

Un outil à visée pédagogique pour discuter de méthodologie

Maud Mediell et Eric Dionne

Université d'Ottawa

Résumé : Dans le cadre de cet article, nous abordons l'importance de la communication de la démarche d'évaluation (processus, méthodologie, résultats et limites de l'évaluation), et ce, dans le but de favoriser l'utilisation des résultats et la mise en œuvre des recommandations. Nous proposons un outil méta-évaluatif formatif à visée principalement pédagogique, centré sur les aspects méthodologiques du processus évaluatif afin d'accompagner les évaluateurs, et plus particulièrement les évaluateurs novices, dans la planification, la mise en œuvre et la communication rigoureuses de la méthodologie. Nous portons particulièrement notre attention sur la communication de la démarche d'évaluation de programme par l'intermédiaire de rapports d'évaluation (techniques et finaux) qui sont généralement le véhicule offrant le maximum d'informations concernant à la fois les résultats et le processus évaluatif. Bien que nous soyons conscients que d'autres moyens de communication existent (ex. : articles publiés dans des revues scientifiques), leur format ne permet pas toujours de fournir l'ensemble des informations inhérentes autant aux résultats qu'à la démarche mise de l'avant par le(s) évaluateur(s) puisqu'on y présente généralement les faits saillants sans s'attarder aux aspects méthodologiques. Puisque notre propos touche spécifiquement aux considérations méthodologiques, il nous apparaissait plus pertinent de nous appuyer sur les informations contenues dans les rapports. Ce choix n'est pas anodin, puisque la qualité d'une évaluation repose, entre autres, sur une méthodologie avérée permettant de fournir les données probantes nécessaires à l'exercice du jugement. Aussi, nous discuterons du rôle de la méta-évaluation (MEV) dans l'enrichissement de la pratique des évaluateurs et la promotion de la qualité d'une évaluation de programme; et nous présenterons un outil méta-évaluatif que nous avons conçu afin de favoriser l'opérationnalisation de méthodologies évaluatives de qualité et la mise en œuvre de pratiques évaluatives efficaces.

Mots clés : communication, méthodologie, méta-évaluation, rapports d'évaluation

Abstracts: In this article, we discuss the importance of communicating the evaluation approach (process, methodology, results, and limits) to promote the use of results and the implementation of recommendations. We present an education-focused, meta-evaluative training tool based on the methodological aspects of the evaluation process, and designed to support evaluators, particularly novice evaluators, in the rigorous planning, implementation and communication of methodology. We are

Auteur-ressource : Maud Mediell, mmedi042@uottawa.ca

focusing on communicating the program evaluation approach through evaluation reports (technical and final) that are usually the means offering the most information on both results and the evaluation process. We realize that there are other means of communication (i.e., journal-published articles), but their format doesn't always allow the provision of all the information relevant to results and the approach chosen by evaluators, since they usually only present highlights and not methodological aspects. Since we zeroed in on methodological considerations, it seemed to us that it would be more relevant to base ourselves on the information included in reports. This is not an innocuous choice, as an evaluation's quality depends, in part, on a recognized methodology that is able to provide the solid evidence needed to exercise judgment. Also, we discuss the role of meta-evaluation (MEV) in enriching evaluation practice and promoting program evaluation's quality. Furthermore, we present a meta-evaluative tool we designed to encourage the operationalization of quality evaluative methodologies and the implementation of effective evaluative practices.

Keywords: *communication, methodology, meta-evaluation, evaluation reports*

PROBLÉMATIQUE

Au fil des dernières décennies, l'éventail des fonctions de l'évaluation de programme s'est grandement élargi. Il ne s'agit plus seulement de porter un regard critique sur un programme, mais d'appuyer les prises de décisions relativement aux résultats d'évaluation, à l'amélioration des programmes ou des organisations concernées et, entre autres, au renforcement du pouvoir d'agir de toutes les personnes concernées par lesdites entités évaluées (parties prenantes primaires et secondaires), ainsi que de la population en générale (Fitzpatrick, 2012). Aussi, au lendemain de l'Année internationale de l'évaluation, dont le thème était « *Evidence for the world we want: using evaluation to improve people's lives—through better policy making* » (EvalYear 2015)¹, la qualité des évaluations de programmes et l'utilisation des résultats demeurent plus que jamais une priorité.

Le mouvement « *evidence-based practices* » qui est apparu dans de nombreux domaines tels que la recherche en sciences sociales et l'évaluation de programme a exercé de fortes pressions sur leurs pratiques et la qualité de leurs études afin de promouvoir, entre autres, l'utilisation des résultats. En évaluation, il s'agit donc de permettre aux parties prenantes clés de baser leurs décisions sur des connaissances ayant émergé de données probantes (Patton, 2005; Oakley, 2002). Les prises de décisions relatives à la vie d'un programme, d'une politique, d'une intervention, etc., sont basées sur l'appréciation de la qualité des connaissances produites par l'évaluation (soient les conclusions et les recommandations), connaissances qui sont évaluées par les décisionnaires, en matière de rigueur, de validité, de fiabilité et de crédibilité (Daigneault, 2008). Aussi, si une évaluation est jugée de mauvaise qualité, les résultats produits par cette dernière ne seront pas utilisés. Si une telle décision est justifiée, il n'en demeure pas moins que cela constitue un gaspillage important de

ressources humaines et financières ainsi qu'une occasion perdue de contribuer à l'amélioration de la pertinence et de l'efficacité d'un programme.

De surcroît, l'utilisation des résultats d'évaluation repose en grande partie sur la communication de l'évaluation, et plus particulièrement, sur le contenu des rapports d'évaluation (techniques et finaux) et leur utilisation par les clients de l'évaluation ainsi que toutes autres personnes concernées par le programme évalué (Torres, 2005; Patton, 2008). En effet, les premières critiques auxquelles les praticiens de l'évaluation ont dû faire face reposaient sur le manque d'utilisation des données d'évaluations, mentionnant notamment le fait que les rapports d'évaluation étaient peu ou pas lus et que les résultats n'étaient pas pris suffisamment en considération par les clients (Alkin, 2005). L'une des raisons expliquant pourquoi certains résultats d'évaluation étaient si peu réinvestis est en lien avec l'objectif même de tout rapport d'évaluation, à savoir que toutes personnes concernées par l'évaluation devraient être en mesure de comprendre le processus évaluatif et les résultats présentés. Aussi, la compréhension des aspects méthodologiques est indispensable pour toute utilisation des données d'évaluation, que ce soit de manière conceptuelle (développer des connaissances) ou instrumentale (apporter les modifications nécessaires afin d'améliorer le programme). Comme nous l'avons vu un peu plus tôt, il s'agit également de donner de la crédibilité aux résultats. En effet, si la démarche méthodologique est pertinente, transparente, et suffisante cela pourra contribuer à la crédibilité des résultats. Face à une telle nécessité, l'étape de communication de l'évaluation (le processus, les résultats et les limites de l'évaluation) constitue une dimension majeure de l'activité de l'évaluateur (Torres, 2009; Patton, 2008, 2013) et elle se doit d'être réalisée de façon adéquate.

La participation des parties prenantes dans le processus évaluatif et la présentation des résultats finaux amorcent un processus d'interprétation chez ces derniers. Ainsi, les arguments évaluatifs (recommandations et conclusions des évaluateurs) sont interprétés et réévalués à un niveau individuel chez les personnes concernées par le programme évalué (Torres, 2009). Les acteurs du programme (particulièrement les parties prenantes clés) construisent leurs propres interprétations au regard de nouvelles informations qu'ils détiennent, et ce, en fonction de leurs expériences personnelles et de leurs connaissances antérieures. De fait, il est important qu'ils reconnaissent la crédibilité des résultats d'évaluation, mais aussi du processus évaluatif, car ils constituent la base de futures actions (Valovirta, 2002; Patton, 2008). Autrement dit, ils doivent être convaincus que les données présentées s'appuient sur une méthodologie non seulement adéquate, mais crédible. En outre, dans le processus de communication de l'évaluation, il est primordial que les évaluateurs fassent un effort de transparence et de clarté, tout particulièrement dans le cadre de la rédaction des différents rapports, s'assurant ainsi d'avoir constamment à l'esprit le public visé (Russ-Eft et Preskill, 2009; Torres, 2005). L'objectif est de procurer suffisamment d'informations claires et valides pour permettre au lecteur de formuler de potentielles conclusions alternatives ou de répliquer l'évaluation dudit programme.

De surcroît, les recommandations présentent dans le rapport d'évaluation et sur lesquelles reposent les prises de décisions sont fondées, en partie, sur les données produites par la méthodologie mise en œuvre par l'évaluateur. Aussi, il ne s'agit pas seulement de s'assurer de réaliser une bonne évaluation conformément aux règles prescrites, il faut garantir la qualité méthodologique de cette dernière. Face à un tel constat, la méta-évaluation (MEV) constitue une méthode efficace afin d'assurer la qualité méthodologique, que celle-ci soit réalisée pendant l'évaluation, au moment de la rédaction du rapport d'évaluation ou après (Madzivhandila, Griffith, Fleming et Nesamvuni, 2010). La MEV est une pratique d'évaluation systématique qui permet de déterminer la qualité du processus et des résultats évaluatifs (Yarbrough, Shulha, Hopson et Caruthers, 2011; Greene, 1992) : elle est formative lorsqu'elle sert à améliorer le processus évaluatif au fil de l'évaluation et elle est sommative lorsqu'elle fournit aux clients, aux publics concernés par le programme ainsi qu'à l'évaluateur des informations en ce qui a trait aux forces et aux faiblesses de l'évaluation (Scriven, 2001, 2005; Stufflebeam, 2001).

Cependant, si son utilité et sa pertinence sont reconnues, la MEV demeure très peu utilisée dans le cadre de la pratique de l'évaluation de programme (Fitzpatrick, 2012; Scriven, 2001; Stufflebeam, 2001). Selon nous, trois raisons peuvent expliquer cette situation. Premièrement, on remarque une accumulation de listes de standards (et de leurs critères) énonçant la qualité en évaluation. En effet, la majorité des associations professionnelles, telles que l'association américaine en évaluation, les sociétés canadienne, européenne, suisse d'évaluation, etc., ont adopté la liste de standards du *Joint Committee on Standards for Educational Evaluation* (2011), et des institutions gouvernementales et non gouvernementales (UNICEF, ONU Femmes, le centre d'excellence en évaluation du secrétariat du trésor du Canada, etc.), ont développé leurs propres lignes directrices (ou modèles) et standards en matière de qualité. Si de telles listes semblent prodiguer un ensemble de normes prescrites en ce qui a trait à la conduite d'une évaluation, elles ne semblent cependant pas être suffisantes pour garantir la qualité d'une évaluation. Deuxièmement, dans la continuité de ce qui vient d'être énoncé, des auteurs tels que Shaw, Greene et Mark (2006) et Cooksy et Caracelli (2005), soulignent le manque de systématisation et de standardisation de la MEV. En effet, de multiples sources semblent offrir des définitions de la qualité en évaluation de programme, guidées par les standards (et leurs critères) de la discipline, qui apparaissent difficiles à opérationnaliser et à utiliser de manière systématique. Une question se pose donc : réalisons-nous régulièrement des méta-évaluations ? Si on se base sur les résultats de la méta-analyse réalisée par Cooksy et Caracelli (2009), hormis les méthodes de gestion et de contrôle de la qualité que l'on retrouve dans des agences gouvernementales telles que le secrétariat du trésor du Canada, l'office gouvernemental des comptes des États-Unis ou de grandes firmes spécialisées en évaluation dans le service public, il semblerait que la MEV ne soit pas encore considérée comme une pratique établie en évaluation. De surcroît, les auteurs ont remarqué un manque de clarté en ce qui a trait à la description des principes de la MEV. Il devient donc impératif que les praticiens et les chercheurs

s'accordent sur la définition des caractéristiques fondamentales de la MEV, et ce, à des fins théoriques, méthodologiques et pratiques. Troisièmement, les méthodes les plus sophistiquées s'avèrent être laborieuses et chronophages (Scriven, 2005), ce qui constitue selon nous un frein supplémentaire à la pratique méta-évaluative. Effectivement, des approches méthodologiques telles que la réplication (un évaluateur externe reproduit l'évaluation avec la même méthodologie) et la triangulation (un évaluateur externe évalue le même programme en ayant recours à une méthodologie différente) sont tout aussi coûteuses en matière de ressources humaines et financières, de temps et d'accès aux données, que l'évaluation elle-même. Cependant, en fonction du type d'évaluation et du type de programme évalué, les parties prenantes ou les évaluateurs de programme ne sont pas nécessairement en mesure d'accorder de telles ressources afin de mettre en œuvre ce genre de MEV. De fait, il incombe de développer des outils pratiques et moins coûteux, afin de faciliter l'implantation de l'exercice méta-évaluatif.

Enfin, House suggérait déjà en 1987 qu'il était temps de développer des méthodes de vérification de la qualité des évaluations et de proposer davantage qu'une liste de standards et de critères pour guider la pratique des évaluateurs. En outre, s'il existe aujourd'hui de nombreuses listes de vérification non méta-évaluatives afin d'accompagner l'évaluateur dans le développement et la mise en œuvre d'une évaluation (ex. : *Utilization-Focused Evaluation Checklist* [Patton, 2013], *Evaluation Design Checklist* [Stufflebeam, 2004], etc.) ainsi que la rédaction du rapport d'évaluation (*Evaluation Report Checklist* [Morin, 2004]), et si ces dernières mentionnent la nécessité de réaliser une MEV à la fin de l'évaluation à l'aide des standards du *Joint Committee on Standards for Educational Evaluation*, elles ne proposent toutefois aucune grille méta-évaluative centrée sur la qualité de la démarche méthodologique. Dans cet ordre d'idée, nous avons développé un outil pratique, à visée pédagogique, qui permet de : (1) réaliser une MEV formative centrée sur le devis d'évaluation, incluant la méthodologie, et la rigueur du processus évaluatif ; (2) fournir, à son utilisateur, une description des concepts qu'il se doit de connaître et de maîtriser afin de développer et de mettre en œuvre les évaluations les plus rigoureuses possible.

Dans la suite de cet article, nous discutons des raisons d'être de la MEV ainsi que de ses différentes méthodes et applications. Puis, nous abordons le processus de création et de validation du contenu de la grille MEV formative que nous avons conçu et qui permet la réalisation d'analyses critiques centrées sur le devis d'évaluation. Enfin, nous présentons ladite grille, ses caractéristiques et ses modalités d'utilisation.

MÉTA-ÉVALUATION ET ANALYSE CRITIQUE CENTRÉE SUR LE DEVIS D'ÉVALUATION

La qualité d'une évaluation constitue l'indicateur qui a l'association statistique la plus forte avec l'utilisation des résultats (Huie Hofstetter et Alkin, 2003; Cousins et Leithwood, 1986). Aussi, la qualité d'une évaluation en général, et sa qualité

méthodologique en particulier, augmenterait les probabilités que les résultats évaluatifs aient une influence sur les processus décisionnels en ce qui a trait à la vie d'un programme (Schwartz et Mayne, 2005a). En effet, la rigueur de la démarche évaluative et la crédibilité des résultats contribuent à la prise de décisions informées et donc davantage rationnelles (Daigneault, 2008). Pour promouvoir la qualité des pratiques évaluatives, Schwartz et Mayne (2005b) proposent une typologie des approches qualité qui offre une description et une classification des mécanismes permettant d'assurer la qualité des évaluations. Aussi, retrouve-t-on quatre catégories, résumées en quatre approches : (1) l'approche structurelle, (2) l'approche formative, (3) l'approche sommative et (4) l'approche systémique. L'approche structurelle porte sur le développement et la proposition de ressources favorisant la réalisation d'évaluations rigoureuses : élaboration de standards et de critères de qualité par les différentes associations professionnelles (AEA, SCÉ, JCSEE, etc.), formation professionnelle continue, certification des évaluateurs, etc. Les approches formatives et sommatives sont des approches qualité de type méta-évaluative : elles prodiguent le recours à des analyses systématiques d'évaluations, et ce, afin de déterminer la qualité de leurs processus et de leurs conclusions (Cooksy et Caracelli, 2005). L'approche formative se penche sur les efforts déployés afin de garantir la qualité des évaluations pendant leur mise en œuvre (c'est-à-dire avant la soumission du rapport final d'évaluation). Elle porte donc sur l'analyse des devis d'évaluation et des rapports techniques, par exemple, soit par un comité directeur, soit dans le cadre d'une auto-évaluation par les évaluateurs, et ce, à l'aide d'une grille, afin de les bonifier (Scriven, 2005). Si elle vise à améliorer l'évaluation, il nous paraît important de mentionner son rôle indéniable dans l'amélioration de la pratique des évaluateurs ainsi que l'enrichissement des connaissances, non seulement des parties prenantes, mais aussi de la discipline et de la communauté évaluatrice. L'approche sommative quant à elle est une approche rétrospective. Elle s'attarde uniquement sur les rapports d'évaluation finaux, ces derniers étant évalués dans une optique de reddition de compte et de responsabilisation. Enfin, l'approche systémique s'exerce dans une logique d'audit et de vérification. Elle ne se concentre pas uniquement sur le rapport d'évaluation, mais sur l'ensemble du processus évaluatif, c'est-à-dire les ressources et la capacité d'évaluation, la planification de l'évaluation, la production des conclusions, etc., soit les mécanismes qui sous-tendent la production et l'utilisation des résultats d'évaluation.

Aujourd'hui, la MEV est l'approche d'assurance qualité la plus discutée parmi les praticiens de l'évaluation de programme (Cooksy et Caracelli, 2005; Scriven, 2005; Stufflebeam, 2001). D'ailleurs, certains standards (et leurs critères) de la profession traitent de l'activité méta-évaluative (JCSSE, 2011). Scriven (2005) décrit cette approche comme l'évaluation d'une évaluation. Il s'agit ainsi d'une évaluation permettant d'apprécier la qualité d'une évaluation, et ce, en recourant aux méthodes évaluatives et en les appliquant à des processus et des rapports d'évaluation. La méta-évaluation vise à déterminer et à contrôler la qualité de l'évaluation (Cooksy et Caracelli, 2005).

L'importance du recours à la MEV, quelle que soit l'approche méta-évaluative utilisée (réplication, triangulation, etc.), est inextricablement liée aux implications éthiques, symboliques et scientifiques de la pratique en évaluation (Scriven, 2005). En effet, il est symboliquement et éthiquement crucial, en tant qu'évaluateurs, d'avoir recours à la MEV dans le cadre de pratiques évaluatives, puisque cela indique qu'en tant que professionnels on fait preuve de réflexivité, que nous nous soumettons nous-mêmes à un processus évaluatif, ce qui peut s'avérer rassurant pour les personnes concernées par le programme que nous sommes en train d'évaluer. Attention, il ne s'agit pas d'une démarche thérapeutique, mais bien de montrer notre engagement dans un processus continu de formation et d'amélioration de la pratique (développement professionnel, mais aussi personnel). En pratiquant la MEV, c'est aussi faire preuve d'engagement scientifique en s'assurant de la qualité de l'approche et de la validité des résultats, le tout dans le cadre d'une pratique reconnue que l'on n'hésite pas à soumettre à un examen critique par les pairs.

La MEV ne prescrit pas le recours à un modèle unique. Celle-ci peut être réalisée en interne ou à l'externe, elle peut être prospective (avant l'évaluation), concomitante (pendant l'évaluation) ou rétrospective (après l'évaluation), elle peut recourir à une ou plusieurs sources et méthodes de collecte de données, et elle peut être formative ou sommative (Cooksy et Caracelli, 2005; Scriven, 2005). Cependant, les approches en MEV demeurent très coûteuses (en matière d'argent, de temps, de ressources humaines, etc.). En outre, il existe d'autres approches certes beaucoup moins coûteuses, mais tout aussi fructueuses et efficaces, qui se centrent sur l'analyse critique du devis d'évaluation ou *evaluation design* (Scriven, 2005). En effet, considéré comme l'empreinte méthodologique d'une évaluation de programme (Rog, 2005), il offre une description précise et complète des composantes du programme, des bornes de l'évaluation (en matière de temps, d'accès aux données, etc.) et de la méthodologie (les sources d'information, les variables et les concepts mesurés qui devraient être déterminés à l'aide de la théorie du programme, les procédures de collecte et d'analyse de données, ainsi que les limites). En focalisant l'analyse critique (ou l'activité méta-évaluative) sur le devis d'évaluation, on se concentre sur les processus méthodologiques qui permettent de déterminer l'opérationnalisation et la mesure des variables, et ce, afin d'assurer la validité, la fiabilité et l'utilité des données recueillies. De surcroît, dans une perspective formative, ce type d'analyse peut être réalisé par l'évaluateur lui-même, et ce, tout au long du processus d'évaluation, ce qui selon nous contribue davantage à son développement professionnel ainsi qu'à sa réflexivité quant à la qualité de sa pratique.

Nous l'avons vu, en matière d'assurance qualité des évaluations de programme, la liste de standards et de critères proposée par le JCSSE (2011) constitue le guide privilégié de la communauté évaluatrice. Selon nous, si à elle seule, cette liste ne constitue pas un outil pratique et systématique, elle est cependant à l'origine de différentes listes de vérification ou d'outils méta-évaluatifs (Scriven, 2005). Stufflebeam (1999) a d'ailleurs été le premier chercheur à proposer un

outil méta-évaluatif, en anglais, à destination des évaluateurs de programme. Il y reprend les quatre standards – utilité, faisabilité, pertinence et précision – (et leurs trente critères au total) du *Joint Committee on Standards for Educational Evaluation*. Chaque standard constitue un axe de la grille, et chacun de leurs critères est étayé sous forme de dix indicateurs de vérification ou *checkpoints* que les évaluateurs doivent cocher s'ils estiment qu'ils l'ont bel et bien effectué. Un score est alors attribué à chaque axe et une moyenne est faite à la fin. Si nous reconnaissons la grande utilité de cette liste de vérification dans le cadre de la réalisation de MEV, nous pensons que cette dernière demeure trop générale et ne constitue pas un outil suffisamment pratique et systématique. En effet, les axes successifs de la liste de vérification de [Stufflebeam \(1999\)](#) ne s'alignent pas avec les étapes d'une évaluation de programme. Or, selon nous, si l'on désire accompagner pleinement les évaluateurs dans leur démarche évaluative (ou l'amélioration de celle-ci), et plus particulièrement les praticiens novices, il est important de faire coïncider la liste des critères avec la tâche qu'ils sont en train d'exécuter (par exemple, choisir le type d'évaluation, développer les questions d'évaluation, sélectionner les sources d'information, etc.). De plus, au regard des enjeux relatifs à la rigueur méthodologique et à l'utilisation des résultats de l'évaluation, nous pensons qu'il est pertinent de proposer un outil méta-évaluatif qui se concentre sur le développement et la mise en œuvre du devis d'évaluation (et donc de la méthodologie évaluative). En effet, une telle grille méta-évaluative pourrait être pertinente en contexte pédagogique puisqu'elle pourrait accompagner, voire guider les praticiens novices et les étudiants au fil de leur évaluation (ou exercice d'évaluation), tout en enrichissant leurs connaissances pratiques et méthodologiques, et en encourageant le développement de nouvelles compétences.

Aussi, nous avons fait le choix de développer une liste de vérification, en français, au regard des standards (et leurs critères) de la discipline, mais aussi des différents éléments qui constituent le devis d'évaluation et chacun de ses axes, à l'aide notamment de listes de vérification existantes dans la littérature scientifique. Enfin, nous sommes en mesure de proposer un outil méta-évaluatif qui permet de s'assurer de la qualité des résultats obtenus sur lesquels se basent bon nombre de décisions et de recommandations, et de fournir, à son utilisateur, une description des concepts à maîtriser afin de développer et de mettre en œuvre une évaluation de programme de qualité.

ÉLABORATION DE LA GRILLE D'ANALYSE CRITIQUE CENTRÉE SUR LE DEVIS D'ÉVALUATION

L'analyse critique ou la MEV d'un devis d'évaluation se base généralement sur les standards du *Joint Committee on Standards for Educational Evaluation* ([Scriven, 2005](#)). Ainsi, l'instrument que nous avons développé a été traduit et adapté de la liste de vérification de la précision des devis d'évaluation de [Sanders et Nafziger \(2011\)](#), qui mobilise, entre autres, les critères du standard « précision » approuvés par le JCSEE. Les auteurs ont proposé des critères tels que la « justesse de la

représentation du programme », la « validité » et la « fiabilité » en ce qui a trait à la précision concernant la collecte et l'analyse des données. Ces derniers coïncident d'ailleurs avec ceux du JCSEE se rapportant à la validité, à la fiabilité et au caractère défendable des informations. En outre, nous avons modifié la grille de vérification de Sanders et Nafziger (2011) au regard des standards SEVAL (Widmer, Lamber et Bachman, 2000). En effet, les auteurs ont proposé de référencer les standards (et leurs critères respectifs) auxquels il est adéquat de se référer en fonction des activités principales d'évaluation telles que la planification de l'évaluation, l'exploitation des informations ou encore la présentation du rapport d'évaluation. À la suite de l'analyse des différentes listes de standards de références SEVAL, nous avons sélectionné les critères relatifs aux différents standards qui peuvent être mobilisés dans le cadre de l'analyse d'un rapport, centrée sur la méthodologie, l'analyse des résultats et les limites. Au regard des différents écrits consultés dans le cadre de l'élaboration de la *grille d'analyse critique centrée sur le devis d'évaluation*, nous avons retenu et adapté les critères du [Tableau 1](#).

Les axes de la *grille d'analyse de rapports d'évaluation* que nous avons développée portent sur la précision concernant :

- la description de la démarche d'évaluation,
- les sources d'information et les procédures d'échantillonnage,
- la description du devis d'évaluation,
- la collecte de données,
- l'analyse des données qualitatives et quantitatives,
- les limites méthodologiques.

Le processus d'étayage des standards et des critères en indicateurs et sous-indicateurs s'est fait à l'appui de différentes listes de vérification non méta-évaluatives dont celles de [Stufflebeam \(2004\)](#) qui portent sur les étapes de l'élaboration du devis d'évaluation, de [Patton \(2013\)](#) qui décrit le processus de préparation et de mise en œuvre d'évaluations centrées sur les résultats, et enfin de [Morin \(2004\)](#) qui traite de la rédaction du rapport d'évaluation. Nous avons également consulté le *Guide pratique pour créer et évaluer une recherche scientifique en éducation* de [Gaudreau \(2011\)](#) ainsi que la liste de vérification MEV de [Scriven \(2011\)](#), qui propose une description des indicateurs essentiels à chacun des cinq concepts fondamentaux garants de la qualité et de la rigueur d'une évaluation : la validité, la crédibilité, la clarté, l'éthique et l'utilité. Enfin, nous avons comparé la déclinaison des indicateurs à ceux de la liste de vérification MEV de [Stufflebeam \(1999\)](#), qui déclinait des standards du *Joint Committee on Standards for Educational Evaluation*.

Afin de nous assurer de la validité du contenu de cette grille MEV, nous avons recruté des experts en mesure et évaluation, en évaluation de programme œuvrant au gouvernement fédéral canadien et en évaluation de programme en milieu éducationnel et communautaire ; soit trois domaines d'expertise en évaluation de programme. Nous avons envoyé un courriel de recrutement à six experts (au total)

Tableau 1. Critères retenus pour la construction de la grille d'analyse critique

La précision	
(Yarbrough, Shulha, Hopson et Caruthers [JCSEE], 2011; Widmer, Lamber et Bachman [SEVAL], 2000)	
<p>P3. Description des objectifs et de la démarche : « Les objectifs, les questions et la démarche de l'évaluation sont documentés avec suffisamment de précision pour qu'on puisse les identifier et les apprécier ».</p> <p>P4. Sources d'informations sûres (défendables) : « Les sources d'information utilisées dans une évaluation sont décrites avec suffisamment de précision pour qu'on puisse apprécier la justesse des informations ».</p> <p>P5. Informations valides : la procédure de récolte d'informations est choisie ou développée puis appliquée de manière à garantir la validité des interprétations dans la perspective de l'objectif donnée ».</p> <p>P6. Informations fiables : « la procédure de récolte d'informations est choisie ou développée puis appliquée de manière à garantir la fiabilité des interprétations dans la perspective de l'objectif donné ».</p> <p>P8 et P9. Analyse des informations qualitatives et quantitatives : « les informations qualitatives et quantitatives d'une évaluation sont analysées de manière appropriée et systématique, afin de répondre effectivement aux questions de l'évaluation ».</p>	<p style="text-align: center;">La Précision concernant la collecte et l'analyse des données</p> <p style="text-align: center;">(Traduit et adapté librement de Sanders et Nafziger, 2011) ⁱⁱ</p> <p>Validité (« objectivity ») : le devis d'évaluation (incluant la méthodologie) contient des procédures permettant de contrôler les biais susceptibles d'affecter les informations collectées ainsi que les inférences émises à partir de l'analyse de ces dernières.</p> <p>Fiabilité (« replicability ») : le devis d'évaluation (incluant la méthodologie) présente des procédures qui permettent de s'assurer que les informations collectées sont exactes.</p> <p>Justesse de la représentation du programme (« representativeness ») : les informations collectées représentent avec précision-justesse le programme évalué. Le devis d'évaluation contient des procédures de contrôle de la validité des inférences faites à partir des données obtenues à l'aide des instruments de collecte de données.</p>
L'utilité	
(JCSEE [2011] et SEVAL [2010])	
<p>U5. Transparence de l'appréciation : « les perspectives et raisonnements sur lesquels se fonde l'interprétation des résultats sont décrits de telle sorte que les bases des jugements de valeur apparaissent clairement ».</p> <p>U6. Caractère complet et clair du rapport : « les rapports d'évaluation décrivent clairement l'objet d'évaluation et son contexte, de même que les objectifs, questions, procédures et résultats d'évaluation, dans le souci de mettre à disposition les informations essentielles sous une forme facilement compréhensible ».</p>	

ⁱⁱ Nous sommes conscients que les intitulés des critères employés par les auteurs peuvent, pour certains, faire écho au courant positiviste-postpositiviste. Aussi, nous avons tenté de les rendre plus généraux. De surcroît, il nous paraît important de souligner que si nous nous sommes inspirés du contenu de ces critères, nous nous sommes assurés de les étayer en indicateurs pertinents pour tous types d'approches évaluatives.

et trois ont accepté de réaliser la validation du contenu de la grille d'analyse de rapports d'évaluation. Le premier expert est professeur agrégé spécialisé en mesure et évaluation, tandis que le second expert travaille dans un département d'évaluation de programme au gouvernement fédéral canadien et participe à la formation des nouveaux employés en partenariat avec la Société canadienne d'évaluation. Enfin, le troisième expert est professeur émérite spécialisé en recherche et en évaluation de programmes éducationnels et communautaires. Au regard de l'analyse de leurs commentaires concernant les indicateurs de la grille et de la complémentarité de ces derniers, nous n'avons pas invité davantage d'experts à participer à la validation du contenu de la grille. Ce processus de validation de contenu de la grille nous a conduits à apporter quelques modifications mineures au niveau de l'étayage des indicateurs et des sous-indicateurs. Aussi, comme nous l'expliquons ci-après, les commentaires des experts nous ont incités à être plus précis quant aux raisons d'être de la grille, à son caractère générique, à sa souplesse, etc.

PRÉSENTATION DE LA GRILLE D'ANALYSE CRITIQUE CENTRÉE SUR LE DEVIS D'ÉVALUATION

À la suite du processus de validation du contenu la grille, il apparaît important d'apporter quelques précisions concernant l'utilisation de cet outil méta-évaluatif. Premièrement, si cette grille est centrée sur la qualité du processus évaluatif, et plus particulièrement la rigueur méthodologique (tel que recommandé par [Scriven \[2005\]](#)), et si elle mobilise des indicateurs relatifs aux concepts de validité, de fiabilité et de fidélité (ex. : contrôle des biais, recours à la triangulation, validation du contenu des instruments, etc.), celle-ci ne concerne pas uniquement les évaluations inscrites dans un paradigme positiviste-postpositiviste. En effet, il s'agit de dimensions primordiales au processus d'assurance et de vérification de la rigueur, de la crédibilité et de l'utilité de l'évaluation et des résultats évaluatifs, et ce, quel que soit le type d'évaluation réalisée (formative ou sommative, à petite ou à grande échelle, qualitative, quantitative ou mixte, etc.). Cependant, nous sommes conscients que certains indicateurs ne sont pas nécessairement pertinents, dépendamment du type d'évaluation réalisée. Aussi, si nous avons la volonté de développer une application numérique de la grille, son format actuel (papier) ne facilite pas la possibilité de cocher la case « N/A » et d'automatiquement passer à une autre section ; ou encore la possibilité de créer des indicateurs alternatifs. L'évaluateur peut, par contre, utiliser la case commentaire afin de faire part du contexte évaluatif, du type d'évaluation réalisée, etc. Deuxièmement, cette grille est à visée formative et non sommative, notre objectif est de sensibiliser les évaluateurs aux démarches permettant de réaliser une évaluation de qualité, et ce, peu importe le contexte évaluatif. Bien entendu, lorsqu'il s'agit, par exemple, de l'évaluation de l'implantation d'un projet pilote, nous savons pertinemment

que certains critères ne s'appliqueront pas et qu'il n'y aura pas, par exemple, la possibilité de trianguler les méthodes. Nous gardons bien à l'esprit que l'objectif d'une évaluation est d'améliorer un programme, mais nous pensons que cette grille permettra aux praticiens d'affiner leur pratique ainsi que le contenu qu'ils rapportent dans le cadre de la communication des résultats de l'évaluation (ex. : rapport final). Enfin, il est important de souligner que cet outil représente une proposition générique, un moyen de discuter de méthodologie, d'offrir l'occasion à ses utilisateurs d'enrichir leurs pratiques méthodologiques et de promouvoir le recours à la MEV formative. Il ne s'agit en aucun cas d'une recette unique et figée, au contraire, nous espérons que les évaluateurs vont se l'approprier et l'amender afin qu'elle corresponde davantage à leur réalité de terrain.

Comme mentionné précédemment, cette grille MEV ([Tableau 2](#)) comprend des critères (ou énoncés) permettant d'évaluer la qualité méthodologique d'une évaluation de programme, au regard des éléments présents, ou non, dans le rapport d'évaluation produit dans le cadre de cette dernière. Ainsi, pour chacun des énoncés de la grille :

- Si un élément ne s'applique pas, l'évaluateur cochera « N/A » et, s'il le désire, il pourra attribuer le score « X » pour le critère concerné,
- Si un élément de la liste est absent du rapport, et ce, sans explication alors qu'il aurait été adéquat et nécessaire de l'aborder, l'évaluateur cochera « non » et, s'il le désire, il pourra attribuer le score « 0 » pour le critère concerné.
- Si un élément de la liste est partiellement présent dans le rapport, alors qu'il aurait été adéquat et nécessaire de l'aborder plus en profondeur, l'évaluateur cochera « Part. » et, s'il le désire, il pourra attribuer le score « 1 » pour le critère concerné.
- Pour les autres éléments qui s'appliquent et qui sont présents dans le rapport d'évaluation, l'évaluateur cochera « oui » et, s'il le désire, il pourra attribuer le score « 2 » pour le critère concerné.

Tout comme la liste de vérification de la rédaction de rapports d'évaluation de [Morin \(2004\)](#), nous avons fait le choix de ne pas opter pour une échelle numérique. En effet, à l'heure actuelle, nous ne sommes pas en mesure de donner un exemple « validé », pour chaque valeur numérique, et ce, pour tous les indicateurs traités. En effet, selon nous, il est préférable de ne pas opter pour un système de notation arbitraire, où chaque utilisateur de la grille aurait sa propre représentation de ce qui constitue un « 1 » ou un « 5 » sur une échelle de Likert (par exemple). De surcroît, cette grille MEV se veut formative et vise avant tout l'accompagnement et l'amélioration des pratiques méthodologiques des évaluateurs de programme, particulièrement celles des évaluateurs novices.

Tableau 2. Grille d'analyse critique centrée sur le devis d'évaluation

Thèmes	Critères principaux	Critères détaillés	Respect des critères			Commentaires
			N/A	Non	Part. Oui	
<p>1. Précision concernant la description de la démarche d'évaluation :</p> <p>[Les objectifs, les questions et la démarche de l'évaluation sont documentés avec suffisamment de précision pour qu'on puisse les identifier et les apprécier. Soit la description détaillée de :</p> <ul style="list-style-type: none"> - l'organisation, - la collecte, - la mise au point, les réajustements, - la saisie et l'exploitation des données.] 	1.1 Description des objectifs de l'évaluation	1.1.1 Le mandat d'évaluation est clairement décrit	N/A	Non	Part.	Oui
		1.1.2 Les objectifs d'évaluation sont clairs	N/A	Non	Part.	Oui
		1.1.3 Les objectifs d'évaluation sont précis	N/A	Non	Part.	Oui
	1.2. Justification du choix du devis d'évaluation	1.2.1 Le type d'évaluation choisi est pertinent pour répondre aux objectifs d'évaluation	N/A	Non	Part.	Oui
		1.2.2 Le choix du type d'évaluation est justifié par l'auteur	N/A	Non	Part.	Oui
		1.2.3 Le devis d'évaluation est fidèle au type d'évaluation choisi	N/A	Non	Part.	Oui
	1.3. Présentation des questions d'évaluation	1.3.1 Les questions d'évaluation sont pertinentes au regard des objectifs d'évaluation	N/A	Non	Part.	Oui
		1.3.2 Les questions d'évaluation sont claires	N/A	Non	Part.	Oui
		1.3.3 Les questions d'évaluation sont précises	N/A	Non	Part.	Oui

1.4 Description de la méthodologie	1.4.1 Les sources d'information et les procédures d'échantillonnage sont décrites et justifiées	N/A	Non	Part.	Oui	Si le devis d'évaluation fait état de plusieurs sources d'information et procédures d'échantillonnage, veuillez utiliser cette colonne (commentaire) pour évaluer la description (et la justification) de chaque source d'information et la procédure d'échantillonnage de cette dernière.
1.4.2 Les procédures de sélection de l'instrumentation sont décrites et justifiées		N/A <td data-bbox="525 522 548 565">Non <td data-bbox="525 453 548 496">Part. <td data-bbox="525 383 548 427">Oui <td data-bbox="525 147 781 392">Si le devis d'évaluation fait état de la sélection de plusieurs instruments, veuillez utiliser cette colonne (commentaire) pour évaluer la description de chaque procédure de sélection des instruments.</td> </td></td></td>	Non <td data-bbox="525 453 548 496">Part. <td data-bbox="525 383 548 427">Oui <td data-bbox="525 147 781 392">Si le devis d'évaluation fait état de la sélection de plusieurs instruments, veuillez utiliser cette colonne (commentaire) pour évaluer la description de chaque procédure de sélection des instruments.</td> </td></td>	Part. <td data-bbox="525 383 548 427">Oui <td data-bbox="525 147 781 392">Si le devis d'évaluation fait état de la sélection de plusieurs instruments, veuillez utiliser cette colonne (commentaire) pour évaluer la description de chaque procédure de sélection des instruments.</td> </td>	Oui <td data-bbox="525 147 781 392">Si le devis d'évaluation fait état de la sélection de plusieurs instruments, veuillez utiliser cette colonne (commentaire) pour évaluer la description de chaque procédure de sélection des instruments.</td>	Si le devis d'évaluation fait état de la sélection de plusieurs instruments, veuillez utiliser cette colonne (commentaire) pour évaluer la description de chaque procédure de sélection des instruments.

(Suite à la page suivante)

Tableau 2. Suite

Thèmes	Critères principaux	Critères détaillés	Respect des critères			Commentaires
			N/A	Non	Part. Oui	
		1.4.3 Les procédures de développement de l'instrumentation sont décrites et justifiées	N/A	Non	Part. Oui	<i>Si le devis d'évaluation fait état du développement de plusieurs instruments, veuillez utiliser cette colonne (commentaire) pour évaluer la description des procédures de développement de chaque instrument.</i>
		1.4.3 Les procédures de collecte de données sont décrites et justifiées	N/A	Non	Part. Oui	<i>Si le devis d'évaluation fait état de multiples procédures de collecte de données, veuillez utiliser cette colonne (commentaire) pour évaluer la description de chaque procédure de collecte de données.</i>
		1.4.4 Les procédures d'analyse de données sont décrites et justifiées	N/A	Non	Part. Oui	<i>Si le devis d'évaluation fait état de multiples procédures d'analyse de données, veuillez utiliser cette colonne (commentaire) pour évaluer la description de chaque procédure d'analyse de données.</i>

2. <u>Précision</u> concernant les sources d'information et les procédures d'échantillonnage [Les sources d'information utilisées dans l'évaluation sont décrites avec suffisamment de précision pour qu'on puisse apprécier la justesse des informations.]	2.1 Description générale des sources d'information	2.1.1 Les critères de sélection des différentes sources d'information sont clairement justifiés	N/A	Non	Part.	Oui
		2.1.2 Les caractéristiques descriptives des sources d'information sont pertinentes au regard des objectifs et des questions d'évaluation	N/A	Non	Part.	Oui
	2.2 Description des procédures d'échantillonnage	2.2.1 L'échantillonnage ou le choix des sources d'information retenues est pertinent par rapport aux objectifs et aux questions d'évaluation	N/A	Non	Part.	Oui
		2.2.2 Le type de procédure d'échantillonnage est précise	N/A	Non	Part.	Oui
		2.2.3 Le processus d'échantillonnage est décrit, et lorsque c'est approprié, la représentativité de l'échantillon est justifiée	N/A	Non	Part.	Oui
		2.2.4 Les contraintes existantes relativement aux processus d'échantillonnage sont décrites	N/A	Non	Part.	Oui
		2.2.5 Les précautions qui ont été prises si l'échantillonnage n'est pas aléatoire sont précisées	N/A	Non	Part.	Oui

(Suite à la page suivante)

Tableau 2. Suite

Thèmes	Critères principaux	Critères détaillés	Respect des critères			Commentaires
			N/A	Non	Oui	
3. Précision concernant la description du devis d'évaluation	3.1 Description de la démarche méthodologique	3.1.1 Le rapport contient les informations nécessaires pour bien rendre compte de la méthodologie	N/A	Non	Part.	Oui
		3.1.2 La démarche méthodologique est organisée	N/A	Non	Part.	Oui
		3.2.1 Le rapport fait état de procédures permettant de vérifier la stabilité des résultats dans le temps (triangulation temporelle [ex. : étude longitudinale])	N/A	Non	Part.	Oui
3.2 Description de procédures d'obtention de données provenant d'angles différents (ex. : recours à la triangulation)		3.2.2 Le rapport fait état de l'implication de plusieurs évaluateurs, intervenant de manière indépendante, au niveau de la collecte et/ou de l'analyse des données (accord inter-juges – triangulation des observateurs)	N/A	Non	Part.	Oui
		3.2.3 Le rapport fait état de procédures d'obtention d'informations auprès de multiples informateurs et/ou diverses sources d'information objectives telles que des documents d'archives (triangulation de sources ou d'informateurs)	N/A	Non	Part.	Oui

	N/A	Non	Part.	Oui
3.2.4 Le rapport fait état du recours à de multiples méthodes de collectes de données, complémentaires ou contrastées (triangulation d'instruments ou de méthodes)	N/A	Non	Part.	Oui
4.1 Description de procédures garantissant la fiabilité des données recueillies [Le devis d'évaluation (incluant la méthodologie) devrait inclure les procédures qui permettent de s'assurer que les informations collectées sont exactes.]	N/A	Non	Part.	Oui
4. <u>Précision</u> concernant la collecte et le traitement des données [- La procédure de récolte d'informations a été développée, puis appliquée de manière à garantir la validité et la fiabilité des interprétations en vue de répondre aux questions d'évaluation. - Les informations qualitatives et quantitatives de l'évaluation sont analysées de manière appropriée et systématique, afin de répondre aux questions d'évaluation.]	N/A	Non	Part.	Oui
4.1.1 Les procédures de collecte de données sont claires	N/A	Non	Part.	Oui
4.1.2 Le rapport fait état de procédures de vérification de la fiabilité des instruments (ex. : validation de contenu, pré-expérimentation, recours à la triangulation, etc.)	N/A	Non	Part.	Oui

Si le devis d'évaluation fait état de plusieurs procédures de collecte de données, veuillez utiliser cette colonne (commentaire) pour évaluer la clarté de chaque procédure de collecte de données.

Si le rapport fait état de plusieurs procédures de vérification de la fiabilité des instruments, veuillez utiliser cette colonne (commentaire) pour évaluer les procédures de vérification de la fiabilité de chaque instrument.

(Suite à la page suivante)

Tableau 2. Suite

Thèmes	Critères principaux	Critères détaillés	Respect des critères			Commentaires
			N/A	Non	Part. Oui	
		4.1.3 Les procédures d'analyse sont suffisamment décrites et justifiées pour que l'on puisse les reproduire et obtenir les mêmes résultats	N/A	Non	Part. Oui	Si le devis d'évaluation fait état de plusieurs procédures d'analyse de données, veuillez utiliser cette colonne (com-mentaire) pour éva-luer la description (et la justification) de chaque procédure d'analyse de données afin de détec-ter s'il est possible de les reproduire et d'obtenir les mêmes résultats.
	4.2 Description de procédures de contrôle des biais au niveau de la collecte et du traitement des données [Le devis d'évaluation (incluant la méthodologie) devrait contenir des procédures permettant de contrôler les biais.]	4.2.1 Les sources d'information sont clairement spécifiées 4.2.2 Les interprétations faites à partir des informations obtenues grâce aux instruments de collecte de données sont fiables 4.2.3 Le rapport fait état de procédures de contrôle des biais	N/A	Non	Part. Oui	
			N/A	Non	Part. Oui	

4.3 Description de procédures de collecte et de traitement de données assurant la fiabilité des informations obtenues vis-à-vis du programme	4.3.1 Le rapport fait état de procédures de contrôle de la validité des interprétations faites à partir des données obtenues à l'aide des instruments de collecte de données	N/A	Non	Part.	Oui
[Les informations collectées devraient représenter avec précision justesse le programme évalué.]	4.3.2 Les instruments de collecte de données sont appropriés pour les objectifs de l'évaluation	N/A	Non	Part.	Oui
	4.3.3 Le devis d'évaluation (incluant la méthodologie) permet de répondre adéquatement aux questions d'évaluation auxquelles il est censé répondre	N/A	Non	Part.	Oui
5. <u>Limites méthodologiques</u> mentionnées et expliquées	5.1 Les limites méthodologiques sont prises en compte (échantillonnage, limites des instruments, biais de l'évaluateur ou du contexte, données insuffisantes, etc.). Le rapport fait état des obstacles techniques et logistiques.	N/A	Non	Part.	Oui

CONCLUSION

Depuis une vingtaine d'années, les praticiens et les chercheurs en évaluation se réfèrent à la liste de standards (et leurs critères) du *Joint Committee on Standards for Educational Evaluation*. Si la qualité et l'utilité de cette dernière ne sont plus à prouver, la standardisation de la pratique de la MEV préconisée par de nombreux auteurs, tels que [Yarbrough, Shulha, Hopson et Caruthers \(2011\)](#), [Cooksy et Caracelli \(2005, 2009\)](#), [Stufflebeam \(2001\)](#), [House \(1987\)](#), demeure difficile à implanter. Aujourd'hui, les préoccupations propres à l'évaluation de programme demeurent les mêmes et abordent les questions de la qualité, de l'utilisation des résultats, des compétences des évaluateurs, de la collaboration avec les parties prenantes, de l'instabilité du contexte évaluatif qui conduit très souvent les évaluateurs à opérer des changements méthodologiques en cours d'évaluation, etc. Ces questions sont intrinsèquement liées, voire dépendantes, et c'est notre rôle en tant que praticiens et chercheurs en évaluation de former et d'outiller les évaluateurs, et ce, afin de surmonter ces questionnements et ces incertitudes. Aussi, c'est dans cet esprit que nous avons décidé de créer cette grille de méta-évaluation centrée sur le devis d'évaluation et de proposer un instrument pratique et accessible, permettant de favoriser la réflexion de l'évaluateur et de vérifier la qualité des évaluations, de la conception à la remise du rapport d'évaluation final.

La conscience de soi en tant qu'évaluateur, la capacité réflexive sur sa propre pratique et le perfectionnement continu constituent le cinquième axe du socle de compétences essentielles proposé par [Stevahn, King, Ghere et Minnema \(2005\)](#). De fait, au fil de ses expériences, l'évaluateur doit être en mesure de faire preuve de réflexivité afin de poursuivre son développement professionnel. Aussi, il semble indispensable de proposer une standardisation et une systématisation des pratiques méta-évaluatives, et ce, à l'aide d'approches et d'outils pratiques, peu coûteux et accessibles afin de garantir la qualité des pratiques en évaluation ainsi que la notoriété de l'évaluation de programme.

Prochaines étapes

Maintenant que le processus de validation du contenu de la grille est terminé, nous allons utiliser cet outil dans le cadre de deux études. La première portera sur l'étude des propriétés pédagogiques de la grille, que nous utiliserons dans des cours de formation en évaluation de programme. Les étudiants seront invités à s'en servir pendant l'élaboration d'un devis d'évaluation de programme, puis dans la réalisation d'analyses critiques de rapports d'évaluation publiés. Aussi, s'agira-t-il d'observer dans quelle mesure le recours à cette grille contribue à l'enrichissement des connaissances théoriques et pratiques des étudiants. La deuxième consistera à conduire l'analyse de devis d'évaluation (incluant la méthodologie), présenté dans des rapports produits dans le cadre d'évaluations en milieu communautaire et éducationnel. L'un des objectifs de cette étude sera de montrer que l'utilisation d'un tel instrument durant la mise en œuvre d'une évaluation et la rédaction

du rapport d'évaluation favorise l'opérationnalisation d'une méthodologie évaluative de qualité. En effet, nous pensons qu'un tel outil permettra aux praticiens de l'évaluation de programme d'assurer la qualité de leurs évaluations tout en renforçant leurs connaissances ayant trait, par exemple, à l'ajustement des procédures méthodologiques en fonction des contraintes de terrain, etc.

NOTE

- 1 *EvalPartners*, le mouvement mondial pour renforcer les capacités nationales d'évaluation, a déclaré l'année 2015, l'Année internationale de l'évaluation (*EvalYear*), à la troisième Conférence internationale sur les capacités nationales d'évaluation qui s'est tenue à São Paulo, au Brésil, en 2013 (<http://mymande.org/evalyear>).

RÉFÉRENCES

- Alkin, M.C. (2005). Utilization of evaluation. Dans S. Mathison (éd.), *Encyclopedia of evaluation* (pp. 435–439). Thousand Oaks, CA: Sage.
- Cooksy, L.J. et Caracelli, V.J. (2005). Quality, context, and use: Issues in achieving the goals of metaevaluation. *American Journal of Evaluation*, 26(1), 31–42. <https://doi.org/10.1177/1098214004273252>
- Cooksy, L.J. et Caracelli, V.J. (2009). Metaevaluation in practice: Selection and application of criteria. *Journal of Multidisciplinary Evaluation*, 6(11), 1–15.
- Cousins, J.B. et Leithwood, K.A. (1986). Current empirical research on evaluation utilization. *Review of Educational Research*, 56(3), 331–364. <https://doi.org/10.3102/00346543056003331>
- Daigneault, P.-M. (2008). L'examen de la qualité des évaluations fédérales : une méta-évaluation réussie? *Canadian Journal of Program Evaluation*, 23(2), 191–224.
- Fitzpatrick, J.L. (2012). An introduction to context and its role in evaluation practice [Context: A framework for its influence on evaluation practice]. *New Directions for Evaluation*, 135, 7–24. <https://doi.org/10.1002/ev.20024>
- Gaudreau, L. (2011). *Guide pratique pour créer et évaluer une recherche scientifique en éducation*. Montréal: Guérin éditeur.
- Greene, J.C. (1992). A case study of evaluation auditing as metaevaluation. *Evaluation and program planning*, 15(1), 71–74. [https://doi.org/10.1016/0149-7189\(92\)90063-Z](https://doi.org/10.1016/0149-7189(92)90063-Z)
- House, E.R. (1987). The evaluation audit. *Evaluation practice*, 8(2), 52–56. [https://doi.org/10.1016/S0886-1633\(87\)80085-5](https://doi.org/10.1016/S0886-1633(87)80085-5)
- Huie Hofstetter, C. et Alkin, M.C. (2003). Evaluation use revisited. Dans T. Kellaghan et D. L. Stufflebeam (éds.), *International handbook of educational evaluation* (pp. 197–222). Dordrecht, The Netherlands: Kluwer Academic. https://doi.org/10.1007/978-94-010-0309-4_14.
- Madzivhandila, T.P., Griffith, G.R., Fleming, E.M. et Nesamvuni, A.E. (2010). *Meta-evaluations in government and government institutions: A case study example from the Australian Centre for International Agricultural Research*. Communication présentée

- la conférence annuelle de l'Australian Agricultural and Resource Economics Society (AARES), Adelaide, Australie.
- Morin, G. (2004). Evaluation report checklist. Récupéré de <https://wmich.edu/sites/default/files/attachments/u350/2014/evaluation-reports.pdf>
- Oakley, A. (2002). Social science and evidence-based everything: The case of education. *Educational Review*, 54(3), 277–286. <https://doi.org/10.1080/0013191022000016329>
- Patton, M.Q. (2005). Misuse of evaluation. Dans S. Mathison (éd.), *Encyclopedia of evaluation* (pp. 255–256). Thousand Oaks, CA: Sage. <https://doi.org/10.4135/9781412950558.n344>.
- Patton, M.Q. (2008). *Utilization-focused evaluation* (4^e éd.) (chap.1–2). Thousand Oaks, CA: Sage.
- Patton, M.Q. (2013). *Utilization-focused evaluation (U-FE) checklist*. Récupéré de https://wmich.edu/sites/default/files/attachments/u350/2014/UFE_checklist_2013.pdf
- Rog, D.J. (2005). Design, evaluation. Dans S. Mathison (éd.), *Encyclopedia of evaluation* (pp. 115–117). Thousand Oaks, CA: Sage.
- Russ-Eft, D. et Preskill, H. (2009). Evaluating the evaluation. Dans D. Russ-Eft et H. Preskill (éds.), *Evaluation in organizations. A systematic approach to enhancing learning, performance, and change* (2^e éd.) (p. 461–474). New York, NY: Basic Books.
- Sanders, J.R. et Nafziger, D.N. (2011). A basis for determining the adequacy of evaluation designs. *Journal of Multidisciplinary Evaluation*, 7(15), 44–78.
- Schwartz, R. et Mayne, J. (2005a). Assuring the quality of evaluative information: Theory and practice. *Evaluation and Program Planning*, 28(1), 1–14. <https://doi.org/10.1016/j.evalprogplan.2004.10.001>
- Schwartz, R. et Mayne, J. (2005b). Does quality matter? Who cares about the quality of evaluative information? Dans R. Schwartz et J. Mayne (éds.), *Quality matters: Seeking confidence in evaluating, auditing, and performance reporting* (pp. 301–322). New Brunswick, NJ: Transaction.
- Scriven, M. (2001). Evaluation: Future tense. *American Journal of Evaluation*, 22(3), 301–307. <https://doi.org/10.1177/109821400102200303>
- Scriven, M. (2005). Metaevaluation. Dans S. Mathison (éd.), *Encyclopedia of evaluation* (pp. 250–252). Thousand Oaks, CA: Sage. <https://doi.org/10.4135/9781412950558.n340>.
- Scriven, M. (2011). *Evaluating evaluations: A meta-evaluation checklist*. Récupéré de http://michaelscriven.info/images/EVALUATING_EVALUATIONS_8.16.11.pdf.
- Shaw, I.F., Greene, J. et Mark, M.M. (2006). On discerning quality in evaluation. Dans I. F. Shaw, J. Greene et M. M. Mark (éds.), *The Sage handbook of evaluation* (pp. 405–419). London: Sage. <https://doi.org/10.4135/9781848608078>.
- Stevahn, L., King, J.A., Ghere, G. et Minnema, J. (2005). Establishing essentials competencies for program evaluators. *American Journal of Evaluation*, 26(1), 43–59. <https://doi.org/10.1177/1098214004273180>
- Stufflebeam, D.L. (1999). *Program evaluation models: Metaevaluation checklist* (based on the program evaluation standards). Récupéré de https://wmich.edu/sites/default/files/attachments/u350/2014/eval_model_metaeval.pdf

- Stufflebeam, D.L. (2001). The Metaevaluation imperative. *American Journal of Evaluation*, 22(2), 183–209. <https://doi.org/10.1177/109821400102200204>
- Stufflebeam, D.L. (2004). *Evaluation design checklist*. Récupéré de <https://wmich.edu/sites/default/files/attachments/u350/2014/evaldesign.pdf>
- Torres, R.T. (2005). Reporting. Dans S. Mathison (éd.), *Encyclopedia of evaluation* (pp. 371–375). Thousand Oaks, CA: Sage. <https://doi.org/10.4135/9781412950558.n484>.
- Torres, R.T. (2009). Communicating and reporting evaluation activity findings and change. Dans D. Russ-Eft et H. Preskill (éds.), *Evaluation in organizations. A systematic approach to enhancing learning, performance and change* (2^e éd.) (p. 399–442). New York, NY: Basic Books.
- Valovirta, V. (2002). Evaluation utilization as argumentation. *Evaluation*, 8(1), 60–80. <https://doi.org/10.1177/1358902002008001487>
- Widmer, T., Lamber, C. et Bachman, N. (2000). *Standards d'évaluation de la Société suisse d'évaluation (Standards SEVAL)*. Récupéré du site de la Société suisse d'évaluation, section Standards SEVAL : <http://www.seval.ch/fr/Standards/index.cfm>
- Yarbrough, D.B., Shulha, L.M., Hopson, R.K. et Caruthers, F.A. (2011). *The program evaluation standards. A guide for evaluators and evaluation users* (3^e éd.). Thousand Oaks, CA: Sage.

PRÉSENTATION D'AUTEUR

Maud Mediell est doctorante à la Faculté d'éducation de l'Université d'Ottawa. Elle détient un Master 2 (professionnel et recherche) en psychologie sociale de la santé ainsi qu'un diplôme d'études supérieures en évaluation de programme. Elle est consultante en évaluation de programme et elle travaille actuellement pour le Centre d'enseignement et d'apprentissage du Collège Algonquin. Elle est également membre du Centre d'appui pédagogique en santé pour la francophonie (CAPSAF) de la Faculté de médecine de l'Université d'Ottawa. Elle s'intéresse particulièrement aux pratiques méthodologiques et à la formation des évaluateurs de programme ainsi qu'à l'évaluation développementale.

Eric Dionne est professeur agrégé à la Faculté d'éducation et à la Faculté de médecine de l'Université d'Ottawa. Il est aussi titulaire de la chaire de recherche en éducation médicale du Centre d'appui pédagogique en santé pour la francophonie (CAPSAF) et de l'Institut du Savoir de l'hôpital Montfort (ISM). Les intérêts de recherche du professeur Dionne touchent à la modélisation des scores aux tests ainsi qu'au développement et à la validation d'outils de cueillette de données afin d'évaluer des apprentissages complexes.