

L'étude du raisonnement dans les pratiques évaluatives

Sheila A. Arens

Volume 29, Number 3, 2006

La relation entre la théorie et la pratique en évaluation de programme : dialogue ET monologue

URI: <https://id.erudit.org/iderudit/1086393ar>

DOI: <https://doi.org/10.7202/1086393ar>

[See table of contents](#)

Publisher(s)

ADMEE-Canada - Université Laval

ISSN

0823-3993 (print)

2368-2000 (digital)

[Explore this journal](#)

Cite this article

Arens, S. A. (2006). L'étude du raisonnement dans les pratiques évaluatives. *Mesure et évaluation en éducation*, 29(3), 45–56.
<https://doi.org/10.7202/1086393ar>

Article abstract

The discipline of program evaluation has invested more energy in developing methodologies during the past twenty years than to understand certain strategic stakes such as the evaluative reasoning which supports the practice. Some authors, to cite only Fournier and Scriven, are particularly interested by the influence of evaluative reasoning on conclusions and justifications but, so far, these concerns do not seem to have aroused much interest. This fact is all the more surprising given that the transition from the description stage to one of judgment constitutes the cornerstone of any good-quality program evaluation. It is from this point of view that this article approaches the question of the reasoning underlying any evaluative process. More precisely, the author studies the way in which evaluators produce and justify their judgment.

L'étude du raisonnement dans les pratiques évaluatives

Sheila A. Arens

Mid-continent Research for Education and Learning

MOTS CLÉS : Évaluation de programme, raisonnement évaluatif, jugement

La discipline de l'évaluation de programme a déployé plus d'énergie à développer les méthodologies au cours des vingt dernières années qu'à comprendre certains enjeux stratégiques, pour ne citer que le raisonnement évaluatif qui soutient la pratique. Certains auteurs, par exemple Fournier et Scriven, se sont plus particulièrement intéressés à l'influence du raisonnement évaluatif sur les conclusions ainsi que sur les justifications mais cela ne semble pas avoir suscité beaucoup d'intérêt jusqu'à ce jour. Ce constat est d'autant plus surprenant que le passage de la description au jugement constitue la pierre angulaire de toute évaluation de programme de qualité. C'est dans cette perspective que le présent article aborde la question du raisonnement sous-jacent à toute démarche évaluative. Plus précisément, l'auteure aborde la façon dont les évaluateurs produisent et justifient les jugements émis.

KEY WORDS : Program evaluation, evaluative reasoning, judgment

The discipline of program evaluation has invested more energy in developing methodologies during the past twenty years than to understand certain strategic stakes such as the evaluative reasoning which supports the practice. Some authors, to cite only Fournier and Scriven, are particularly interested by the influence of evaluative reasoning on conclusions and justifications but, so far, these concerns do not seem to have aroused much interest. This fact is all the more surprising given that the transition from the description stage to one of judgment constitutes the cornerstone of any good-quality program evaluation. It is from this point of view that this article approaches the question of the reasoning underlying any evaluative process. More precisely, the author studies the way in which evaluators produce and justify their judgment.

Note – Cet article a été traduit et adapté par M. Hurteau et S. Mongiat. Toute correspondance peut être adressée par courriel soit à l'auteure [sarens@mcrel.org], soit à la responsable de la traduction [hurteau.marthe@uqam.ca].

PALAVRAS-CHAVE: Avaliação de programas, raciocínio avaliativo, juízo

A disciplina de avaliação de programas despendeu mais energia a desenvolver as metodologias ao longo dos últimos vinte anos do que a compreender determinados desafios estratégicos, desde logo o raciocínio avaliativo que sustenta a prática. Alguns autores, por exemplo Fournier e Scriven, já não se interessam pela influência que exerce o raciocínio avaliativo sobre as conclusões, assim como com as justificações. Porém, isto não parece ter despertado muito interesse até ao momento. Esta constatação é tanto mais surpreendente quanto a passagem da descrição ao juízo constitui a pedra angular de qualquer avaliação de programas de qualidade. É nesta perspectiva que o presente artigo aborda a questão do raciocínio subjacente a qualquer procedimento de avaliação. A autora aborda, mais precisamente, a forma como os avaliadores produzem e justificam os juízos emitidos.

Introduction

L'évaluation de programme se définit comme une activité scientifique visant à établir la valeur ou le mérite d'un programme en s'appuyant sur des conclusions précises, justifiables, acceptables et légitimes (House, 1993), en bref, valides (Fournier & Smith, 1993). Au cours du dernier siècle, son développement s'est inspiré de plusieurs courants de pensées (Sirotnik & Oakes, 1990) qui ont provoqué des changements tant structuraux que conceptuels et philosophiques. Il n'en demeure pas moins que si le raisonnement conduisant à des conclusions évaluatives demeure une préoccupation actuelle, comme en témoignent les travaux de Fournier (1995), plusieurs questionnements s'y référant demeurent toujours sans réponse. Citons, à titre d'exemple, celui concernant les modalités qui permettent d'obtenir des données de qualité ainsi que leur transformation en évidences et l'agencement de ces évidences en vue de formuler des conclusions évaluatives. À ce jour, ce questionnement demeure une préoccupation importante. La présente étude s'inscrit dans ce créneau de réflexion en tentant de comprendre la pratique évaluative et, plus spécifiquement, de quelle façon le jugement de valeur, ou les conclusions évaluatives, s'avèrent-ils valides. Cet examen du raisonnement évaluatif permettra à l'auteur d'analyser :

- la façon dont les évaluateurs présentent leurs conclusions évaluatives (catégorisation des évaluations) ;
- les types d'évidences et de déclarations présentés pour les appuyer (les constituants de son argumentation) et, finalement,
- si celles-ci sont justifiées et cautionnées (examen de leur validité).

Le cadre conceptuel

La validité d'un jugement de valeur en évaluation de programme repose sur la qualité de son argumentation. Cette dernière s'appuie sur le raisonnement évaluatif qui met en relation les évidences (résultats d'analyse) et les déclarations (tirées de ces résultats) (Fournier, 1995). Le jugement évaluatif repose sur deux processus intimement liés, le premier qui est basé sur des évidences et le second basé sur des inférences. Pour cette auteure, l'étude de ces deux processus permet de considérer le raisonnement évaluatif comme un moyen de discriminer les arguments pauvres des riches. Le raisonnement évaluatif – qui assure une argumentation solide – ainsi que ces deux processus seront développés dans le point présent, et ce, afin de comprendre comment un jugement de valeur valide est produit.

Le raisonnement évaluatif

Pour Fournier (1995) et Valovirta (2002), le raisonnement évaluatif consiste à tirer des conclusions à partir d'évidences qui sont elles-mêmes reliées par un argumentaire. L'auteure s'appuie sur les travaux de Toulmin (1958) pour expliquer la démarche inhérente au raisonnement, tel qu'illustrée à la figure 1.

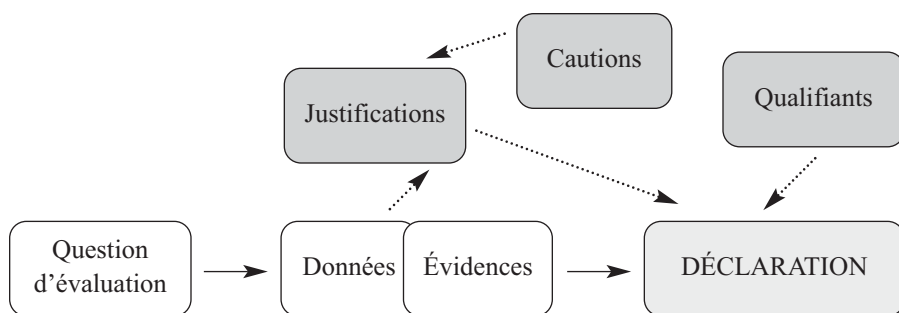


Figure 1 *La démarche sous-jacente au raisonnement évaluatif.*

La justification prend la forme de suppositions ou de prémisses qui relient logiquement une conclusion évaluative à l'évidence sur laquelle elle s'appuie. La caution, quant à elle, n'est utilisée que dans le cas où l'évaluateur a l'impression que son jugement (déclaration) pourrait être remis en question (Booth, Colomb & Williams, 2003). Elle peut prendre la forme de règles, lois, principes, formules mathématiques (Toulmin, 1958) ou s'appuyer sur l'expérience, l'autorité, un système de croyance ou de connaissance, une entente

générale, méthodologie, ou tout simplement la foi (Booth, Colomb & Williams, 2003). Finalement, le qualifiant consiste en un énoncé qui fournit les conditions dans lesquelles la justification peut-être considérée comme valide (quand, lorsque, à l'exception, etc.). Toujours selon Toulmin (1958), la justification et le qualifiant constituent un soutien additionnel auquel on a recours lorsque les justifications ne semblent pas suffisamment fondées ou sont susceptibles d'être remises en question. Tel que mentionné, cette démarche s'appuie sur deux types de processus (Fournier, 1995). Les deux processus sous-tendant le raisonnement évaluatif sont maintenant développés.

Le processus basé sur les évidences

Ce premier processus se réfère aux composantes suivantes du raisonnement évaluatif:

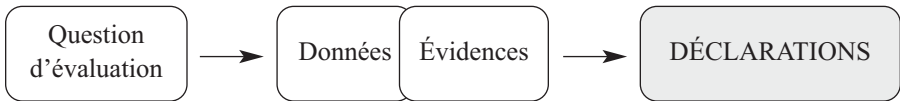


Figure 2 *Le processus basé sur les évidences*

Le processus basé sur les évidences consiste à recueillir et à rassembler les données en évidences. Il décrit la relation logique qui existe entre la question d'évaluation, à l'origine de la démarche, la prise de données qui construit des évidences sur lesquelles s'appuieront l'éventuel jugement de valeur ou les conclusions évaluatives.

Second processus basé sur les inférences

Le processus basé sur les inférences est celui à l'intérieur duquel il faut faire et supporter des inférences. Alors que le processus précédent s'appuie sur des faits, le présent processus est d'un autre ordre puisqu'il implique une opération mentale qui consiste à généraliser, et ce, en s'appuyant sur une information relativement limitée. Il est alors possible de produire des déclarations concernant un aspect du programme et d'autres, plus globales, qui résumeront toutes ces dernières déclarations pour produire une déclaration synthèse. Cette déclaration synthèse peut également s'appeler la conclusion évaluative, qui consiste à porter un jugement global sur le programme ou sur une de ses composantes. Il existe une variété d'approches pour arriver à des déclarations synthétiques. Pendant que certains considèrent des moyens d'association quantitative méta-analytique (ex., ceux utilisés pour What Works Clearinghouse,

2003) utiles pour synthétiser des déclarations évaluatives, d'autres sont plus sceptiques concernant la possibilité que ces méthodes traditionnelles puissent être défendables parmi les multiples déclarations divergentes liées à des valeurs et à des opinions différentes.

Ce traitement intellectuel, qui est difficilement retraceable, constitue en quelque sorte le cœur de l'argumentaire sur lequel s'appuie le passage des évidences aux conclusions évaluatives. Il existe plusieurs façons d'inférer une déclaration à partir d'évidences mais aucune d'elles ne semble réellement avoir été validée à ce jour (Schwandt, 2004). Ainsi, ce second processus, qui s'appuie sur le premier, implique plus spécifiquement les composantes suivantes du raisonnement évaluatif :

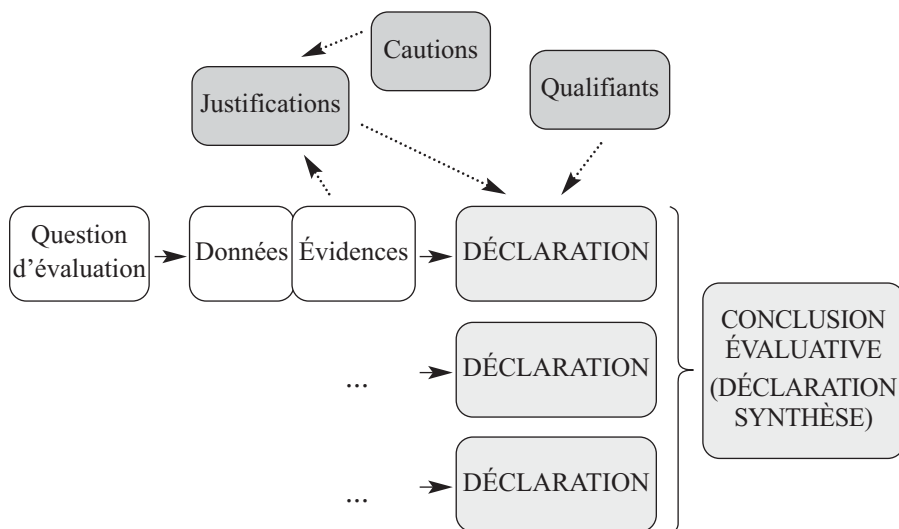


Figure 3 *Le processus basé sur les inférences*

Ainsi, les conclusions évaluatives constituent-elles l'aboutissement logique du processus et revêtent-elles plus spécifiquement la forme d'un jugement de valeur dans le contexte de l'évaluation de programme. Plusieurs auteurs (House 1995; House & Hawe, 1999; Scriven, 1991; What Works Clearinghouse, 2003) suggèrent des modalités pour y parvenir mais aucune ne semble efficace (Schwandt, 2002). Cela ne veut pas dire pour autant que les conclusions évaluatives doivent être considérées comme invalides : il revient à l'évaluateur de relater les faits concernant le programme de telle sorte que les personnes intéressées le « reconnaissent » et, par le fait même, accordent une crédibilité au jugement qui émane de la démarche. C'est ce qui fait dire à Biesta et

Burbules (2006) que le jugement concernant un programme est toujours relatif à une situation donnée et qu'il est donc provisoire. Il en est de même pour les recommandations qui en découlent (Fournier & Smith, 1993). Elles vont même parfois au-delà des évidences et font appel à des extrapolations qui requièrent une connaissance approfondie du programme (Weiss, 1998).

L'argumentation requise pour assurer un jugement valide étant issue du raisonnement évaluatif, la présente étude analyse plus spécifiquement le recours à celui-ci dans le cadre de pratiques évaluatives.

Méthodologie

La présente étude constitue une description des pratiques évaluatives en s'appuyant sur un cadre conceptuel circonscrit.

Échantillonnage

Pour ce faire, il a été décidé d'analyser des rapports d'évaluation jugés comme des «pratiques exemplaires» par le *American Journal of Evaluation*. Cinq des sept rapports ont été retenus, les deux écartés ne fournissant pas l'ensemble de l'information requise par la présente démarche.

Méthode

L'étude constitue une analyse du contenu des différents rapports à partir d'une grille élaborée à cette fin. L'élaboration de cette grille fait appel aux principes émis par Neuendorf (2001), et elle inclut les éléments suivants :

- Éléments liés au processus basé sur les évidences et, plus spécifiquement, la description :
 - du programme,
 - des questions,
 - des méthodes
 - et des données ainsi que des évidences recueillies.
- Éléments liés au processus basé sur les inférences et, plus spécifiquement :
 - le type et la structure de l'argumentation :
 - la façon dont les évidences sont positionnées en relation avec les déclarations,
 - dans quelle mesure les évidences supportent les déclarations et la façon dont elles y parviennent,
 - si les qualifiants sont évidents ;

- l'enlignement des critères, des questions d'évaluation et des objectifs de l'évaluation,
- la façon de parvenir aux conclusions, leur justification et cautionnement,
- la présence de recommandations qui sont en lien avec l'argumentation et les évidences fournies.

Les données ont été générées par un processus itératif d'analyses de contenu au moyen de la grille.

Résultats

Plusieurs constats sont issus de la présente étude.

Premier constat d'ordre général: La démarche permettant d'arriver à des déclarations synthétiques ou à des conclusions évaluatives à partir des évidences n'est pas clairement saisissable à l'intérieur des rapports d'évaluation. Sa justification est douteuse, ce qui laisse place à l'interprétation du lecteur.

S'il existe une variété d'approches documentées pour synthétiser les différentes déclarations en une conclusion globale (les moyens d'association quantitative, méta-analytique comme ceux utilisés pour What Works Clearinghouse, (2003), il n'en demeure pas moins que les évaluateurs ne les ont pas utilisées. En effet, elles ne sont aucunement documentées, et ce, même si de nombreux auteurs, dont Scriven, soulignent l'importance de décrire de quelle façon les résultats sont combinés pour produire des conclusions. Les évaluateurs devraient toujours être en mesure d'expliquer de quelle façon les résultats ont permis de générer les conclusions. Cependant, nous remarquons dans les faits que la démarche évaluative n'est pas linéaire, ce qui implique que les évaluateurs recueillent des évidences, font des déclarations, élaborent une synthèse (jugement) qui repose autant sur les données prévues au point de départ que sur celles qui sont recueillies en cours de route. Ce canevas du raisonnement n'étant pas définitif avant la fin de la démarche, il ne peut donc être communiqué et le lecteur doit ainsi se fier au fait que l'évaluateur a pris en considération l'ensemble des évidences pour émettre ses conclusions, une reproduction de la démarche à des fins d'assurer la fidélité étant impossible dans les circonstances.

Deuxième constat : Si les critères peuvent être aisément retracés, les standards, quant à eux, sont la plupart du temps absents ou ils prennent la forme de standards comparatifs qui s'expliquent par les choix méthodologiques effectués.

Ainsi, même si les critères ne sont pas formellement établis, ils peuvent être retracés au moyen des questions et des visées de l'évaluation qui sont clairement énoncées. Ceci implique cependant une transformation du processus évaluatif, tel qu'illustré à la figure 4. Le lecteur constatera que les questions sous-jacentes encadrent la démarche, créant ainsi une profonde mutation du processus.

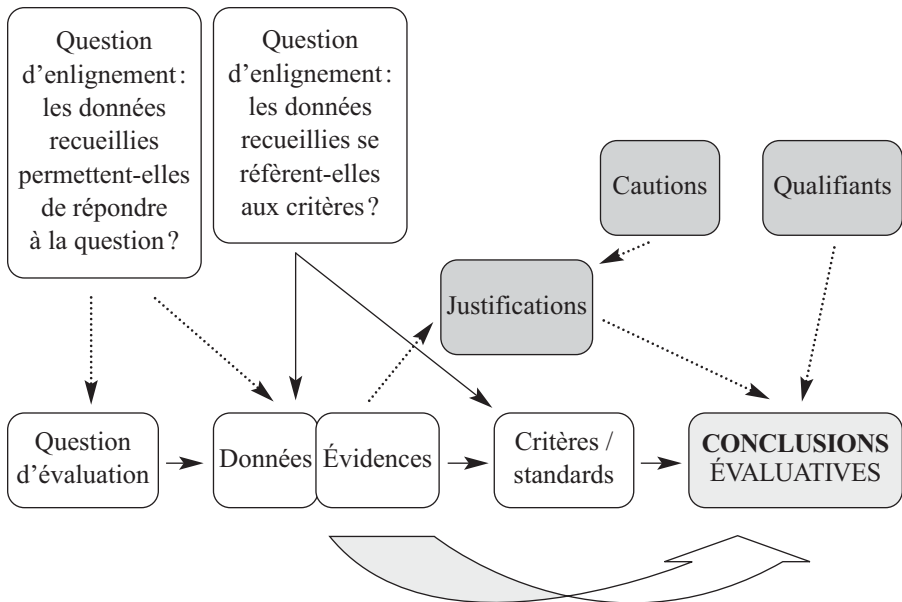


Figure 4 *Version révisée de la démarche sous-jacente au raisonnement évaluatif.*

Les standards sont donc absents, ou ils prennent la forme de standards comparatifs si l'évaluateur compare les résultats obtenus avec ceux d'un autre groupe (schémas expérimentaux ou quasi-expérimentaux). Cependant, dans ces cas, les standards ne spécifient pas toutes les précisions méthodologiques supplémentaires concernant les détails des tests et le seuil de signification statistique requis pour faire pencher le jugement d'un côté ou de l'autre. Ce

constat a un impact direct sur la crédibilité des déclarations et du jugement qui découlent de cette démarche. Aussi, lorsque la norme est inexistante, les évaluateurs sont réduits à qualifier la performance du programme au moyen d'expressions telles «la plupart du temps», «certains», «plusieurs». Non seulement est-il difficile de porter un jugement crédible à partir de qualificatifs aussi flous mais, en plus, l'information est insuffisante pour étoffer le dit jugement. En effet, nous pourrions nous questionner, à titre d'exemple, si la différence est réellement significative sur le terrain et s'il existe des effets secondaires importants qui devraient être pris en considération pour formuler le jugement ainsi que les recommandations.

L'ensemble des considérants sous-jacents au deuxième constat a donc un impact sur la crédibilité aux déclarations formulées et, par le fait même, sur la crédibilité du jugement qui en résulte.

Troisième constat: Les recommandations s'appuient davantage sur des justifications basées sur des autorités dans la matière que sur des évidences issues de déclarations. Les recommandations qui en résultent sont donc plus des faits qui suggèrent que les changements recommandés entraîneront automatiquement une amélioration sans être basés sur les résultats de l'évaluation.

Les recommandations ne se sont cependant pas révélées supportées par les évidences ni par les déclarations qui ont servi à générer la conclusion évaluative. Ce constat semble inquiétant puisque les recommandations sont des déclarations possédant une grande force et qu'elles sont à l'origine des changements structureaux relatifs aux programmes. Certains rapports, provenant de cas exemplaires, amènent même le lecteur à se demander si les changements suggérés seraient bénéfiques pour le programme, tant ils ne sont pas appuyés.

Les recommandations devraient être directement issues des déclarations. Dans les faits cependant, les recommandations vont au-delà de l'information générée pour plutôt s'appuyer sur l'expérience et l'expertise des évaluateurs; cela demeure tolérable si on accepte que le processus sous-jacent aux recommandations s'appuie sur un processus qui consiste à «extrapoler au-delà des évidences» (Weiss, 1998).

Forces et limites de l'étude

Il faut reconnaître que la présente étude constitue un effort de conceptualisation de la démarche évaluative, ce qui est tout aussi important que rare dans la pratique. De plus, cette étude descriptive s'appuie sur un solide cadre théorique bien documenté. Cependant, nous ne pouvons passer sous silence le faible nombre des cas ($n = 35$) ainsi que le fait que l'auteure a effectué seule les analyses de contenu, sans effectuer une quelconque forme d'agrément inter-juge qui aurait assuré une plus grande crédibilité aux données générées ainsi qu'à leur analyse. Finalement, la plus grande difficulté à générer des résultats significatifs réside probablement dans le type de méthode retenue. En effet, l'objet de l'étude, à savoir le raisonnement évaluatif, aurait probablement été plus facile à cerner par des entretiens avec les évaluateurs, ce qui aurait facilité la tâche de retracer leur raisonnement à travers l'étude des rapports.

Discussion

La pratique évaluative, du moins dans ce qui se dégage de l'analyse des rapports, ne semble pas être soutenue par un processus explicite, compte tenu de l'absence de plusieurs des composantes clés du modèle présenté à la figure 1, de la difficulté à classer certaines données et de l'absence tant des justifications que des cautions et des qualifiants. Même s'il existe toujours des données qui viennent cautionner en quelque sorte les déclarations et la conclusion évaluative, ces éléments n'ont pas été assez documentés théoriquement pour que nous soyons en mesure de les cerner à l'intérieur des différents types d'enquêtes. Si pour Valovirta (2002) «[...] l'absence de standard suggère que la pratique repose sur des standards implicites et que les jugements sont implicites, remettant en question la logique de l'évaluation», il semblerait, en bout de piste, que le lecteur doit se fier à ce qui lui est proposé de la part de l'évaluateur. Ce constat rejoint celui effectué par House (1995) à l'effet que les évaluateurs produisent une reddition la plus cohérente possible avec les évidences dont ils disposent.

Conclusion

Si le raisonnement évaluatif doit être compris comme un processus ouvert qui se transforme selon les questions posées, le contexte de l'évaluation, etc., il peut fournir aux évaluateurs une façon d'asseoir leur pratique. Cependant, lorsqu'il est construit sous la forme d'un processus linéaire, il y a des chances que cela soit considéré comme un *check list* désincarné qui s'avère peu utile pour bonifier la pratique.

Il semblerait que le raisonnement, dans la pratique, soit loin de la version idéalisée et que plus de recherche soit requise pour comprendre comment les praticiens conçoivent leur pratique et comment ils prennent des décisions pour effectuer le passage des évidences aux conclusions. Enfin, même les schémas les plus rigoureux impliquent un jugement, proclame Rog (1995), tout en ajoutant que «la manière de générer des jugements de façon à ce qu'ils soient le résultat d'un raisonnement en accord avec la logique de l'étude n'est pas encore fournie aux évaluateurs».

RÉFÉRENCES

- Biesta, G. J., & Burbules, N. C. (2003). *Pragmatism and educational research*. Lanham, MD: Rowman & Littlefield.
- Booth, W. C., Colomb, G. G., & Williams, J. M. (2003). *The craft of research* (2^e éd.). Chicago: University of Chicago Press.
- Fournier, D. M. (1995). Establishing evaluative conclusions : A distinction between general and working logic. *New Directions for Evaluation*, 68, 15-32.
- Fournier, D. M., & Smith, N. L. (1993). Clarifying the merits of argument in evaluation practice. *Evaluation and Program Planning*, 16(4), 315-323.
- House, E. R. (1993). *Professional evaluation : Social impact and political consequences*. Thousand Oaks, CA: Sage.
- House, E. R. (1995). Putting things together coherently: Logic and justice. *New Directions for Evaluation*, 68, 33-48.
- House, E. R., & Howe, K. R. (1999). *Values in evaluation and social research*. Thousand Oaks, CA: Sage.
- Neuendorf, K. A. (2001). *The content analysis guidebook*. Thousand Oaks, CA: Sage.
- Rog, D. J. (1995). Reasoning in evaluation: Challenges for the practitioner. *New Directions for Evaluation*, 68, 93-100.
- Schwandt, T. A. (2002). *Evaluation practice reconsidered*. New York: Peter Lang.
- Schwandt, T. A. (2004). What methods best support evaluative judgments? In P. Brandon (dir.), *What methods best support evaluative judgments?* Séance plénière de la réunion de l'American Evaluation Association, Atlanta, GA, novembre.
- Scriven, M. (1991). *Evaluation thesaurus* (4^e éd.). Newbury Park, CA: Sage.

- Scriven, M. (1995). The logic of evaluation and evaluation practice. *New Directions for Program Evaluation*, 68, 49-70.
- Sirotnik, K., & Oakes, J. (1990). Evaluation as critical inquiry. *New Directions for Program Evaluation*, 45, 37-59.
- Toulmin, S. (1958). *The Uses of Argument*. Cambridge: Cambridge University Press.
- Valovirta, V. (2002). Evaluation utilization as argumentation. *Evaluation*, 8(1), 60-80.
- Weiss, C. (1998). *Evaluation: Methods for studying programs and policies* (2^e éd.). Saddle River, NJ: Prentice Hall.
- What Works Clearinghouse (2003). *Introduction to the What Works Clearinghouse Evidence Report Process and the Role of Scientific Standards*. Extrait de [<http://www.w-w-c.org/july2003.html>], le 28 juillet 2003.