

Hurteau, M., Houle, S., & Guillemette, F. (Éds.). (2012). *L'évaluation de programme axée sur le jugement crédible*. Québec, QC : Presses de l'Université du Québec, 200 pages. Disponible en livre broché (ISBN 978-27605-3548-0) et numérique (ISBN 978-2-7605-3549-7).

Compte-rendu par Pierre-Marc Daigneault

L'évaluation de programme consiste à porter un jugement sur la valeur d'un programme à l'aide d'une méthodologie rigoureuse afin d'éclairer la prise de décision (Fournier, 2004; Jacob, 2010; Organisation de Coopération et de Développement Économiques, 1999; Secrétariat du Conseil du trésor du Québec, 2002). L'attention portée à la valeur intrinsèque (*merit*) et extrinsèque (*worth*) d'un objet—qu'il s'agisse d'une politique, d'un programme, d'un produit, d'un processus ou encore d'un membre du personnel—constitue ainsi un attribut fondamental de la pratique évaluative. Malgré l'importance du jugement en évaluation, il n'y a eu jusqu'ici que peu de recherches sur ce thème, exception faite des travaux théoriques portant sur la logique de l'évaluation par Michael Scriven et sur la distinction entre une conception descriptive et prescriptive des valeurs (Shadish, Cook, & Leviton, 1991). Avec leur programme de recherche sur le jugement en évaluation de programme, Marthe Hurteau et ses collègues (e.g., voir Hurteau, Houle, & Mongiat, 2009; Hurteau, Lachapelle, & Houle, 2006; Hurteau, Valois, & Bossiroy, 2011) contribuent à pallier cette lacune. *L'évaluation de programme axée sur le jugement crédible*, le dernier-né de ce programme de recherche, arrive à point nommé, soit au moment où l'enjeu de la qualité et des « données probantes » acquiert de plus en plus d'importance dans le domaine de l'évaluation (Daigneault, 2010). Cet ouvrage constitue en outre une réponse fort bienvenue aux appels des dernières années en faveur de davantage de recherches empiriques sur l'évaluation (e.g., voir Christie, 2003; Mark, 2001).

Comme son nom l'indique, *L'évaluation de programme axée sur le jugement crédible* s'attaque à la question importante (et complexe) de déterminer ce qui constitue un jugement crédible en évaluation de programme. Au risque de simplifier outre mesure l'argument des auteurs, on considère qu'un jugement est *crédible* dans la mesure où il est perçu à la fois comme juste (vrai, valide) et acceptable (recevable, digne de confiance) par les parties prenantes (pp. 50 et 78). La

modélisation des processus de production et de « crédibilisation » du jugement qui est articulée dans l'ouvrage en constitue sans aucun doute la contribution principale. Six principes sous-tendent cette modélisation : (a) les processus de production et de crédibilisation se déploient en synchronicité; (b) le jugement doit s'appuyer sur des données valides; (c) l'évaluateur doit faire preuve de flexibilité et d'adaptabilité dans sa démarche; (d) une argumentation solide est essentielle parce qu'elle établit le lien entre le jugement et les données; (e) les parties prenantes doivent être mises à contribution tout au long de la démarche jusqu'à la formulation du jugement; et (f) l'évaluateur doit posséder des compétences particulières qui vont au-delà des référentiels proposés au cours des dernières années (p. 80). Cette modélisation est fondée sur une démarche qualitative inductive mais rigoureuse qui a permis de générer des données et des construits riches de sens. Un échantillon intentionnel composé de 19 professionnels provenant de différents domaines (gestion, formation continue, santé et services sociaux, services d'urgence et d'enquête, justice) et ayant recours au *jugement* (au sens général du terme) dans le cadre de leur pratique a ainsi été formé. Des entrevues approfondies ont été réalisées auprès de ces répondants afin de comparer des situations où ceux-ci ont fait preuve d'un « bon » et d'un « moins bon » jugement. Les entrevues ont été enregistrées, transcrites, et analysées thématiquement par les chercheurs de manière individuelle. Les résultats ont fait l'objet de discussion lors d'une séance de mise en commun et ont fourni le support nécessaire à la modélisation présentée dans l'ouvrage. Cette dernière a ensuite été validée par une douzaine de théoriciens et de praticiens en évaluation de programme du Québec et des États-Unis afin de s'assurer de leur pertinence pour le domaine.

L'ouvrage compte neuf chapitres qui sont structurés en deux parties. Dans l'introduction, Marthe Hurteau et Sylvain Houle relatent l'évolution de leur programme de recherche sur le jugement crédible. Les trois contributions de la première partie se concentrent ensuite sur les fondements de la pratique évaluative. Dans le premier chapitre, Jody Fitzpatrick effectue un tour d'horizon sur l'évaluation qui traite de la définition de l'évaluation, de ses types et visées, des théories évaluatives, et des étapes de la démarche. Michael Schleifer et Marthe Hurteau procèdent à l'analyse des concepts de jugement et de crédibilité en prenant soin de distinguer le premier du raisonnement et de la prise de décision et le second de l'argumentation, de la conviction, et de la persuasion. Au troisième chapitre, Jennifer C. Greene soutient la thèse selon laquelle la production de données crédibles

en évaluation, loin de reposer uniquement sur une méthodologie rigoureuse, doit, par le biais de processus inclusifs et dialogiques, prendre en considération les préoccupations, problèmes, et questions des parties prenantes, en particulier celles qui sont les moins bien servies par le programme.

Les quatre contributions de la seconde partie s'attaquent ensuite au thème central de l'ouvrage, soit les processus de production et de crédibilisation du jugement en évaluation. Au chapitre quatre, Marthe Hurteau et ses collègues présentent leurs résultats de recherche sous la forme d'une modélisation fondée sur six principes (voir supra). Sheila A. Arens explore ensuite la relation entre validité, contexte, et contingences en évaluation. Elle affirme que l'évaluateur doit, afin de produire un jugement crédible, tenir compte du contexte et des contingences à toutes les étapes du processus plutôt que d'ignorer ceux-ci ou de chercher à les contrôler. Au chapitre six, Marthe Hurteau, Sylvain Houle, et Marie-Pier Marchand s'attardent à l'enjeu fondamental de la sélection des parties prenantes à l'évaluation et proposent une liste de critères à cette fin. Le chapitre sept, rédigé par J. Bradley Cousins et Lyn M. Shulha, porte sur l'établissement de seuils de réussite (ou standards de performance) dans le cadre de l'évaluation collaborative. Marie-Pier Marchand procède au chapitre huit à une analyse très intéressante des implications de la modélisation des processus de production et de crédibilisation du jugement pour les compétences des évaluateurs. Le chapitre neuf et la conclusion de l'ouvrage, tous deux rédigés par Sylvain Houle et Marthe Hurteau, visent respectivement à synthétiser les différentes contributions en prenant pour appui les étapes de la démarche évaluative et à tirer les leçons de la réalisation de cette étude.

L'évaluation de programme axée sur le jugement crédible représente sans aucun doute une contribution remarquable à la littérature francophone sur l'évaluation. L'étude empirique des processus de production et de crédibilisation du jugement ainsi que la modélisation qui en découle constituent une avancée significative non seulement pour le domaine de l'évaluation mais également pour d'autres disciplines appliquées telles que les sciences infirmières et le service social. Les citations des répondants qui sont présentées au fil des différents chapitres illustrent en outre de manière fort à propos les principes de la modélisation. J'aurais toutefois souhaité retrouver plus de détails dans l'ouvrage concernant le processus de recherche lui-même, notamment l'approche de théorisation ancrée (*grounded theory*) adoptée par les chercheurs et le fait que ceux-ci

aient été guidés par un spécialiste en la matière (Marthe Hurteau, communication personnelle, 12 octobre 2012). Chacun des chapitres du collectif apporte par ailleurs une contribution particulière à la littérature francophone en évaluation. D'une part, certains chapitres (un, six, et sept) ne présentent certes pas de résultats empiriques ou théoriques originaux mais fournissent en revanche des synthèses pertinentes, à jour, et en français que j'utiliserai certainement dans le cadre de mon enseignement. D'autre part, d'autres chapitres (trois, quatre, cinq, et huit) présentent des thèses audacieuses et des analyses stimulantes. Dans tous les cas, les sources citées dans les différents chapitres du collectif sont pertinentes et à jour. Bien que la littérature francophone en évaluation aurait pu être davantage mobilisée, il faut rappeler que plusieurs des chapitres ont été rédigés en anglais par des chefs de file nord-américains du domaine de l'évaluation avant d'être traduits (je note au passage que les traductions laissent parfois à désirer).

Ce collectif comporte par ailleurs quelques faiblesses qu'il est utile de mentionner. D'abord, le public auquel s'adresse l'ouvrage reste nébuleux : s'agit-il des étudiants, des praticiens, ou encore des chercheurs du domaine ? Plusieurs des contributions se situent à un niveau d'analyse assez élevé et contiennent rarement des lignes directrices claires pour la pratique évaluative. Ce livre est par conséquent davantage susceptible d'être utile aux étudiants et aux chercheurs du domaine qu'aux évaluateurs de programme. Ensuite, le fil conducteur que constitue le jugement crédible n'est pas toujours aussi clair et affirmé qu'on pourrait le souhaiter d'un chapitre à l'autre. Cela participe à mon avis de deux phénomènes. D'un côté, si les différentes acceptions des concepts de *jugement* et de *crédibilité* sont recensées au chapitre deux (e.g., jugement comme produit et comme compétence; crédibilité comme vérité et comme dialogue), il manque une prise de position claire par rapport à celles-ci (un problème dont souffre également le chapitre six). D'un autre côté, quoique pertinents, certains des chapitres du collectif n'ont recours au concept de crédibilité que de manière « allégée » (chapitre cinq) ou ne s'y réfèrent pas du tout (chapitre sept). À cet égard, un cadre conceptuel clair et appliqué de manière plus cohérente aurait contribué à rehausser la clarté des différentes contributions et à stimuler le dialogue entre leurs auteurs.

Ensuite, certaines questions auraient mérité un traitement plus approfondi dans ce collectif. La première est le postulat selon lequel des données valides sont une condition nécessaire mais non suffisantes à

un jugement crédible. Or, il y a souvent des arbitrages entre validité et acceptabilité des conclusions de l'évaluation. Certaines parties prenantes n'ont en effet pas les compétences requises pour apprécier la validité des résultats d'une évaluation. D'aucuns pourraient par exemple considérer que des entrevues semi-structurées menées auprès d'acteurs-clés d'un programme sont suffisamment crédibles pour juger de l'impact de celui-ci. Même si les auteurs spécifient que l'acceptabilité « ne devrait surtout pas être interprétée comme le fait de vouloir “plaire” au client » (p. 50), il n'en demeure pas moins que l'acceptabilité est subjective et dépend des connaissances et des compétences des « consommateurs » de l'évaluation. Deuxièmement, aussi désirable que soit la participation des parties prenantes à l'évaluation pour des raisons pratiques, politiques, ou épistémologiques (Cousins & Earl, 1992; Weaver & Cousins, 2004)—ce qui incidemment n'est pas une nouvelle façon de concevoir la contribution des parties prenantes, contrairement à ce qui est affirmé en conclusion du collectif (p. 194)—je doute en revanche qu'elle soit *essentielle* à la production d'un jugement crédible. Certaines évaluations non-participatives peuvent en effet générer des conclusions crédibles et, à l'inverse, le fait d'impliquer les parties prenantes dans l'évaluation peut parfois nuire à la crédibilité des conclusions (Jacob & Ouvrard, 2009). L'argument aurait en effet gagné en crédibilité si l'analyse des concepts sous-jacents avait été plus systématique (e.g., sur ce point, voir Daigneault & Jacob, 2009, 2012).

En définitive et malgré les quelques faiblesses soulevées, *L'évaluation de programme axée sur le jugement crédible* est sans aucun doute un ouvrage fort pertinent et stimulant. Avec ce collectif, Marthe Hurteau, Sylvain Houle, et François Guillemette ont enrichi de manière significative la littérature francophone sur l'évaluation.

RÉFÉRENCES

- Christie, C. A. (2003). What guides evaluation? A study of how evaluation practice maps onto evaluation theory. *New Directions for Evaluation*, 97, 7–35. doi:10.1002/ev.72
- Cousins, J. B., & Earl, L. M. (1992). The case for participatory evaluation. *Educational Evaluation and Policy Analysis*, 14(4), 397–418.
- Daigneault, P.-M. (2010). L'examen de la qualité des évaluations fédérales : une méta-évaluation réussie ? *Revue canadienne d'évaluation de programme*, 23(2 [automne 2008]), 191–224.

- Daigneault, P.-M., & Jacob, S. (2009). Toward accurate measurement of participation: Rethinking the conceptualization and operationalization of participatory evaluation. *American Journal of Evaluation*, 30(3), 330–348.
- Daigneault, P.-M., & Jacob, S. (2012). Les concepts souffrent-ils de négligence bénigne en sciences sociales? Éléments d'analyse conceptuelle et examen exploratoire de la littérature francophone à caractère méthodologique. *Information sur les sciences sociales*, 51(2), 188–204. doi:10.1177/0539018412437107
- Fournier, D. M. (2004). Evaluation. Dans S. Mathison (Éd.), *Encyclopedia of evaluation*. Thousand Oaks, CA : Sage.
- Hurteau, M., Houle, S., & Mongiat, S. (2009). How legitimate and justified are judgments in program evaluation? *Evaluation*, 15(3), 307–319. doi:10.1177/1356389009105883
- Hurteau, M., Lachapelle, G., & Houle, S. (2006). Comprendre les pratiques évaluatives afin de les améliorer : la modélisation du processus spécifique à l'évaluation de programme. *Mesure et évaluation en éducation*, 29(3), 27–44.
- Hurteau, M., Valois, P., & Bossiroy, A. (2011). Jugement crédible en évaluation de programme: définition et conditions requises. *Revue canadienne d'évaluation de programme*, 25(2 [automne 2010]), 83–101.
- Jacob, S. (2010). Évaluation. Dans L. Boussaguet, S. Jacquot, & P. Ravinet (Éds.), *Dictionnaire des politiques publiques* (pp. 257–265). Paris : Les Presses de Science Po.
- Jacob, S., & Ouvrard, L. (2009). L'évaluation participative. Avantages et difficultés d'une pratique innovante. *Cahiers de la performance et de l'évaluation*, 1. Québec : Laboratoire de recherche PerfEval, Université Laval.
- Mark, M. M. (2001). Evaluation's future: Furor, futile, or fertile? *American Journal of Evaluation*, 22(3), 457–479.
- Organisation de Coopération et de Développement Économiques, Service de la gestion publique, Comité de la gestion publique. (1999). *Vers de meilleures pratiques de l'évaluation : guide de meilleures pratiques à suivre pour l'évaluation et guide de référence*. Paris : Auteur.

Secrétariat du Conseil du trésor. Sous-secrétariat aux politiques budgétaires et aux programmes. (2002). *L'évaluation de programme : document destiné aux dirigeants et dirigeantes de ministères et d'organismes*. Québec, QC : Auteur.

Shadish, W. R., Cook, T. D., & Leviton, L. C. (1991). *Foundations of program evaluation: Theories of practice*. Newbury Park, CA : Sage.

Weaver, L., & Cousins, J. B. (2004). Unpacking the participatory process. *Journal of MultiDisciplinary Evaluation*, 1, 19–40.