hurteau, M. & Houle, S. (2005). Évaluation de programme et recherche évaluative : des activités distinctes. *Mesure et évaluation en éducation*, 28(3), 83–95. <https://doi.org/10.7202/1087032ar>

Article abstract

The eighties were rich in writings, which made distinctions between research, evaluative research and program evaluation. There was then a consensus, which is now called in question for various considerations. The present article aims to establish distinctions between program evaluation and evaluative research and thus to contribute to the development of a theoretical, methodological and empirical corpus harmonized which is necessary so that the evaluation of program develops like a scientific discipline (Dubois & Marceau, 2005). If these activities have several similarities – their reference to a theory of the program, the paradigms and methodologies – which moreover contribute to maintain confusions, their specific aims still offer the best perspective to understand in what manner they differ.

Évaluation de programme et recherche évaluative : des activités distinctes

Marthe Hurteau

Sylvain Houle

Université du Québec à Montréal

MOTS CLÉS: Évaluation de programme, recherche évaluative

Les années 80 ont produit de nombreux écrits traitant de la distinction entre la recherche, la recherche évaluative et l'évaluation de programme. S'il se dégageait alors un consensus, il est maintenant remis en question à cause d'enjeux de diverses natures. Le présent article vise à établir les distinctions plus particulièrement entre l'évaluation de programme et la recherche évaluative et ainsi contribuer au développement d'un corpus théorique qui est requis pour que l'évaluation de programme s'établisse comme une discipline scientifique (Dubois & Marceau, 2005). Si ces activités revêtent plusieurs similitudes – leur référence à une théorie du programme; les paradigmes et les méthodes – qui contribuent d'ailleurs à maintenir les confusions, il demeure que leurs visées réciproques offrent la meilleure perspective pour comprendre en quoi elles se distinguent.

KEY WORDS: Program evaluation, evaluative research

The eighties were rich in writings, which made distinctions between research, evaluative research and program evaluation. There was then a consensus, which is now called in question for various considerations. The present article aims to establish distinctions between program evaluation and evaluative research and thus to contribute to the development of a theoretical, methodological and empirical corpus harmonized which is necessary so that the evaluation of program develops like a scientific discipline (Dubois & Marceau, 2005). If these activities have several similarities – their reference to a theory of the program, the paradigms and methodologies – which moreover contribute to maintain confusions, their specific aims still offer the best perspective to understand in what manner they differ.

PALAVRAS-CHAVE: Avaliação de programas, investigação avaliativa

Os anos 80 foram ricos em escritos sobre a distinção entre a investigação, a investigação avaliativa e a avaliação de programas. O consenso então encontrado é, agora, posto em causa por desafios de naturezas diversas. O presente artigo visa estabelecer as distinções nomeadamente entre a avaliação de programas e a investigação avaliativa e, assim, contribuir para o desenvolvimento de um corpus teórico que é exigido para que a avaliação de programas se implante como uma disciplina científica (Dubois & Marceau, 2005). Ainda que ambas as actividades tenham várias similitudes – a sua referência a uma teoria do programa; os paradigmas e as metodologias – que, aliás, contribuem para manter as confusões, continua a ser válido que são os objectivos recíprocos que oferecem a melhor perspectiva para compreender em que é que as abordagens se distinguem.

Introduction

L'évaluation de programme a acquis ses lettres de noblesse à titre de discipline au cours des années 80, ce qui a favorisé le développement de son propre corpus théorique et méthodologique. Parmi les nombreux écrits de la documentation actuelle, certains se sont plus particulièrement attardés à différencier l'évaluation de programme de la recherche et de la recherche évaluative (Borich & Jemelka, 1982; Gephart, Ingle & Saretsky, 1973; Hurteau & Nadeau, 1987; Levin-Rozalis, 2003; Worthen & Sanders, 1973). Si la distinction entre l'évaluation de programme et la recherche s'avère évidente, une confusion persiste jusqu'à ce jour entre l'évaluation de programme et la recherche évaluative; cette confusion est devenue de plus en plus perceptible à travers les années, ce qui a conduit les chercheurs et les praticiens à utiliser ces termes de façon indifférenciée. Certains enjeux contribuent au maintien de cette situation, pour ne citer que ceux liés à l'obtention de fonds, qui incitent des chercheurs à qualifier leur démarche de «recherche évaluative» pour ainsi augmenter leurs chances de publier dans des revues scientifiques et d'obtenir des subventions traditionnellement allouées à la «recherche scientifique».

Note des auteurs: Toute correspondance peut être adressée par courriel à l'adresse suivante: [hurteau.marthe@uqam.ca]. Nous tenons à remercier monsieur Réjean Auger, professeur titulaire, ainsi que monsieur Renald Legendre, professeur émérite, tous deux rattachés au département d'éducation et de pédagogie de la Faculté des sciences de l'éducation (Université du Québec à Montréal). Leurs précieux conseils ainsi que les réflexions qu'a suscitées leur lecture de l'article auront contribué à le bonifier.

Le présent article vise à établir une distinction entre l'évaluation de programme et la recherche évaluative afin d'améliorer les pratiques et, à plus long terme, à contribuer à l'établissement d'un corpus théorique, méthodologique et empirique *harmonisé*, qui est préalable au développement de l'évaluation de programme comme discipline scientifique (Dubois & Marceau, 2005).

D'entrée de jeu, il appert que *les visées* respectives de ces deux activités offrent la perspective la plus intéressante pour les distinguer, comme en fait foi la figure 1.

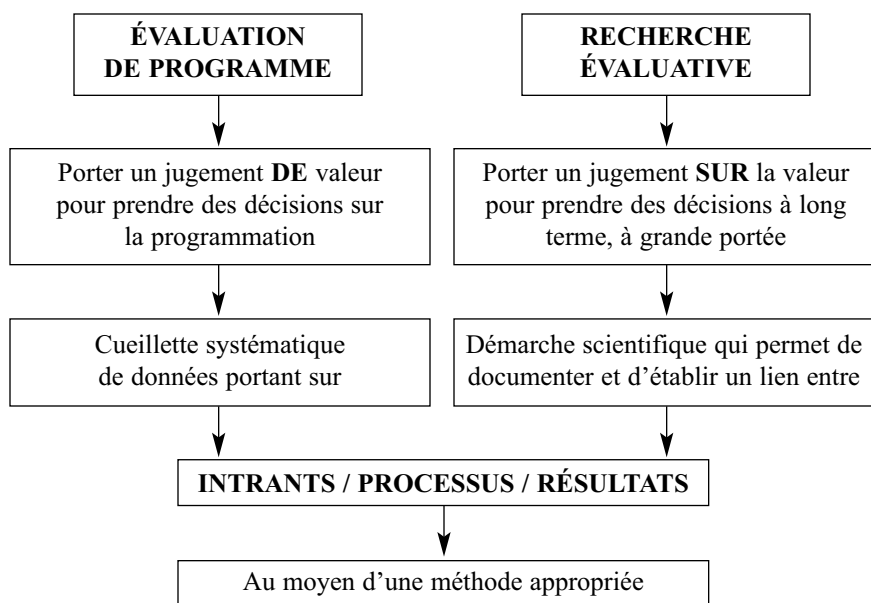


Figure 1. *Distinction entre l'évaluation programme et la recherche évaluative.*

Plus précisément, l'évaluation de programme est associée aux activités de gestion d'un programme. Elle a pour but d'éclairer les prises de décision alors que la recherche évaluative vise, quant à elle, à favoriser une compréhension en profondeur d'un programme afin d'envisager des décisions à grande portée et à long terme (généralisation). C'est d'ailleurs ce qui amène Gaudreau (1997) à avancer que l'évaluation de programme porte un jugement **DE** valeur sur un programme (fait-il ce qu'il devrait faire dans l'immédiat?) alors que la recherche évaluative porte un jugement **SUR** la valeur d'un programme (est-il suffisamment « bon » pour être généralisé?). Cette idée va être reprise et développée dans le texte.

Nous proposons au lecteur d’amorcer le présent exposé par une brève présentation du concept «programme» puisqu’il constitue l’objet de ces démarches. Nous décrirons par la suite chacune des activités de façon à mieux saisir leurs visées et nous terminerons en abordant les similitudes entre les deux activités puisqu’elles contribuent, dans une certaine mesure, à maintenir la confusion entre les termes.

Le concept de «programme»

Quelle que soit l’origine disciplinaire d’un programme, ce dernier revêt une finalité universelle qui consiste à améliorer le bien-être de clientèles ciblées (ex. : prévention d’une épidémie, réduction de la pauvreté, développement d’habiletés ou de compétences fondamentales ou professionnelles, etc.). Sa survie est directement liée aux deux fonctions qui lui sont allouées, à savoir, assurer la transformation des intrants en résultats souhaitables et maintenir des liens avec son environnement afin de bénéficier de ses ressources et de son soutien. Un programme constitue ainsi un système ouvert et systémique composé des dimensions qui lui sont propres – intrants, processus et résultats, environnement et rétroactions. La figure 2, inspirée des travaux de Chen (2005), illustre notre propos.

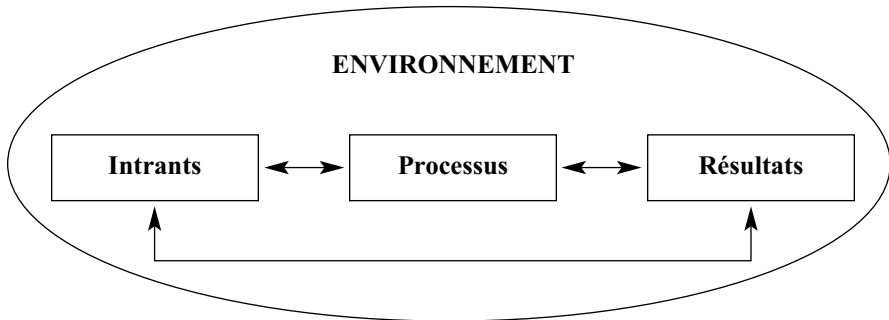


Figure 2. *Un programme : un système.*

Plus spécifiquement, les intrants font référence aux clients et à leurs besoins – qui sont transformés en objectifs – ainsi qu’aux ressources qui proviennent de l’environnement et qui sont essentielles au bon fonctionnement de la structure (personnel, équipement, finances, technologie, etc.). Le processus constitue cependant l’âme d’un programme puisqu’il assure la transformation annoncée en offrant des interventions – activités, traitements,

services – à la clientèle. En troisième lieu, les résultats (impacts, effets) rendent compte du degré de transformation atteint. L’environnement, quant à lui, se réfère autant aux ressources, tel que mentionné plus haut, qu’à l’ensemble des facteurs qui impose des balises, des contraintes au programme – normes sociales, les structures politiques et économiques, la législation, les groupes de pression, les intérêts de la société. Il appert que la définition même de l’environnement ne peut qu’alimenter une ambivalence des rapports entre ce dernier et les intrants. Finalement, le système de rétroactions permet d’apporter les ajustements requis au programme au fur et à mesure afin d’éviter qu’il ne se détériore. Cette composante introduit la notion d’évaluation de programme puisqu’elle consiste justement à fournir de l’information pertinente et adéquate pour alimenter le système de rétroactions.

L’évaluation de programme

L’évaluation de programme consiste en une cueillette systématique de données ayant pour objet d’étude le système de rétroactions et qui est effectuée en vue de porter un jugement de valeur, améliorer ou prendre des décisions concernant sa programmation future (Patton, 2002; Patton, in Alkin, 1990; Scriven, 1991; Stake, 2004). Tel que mentionné, elle s’inscrit dans une perspective de gestion d’un programme et il est plus aisé de comprendre la nature de ses visées en se penchant sur les auditoires cibles auxquels elle s’adresse ainsi que sur la nature même des jugements qu’elle génère.

Les auditoires cibles

L’évaluation de programme prend en considération les préoccupations des différents acteurs (*stakeholders*), tels les gestionnaires, les intervenants au sein du programme, les clientèles à qui s’adresse le programme, etc. C’est ce qui fait dire à Allard (1991) qu’elle opte pour une perspective qui se situe entre la science et la médiation.

La nature des jugements

Le concept de jugement formatif et sommatif introduit par Scriven (1991) s’avère une contribution inestimable dans le cadre de l’évaluation de programme puisqu’il qualifie la nature même des jugements qui émanent d’une telle démarche. Ainsi, le jugement formatif prend place *durant* le développement ou la mise en œuvre d’un programme et il permet de lui fournir une rétroaction constante ainsi que de cerner les améliorations qui devraient être apportées au programme en vue d’assurer ultimement son efficacité. Une fois

les activités du programme stabilisées ou complétées, le jugement porte alors sur son mérite ou sa valeur et revêt une portée sommative, une forme de sanction. Dans ce contexte, l'évaluation de programme vise avant tout à porter un jugement DE valeur sur les effets du programme, par exemple, pour prendre une décision sur les suites à lui donner. Il faut préciser qu'elle présume alors du lien logique qui unit les dimensions du programme – intrants, processus et résultats – pour attribuer les effets observés à l'intervention qui vient d'avoir lieu sans pour autant en faire la démonstration.

Cette activité s'adresse donc aux personnes qui gravitent directement autour du programme et elle génère de l'information pertinente pour prendre des décisions adéquates le concernant.

La recherche évaluative

La recherche évaluative vise à documenter en profondeur un programme, à démontrer le lien de causalité entre les impacts observés et l'intervention – processus – afin, ultimement, d'envisager une éventuelle généralisation ou application sur une grande échelle de ce programme (Alkin, 1990; Caro, 1971; Constandriopoulos, Champagne, Denis & Avargues, 2000; Patton, 2002; Rutman, 1977; Rossi & Freeman, 1985; Stake, 2004; Suchman, 1969; Weiss, 1972). La mise au point d'un médicament peut illustrer le propos et la transposition peut aisément être effectuée, par la suite, à tout type de programme. En effet, les chercheurs visent à établir, dans un premier temps, un lien direct et hors de tout doute entre la prise de médicament et les effets perçus. Par la suite, ils s'assurent que les effets observés par la prise du médicament s'appliquent à la population en général. Des résultats portent ainsi un jugement SUR la valeur du médicament et ils doivent être positifs dans les deux cas pour permettre sa mise en marché.

Les auditoires cibles

Par sa fonction même, la recherche évaluative s'éloigne des intérêts des acteurs qui gravitent directement autour du programme, de l'intervention, pour plutôt desservir les intérêts de preneurs de décision tels que les hautes directions, les fonctionnaires, les politiciens, etc.

La nature des jugements

La recherche évaluative s'intéresse à la faisabilité du programme (les ressources disponibles par rapport aux ressources requises, les coûts impliqués, les éventuels impacts inattendus et les effets pervers, les variables intermédiaires, etc.) mais, le plus important, dans la perspective d'acquérir une compréhension approfondie du programme, d'assurer sa cohérence interne – causalité – ainsi que sa validité externe – généralisation. Ces deux notions sont maintenant développées.

La causalité

La validité interne assure effectivement la cohérence logique du programme (*liens de causalité entre les différentes composantes du programme et l'impact à long terme – changement de condition de la population visée par le programme*). Cette démarche essentielle est préalable à toute généralisation ou *généralisabilité* des résultats. Dans cette perspective, les intrants et le processus du programme sont envisagés comme des variables indépendantes qui expliquent les impacts observés, eux-mêmes considérés comme la variable dépendante (Blalock, 1999). Rossi, Lipsey et Freeman (2004) soulignent l'importance de s'assurer que le programme est suffisamment implanté afin de se prêter à un tel exercice, au risque de commettre un impair dans les conclusions qui pourraient émaner d'une telle démarche.

La généralisation ou généralisabilité des résultats

Alors que la généralisation est établie au moyen d'indices statistiques inférentiels, la *généralisabilité*, quant à elle, porte un regard phénoménologique sur les qualités de transférabilité du programme (*capacité de généralisation, de généralisabilité* – traduction libre du terme *generalisability*). Dans un cas comme dans l'autre, la recherche évaluative vise à garantir que les résultats générés concernant un programme peuvent être considérés représentatifs et, par le fait même, potentiellement généralisables (Stake, 2004).

Similitudes entre les deux activités

Si l'évaluation de programme et la recherche évaluative constituent des activités distinctes, il n'en demeure pas moins qu'elles revêtent certaines similitudes qui se résument dans les trois points suivants: leur référence à une théorie du programme, les paradigmes au sein desquels ces démarches s'inscrivent ainsi que les considérations méthodologiques. La théorie du programme

ne sera pas développée davantage puisqu'elle l'a été dans le cadre du premier point portant sur le programme. Mentionnons toutefois que le modèle logique constitue l'application la plus reconnue d'une théorie du programme (Chen, 2005 ; McLaughlin & Jordan, 1991 ; Wholey, 1983, 1987).

Les paradigmes

Contrairement au point de vue adopté par Gaudreau (1997), nous croyons que les paradigmes sont communs aux deux activités puisqu'ils offrent une façon de capter la réalité. Le présent propos sera illustré au moyen des deux paradigmes les plus courants, soit le positiviste classique et postpositiviste, et l'interprétatif. Le tableau 1, dont le contenu s'inspire des auteurs tels que Lecomte et Rutman (1982), Guba et Lincoln (1989) et Lincoln (1990), met en évidence ces distinctions.

Tableau 1
Distinctions entre les paradigmes positiviste classique et postpositiviste et le paradigme interprétatif

<i>Dimensions</i>	<i>Paradigmes positiviste classique et postpositiviste</i>	<i>Paradigme interprétatif</i>
Délimitation des phénomènes à évaluer	Élémentaliste	Holistique
Perception de la réalité à évaluer	Vision consensuelle de la réalité	Réalité qui se négocie jusqu'à l'obtention d'un consensus
Ampleur de la généralisation	Objective	Subjective
Perspective du chercheur	«Extrospective» : Se concentre sur la correspondance entre les résultats et la réalité	Introspective : Découvre et vérifie continuellement la réalité

Ainsi, les paradigmes positiviste classique et postpositiviste considèrent la réalité comme pouvant être décrite comme des évidences clairement énoncées alors que le paradigme interprétatif adopte plutôt une perspective non déterministe qui reconnaît que les buts et les objectifs des programmes peuvent être contradictoires, que les effets sont rarement clairs, que l'hétérogénéité structurelle et contextuelle des programmes rend difficile leur utilisation directe dans d'autres sites (Shadish, Cook & Leviton, 1991). Le lecteur comprendra qu'il ne peut exister un « bon » paradigme mais plutôt un paradigme qui est

approprié et justifié par rapport à un contexte précis. Ce constat nous amène à nous demander quels auraient été les tenants et aboutissants des évaluations de programme comme *Push-Excel*, *Head Start* et *Follow Through*, qui se sont d'ailleurs révélées un échec monumental, si elles s'étaient inscrites dans une perspective interprétative plutôt que positiviste. En effet, ces programmes, qui visaient à lutter contre la pauvreté en offrant des interventions compensatoires, étaient en phase de développement lorsqu'ils ont fait l'objet d'une évaluation. Par la suite, les variables retenues pour appuyer le jugement sommatif – résultats à des épreuves standardisées en anglais et en mathématiques – se sont révélées peu significatives par rapport à d'autres qui ont émergé en cours de route, pour ne nommer que l'établissement de liens de confiance entre les parents et l'école et la création d'un tissu social important au sein des communautés.

Méthode

Même si les auteurs considèrent la méthode comme une entité indépendante des paradigmes (Cook & Reichardt, 1979; Guba & Lincoln, 1989; Schwandt, 1990), il n'en demeure pas moins que ces concepts sont étroitement liés dans la pratique. Stake (2004) adopte une position qui s'inscrit dans une perspective positiviste par rapport à la recherche évaluative en avançant que celle-ci requiert la mise en place d'une instrumentation qui répond à des conditions spécifiques :

- avoir recours à un protocole expérimental qui inclut la population des cohortes, les modalités d'application, les lieux d'application, etc., du programme. Cette population doit se traduire aussi par un nombre d'observations suffisamment grand pour répondre aux exigences docimologiques afin de réaliser des analyses inférentielles ;
- avoir recours à un protocole expérimental qui assure la validité interne (lien causal) et la validité externe (généralisation) des résultats.

Cette position s'inscrit dans une perspective positiviste, de confirmation, et d'autres approches méthodologiques pourraient être plus appropriées dans une perspective de développement.

De façon générale, la méthode doit tenir à ce que la crédibilité de toute démarche réside en sa capacité de se conformer à des standards élevés (Boruch, Synder & DeMoya, 2000). Dans cette perspective, les approches quantitatives ou qualitatives et le recours à des devis expérimentaux ou non, à des questionnaires ou à des entrevues ouvertes, sont autant d'options qui s'offrent et qui doivent être envisagées à la lumière de l'information requise et du contexte.

En conclusion, le traitement des similitudes entre ces deux activités permet de comprendre l'émergence de certaines confusions, la distinction entre ces deux activités se référant à leurs visées respectives et non pas à des considérations méthodologiques.

Discussion

Alors que l'évaluation de programme s'inscrit dans le cadre des activités de gestion d'un programme, la recherche évaluative vise à documenter le programme (recherche) afin de confirmer sa contribution significative (lien de causalité entre le processus et les résultats générés) et de déterminer dans quelle mesure il devient possible de l'implanter à une grande échelle, à long terme (généralisation ou *généralisabilité*). La nature même de ces visées va nécessairement avoir un impact significatif sur la façon d'envisager le mandat et de l'accomplir. Illustrons notre propos au moyen de deux scénarios élaborés à partir d'une même intervention qui vise à intégrer des enfants présentant des troubles de comportement dans les classes régulières. Certaines de leurs caractéristiques sont résumées dans le tableau 2 afin d'en faciliter la compréhension.

Tableau 2

Distinctions entre les deux scénarios d'une même intervention portant sur l'intégration d'enfants

<i>Dimensions</i>	<i>Projet éducatif</i>	<i>Projet pilote</i>
Visées	Atteindre les cibles établies dans la planification stratégique	Établir la pertinence de l'intervention pour l'implanter sur le plan provincial
Acteurs	L'ensemble du personnel de l'école, les enfants et les parents visés ainsi que les commissions scolaires	Les répondants du ministère de l'Éducation et des commissions scolaires
Méthode	Documenter les indicateurs de performance établis dans le cadre de la planification stratégique	Assurer les validités interne (causalité) et externe (généralisation)
Ampleur de la démarche	Locale	Provinciale

Le premier cas s'inscrit dans la perspective d'un projet éducatif au sein d'une école qui requiert une évaluation axée sur les résultats, à savoir le niveau d'atteinte des cibles établies dans le cadre de la planification stratégique. Le second cas, quant à lui, s'inscrit dans le cadre d'un projet pilote

qui vise à établir une intervention qui pourrait être appliquée dans l'ensemble des commissions scolaires de la province. Le lecteur sera à même de constater que la première démarche s'inscrit dans la perspective d'une évaluation de programme alors que la seconde s'oriente vers la recherche évaluative. Si cet exemple fait appel à des positions extrêmes afin de mettre en évidence les distinctions, il est aussi évident que certaines évaluations de programme peuvent revêtir des caractéristiques qui les apparentent à la recherche évaluative, et vice-versa. Toujours dans cet esprit, l'importance de l'objet à évaluer n'influence pas la nature des activités. Ainsi, l'État québécois entreprend actuellement une évaluation sur sa réforme en éducation et, même si cette opération est d'envergure puisqu'elle implique l'ensemble des commissions scolaires de la province, elle s'inscrit toujours dans la perspective d'une évaluation de programme parce qu'elle vise à déterminer le niveau d'atteinte des effets escomptés. Par contre, si la réforme avait préalablement fait l'objet d'un projet pilote afin d'établir un lien de causalité entre les changements introduits sur le plan des interventions pédagogiques et les effets observés, avant d'être introduite sur l'ensemble du territoire québécois, nous aurions alors pu effectivement parler d'une recherche évaluative.

Conclusion

Le présent article visait à établir la distinction entre l'évaluation de programme et la recherche évaluative. En aucun cas, il n'a eu la prétention de porter un jugement sur leur contribution respective dans la pratique puisque, chacune à leur façon, elles fournissent une réponse appropriée et impérative à des besoins d'information distincts.

Le développement d'un corpus théorique, méthodologique et empirique *harmonisé* en évaluation de programme ne peut tolérer des ambiguïtés en ce qui concerne les termes de référence. Dans ce sens, il est important d'établir une distinction entre l'évaluation de programme et la recherche évaluative, ou d'effectuer un choix entre les deux terminologies, parce qu'il devient absurde de faire référence à une même réalité en ayant recours à deux termes différents. Le présent exercice, tout comme celui effectué par Gaudreau (1997), constitue un effort pour distinguer ces termes afin de bonifier le corpus théorique disciplinaire. Il aura permis de mettre en évidence leur contribution respective, ces deux activités se référant à des pratiques distinctes dont le meilleur élément discriminant demeure leurs visées respectives.

RÉFÉRENCES

- Allard, D. (1991). La régulation de la science et de la technique dans les modèles d'évaluation. In J. Légaré & A. Demers (éds), *L'évaluation sociale: savoirs, éthique, méthodes* (pp. 239-272). Laval (Québec): Éditions du Méridien.
- Alkin, C.A. (1990) (éd.). *Debates on evaluation*. Newbury Park: Sage Publications.
- Blalock, A.B. (1999). Evaluation research and the performance management movement. From estrangement to useful integration? *Evaluation*, 5, 117-149.
- Borich, G.P., & Jemelka, R.D. (1982). *Program and systems, An evaluation perspective*. The Educational Technology Series. Toronto: Academic Press.
- Boruch, R., Synder, B., & DeMoya, D. (2000). The importance of randomized field trials. *Crime and Delinquency*, 46(2), 156-180.
- Caro, F.G. (éd.) (1971). *Readings in evaluation research*. New York: Russell Sage Foundation.
- Chen, H.-T. (2005). *Practical program evaluation*. Thousand Oaks: Sage Publications.
- Cook, T.D., & Reichardt, D.T. (éds) (1979). *Qualitative and quantitative methods in evaluation research*. Beverly Hills: Sage Publications.
- Constandriopoulos, A.-P., Champagne, F., Denis, J.-L., & Avargues, M.-C. (2000). L'évaluation dans le domaine de la santé: concepts et méthodes. *Revue d'épidémiologie et santé publique*, 48, 517-539.
- Dubois, N., & Marceau, R. (2005). Un état des lieux théoriques de l'évaluation: une discipline à la remorque d'une révolution scientifique qui n'en finit pas. *La Revue canadienne en évaluation de programm*, 20(1), 1-36.
- Gaudreau, L. (1997). De l'évaluation de programme à la recherche évaluative. In L. Sauvé, A.T. Barba, M. Sato & E.Y. Castillo (éds), *De la evaluacion de programas a la investigacion evaluative* (pp. 289-296). Actas des seminario inernacional de investigacion-formacion EDAMAZ, Educacion ambiental en Amazonia. Montréal: CIRADE, Université du Québec.
- Gephart, W.J., Ingle, R.B., & Saretsky, G. (1973). *Similarities and differences in the reseach and evaluation processes*. Bloomington, Ind: Phi Delta Kappa.
- Guba, E.G., & Lincoln, Y.S. (1989). *Fourth generation evaluation*. Newbury Park: Sage Publications.
- Hurteau, M., & Nadeau, M.A. (1987). Évaluation de programmes, recherche évaluative et recherche. *Revue canadienne de psycho-éducation*, 12(3), 404-416.
- Lecomte, R., & Rutman, L. (1982). *Introduction aux méthodes de recherche évaluative*. Québec: Presses de l'Université Laval.
- Levin-Rozalis, M. (2003). Evaluation and research: Differences and similarities. *La Revue canadienne en évaluation de programme*, 18(2), 1-31.
- Lincoln, Y.S. (1990). The making of a constructivist: A remembrance of transformations in the past. In E.G. Guba (éd.), *The Paradigm Dialog* (pp. 46-66). Newbury Park: Sage Publications.
- McLaughlin, J.A., & Jordan, G.B. (1991). Logic models: A tool for telling your performance story. *Evaluation and Program Planning*, 22(1), 65-72.
- Patton, M.Q. (2002). *Qualitative research evaluation methods* (third edition). Thousand Oaks: Sage Publications.

- Rossi, P., & Freeman, H. (1985). *Evaluation. A systematic approach* (third edition). Beverly Hills, CA: Sage Publications.
- Rossi, P., Lipsey, M.W., & Freeman, H. (2004). *Evaluation. A systematic approach* (seventh edition). Thousand Oaks: Sage Publications.
- Rutman, L. (1977). *Evaluation research methods: A basic guide*. Beverly Drive: Sage Publications.
- Schwandt, T.R. (1990). Paths to inquiry in the social disciplines: Scientific, constructivist, and critical theory methodologies. In E.G. Guba (éd.), *The paradigm dialog* (pp. 258-276). Newbury Park: Sage Publications.
- Scriven, M. (1991). *Evaluation thesaurus*. (fourth edition). Newbury Park: Sage Publications.
- Shadish, W.E., Cook, T.D., Leviton, L.C. (1991). *Foundations of program evaluation. Theories of practice*. Newbury Park: Sage Publications.
- Stake, R.E. (2004). *Standards-based and responsive evaluation*. Thousand Oaks: Sage Publications.
- Suchman, E. A. (1969). Evaluating educational programs. *The Urban Review*, 3(4), 15-17.
- Weiss, C.H. (1972). *Evaluation research. methods for assessing program effectiveness*. Englewood Cliffs, NJ: Prentice-Hall Inc.
- Wholey, J.S. (1983). *Evaluation and effective public management*. New York: Little Brown.
- Wholey, J.S. (1987). Evaluability assessment: Developing program theory. In L. Bickman (éd.), *Using program theory in evaluation*. New directions for program evaluation (33). San Francisco: Jossey-Bass.
- Worthen, B.L., & Sanders, J.R. (1973). *Educational evaluation. Theory and practice*. Belmont, California: A. Charles & A. Jones Publication.