

Dialogues entre théories spontanées et théories académiques de l'évaluation

Claire Tourmen and Nathalie Droyer

Volume 36, Number 3, 2013

URI: <https://id.erudit.org/iderudit/1025739ar>

DOI: <https://doi.org/10.7202/1025739ar>

[See table of contents](#)

Publisher(s)

ADMEE-Canada - Université Laval

ISSN

0823-3993 (print)

2368-2000 (digital)

[Explore this journal](#)

Cite this article

Tourmen, C. & Droyer, N. (2013). Dialogues entre théories spontanées et théories académiques de l'évaluation. *Mesure et évaluation en éducation*, 36(3), 25–49. <https://doi.org/10.7202/1025739ar>

Article abstract

Evaluation theories: from whom and about what? We will follow a hypothesis from Shadish, Cook and Leviton (1991). According to them, there are “spontaneous theories” of evaluation among practitioners and “academic theories” among the literature. In this paper, we will try to describe, compare and discuss these two kinds of theories. We will first examine the “spontaneous theories” of evaluation. Indeed, we found traces of them in the acts and discourses of a sample of professional evaluators that we met during a doctoral research. We will describe the form and content of these theories before examining their origin and development. We will then describe the “academic theories” of evaluation developed in the literature (form, content, origin and development). We will finally discuss the advantages and limits of each kind of theory and try to understand how they circulate and how they can enrich each other.

Dialogues entre théories spontanées et théories académiques de l'évaluation

Claire Tourmen
Nathalie Droyer

AgroSup Dijon, France

MOTS CLÉS: évaluation de programme, théorie, activité, expérience

Théories de l'évaluation, oui, mais de qui et sur quoi? Les auteurs partiront de l'hypothèse de Shadish, Cook, et Leviton (1991) selon laquelle il existerait à la fois des « théories spontanées » de l'évaluation chez les praticiens, ainsi que des « théories académiques » de l'évaluation dans la littérature. Cet article vise à mieux décrire, comparer et discuter les différentes théories de l'évaluation sur la base de cette distinction. Seront examinées tout d'abord les « théories spontanées » de l'évaluation, à travers les traces décelées chez des évaluateurs rencontrés à l'occasion d'une thèse. Les théories (formes, contenus) des praticiens seront analysées pour tenter d'en comprendre l'origine et le développement. Par la suite, cet article montrera qu'il existe aussi des tentatives d'avancer sur des « théories académiques » de l'évaluation dans la littérature. La forme, le contenu et le mode de constitution de ces théories seront réexaminés. Enfin seront discutés les atouts et les limites de chaque forme de théorisation de l'évaluation pour aboutir sur les liens qui peuvent unir les théories spontanées et les théories académiques de l'évaluation, en tentant de comprendre comment elles circulent et peuvent potentiellement s'enrichir les unes et les autres.

KEY WORDS: program evaluation, theory, activity, experience

Evaluation theories: from whom and about what? We will follow a hypothesis from Shadish, Cook and Leviton (1991). According to them, there are "spontaneous theories" of evaluation among practitioners and "academic theories" among the literature. In this paper, we will try to describe, compare and discuss these two kinds of theories. We will first examine the "spontaneous theories" of evaluation. Indeed, we found traces of them in the acts and discourses of a sample of professional evaluators that we met during a doctoral research. We will describe the form and content of these theories before examining their origin and development. We will then describe the "academic theories" of evaluation developed in the literature (form, content, origin and development). We will finally discuss the advantages and limits of each kind of theory and try to understand how they circulate and how they can enrich each other.

PALAVRAS-CHAVE: avaliação de programas, teoria, atividade, experiência.

Teorias da avaliação, sim, mas de quê e sobre quê? Os autores partirão da hipótese de Shadish, Cook e Leviton (1991), segundo a qual haveria, ao mesmo tempo, “teorias espontâneas” de avaliação entre os práticos e “teorias acadêmicas” de avaliação na literatura. Este artigo visa descrever, compara e discutir as diferentes teorias da avaliação a partir desta distinção. Serão examinadas, desde logo, as “teorias espontâneas” de avaliação através dos traços revelados por avaliadores no âmbito de uma tese de doutoramento. As teorias (formas, conteúdos) dos práticos serão analisadas para tentar compreender a sua origem e desenvolvimento. De seguida, o artigo mostrará que existem também tentativas de avançar sobre as “teorias acadêmicas” de avaliação na literatura. A forma, o conteúdo e o modo de constituição destas teorias serão reexaminados. Finalmente, discutir-se-ão as vantagens e os limites de cada forma de teorização de avaliação para encontrar as ligações que podem unir as teorias espontâneas e as teorias acadêmicas da avaliação, tentando compreender como circulam e como podem, potencialmente, enriquecer-se mutuamente.

Note des auteurs – La correspondance liée à cet article peut être adressée à Claire Tourmen, Maître de conférences en sciences de l'éducation, AgroSup Dijon, 26 boulevard Docteur Petitjean 21000, Dijon, France, téléphone : +33 3 80 77 25 25, ou par courriel à l'adresse suivante : [Claire.tourmen@educagri.fr].

Des pratiques d'évaluation de programme – ou de politiques publiques¹ – se développent depuis les années 1960 en Amérique du Nord, et depuis les années 1990 en France et ailleurs dans le monde (Furubo, Rist, & Sandahl, 2002). Le développement de ces pratiques est accompagné par une abondante littérature internationale qui ne cesse d'augmenter – revues scientifiques, ouvrages, guides, etc. –, littérature dont la tonalité reste majoritairement prescriptive. Il est de coutume de parler, notamment en Amérique du Nord, de l'existence de différentes « théories » de l'évaluation, qui constitueraient autant d'approches, « modèles » ou « paradigmes » concurrents de cette activité (Alkin & Christie, 2004 ; Patton, 1986).

Toutefois, dès le début des années 1990, des auteurs américains, Shadish, Cook, et Leviton (1991), se sont interrogés sur la possibilité d'une « théorie » de l'évaluation dans un sens différent : celui d'un ensemble de lois visant à décrire et comprendre les phénomènes évaluatifs et qui pourraient fonder les pratiques. Ils l'ont fait à travers un ouvrage au titre éloquent : *Foundations of Program Evaluation: Theories of Practice*. Une telle théorie permettrait de dépasser, selon leur propos, les « listes athéoriques de méthodes ». L'évaluation de programme peut dès lors se décrire, depuis cet ouvrage, comme un ensemble de pratiques en quête d'une théorie pour les fonder, aucune théorie unificatrice n'existant à ce jour, comme le rappelle Figari dans ce numéro, malgré plusieurs tentatives de théorisation de l'évaluation. Ce numéro spécial de la revue témoigne de la volonté des rédacteurs d'y voir plus clair et de tenter d'avancer sur cette question épineuse. Donaldson et Lipsey (2008) affirmaient en effet en 2008 que « la nature et le rôle de la théorie en évaluation sont souvent un sujet controversé » (p. 57).

Réfléchir à la possibilité d'une théorie de l'évaluation pose plusieurs questions : théorie de qui et de quoi ? Quel rapport avec les pratiques ? Comment se développe-t-elle ? Les auteurs se garderont ici d'adopter une conception figée et prescriptive des théories – comme autant de modèles à suivre par les praticiens, la « théorie » étant alors opposée à la « pratique » (Schwandt, 2008) – pour au contraire reprendre une conception épistémologique du terme telle que développée dans le travail fondateur de

Shadish et al. (1991). Ils définissent une théorie comme un ensemble de «concepts», de «savoirs» – aussi appelés «connaissances» – et de «lois» permettant de «décrire des phénomènes» et d'en prédire l'état et l'évolution, définition qui guidera la rédaction de cet article. Le postulat d'une différence radicale ou de nature entre théories et pratiques est rejeté, faisant place à l'hypothèse d'un continuum entre différentes formes de connaissances, en circulation entre les univers individuels et sociaux (Rabardel, 1995; Valsiner, 2007), dans une épistémologie d'inspiration pragmatiste (Dewey, 1993). S'inscrivant dans cette tradition, Shadish et al. (1991) distinguent deux types de théories: «spontanées» (celles des évaluateurs) et «académiques» (énoncées dans la littérature sur l'évaluation), distinction reprise ici car elle semble féconde pour comprendre les multiples réalités qui se cachent sous les usages tout aussi multiples de ce terme. Cette distinction est proche d'une autre distinction faite dans certains travaux de psychologie et en didactique professionnelle, témoin la distinction que fait Vergnaud (2001) entre les «formes prédicatives» de la connaissance (énoncées dans la littérature savante) et les «formes opératoires» (mobilisées en acte par les personnes). Les théories ont finalement des auteurs (qui les énoncent?), des formes d'énonciation (à quoi les reconnaît-on?), des contenus (de quoi parlent-elles?) et une origine (comment sont-elles développées? Comment évoluent-elles?) qui serviront de critères pour leur analyse comparée.

À partir d'une étude de l'activité des évaluateurs en situation de travail observée en administration et cabinet conseil (Tourmen, 2007), mais aussi d'une revue de littérature en éducation, le présent article examinera, dans un premier temps, les traces de théories spontanées de l'évaluation déjà observées, avant d'examiner, à rebours, les théories académiques de l'évaluation qui sont énoncées dans la littérature. Seront comparées les deux formes de théories selon les critères suivants :

- 1) auteur (qui?);
- 2) forme d'énonciation (comment?);
- 3) contenu (sur quoi?); et
- 4) origine et évolution (d'où?);

ce qui permettra d'en discuter les atouts et limites respectifs. Un tableau viendra synthétiser les propos exposés, et suivra une conclusion sur les liens entre les deux types de théories de l'évaluation en questionnant les conditions de leur possible enrichissement mutuel.

Les théories spontanées des évaluateurs

Qui? Les évaluateurs, théoriciens spontanés de l'évaluation

Chose n'est pas coutume, pour avancer sur la question des théories, il faut tout d'abord plonger dans celle des pratiques². Le présent article s'intéresse d'abord aux évaluateurs au travail. Que font-ils et que pensent-ils? Comme toute activité, l'évaluation peut se décrire comme une suite d'actions observables mais aussi de pensées et de raisonnements qui l'orientent en même temps qu'ils sont orientés par elle (Vergnaud, 1996). Vingt-quatre évaluateurs de programme professionnels ou débutants³ ont été rencontrés à l'occasion d'un travail de thèse (Tourmen, 2007, 2009) qui visait à observer et analyser leur activité en situation réelle. Cette recherche, en puisant dans les méthodes d'analyse du travail développées en psychologie du travail (Leplat, 1997), a permis de réaliser une analyse des tâches prescrites puis une analyse de l'activité d'un échantillon d'évaluateurs professionnels: 14 ont été rencontrés lors d'entretiens exploratoires semi-directifs portant sur leur travail, neuf ont accepté de se faire observer en train de travailler et quatre se sont livrés à une étude de cas expérimentale suivie par des entretiens cliniques (pour plus de détail sur les méthodes et l'échantillon, voir Tourmen, 2007, 2009).

L'analyse a montré que les évaluateurs professionnels conçoivent et négocient des démarches d'évaluation, rédigent des cahiers des charges (aussi appelés «mandats d'évaluation» au Québec), des référentiels et des rapports, recueillent et analysent des données, produisent des jugements. Pour faire ceci, ils diagnostiquent l'état de chaque situation dans laquelle ils se trouvent et raisonnent pour s'y adapter, en anticipant et sous-pesant les conséquences de leurs choix par leurs raisonnements. Ils raisonnent donc en cours d'activité, notamment sous forme de raisonnements conditionnels de type «*si... alors...*» (Vergnaud, 2001), comme le montre l'exemple de ce responsable d'évaluation en Conseil Régional qui énonce le raisonnement qu'il tient au moment où il rédige des questions évaluatives:

... les questions évaluatives vont tomber chez les élus (rires), si je commence à poser des problèmes de pilotage, à dire y'a des conflits entre les uns et les autres, euh... on risque de me rayer la question alors qu'elle est centrale⁴,

... ou encore au moment où il compose un comité de pilotage pour une évaluation:

soit j'écarte les élus, par principe parce que on peut craindre une pression sur l'évaluation et ce qu'elle produit, ou parce qu'on va avoir une réaction de rejet par un certain nombre d'acteurs qui vont dire c'est encore une évaluation qui est contrôlée par les élus donc elle est pas valable, bon, mais dans ce cas-là je peux avoir l'effet inverse que je crains beaucoup qui est le désintérêt des élus et la non-appropriation des résultats, l'expérience prouve que quand j'ai des élus membres de l'instance, j'ai une meilleure appropriation et une meilleure utilisation des résultats, donc je garde l'élu (Tourmen, 2007, p. 294).

La description de leur activité et de leurs raisonnements en cours d'action (voir Tourmen, 2009) ne sera pas développée davantage ici.

Comment? Formes de théories spontanées de l'évaluation

Les données récoltées laissent apparaître des traces de ce qui a été appelé, en reprenant le terme de Shadish et al. (1991), des « théories spontanées » de l'évaluation, théories décelées chez tous les évaluateurs rencontrés. Ainsi, le professionnel cité ci-dessus a mobilisé dans son raisonnement un élément de théorie – ou représentation – qu'il dit fondé sur l'expérience⁵ : « l'expérience prouve que quand j'ai des élus membres de l'instance, j'ai une meilleure appropriation et une meilleure utilisation des résultats » et c'est la convocation de cette proposition assertive qui lui permet de trancher entre plusieurs possibilités pratiques – à savoir inviter ou non l'élu en charge de la politique évaluée dans le comité d'évaluation qu'il est en train de constituer. Cet exemple n'est pas isolé : en effet, les professionnels expérimentés observés mobilisent et énoncent tous quantité de ces représentations,

- soit sous une forme conditionnelle, ce que Vergnaud (2001) appelle des « règles d'action » du type « si..., alors... », comme dans l'exemple ci-dessus (« si je commence à poser des problèmes de pilotage ... on risque de me rayer la question »). Les propositions conditionnelles peuvent être spécifiques à une situation ou généralisées à un ensemble de situations ;
- soit sous une forme assertive, ce que Vergnaud (2001) appelle des « théorèmes en actes » ou « propositions tenues pour vraies sur le réel », par exemple : « la facilité des experts, c'est de satisfaire le commanditaire », ou « toute chose nouvelle dans un paysage urbain effraie et on a tendance à la détruire », ou encore « avec des questions évaluatives précises, les critères découlent assez logiquement, simplement il faut se

poser la question de la faisabilité»⁶. On les appellera aussi «croyances» ou «connaissances en actes» (Vergnaud, 2001). Elles aussi peuvent être plus ou moins généralisées.

Selon l'étude philosophique des croyances, l'une et l'autre forme d'énonciation seraient comme les deux faces d'une même pièce (croyance), toutes deux pouvant être analysées comme des «dispositions à l'action» (Peirce, 1931): «dire que le sucre est soluble c'est dire que s'il était plongé dans l'eau, il se dissoudrait» (Engel, 1995, p. 25).

Ces éléments de théorie (ou, si l'on reprend les termes de Vergnaud, 2001), ces «invariants opératoires» constitutifs des schèmes organisateurs de l'activité et donc de l'expertise) irriguent et orientent les raisonnements tenus en cours d'action en permettant de générer des inférences⁷ en situation. Celles-ci permettent le diagnostic d'une situation donnée, le choix d'actions adaptées et l'anticipation des conséquences, éléments caractéristiques de l'activité experte selon des travaux en psychologie de l'expertise (Chi, Glaser, & Farr, 1988). Selon des travaux encore plus récents en psychologie cognitive, l'esprit humain tenterait en permanence de déterminer la probabilité des événements sur la base de ses observations minimales, utilisant des formes élaborées de raisonnement probabiliste de type bayésien (Gopnik & Tenenbaum, 2007; Jaynes, 2003; Yang & Shadlen, 2007) permettant de prédire un événement en fonction de la connaissance de ses occurrences passées mais aussi de l'état de la situation présente. Si ces éléments de théorie sont énoncés sous une forme assertive ou conditionnelle par les professionnels, ils peuvent aussi être inférés – prudemment – par le chercheur à partir de l'observation des actions des professionnels qui portent la trace de ce qui les a orientées. Il est important de noter que ces éléments de théorie sont en acte, non complètement explicites ni conscients pour les acteurs eux-mêmes, ni nécessairement formulés pour soi-même ou pour d'autres.

Le phénomène n'est pas nouveau ni isolé: les psychologues ont en effet depuis longtemps mis à jour le fait que toute action s'accompagne de formes – même non conscientes et parcellaires – de conceptualisations (Vergnaud, 1996), qu'il existe des formes stabilisées de connaissances (Richard, 1990), aussi appelées croyances, systèmes de représentations ou encore bases d'orientation (Savoyant, 1979) qui orientent l'activité et s'y développent dans un même mouvement. Cet ensemble de ressources difficile à caractériser et à nommer est parfois qualifié de «savoirs d'action» ou

«d'expérience» (Barbier, 1996). Autrement dit, il n'existerait pas d'actions efficaces sans une théorie (au sens d'un système de croyances/représentations et de concepts) sous-jacente de l'activité et des situations sur lesquelles et dans lesquelles agir. Les théories agissent comme autant de systèmes d'interprétation des situations permettant de s'y adapter.

Les évaluateurs de programme ne font pas exception à la règle, comme l'ont déjà observé les chercheurs américains Schwandt (2008) et Donaldson et Lipsey (2008), mettant à jour l'existence et l'importance d'un système de «savoirs pratiques» («*practical knowledge*») utilisé par les évaluateurs dans leur activité quotidienne et qui orienterait à la fois leur manière de «lire les situations d'évaluation» (Schwandt, 2008) et leurs choix d'action. Une des difficultés, soulignée par Schwandt (2008), est que ce type de savoirs est peu formalisé, souvent peu conscient pour les individus eux-mêmes, ce qui le rend difficile à exploiter et à étudier.

La sociologue française Demailly (2001) observait elle aussi que les évaluateurs mobilisent des «théories sociologiques spontanées» qui soutiennent l'analyse des systèmes évalués, comme celle qu'elle donne en exemple: «Des pratiques d'évaluation postulent par exemple que l'action organisée peut être pensée comme lien simple et instrumental entre des objectifs clairs et des résultats mesurables, ce qui est une vision très administrative de la réalité sociale» (p. 340). Shadish et al. (1991) avaient finalement qualifié ces éléments de «théorie spontanée» des évaluateurs: «La plupart des praticiens de l'évaluation sont des théoriciens spontanés de l'évaluation. Ils pensent à ce qu'ils font, savent quelle méthode utiliser dans quelle situation, pèsent les avantages et les inconvénients de leurs choix et apprennent de leurs réussites et de leurs échecs des évaluations passées» (p. 35). Les observations corroborent ces hypothèses: les évaluateurs rencontrés mobilisent bien des théories spontanées sur leur environnement de travail.

Sur quoi? Contenus des théories spontanées de l'évaluation

Sur quoi ces théories portent-elles? Grâce aux méthodes d'analyse du travail utilisées pour à la fois observer et faire expliciter les acteurs au travail (d'après la méthode des entretiens d'explicitation développée par Vermersch, 1994), les auteurs du présent article ont pu en savoir plus. Ainsi, les évaluateurs rencontrés théorisent sur les trois principales variables de situations auxquelles ils sont attentifs dans leurs diagnostics

au moment de la conception d'une évaluation et auxquelles ils cherchent à s'adapter (en même temps qu'ils cherchent à en anticiper et en influencer l'évolution) :

1. **Les stratégies des acteurs autour de l'évaluation**, qui donnent lieu à de fréquents diagnostics orientés par des théories sur la manière dont les différents acteurs agissent et réagissent à l'évaluation, leurs attentes explicites et implicites, la manière dont ils peuvent en utiliser les résultats ou les rejeter... C'est par exemple un responsable d'évaluation en ministère qui énonce que :

On a grosso modo trois pôles et on se situe entre ces trois pôles: on a celui des décideurs qui veut que ça soit vite, bien et clair, on a le pôle des opérateurs qui veulent l'évaluation pour pouvoir continuer à financer d'autres choses, donc l'évaluation alibi parfois, et puis on a le poids des consultants qui vous promettent la lune et puis vous donnent un pâté informe après et donc il faut naviguer entre ces trois pôles sans oublier de satisfaire les gens (Tourmen, 2007, p. 214);

2. **Les programmes publics**, qui donnent lieu à de tous aussi fréquents diagnostics orientés par des théories sur la façon dont les programmes publics (d'un domaine ou de plusieurs domaines d'action publique) «marchent» ou «fonctionnent», les façons dont ils sont conçus, mis en œuvre et les effets qu'ils peuvent ou non produire... C'est par exemple une responsable d'évaluation dans un Conseil Régional qui énonce, à propos des objectifs des programmes, que :

Il y a un certain nombre d'objectifs, ou d'intentions ou d'ambitions qui sont cités les uns à la suite des autres et qui vont pour autant pas se dérouler dans le même temps, hein, ou qui vont arriver les uns après les autres. Ça c'est effectivement de notre travail, nous quand on est avec les services, pour à partir de ces documents qu'ils ont quand même contribué à écrire, eh bien quelque part d'essayer de faire une cartographie un peu logique de ce qu'ils avaient essayé de vendre à leurs politiques ou de ce qu'ils avaient compris que demandaient les politiques (Tourmen, 2007, p. 203);

3. **Les moyens des évaluations** et la manière dont différentes méthodes d'évaluation peuvent produire différents résultats, leur coût et le temps nécessaire pour les déployer, leurs biais, «solidité» et leur pertinence... C'est par exemple un évaluateur en cabinet de conseil privé (sociologue de formation) qui énonce que «s'il n'y a pas de chiffres, les économistes pensent que ça ne vaut rien. Ils ne font pas confiance aux méthodes qualitatives» (Tourmen, 2009, p. 251).

Ces trois composantes des théories spontanées des évaluateurs peuvent être comparées aux cinq dimensions d'une théorie de l'évaluation qui ont été proposées par Shadish et al. (1991):

- 1) «la programmation sociale»;
- 2) «l'usage des connaissances»;
- 3) «l'attribution de valeur»;
- 4) la «construction des connaissances»; et
- 5) «la pratique de l'évaluation» (p. 35).

Nos résultats suggèrent une autre façon de regrouper les mêmes éléments, en accord avec les trois principales variables de situation identifiées comme critiques dans l'activité des évaluateurs observés (les stratégies des acteurs, les programmes publics et les moyens alloués, voir ci-dessus).

D'où? Origines et développement des théories spontanées de l'évaluation

Comment ces théories se sont-elles construites? Fidèles en cela à une tradition de pensée pragmatiste en Amérique du Nord (Dewey, 1993), Shadish et al. (1991) postulaient que les évaluateurs théorisent en tirant des leçons de leur expérience. Selon eux, ils «apprennent de leurs réussites et de leurs échecs des évaluations passées» (p. 35). Nos observations vont à nouveau dans le même sens: d'une part, des évaluateurs aux profils différents (en termes de formation et d'expériences) de notre échantillon ont énoncé des éléments de théories pour partie différents, notamment des conceptions différentes des attentes des acteurs et des techniques les plus crédibles pour évaluer. Elles sembleraient ainsi liées à l'expérience de chacun, ce qui va dans le sens des observations de Fitzpatrick (2004) qui a observé de telles variations individuelles à la fois dans les pratiques et dans les discours d'évaluateurs nord-américains, mettant notamment à jour des conceptions très différentes de ce qui assure la légitimité (crédibilité, validité, utilité, etc.) des résultats d'une évaluation; de plus, la comparaison systématique des évaluateurs débutants et expérimentés de l'échantillon a permis de montrer que les expérimentés énonçaient davantage de règles d'action et de théorèmes en acte que les novices, à tâche et consigne égales (Tourmen, 2007).

Il est possible de retrouver ici un phénomène étudié en psychologie cognitive, soit que l'expérience amène à raisonner par analogie avec des problèmes et des situations déjà rencontrés (De Montmollin, 1997). L'analogie est une forme de « raisonnement par ressemblance » qui consiste à étendre les propriétés d'une situation à une autre situation et qui permet d'élaborer des hypothèses sur une nouvelle situation, hypothèses que l'activité à venir pourra confirmer ou infirmer. Ces hypothèses, si elles sont vérifiées, peuvent devenir de véritables théories – ou réseaux de « propositions tenues pour vraies sur le réel » et de « règles d'action » (Vergnaud, 1996) – que les professionnels se constituent et appliquent à l'ensemble d'une classe de situations, dans un processus de « généralisation » au cœur des processus d'apprentissage par l'expérience. Une des limites au processus reste la difficulté potentielle d'observer les résultats des évaluations de programme, qui ont lieu sur du long terme (parfois plusieurs années) et dont les effets sont parfois difficiles à estimer (la satisfaction réelle d'un commanditaire?) voire inaccessibles (les effets à long terme sur un processus de décision publique?). Reste la question de la validité d'une telle théorie, de son degré de généralité ou de vraisemblance. Une partie des théories spontanées des évaluateurs ont aussi pu être verbalisées, partagées et transmises au sein d'un groupe de travail, lors d'un tutorat avec un plus expérimenté ou même en formation, comme observé dans le travail de thèse. Certaines deviennent alors des croyances partagées dans une culture ou micro-culture de travail donnée. Au risque aussi de généraliser sur des cas particuliers, de se contenter de suivre le sens commun ou de s'enfermer dans des théories limitées et limitantes de l'évaluation. Par exemple, que penser d'une telle théorie énoncée par une professionnelle, « Les élus ne lisent jamais les rapports d'évaluation » (Tourmen, 2007, p. 251): est-elle trop générale? Sur quoi est-elle fondée? Peut-elle être discutée?

Il est donc possible de conclure provisoirement que les évaluateurs sont des théoriciens spontanés de l'évaluation, au sens où ils théorisent, à partir de leur pratique et des échanges avec des pairs, à la fois l'influence des contextes sur l'évaluation, l'enchaînement des choix d'action possibles et leurs conséquences probables/possibles, et que ce corpus de théories (ou de « savoirs pratiques ») organise leurs actions futures, en même temps qu'il pourra y évoluer. Ces éléments de théorie sont-ils pour autant différents des écrits sur l'évaluation, où sont aussi présentes certaines formes de théorisation de l'évaluation?

Théories académiques de l'évaluation

Qui? Théories des chercheurs et praticiens de l'évaluation

Les auteurs vont à présent se pencher sur les tentatives de théorisation présentes dans la littérature en évaluation, à la fois dans des ouvrages, des articles scientifiques et des guides pratiques, théories qualifiées d'«académiques» par Shadish et al. (1991), au croisement de plusieurs disciplines. Le travail de thèse a en effet été l'occasion de réaliser une revue de littérature croisée sur l'évaluation de programme et l'évaluation en éducation, sur le plan international (Tourmen, 2007), puisant aussi bien dans les ouvrages et revues spécialisées dans le domaine de l'évaluation qu'en sciences de l'éducation, en sciences politiques et en sociologie. Il ressort que les écrits sur l'évaluation sont majoritairement prescriptifs et méthodologiques, présentant l'évaluation comme un ensemble d'étapes à suivre incluant la présentation de différentes méthodes pour mener la démarche évaluative mais aussi pour recueillir et traiter des données sur le déroulement et les impacts des programmes publics. Certains textes, moins nombreux, sont davantage descriptifs; ils prennent alors souvent la forme d'études de cas présentant l'évaluation de programmes particuliers et les résultats obtenus, cette présentation étant l'occasion de discuter, par exemple, une méthode innovante ou la validité de certains indicateurs, entre autres choses. D'autres textes s'intéressent enfin aux contextes d'évaluation, à leur émergence historique, par exemple, et aux activités évaluatives. Il faudra y revenir. Les auteurs de ces textes sont en majorité des chercheurs, même s'il y a aussi des fonctionnaires, des consultants, des enseignants qui conceptualisent à partir de leurs pratiques.

Comment? Formes de théorisations académiques de l'évaluation

Dans cette littérature, il est de coutume de distinguer les «pratiques» d'évaluation des «théories» de l'évaluation (Schwandt, 2008), les deux étant de prime abord opposées. On trouve ainsi le terme de «théorie» utilisé dans un premier sens, très courant en Amérique du Nord: les auteurs font référence à l'existence d'une pluralité de «théories» de l'évaluation ayant émergé chez des auteurs distincts ou à des périodes distinctes. La théorie renvoie ici à un «modèle» prescriptif défini par un auteur ou un courant d'évaluation, ce modèle préconisant pourquoi et comment évaluer (finalités, étapes à suivre, démarches et méthodes recommandées, etc.).

Une des formalisations les plus connues de ces théories est le schéma en forme d'arbre d'Alkin et Christie (2004), dit des « racines de l'évaluation », qui distingue par exemple l'*Utilization-Focused Evaluation* (Patton, 1986) de l'*Empowerment evaluation* (Fetterman, Kaftarian, & Wandersman, 1996) et de la *Theory-driven evaluation* (Chen, 1990), etc. De fréquentes distinctions sont également faites entre types d'évaluations selon les fonctions principalement visées: fonction de contrôle ou de régulation (Vial, 2001), fonction sommative ou formative (Scriven, 1967), posture de contrôleur ou posture d'« ami critique » (Jorro, 2006), etc. De nombreux auteurs ont tenté de regrouper ces approches en « paradigmes » qui se seraient succédés historiquement avant de cohabiter (voir par ex., Cardinet, 1986b; De Ketele, 1993; Guba & Lincoln, 1989; Monnier, 1992; Nadeau, 1988; Patton, 1986; Rodrigues, 2002; Stufflebeam, 1980), se différenciant avant tout sur le degré d'implication des porteurs d'intérêt et les méthodes mises en avant (Christie, 2003; Rodrigues, 2002). Différentes définitions de l'évaluation sont aussi promues par ces approches (par ex., évaluation mesure/jugement). Patton (1986) alla jusqu'à parler d'une « guerre des paradigmes » qui agiterait le champ professionnel et académique des évaluateurs. Ces théories sont en effet en concurrence, car elles sont souvent normatives et prescriptives: elles constituent différents modèles de la manière dont les évaluations devraient être conduites, avec leur lot de prescriptions pratiques. Comme l'ont déjà noté Fitzpatrick, Christie, et Mark (2008), « Les évaluateurs lisent et écrivent beaucoup au sujet des théories, approches et modèles de l'évaluation – à savoir comment l'évaluation devrait être pratiquée » (p. 355). Alkin (2004) reconnaissait lui aussi que les théories de l'évaluation offraient avant tout un ensemble de prescriptions, « interdictions » et « principes d'action » spécifiant « ce qu'est une bonne évaluation » et « comment elle devrait être menée ».

Shadish et al. (1991) furent parmi les premiers à utiliser le terme « théorie » dans un autre sens. Selon eux,

une théorie désigne un corps de savoirs qui organise, catégorise, décrit, prédit, explique et aide à comprendre et contrôler un sujet. Les théories le font de différentes façons, en cherchant des lois invariantes, utilisant des définitions et des axiomes pour déduire des propositions testables, et décrivant des processus causaux qui expliquent certaines relations (p. 30).

Ils ont ainsi évoqué la possibilité d'une théorie générale et unifiée de l'évaluation, dans un autre sens du terme, dépassant ainsi les «listes athéoriques de méthodes» et la «guerre des paradigmes» dont parlait Patton (1986). Une telle théorie n'aurait pas vocation à être prescriptive – même si elle pouvait déboucher sur des principes d'action, – elle viserait avant tout à comprendre, à décrire et à prédire les phénomènes impliqués dans des démarches d'évaluation et les lois qui les gouvernent. Toujours selon Shadish et al. (1991), «la théorie d'évaluation idéale (jamais atteignable) décrirait et expliquerait pourquoi certaines pratiques d'évaluation produisent certains résultats selon les situations que les évaluateurs rencontrent» (p. 31). Une telle théorie fonderait le choix des méthodes utilisées : «Une théorie de l'évaluation nous dit quand, où et pourquoi une méthode devrait être préférée à d'autres ..., leurs apports et leurs limites» (p. 34). Enfin, selon eux, en 1991, une telle théorie était loin d'être achevée.

Faudrait-il même une théorie générale de l'évaluation ? Cette question est débattue, certains auteurs en dressent les avantages, affirmant par exemple qu'elle constituerait un langage commun facilitant la communication entre évaluateurs mais aussi avec les non-spécialistes, qu'elle aiderait les praticiens à comprendre ce qu'ils font, et que cette base de connaissances aiderait à stabiliser une profession d'évaluateur et une identité professionnelle (Mark, 2005). Ce dernier auteur souligne toutefois qu'une telle théorie ne résoudrait pas toutes les difficultés pratiques rencontrées par les évaluateurs mais qu'elle pourrait servir à guider les jugements. Comme dans le cas des théorisations spontanées, ces théorisations dites «académiques» sont composées d'un système d'assertions. La littérature en est pleine, les auteurs du présent article n'en ont repris ici que quelques-unes à titre d'exemple : «Comme les données de l'évaluation, les critères ou les normes de jugement n'existent pas à l'état naturel ; ils sont toujours le résultat de processus spécifiques conduisant à leur élaboration» (Barbier, 1985, p. 77) ; c'est la construction des «référents» qui guide celle des «référés» (Figari, 1994, p. 48) ; «Aucune approche n'est vraie ou fausse, meilleure ou pire dans l'absolu» (Patton, 1986, p. 181), ou «Plusieurs facteurs peuvent influencer l'usage de l'évaluation au Congrès mais la politique est toujours un facteur dominant» (Patton, 2008, p. 20), ou encore «L'usage de l'évaluation dépend de l'engagement à accepter le test de réalité» (Patton, 2008, p. 47). De ces assertions peuvent dériver des règles

d'action (si je veux faire ceci, alors il vaut mieux commencer par là ou si je fais ça, alors il va se passer ça...), ce qui confère à donner à ces théories un caractère potentiellement opératoire.

Sur quoi? Contenus des théories académiques

Si, comme évoqué brièvement, les théories au sens de «modèles» ou «paradigmes» de l'évaluation portent sur les finalités, étapes et méthodes du processus, sur quoi une théorie dans le second sens du terme – un corps de savoirs décrivant les phénomènes évaluatifs – porterait-elle? Selon Shadish et al. (1991), une telle théorie pourrait porter sur :

- les contextes des évaluations ;
- leur déroulement ;
- leurs effets.

Selon Shadish et al. (1991), «Elle clarifierait (a) les activités, processus et buts des évaluations ; (b) expliquerait les relations entre les activités évaluatives et les processus et buts qu'elles favorisent ; et (c) ferait des tests empiriques pour identifier ceux qui sont en conflit avec la recherche ou certains savoirs en évaluation» (p. 31). Et d'ajouter :

une théorie de l'évaluation inclut des questions de méthodes, mais pas juste de méthodes. Sa théorie inclut un vaste champ de décisions sur la forme, la conduite et les effets des évaluations. Pour aider les évaluateurs dans le choix des méthodes, il est nécessaire de discuter de philosophie des sciences, de politiques publiques⁸, de théories de la valeur et de théories de l'usage (p. 31),

selon les cinq composantes qu'ils donnent à une théorie et qui ont été précédemment citées. Sans cela, l'évaluation ne serait qu'une méthodologie de sciences sociales appliquée à certains types d'objets. Ils pensent au contraire que l'évaluation «est en train de développer un corps unique de savoirs qui la différencie d'autres spécialités pourtant proches» (p. 31). Une théorie générale de l'évaluation serait donc possible. Tentons d'aller plus loin.

Si, comme le souligne Figari dans ce numéro, aucune théorie intégratrice de l'évaluation (quels que soient ses objets) n'existe à ce jour, plusieurs auteurs, issus de différentes disciplines⁹ et notamment en sciences de l'éducation, ont tenté d'avancer dans ce sens. Ils ont posé les bases de théorisations qui décrivent ce que serait l'acte d'évaluer et les conditions de réussite et d'échec des évaluations. Il est possible, par exemple, de trouver une trace d'une telle volonté de théorisation de l'acte évaluatif chez Scriven qui a décrit la «logique de l'évaluation» en 1980 en quatre étapes deve-

nues célèbres : « Le type d'évaluation le plus commun implique de déterminer des critères de mérite (généralement à partir d'une évaluation des besoins), de poser des standards de performance ... et de déterminer la performance » (p. 117) par une mesure, avant de synthétiser les résultats dans un jugement de valeur. Des auteurs en éducation comme Bonniol (1981), Barbier (1985), Cardinet (1986a), Hadji (1989), Figari (1994, voir aussi Tourmen & Figari, 2006), Hurteau, Houle, et Mongiat (2009), ont aussi tenté de décrire les actes constitutifs ou « invariants » de toute évaluation, quel que soit son objet. Ils en discutent l'ordre, le déroulement, l'outillage et les effets, en s'inspirant des théorisations de la mesure réalisées dans les années 1970 en docimologie (Bonboir, 1972) mais en y incluant une théorisation du jugement. Ils définissent par exemple l'acte d'évaluation comme une production de jugements issus de la comparaison entre des « référés » (données sur l'état d'un objet) et des « référents » (systèmes de références permettant de leur donner un sens et une valeur), les évaluateurs devant identifier et souvent reconstruire référents et référés (Barbier, 1985), dans ce que Figari (1994) appelle un processus continu de « référentialisation ». Hurteau et al. (2009) utilisent un modèle proche de l'activité évaluative pour examiner le mode de construction et de justification des jugements (sous l'angle de leur « crédibilité ») dans des rapports d'évaluation. Des concepts propres au domaine de l'évaluation sont énoncés et discutés dans ces travaux : critère, indicateur, référent, étalon, standard, référentiel, mesure, jugement, etc., qui pourraient constituer autant de concepts centraux dans une théorie de l'évaluation.

D'autres travaux vont au-delà de l'étude du déroulement de toute activité évaluative et tentent de théoriser et de documenter la pertinence des activités d'évaluation de programme en fonction des contextes. Rossi et Freeman (1993) ont ainsi identifié des méthodes plus efficaces que d'autres quand il s'agit d'évaluer des programmes établis, nouveaux ou à modifier. Patton (1986) distingue également des types d'évaluation plus ou moins adaptés aux usages et aux types de programmes, en concluant « une évaluation réussie (qui est pratique, éthique, utile et pertinente) est due aux conditions particulières d'une situation – un mélange d'acteurs, de jeux politiques, d'histoire, de contextes, de ressources, de contraintes, de valeurs, de besoins, d'intérêts et de chance » (p. 315). D'autres auteurs tentent de théoriser les conditions de réussite et d'échec des évaluations à partir de l'étude de leurs effets positifs et pervers (Demailly, 2001), dans une démarche appelée « méta-évaluation critique » (Demailly, 2006). Par

exemple, selon elle, les conditions de réussite d'une évaluation sont les mêmes que celles d'une innovation réussie, notamment la participation des acteurs concernés, la poursuite d'objectifs « démocratiques » et la présence d'un chef de projet ayant « la capacité de convaincre » (Demaillly, 2001, p. 28). Elle développe également des exemples inverses d'évaluation ayant eu des effets négatifs sur les systèmes concernés (voir aussi Tourmen & Demaillly, 2012). Dans la même lignée, Hurteau (2008) a étudié le facteur « implication des partenaires dans l'évaluation » comme source de réussite (assurer la crédibilité et l'utilité des résultats) mais aussi source de biais et de remises en cause possibles. Cette approche repose sur un postulat de recherche clairement formulé par Demaillly (2001) : « L'évaluation peut être la meilleure ou la pire des choses » (p. 30).

D'où? Origines et développement des théories académiques

Les théories académiques sont le fruit de travaux de recherche, incluant des travaux dits de « méta-évaluation » (Demaillly, 2006), mais aussi de réflexions de praticiens à partir de leur expérience. Elles sont toutefois éclatées dans différentes disciplines, différentes revues, chez différents auteurs à tel point qu'une synthèse exhaustive et pluridisciplinaire semble aujourd'hui difficile. Une des critiques qui peuvent être adressées aux théories académiques est qu'elles ne traitent que de ce que les auteurs auront choisi de traiter, pouvant ainsi méconnaître des pans entiers de préoccupation pourtant pris en charge par les théories spontanées mobilisées dans l'action. Ainsi, les lectures faites par les rédacteurs du présent article montrent que des considérations liées à la « faisabilité » des évaluations et à leur « crédibilité », très développées dans les théories spontanées des évaluateurs (Tourmen, 2009), sont peu présentes dans les théories académiques de l'évaluation, qui mettent davantage l'accent sur des considérations liées à la validité instrumentale et technique des évaluations. Shadish et al. (1991) avaient aussi noté qu'« aucune des théories que nous avons étudiées ne traite explicitement des cinq composantes » (p. 35) qu'ils ont identifiées, chaque théoricien proposant des « hypothèses » développant davantage l'une ou l'autre des cinq dimensions selon ses travaux. Ainsi, Scriven s'intéresse à la construction du jugement, tandis que Campbell s'intéresse davantage à la construction du savoir (Shadish et al., 1991). Le caractère académique de ces théories leur confère toutefois, *a priori*, un degré de généralité et de crédibilité plus important que les théories issues de l'expérience, même si, en sciences sociales comme en évaluation, plu-

sieurs modes de construction de la légitimité du savoir peuvent exister voire s'opposer. Shadish et al. (1991) déploreraient d'ailleurs que les théories de l'évaluation reposent sur des corps d'assertions qui ne soient pas davantage testés et enrichis par des données issues du terrain, ce qui les rendrait plus solides. Ils appellent de leurs vœux un plus grand va-et-vient entre spéculations et données. Tous ces efforts vont donc dans le sens de la constitution d'un corps de savoirs propre à l'évaluation, bien que cette construction soit éclatée, ce qui rend l'accumulation des savoirs difficile. Une théorie ou une «méta-théorie» de l'évaluation, qui dépasserait l'exposé des méthodes et tenterait d'intégrer leurs apports dans une réflexion sur les contextes, le déroulement et les conditions de réussite et d'échec des évaluations, reste donc à construire.

Synthèse

Le tableau suivant récapitule ce qu'a produit l'examen des deux formes de théorisation de l'évaluation.

Tableau 1

Synthèse comparant les deux formes de théorisation de l'évaluation

Critères de comparaison	Théories spontanées de l'évaluation	Théories académiques de l'évaluation
Théories de qui? Leurs auteurs	Praticiens de l'évaluation	Littérature sur l'évaluation : chercheurs et praticiens réflexifs
Comment ?	<p>Systèmes de représentations en actes :</p> <ul style="list-style-type: none"> • Théorèmes en actes aussi appelés « propositions tenues pour vraies sur le réel » ou croyances/connaissances énoncées sous forme assertive. • Raisonnements conditionnels (« si... alors ») aussi appelés règles d'action visant l'adaptation, le diagnostic et l'anticipation. 	<p>Stocks d'énoncés :</p> <ul style="list-style-type: none"> • Prescriptifs – ou « modèles » de l'évaluation en concurrence qui prescrivent son déroulement (finalités à viser, étapes à suivre, méthodes à privilégier...). • Descriptifs ou tentatives de théorisation visant à décrire et prédire les phénomènes évaluatifs (déroulements, origines, effets et lois qui les gouvernent) et d'en déterminer les conditions de réussite et d'échec.

Sur quoi? Contenu des théories	Représentations qui portent sur : • Les programmes publics • Les stratégies des acteurs autour de l'évaluation • Les moyens pour évaluer et permettent de diagnostiquer le contexte, de déduire l'enchaînement des actions et d'en anticiper les conséquences possibles.	Énoncés qui portent sur : • Les situations d'évaluation • Le déroulement des activités évaluatives • Leurs effets selon les contextes à travers un système de concepts propres à l'évaluation (critères, indicateurs, référents...).
D'où viennent-elles? Origines évolutions des théories	Représentations développées par l'expérience : • Activité de travail et observation de ses résultats (leçons tirées) • Transmission directe sur les lieux de travail (orale) ou indirecte (lectures, conférences...).	Énoncés issus soit : • De recherches (énoncé et/ou test d'hypothèses) • De «méta-évaluations critiques» (Demailly, 2006) • De conceptualisations de l'expérience.
Discussion : Atouts possibles	Représentations : • Opératoires : efficaces pour traiter les problèmes du travail • Pertinentes : proches des préoccupations quotidiennes du travail et de la spécificité de ses contextes • Peuvent être transmises sur les lieux de travail ou en formation (par tutorat par exemple).	Énoncés : • Grand degré de validité (généralité) quant aux phénomènes étudiés • Possibilité de test, vérification, validation ou de rejet pour fonder les pratiques sur des lois éprouvées • Concepts scientifiques faisant système dans une discipline • Corps unique de savoirs/concepts propres à la communauté des évaluateurs.
Discussion : Limites possibles	Représentations : • Peu conscientes, peu formalisées • Parfois difficiles à transmettre/ verbaliser • Degré de validité limité, risque de généraliser sur des cas particuliers • Concepts quotidiens	Énoncés : • Parfois éloignés des préoccupations du travail, ils risquent d'être trop généraux par rapport aux contextes de travail et à l'activité, difficiles à utiliser • En concurrence – lequel choisir ou croire?

- Possibilité de réutiliser des croyances issues du sens commun et non mises en doute
 - Résultats du travail parfois longs/difficiles à observer, d'où une difficulté à théoriser à partir de l'expérience.
 - Souvent restés au statut d'hypothèses théoriques, peu de vérifications empiriques (Shadish et al., 1991)
 - Prescriptifs, moins souvent descriptifs
 - Éclatés entre différentes disciplines, revues, collections, auteurs... manque d'unité, accumulation de savoirs difficile
 - Pans entiers de préoccupation pratiques négligés et non traités.
-

Conclusion

Les auteurs de cet article ont volontairement séparé les théories «spontanées» et les théories «académiques» de l'évaluation afin de pousser leur examen comparatif le plus loin possible, en prenant à la lettre la distinction de Shadish et al. (1991). Toutefois, le temps est venu, en conclusion, de nuancer et de discuter cette distinction. En effet, les théories «académiques» ne sont pas totalement déconnectées des pratiques et encore moins des théories dites «spontanées» des évaluateurs. Au contraire, même si elles sont construites de manière différente, elles peuvent traiter en partie des mêmes objets voire s'influencer les unes et les autres.

Les auteurs du présent article ont en effet constaté que les théories académiques peuvent traiter d'objets également pris en charge par les théories spontanées. Les deux s'intéressent, certes de manière parcellaire et spécifique,

- 1) aux situations ou contextes d'évaluation ;
- 2) au déroulement des évaluations ; et
- 3) à leurs effets possibles.

Les deux peuvent donc dire quelque chose des conditions de réussite et d'échec des évaluations selon les contextes, ce qui crée un champ de préoccupation commun potentiellement fécond pour un dialogue entre praticiens et théoriciens. Ainsi, des éléments de théorisation académiques peuvent-ils venir renforcer – ou au contraire contredire – des assertions des

évaluateurs de terrain? Les théories spontanées des évaluateurs peuvent ainsi être validées ou invalidées par les théories académiques, du moins mises en discussion, nourries, mises en mouvement. À l'inverse, les théories des évaluateurs peuvent aussi venir nourrir les théories académiques, en fournissant des assertions qu'il s'agira de tester, comme l'énonçaient déjà Shadish et al. (1991). Selon eux, «les concepts pragmatiques développés dans la pratique constituent certainement la base la plus importante pour les théories académiques» (p. 35). Par exemple, que dire d'une telle assertion formulée par une responsable d'évaluation en Conseil Régional: «Les meilleures études sont toujours faibles dans leur recommandations» (Tourmen, 2007, p. 226)? Des recherches permettent-elles de tester ces affirmations, de les discuter, d'en savoir plus sur leurs conditions de validité? Voilà de quoi renouveler l'agenda de la recherche en évaluation, en identifiant et en testant des théories issues de l'expérience

De plus, des énoncés théoriques – concepts, représentations, croyances, connaissances, etc. – circulent entre les deux univers à tel point qu'il est parfois illusoire de vouloir distinguer leur origine. Les théories académiques sont en partie envahies d'éléments de théories spontanées que les praticiens réflexifs – voire les chercheurs – reprennent à bon compte, valident ou simplement énoncent. À l'inverse, les évaluateurs praticiens sont baignés dans un univers sémiotique d'instruments, de guides et d'ouvrages où des concepts scientifiques circulent à l'occasion d'échanges, de colloques, de discussions entre pairs et avec des chercheurs, qui restent nombreux à pratiquer l'évaluation de programme. Si l'on va au bout du raisonnement, l'expérience prend elle-même place dans des environnements saturés de discours et de prescriptions porteurs des mêmes concepts que ceux qui peuvent être travaillés dans la littérature, les théories «spontanées» n'ayant alors plus grand-chose de spontané!

Le propos des auteurs de cet article était finalement de montrer que des éléments de théories (concepts, croyances et connaissances) circulent entre univers collectifs et individuels (pour reprendre la distinction du sociologue Simmel, voir Valsiner, 2007) et entre univers académiques et univers de travail... Le mouvement de développement des théories va alors sans cesse de l'univers collectif vers les individus (appropriation) et des individus vers les univers collectifs (transmission) dans un va-et-vient créatif permanent entre les individus et les cultures (Valsiner, 2007). Tout dispositif favorisant l'expression et la circulation des théories entre les univers

individuels et collectifs (comme du tutorat, des groupes de discussion, des sessions de formation mettant les deux en discussion, des réseaux mêlant chercheurs et praticiens, etc.) ne peuvent dès lors qu'être encouragés dans l'espoir de favoriser ces processus d'apprentissage et de germination des concepts et théories individuels et collectifs.

NOTES

1. Le premier terme est celui qui est largement utilisé en Amérique du Nord (y compris au Québec) tandis que le second est davantage utilisé en France. Les subtiles distinctions qui sont parfois données ne font pas l'objet du présent article. L'intérêt se porte ici sur les programmes et sur les politiques de tous les secteurs d'action publique, le secteur éducatif – dont s'occupe en priorité cette revue - étant bien sûr inclus.
2. Sauf dans une épistémologie pragmatiste où ce qui est premier ce sont les interactions entre des organismes et leur environnement, les connaissances et théories étant avant tout des outils permettant l'adaptation et la résolution des situations embrouillées, voir Dewey (1993).
3. Vingt-quatre professionnels de l'évaluation ont été objet de l'enquête : six évaluateurs en cabinets de conseil privé (dont deux débutants), ainsi que 16 responsables d'évaluation en administration (dont trois débutants) et deux étudiants en Master d'évaluation.
4. Toutes les citations en italique sont des extraits intégralement retranscrits des propos des professionnels rencontrés ou observés durant la thèse.
5. La sienne ? Celle des autres ? Il ne le dit pas.
6. Les deux premières assertions ont été énoncées par un responsable d'évaluation en Ministère, la troisième par un responsable d'évaluation en Conseil Régional.
7. L'inférence est un processus commun de la pensée qui consiste à tirer des conclusions à partir de prémisses ou d'observations, la pensée procédant par inférences successives (Dewey, 1993 ; Valsiner, 2007).
8. La question du statut et du rôle des théories des programmes/politiques publiques dans l'évaluation est fréquemment débattue dans la littérature, et fait l'objet d'une recherche en cours de la part des auteurs. Elle ne sera pas discutée ici.
9. Les rédacteurs du présent article ont fait le choix de ne pas séparer les auteurs selon leurs disciplines et leurs objets, faisant l'hypothèse, en suivant les travaux de Shadish et al. (1991), qu'une théorie de l'évaluation ne pourrait être que pluridisciplinaire. La possibilité d'étendre une théorie de l'évaluation issue d'un domaine (en éducation, par exemple) à d'autres champs et objets fait toutefois débat au sein de la communauté des évaluateurs.

RÉFÉRENCES

- Alkin, M. C. (2004). *Evaluation roots*. Thousand Oaks, CA: Sage.
- Alkin, M. C., & Christie, C. A. (2004). An evaluation theory tree revisited. In M. C. Alkin (Ed.), *Evaluation Roots* (pp. 12-66). Thousand Oaks, CA: Sage.
- Barbier, J. M. (1985). *L'évaluation en formation*. Paris, France: PUF.
- Barbier, J. M. (1996). *Savoirs théoriques et savoirs d'action*. Paris, France: PUF.
- Bonboir, A. (1972). *La docimologie*. Paris, France: PUF.
- Bonniol, J. J. (1981). *Déterminants et mécanismes des comportements d'évaluation d'épreuves scolaires* (Thèse de doctorat non publiée). Université Michel de Montaigne Bordeaux III, France.
- Cardinet, J. (1986a). *Pour apprécier le travail des élèves*. Bruxelles, Belgique: De Boeck-Wesmael.
- Cardinet, J. (1986b). *Évaluation scolaire et mesure*. Bruxelles, Belgique: De Boeck.
- Chen, T. H. (1990). *Theory-driven evaluations*. Newbury Park, CA: Sage.
- Chi, T. H., Glaser, R., & Farr, M. J. (1988). *The Nature of Expertise*. Hillsdale, NJ: Laurence Erlbaum associates.
- Christie, C. A. (2003). What guides evaluation? A study of evaluation practice maps onto evaluation theory. In C. A. Christie (Ed.) *The practice-theory relationship in evaluation. New Directions for Evaluation* (pp. 7-35). San Francisco, CA: Jossey-Bass.
- De Ketele, J. M. (1993). L'évaluation conjugée en paradigmes. *Revue française de pédagogie*, 10, 59-80.
- Demaillly, L. (2001). *Évaluer les politiques éducatives*. Bruxelles, Belgique: De Boeck.
- Demaillly, L. (2001). Les évaluations servent-elles à quelque chose? In G. Figari & M. Achouche (dir), *L'activité évaluative réinterrogée* (pp. 340-355). Bruxelles, Belgique: DeBoeck.
- De Montmollin, M. (1997). *Vocabulaire de l'ergonomie*. Toulouse, France: Octarès.
- Dewey, J. (1993). *Logique, théorie de l'enquête*. Paris, France: PUF.
- Donaldson, S. I., & Lipsey, M. W. (2006). Roles for theory in contemporary evaluation practice: Developing practical knowledge. In I. Shaw, J. Greene, & M. Mark (Eds.), *Handbook of Evaluation* (pp. 56-76). London, UK: Sage.
- Engel, P. (1995). Les croyances. In D. Kambouchner (Ed.), *Notions de philosophie, II* (pp. 9-101). Paris, France: Folio.
- Fetterman, D. M., Kaftarian, S. J., & Wandersman, A. (1996). *Empowerment evaluation. Knowledge and tools for self-assessment and accountability*. Thousand Oaks, CA: Sage.
- Figari, G. (1994). *Évaluer: Quel référentiel?* Bruxelles, Belgique: De Boeck.
- Fitzpatrick, J. L. (2004). Exemplars as case studies: Reflections on the link between theory, practice and context. *American Journal of Evaluation*, 25, 541-559. doi:dx.doi.org/10.1177%2F109821400402500409
- Fitzpatrick, J. L., Christie, T. A., & Mark, M. M. (2008). *Evaluation in action: interviews with expert evaluators*. Thousand Oaks, CA: Sage.

- Furubo, J. E., Rist, R. C., & Sandahl, R. (2002). *International Atlas of Evaluation*. New Brunswick, NJ: Transaction Publishers.
- Gopnik, A., & Tenenbaum, J. (2007). Bayesian networks, Bayesian learning and cognitive development. *Developmental Science (special section on Bayesian and Bayes-Net approaches to development)*, 10(3), 281-287. Gopnik et Tenebaum 2007. doi:dx.doi.org/10.1111%2Fj.1467-7687.2007.00584.x
- Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Newbury Park, CA: Sage Publications.
- Hadji, C. (1989). *L'évaluation, règles du jeu*. Paris, France: ESF.
- Hurteau, M. (2008). L'implication des détenteurs d'enjeux (stakeholders) au sein de la démarche d'évaluation de programme: problème et/ou solution? *Mesure et évaluation en éducation*, 31(3), 63-76. doi:dx.doi.org/10.1177%2F1356389009105883
- Hurteau, M., Houle, S., & Mongiat, S. (2009). How legitimate and justified are judgments in program evaluation? *Évaluation*, 15(3), 307-319.
- Jaynes, E. T. (2003). *The logic of science*. Cambridge, UK: Cambridge University Press.
- Jorro, A. (2006). Devenir ami critique. Avec quelles compétences et quels gestes professionnels? *Mesure et évaluation en éducation*, 29(1), 31-44.
- Leplat, J. (1997). *Regards sur l'activité en situation de travail. Contribution à la psychologie ergonomique*. Paris, France: PUF.
- Mark, M. M. (2005). Evaluation theory or what are evaluation methods for? *The evaluation exchange*, 11(2), 2-3.
- Monnier, E. (1992). *Évaluations de l'action des pouvoirs publics*. Paris, France: Economica.
- Nadeau, M. A. (1988). *L'évaluation de programme, théorie et pratique*. Québec, QC: Les Presses de l'Université Laval.
- Patton, M. Q. (1986, 2^e éd., 2008). *Utilization Focused Evaluation*. Thousand Oaks, CA: Sage.
- Peirce, C. S. (1931). *Collected papers of Charle Sanders Peirce* (Vol. 2). Cambridge, MA: Harvard University Press.
- Rabardel, P. (1995). *Les hommes et les technologies: approche cognitive des instruments contemporains*. Paris, France: A. Colin.
- Richard, J. F. (1990). *Les activités mentales. Comprendre, raisonner, trouver des solutions*. Paris, France: Armand Colin.
- Rodrigues, P. (2002). *Avaliação da formação pelos participantes em entrevista de investigação*. Porto, Portugal: Fundação Calouste Gulbenkian.
- Rossi, P. H., & Freeman, H. E. (1993). *Evaluation: a systematic approach* (5th ed). Newbury Park, CA: Sage publications.
- Savoyant, A. (1979). Éléments d'un cadre d'analyse de l'activité: quelques conceptions essentielles de la psychologie soviétique. *Cahiers de psychologie*, 22, 248-252.
- Schwandt, T. (2008). The relevance of practical knowledge traditions to evaluation practice. In N. L. Smith & P. R. Brandon (Eds.), *Fundamental issues in evaluation* (pp. 29-40). New York, NY: The Guilford Press.
- Scriven, M. S. (1967). The methodology of evaluation. In R. Tyler, R. Gagne, & M. Scriven (Eds.) *Perspectives of curriculum evaluation* (pp. 39-83). Chicago, IL: Rand McNally.

- Scriven, M. S. (1980). *The logic of evaluation*. Inverness, CA: Edgepress.
- Shadish, W. R., Cook, T. D., & Leviton, L. C. (1991). *Foundations of Program Evaluation Theories of Practice*. Newbury Park, CA: Sage publications.
- Stufflebeam, D. (1980). *L'évaluation en éducation et la prise de décision*. Ottawa, Ont: NHP.
- Tourmen, C., & Figari, G. (2006). La référentialisation : une façon de modéliser l'évaluation de programme, entre théorie et pratique. Vers une comparaison des approches en France et au Québec. *Mesure et évaluation en éducation*, 29(3), 5-25.
- Tourmen, C. (2007). *Les compétences des évaluateurs, le cas des évaluateurs de politiques publiques* (Thèse de doctorat non publiée). Université de Grenoble II, France.
- Tourmen, C. (2009). Evaluators' Decision Making: Between Theory, Practice and Experience. *American Journal of Evaluation*, 30(1), 7-30. doi:dx.doi.org/10.1177%2F1098214008327602
- Tourmen, C., & Demailly, L. (2012). Évaluations en tensions : pour une critique documentée et constructive de l'évaluation. In Baron, G., & Matyjasik, N. (Eds.) *L'évaluation des politiques publiques. Défis d'une société en tension* (pp. 51-67). Paris: L'Harmattan, Société Française de l'Évaluation.
- Valsiner, J. (2007). *Culture in minds and societies. Foundations of cultural psychology*. New Delhi, India: Sage.
- Vergnaud, G. (1996). Au fond de l'action, la conceptualisation. In J. M. Barbier (Ed.), *Savoirs théoriques et savoirs d'action* (pp. 275-292). Paris, France: PUF.
- Vergnaud, G. (2001, Mai). Forme prédicative et forme opératoire de la connaissance. In J. Portugais (Ed.), *La notion de compétence en enseignement des mathématiques, analyse didactique des effets de son introduction sur les pratiques de la formation*. Actes du Colloque GDM, Montréal, Canada.
- Vermersch, P. (1994). *L'entretien d'explicitation en formation initiale et en formation continue*. Paris, France: Esf.
- Vial, M. (2001). Évaluation et régulation. In G. Figari & M. Achouche (Eds.) *L'activité évaluative réinterrogée* (pp.68-78). Bruxelles, Belgique: De Boeck.
- Yang, T., & Shadlen, M. N. (2007). Probabilistic reasoning by neurons. *Nature*, 447, 1975-1080. doi:dx.doi.org/10.1038%2Fnature05852

Date de réception : 5 avril 2013

Date de réception de la version finale : 4 septembre 2013

Date d'acceptation : 15 septembre 2013