



HAL
open science

Mobiliser la méta-évaluation pour mieux maîtriser la qualité des connaissances issues des évaluations dans la décision publique

Maurice Baslé, Marielle Berriet-Sollic, Catherine E. Laurent, Nadia Okbani, Frédéric Zahm

► To cite this version:

Maurice Baslé, Marielle Berriet-Sollic, Catherine E. Laurent, Nadia Okbani, Frédéric Zahm. Mobiliser la méta-évaluation pour mieux maîtriser la qualité des connaissances issues des évaluations dans la décision publique. Journées françaises de l'évaluation (SFE), Société Française de l'Evaluation (SFE). FRA., Jun 2013, Grenoble, France. hal-00863836

HAL Id: hal-00863836

<https://hal.science/hal-00863836>

Submitted on 19 Sep 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Mobiliser la méta-évaluation pour mieux maîtriser la qualité des connaissances issues des évaluations dans la décision publique

Mobilize the Meta-evaluation for better using
the quality of knowledge from evaluations in public decision

Synthèse du séminaire de recherche organisé
par le Réseau des Chercheurs en Evaluation des Politiques Publiques (RCEPP-SFE)
Paris - 18 avril 2013.

Communication présentée à la pré-session du RCEPP
et au café forum des Journées françaises de l'Évaluation (SFE) 6 et 7 juin 2013 □Grenoble



Auteurs : Maurice Baslé, Marielle Berriet-Sollic, Catherine Laurent, Nadia Okbani et Frédéric Zahm*
*auteur correspondant : frederic.zahm@irstea.fr co-animateur RCEPP-SFE

Mobiliser la méta-évaluation pour mieux maîtriser la qualité des connaissances issues des évaluations dans la décision publique.

Auteurs :

- **Maurice Baslé**, Professeur émérite d'économie (Université de Rennes 1 - CREM-CNRS)
- **Marielle Berriet-Sollicec**, Professeure d'économie (AgroSup Dijon - UMR 1041 CESAER)
- **Catherine Laurent**, Directrice de recherche en économie (INRA - UMR 1048 SADAPT)
- **Nadia Okbani**, doctorante en Sciences Politiques (Centre Emile Durkheim - Sciences Po Bordeaux)
- **Frédéric Zahm***, Chercheur en agro-économie (Irstea, Bordeaux, unité ADBX)

*correspondant et animateur du séminaire

Résumé : cette communication est une synthèse du séminaire de recherche sur le concept de Méta-évaluation organisé le 18 avril 2013 par le réseau *Réseau des Chercheurs en Evaluation des Politiques Publiques (RCEPP-SFE)*. Les deux premiers points visent à préciser les deux concepts discutés : (1) méta-évaluation et (2) relation entre connaissances acquises et forme de preuves lors du processus de décision publique issue d'une évaluation. Le troisième point est un résumé de deux études de cas de méta-évaluation présentés lors du séminaire.

Abstract : this paper is a synthesis of the research seminar on the concept of Meta-evaluation (the 18th of April 2013) organised by the Network of Researchers on Evaluation of French Society of Evaluation (RCEPP-SFE). The first two points are intended to clarify the two concepts discussed: (1) meta-evaluation and (2) the relationship between knowledge and forms of evidence in the process of public decision after evaluation. The third point is a summary of two case studies of meta-evaluation presented at the seminar.

Mots-clés : Méta-Evaluation, Politiques publiques, Référentiel, Standards, Qualité totale, Super-critères de méta-Evaluation, Apprentissage et capitalisation des connaissances évaluatives, Evidence-based policy, Connaissances scientifiques, Preuves.

Key words : Meta-evaluation, Public policies, Quality, Evidence-based policy, Scientific Knowledge

Ce texte collectif résulte pour partie de la reprise de communications de M. Baslé (pour le point 1), M. Berriet Sollicec et C. Laurent (pour le point 2), Jacques Toulemonde (pour le point 3.1) et Nadia Okbani (pour le point 3.2). Certaines parties sont en cours de publication. **Tous droits réservés.**

Dans des travaux en évaluation de politiques publiques, il existe un réel besoin d'augmenter la connaissance sur les travaux de recherche qui s'intéressent à la formalisation et à la stabilisation d'une démarche de méta-évaluation. Dans un processus de méta-évaluation, il y va de "la qualité de la connaissance utile à l'action publique". La méta-évaluation renvoie au processus de la formation d'un jugement de valeur sur une ou plusieurs évaluations réalisées sur une thématique, un projet ou une politique donnée en s'intéressant non pas aux seuls critères de qualité vu par les standards professionnels mais en interrogeant la question de l'utilité de l'évaluation du point de vue des utilisateurs, et de la société en général. La qualité d'une évaluation ne se résume pas à sa seule cohérence conceptuelle; il importe également de prendre en compte les finalités sociales de l'évaluation. Discuter d'un tel concept implique de revenir sur les différents sens donnés dans la littérature scientifique pour en préciser le périmètre. La méta-évaluation est un instrument privilégié de la recherche en évaluation : elle va au delà d'une analyse des résultats mais porte sur la qualité des pratiques, les méthodes, les logiques théoriques, et aussi et surtout sur le caractère utilisable des évaluations.

Les deux premiers points visent à préciser les deux concepts discutés : (1) méta-évaluation et (2) relation entre connaissances acquises et forme de preuves lors du processus de décision publique issue d'une évaluation. Le troisième point est un résumé de deux études de cas de méta-évaluation présentés lors du séminaire.

Programme

Introduction du séminaire, Frédéric Zahm (Irstea Bordeaux et co-animateur du réseau des chercheurs en évaluation des politiques publiques de la SFE).

Que recouvre la méta-évaluation, comment la définir ? Un retour sur les critères de méta-évaluation. Maurice Baslé, Professeur émérite d'économie (Université Rennes 1, CREM-CNRS)

Méta connaissance et niveaux de preuve Marielle Berriet-Sollicec, Professeur d'économie (AgroSup Dijon – UMR 1041 CESAER) et Catherine Laurent, Directrice de recherche en économie (INRA, UMR 1048 SADAPT)

Vers une méta-évaluation de l'évaluation des expérimentations du RSA : une méthode scientifique ou une stratégie politique ? Nadia Obkani (Centre Emile Durkheim - Science Po Bordeaux)

Évaluer l'évaluation sur le critère de son utilisation par les décideurs politiques Jacques Toulemonde (Eureval et chercheur associé IEP- Lyon)

Table ronde avec les organisateurs - synthèse de la journée et perspectives

1. Que recouvre le terme de méta-évaluation dans la littérature ?

La littérature scientifique ne propose pas de définition universellement reconnue ou stabilisée du concept de méta-évaluation.

1.1 Quatre sens différents donnés au concept de méta-évaluation dans la littérature

Quatre approches du concept et de la pratique de la méta-évaluation sont distinguées dans la littérature et rappelées ci-dessous de façon synthétique.

❖ **Approche 1** : la méta-évaluation est définie comme "l'évaluation" **ex post** des forces et des faiblesses d'une évaluation.

Il s'agit d'une activité d'évaluation d'une évaluation (Scriven, 1969; Cook & Gruder, 1978). Stufflebeam (2001) définit la méta-évaluation comme le « *process of delineating, obtaining, and applying descriptive information and judgmental information about the utility, feasibility, propriety, and accuracy of an evaluation and its systematic nature, competent conduct, integrity/honesty, respectfulness, and social responsibility - to guide the evaluation and/or report its strengths and weaknesses* ». Dans cette approche (la plus ancienne), la méta-évaluation est une espèce de "niveau 2" de l'évaluation pour apprécier la qualité d'une évaluation.

Au plan opérationnel, de nombreux critères de qualité sont mis en avant dans la littérature pour caractériser la qualité du processus évaluatif. Scriven (2010) retient les 6 critères suivants : la validité, la clarté, la crédibilité, l'éthique, le ratio coût-utilité et la capacité à généraliser les résultats. Le Conseil Supérieur de l'Évaluation (CSE, 1996) propose les 5 critères suivants : Utilité-pertinence, Fiabilité, Objectivité, Possibilité de généralisation et Transparence.

❖ **Approche 2** : la méta-évaluation peut être définie comme une synthèse **ex post de plusieurs** évaluations se rapportant à une même politique ou à ses programmes.

Dans ce processus, il s'agit d'une évaluation ex-post de plusieurs évaluations. C'est une forme de compilation de ce qui a été produit dans différentes évaluations déjà réalisées. On en espère une « super-connaissance » et un bilan vraiment complet. Ce bilan peut être une capitalisation des résultats utilisables, une progression de la connaissance évaluative utilisable. La méta-évaluation de synthèse ex post fait appel à des super-critères de qualité au regard desquels il faut évaluer cette fois de manière systématique **plusieurs** évaluations précédemment effectuées. Ces super-critères sont assez proches de ceux présents dans les guides d'évaluation et dans ce que les sociétés professionnelles d'évaluateurs ont posé comme standards et déontologie pour les métiers des évaluateurs (voir par exemple Stufflebeam, 1999).

❖ **Approche 3** : la méta-évaluation peut aussi se définir comme un projet **ex ante** de mise en place des évaluations pour des **programmes identiques** mais déployés dans le temps ou dans différents contextes.

Dans ce processus, il est alors prévu de produire des données brutes qui peuvent être ensuite recoupées et analysées et il semble possible de tester la validité externe de résultats concernant les processus qui marchent et les processus qui ne marchent pas. Il s'agit d'un projet approfondi de supervision d'évaluations tentant de produire à plusieurs reprises la même évaluation de la même politique et d'en retirer plus de connaissances de ce qui « fonctionne » ou pas dans les processus et les chaînes logiques « réalisations-résultats et impacts ». La force de la preuve d'un « mécanisme » ou « processus » qui fonctionne ou non doit en être renforcée. C'est la tâche des « knowledge capitalisation units » de capitaliser (voir AFD. <http://www.afd.fr/lang/en/home/recherche/evaluation-capitalisation/evaluation-et-capitalisation>).

❖ **Approche 4** : cette dernière approche s'inscrit comme une variante de l'approche 3 décrite ci dessus. Il s'agit aussi d'un projet approfondi de supervision d'évaluations. Cependant, cette fois ci les programmes évalués sont des expériences avec contrefactuel (ou placées autant que possible sous contrôle), "modulo" les limites dues à la construction des conditions expérimentales (« artificielles »). Cette approche de méta-évaluation a comme objectif de tester la validité externe de résultats concernant les dispositifs qui marchent de ceux qui ne marchent pas.

1.2 Quels enseignements tirés de ces différents approches ?

Les quatre approches identifiées se distinguent selon qu'une évaluation ou plusieurs sont passées en revue ou analysées ou synthétisées. Elles se distinguent également selon le temps de la méta-évaluation (conçue **ex ante**, déployée volontairement **pendant**, inventée **ex post** □). Dans une acception assez particulière (approche 4), la méta-évaluation s'appuie sur le déploiement conçu à l'avance d'expérimentations en parallèle : dans ce cas, les préoccupations de la méta-analyse rejoignent celles que l'on trouve dans certains cas dans l'activité de méta-évaluation.

Trois remarques peuvent être mises en avant :

1. La littérature n'est pas centrée sur méta-évaluation des politiques publiques mais sur la méta-évaluation. C'est une approche souvent très générale de la réflexivité ou du retour sur les pratiques d'évaluation, en prévoyant ce retour ex ante, en l'inventant ex post ou en proposant de faire une méta - évaluation chemin faisant.
2. La position réflexive de la méta-évaluation est semblable à celle de l'activité de Recherche en Sciences Humaines et Sociales (SHS) et aussi à celle de l'évaluation de cette recherche en SHS. La Recherche pour l'évaluation utilise les apports des méta-évaluations.
3. Les questions d'arbitrage entre ressources consacrées aux évaluations et budget consacré à la méta-évaluation ne sont pas souvent évoquées. Il semble pourtant a priori important de réfléchir, côté financeur, à une réserve budgétaire spécifiquement dédiée d'emblée ou plus tard au lancement de la méta-évaluation.

En définitive, dans tous les cas, la méta-évaluation consiste à renforcer les connaissances évaluatives de toutes natures en prenant un point de vue réflexif (interne ou/et externalisé) en vue d'une meilleure compréhension des résultats, de leur clarification, en vue aussi de déterminer la confiance que l'on peut accorder aux résultats produits et de leur « caractère utilisable » dans les politiques publiques à venir. En ce sens, l'activité de méta-évaluation correspond à un mode d'apprentissage permanent dans l'activité évaluative : apprentissage individuel, apprentissage organisationnel, apprentissage dans la production et le management des connaissances obtenues en retour de savoirs évaluatifs, apprentissage des politiques. L'activité de méta-évaluation combine des expertises, des approches et multiplie les points de vue, ce qui en fait une activité transdisciplinaire.

2. Méta connaissance, qualité et niveaux de preuves d'une évaluation

Dans le processus d'évaluation au service de la décision publique, les différents utilisateurs finaux (décideurs publics, praticiens, autres chercheurs, citoyens...) sont confrontés à de grandes difficultés pour savoir à la fois quelles sont les connaissances disponibles mais aussi en maîtriser l'accès et l'usage pour appuyer la décision publique qui s'appuie sur un processus de connaissances ayant valeur de preuves (Evidence-based decision). Le processus pour établir un état de l'art systématique sur une question posée passe nécessairement par une analyse de la littérature sur un sujet donné. Il importe que ces procédures

soient explicites pour que ces états de l'art puissent être en permanence complétés et actualisés. Toutefois, ce processus d'état de l'art reste trop souvent insuffisant compte tenu du nombre et de l'hétérogénéité des productions scientifiques disponibles. Les chercheurs confrontés à cette réalité observent "*qu'il faut plusieurs semaines pour préciser la question de départ et les critères de qualité qui seront retenus, établir des requêtes satisfaisantes, délimiter l'univers des documents à explorer et en tirer des conséquences sur les limites de validité de l'exercice*" (Laurent *et al.*, 2009). C'est dans ce contexte que les travaux sur la méta-évaluation s'inscrivent et participent d'un large mouvement qui vise à produire des méta-connaissances pour tous ces différents utilisateurs finaux (décideurs publics, praticiens, autres chercheurs, citoyens...). Un des objectifs visés est une meilleure circulation de ces connaissances dans le champ des informations disponibles. C'est ainsi qu'autour des approches en termes d'*Evidence-based decision*, des travaux conduits depuis une vingtaine d'années ont débouché sur une boîte à outils méthodologique destinée à favoriser un usage judicieux des connaissances disponibles aussi bien pour concevoir que mettre en œuvre ou évaluer des programmes d'action publique : méta-analyses, grille d'analyse de la qualité des preuves, méthodes d'états de l'art systématique, etc.

Afin de choisir **la méthode d'évaluation la plus appropriée**, les décideurs publics accordent désormais une grande importance à ce que signifie une « donnée probante » ou « evidence ».

2.1 Qu'entend-on par le terme preuve ("donnée probante") dans un processus d'évaluation de politique ?

Ce concept d'*Evidence* en anglais n'a pas de traduction directe en français. Il se rapporte certes le plus souvent à la notion de preuve avec un fort contenu empirique. Si le principe de corroboration aux faits s'impose, par contre, ce terme renvoie à une connaissance acquise non pas uniquement à partir de la recherche suivant des méthodes scientifiques ou de différentes formes d'investigations systématiques (statistiques) mais aussi de tous types d'informations telles que des opinions faisant autorité, des résultats ou faits issus de savoir d'expérience et de savoir gestionnaires. Le mot *Evidence* n'est donc pas synonyme de « connaissance scientifique » mais renvoie à un faisceau d'indices ou d'éléments de preuves concordantes.

Dans le domaine de l'évaluation des politiques publiques, l'expression *evidence-based policy* se caractérise par différentes traductions qui ne renvoient pas au même sens : « *politique fondée sur les preuves* » ou « *politique fondée sur des données probantes* », ou « *politique fondée sur les niveaux de preuves* », ou encore sur des « *évidences factuelles* » (Laurent *et al.*, 2009). Chacune de ces traductions met l'accent sur une préoccupation principale :

- "Preuve", "données probantes" pour marquer la nécessité "*de disposer de connaissances fiables pour étayer les actions et garantir autant que possible leur efficacité au regard des objectifs poursuivis*",
- "Niveaux de preuve" pour marquer le fait "*qu'il convient de hiérarchiser les connaissances disponibles*",
- "Évidences factuelles", "réseaux empiriques" "*pour rappeler que les «évidences/preuves» sont fondées sur des données empiriques et corroborées par les faits*" (Laurent *et al.*, 2009).

2.2 Choisir sa méthodologie d'évaluation en tenant compte de la qualité des preuves

Lorsque l'on s'intéresse à la question d'une preuve fiable (crédible) pour l'évaluation d'une politique, il est utile de mobiliser les travaux qui renvoient à la « *théorie des preuves pour l'évaluation* ». Cette dernière distingue dans le processus de formalisation de la preuve deux niveaux complémentaires : d'une part, les types de preuve mis en avant et d'autre part, le niveau de la preuve qui permet de hiérarchiser les connaissances disponibles.

Il est nécessaire de clairement différencier dans l'analyse des connaissances mobilisées **le type de preuve** de celui sur les **niveaux de preuve**.

❖ Trois grands types de preuves pertinentes qu'il convient de distinguer

Les travaux de Laurent *et al.* (2009) à la suite notamment de Shaddish *et al.* 2002; Illari 2011; Cartwright et hardie 2012, ont permis de mettre en évidence trois grands différents types de preuves pertinentes mobilisables pour concevoir, mettre en œuvre et évaluer les politiques.

Les différents types de preuves

1. Des "**preuves de présence**" : il s'agit de la mise en évidence de la présence d'une chose (tels que des inventaires biologiques pour la biodiversité par exemple).
2. Des "**preuves de mécanisme**" d'un phénomène: quand sont connues les entités ou les activités qui produisent le phénomène concerné et leur organisation pour produire ce phénomène. Par exemple : mécanismes par lesquels un remembrement foncier va avoir un impact sur la biodiversité en favorisant la circulation de certaines espèces.
3. Des "**preuves d'effet**" (evidence of difference-making). Il est possible de distinguer (i) les "**preuves d'efficacité**" : mise en évidence qu'une intervention produit un effet attendu, indépendamment de la connaissance des raisons de cette efficacité (par exemple : effet d'un soutien au revenu agricole sur le volume de production final) et (ii) les "**preuves de dangerosité/innocuité**" : mise en évidence des effets négatifs d'un produit, d'une action (l'action est dépourvue d'effets négatifs (par exemple, telle subvention pour le soutien à la biodiversité en agriculture n'a pas d'effets spécifiquement défavorables sur les exploitations agricoles les plus petites).

❖ Une utilisation des connaissances à différencier selon le niveau de preuves disponibles

Cette réflexion sur la différenciation des connaissances ne se borne pas à distinguer différents types de savoirs et de preuves. Une part importante de la réflexion concerne aussi la hiérarchisation des niveaux de preuves, pour un même type de preuve. Ces différents niveaux de preuve renvoient aux méthodes de recueil et de traitement de ces données (telles que les études de cas, la méthode de double-différence, les essais Randomisés contrôlés, etc.. □) et vont permettre de juger de la fiabilité des preuves. Divers principes de hiérarchisation coexistent qui font l'objet de nombreuses discussions. On peut par exemple distinguer les différents niveaux suivants :

1. Preuve issue d'opinion d'autorités respectées fondées sur l'expérience clinique, des études descriptives ou des rapports de comités d'experts
2. Preuve issue de comparaisons historiques,
3. Preuve issue d'une cohorte ou d'une étude cas-témoins,
4. Preuve issue d'au moins un essai randomisé contrôlé.

Les niveaux de preuve ne sont que l'une des dimensions de la qualité des preuves, et il importe de ne pas privilégier les preuves de haut niveau au détriment des preuves socialement pertinentes et de type adéquat pour la question traitée.

❖ Que permet au final une telle distinction entre type de preuves et niveaux de preuves ?

La distinction entre "**types de preuves**" et "**niveaux de preuves**" permet au final de raisonner l'adéquation entre *types de « preuves »* et les différents *objectifs d'une évaluation* (apprendre, mesurer, comprendre).

- Pour mesurer l'impact et non les mécanismes de la politique => mobiliser des « preuves » d'efficacité
- Pour comprendre les mécanismes de l'intervention publique => mobiliser des « preuves » de mécanisme
- Pour apprendre sur la politique, ses programmes => tous les types de « preuves » sont complémentaires

3. Deux études de cas de méta-évaluation

3.1 Une méta-évaluation des politiques de formation et d'accompagnement des demandeurs d'emploi

L'étude de cas (présentée par Jacques Toulemonde) concerne la formation et l'accompagnement des demandeurs d'emploi, plus précisément lorsque ces actions sont cofinancées par le Fonds Social Européen (CE, 2010). Dans les années 2007-2009, la Commission européenne a encouragé les autorités nationales gestionnaires de ces actions à en évaluer l'impact de façon aussi rigoureuse que possible. Cette dynamique s'est traduite par la réalisation d'une dizaine d'évaluations utilisant des approches quasi-expérimentales dans différents états membres. L'étude de cas présente une évaluation de ces évaluations réalisée en 2010 par une équipe internationale à la demande de la Commission européenne.

La question posée était de savoir si une plus grande rigueur conduisait à une plus grande utilisation. La méthode a combiné une analyse « *basée sur la théorie* » (SFE, 2010) et une série de quatre études de cas couvrant la Belgique, la Hongrie, l'Italie et la Pologne. La « théorie » testée avait la forme d'une chaîne causale composée des maillons suivants :

1. Les évaluations répondent à une demande politique
2. Elles sont robustes et pertinentes
3. Elles produisent des connaissances nouvelles et pertinentes
4. Ces connaissances sont mises à disposition des personnes conseillant les décideurs
5. Ces personnes sont convaincues
6. Les connaissances produites sont utilisées dans des décisions

L'analyse des évaluations conduites dans les quatre pays étudiés confirme à peu près les hypothèses 1 à 5 ci-dessus. Par contre, elles réfutent la dernière en ce sens qu'il n'y a pas eu d'utilisation directe des connaissances acquises, même dans les meilleurs contextes. Ainsi, plusieurs évaluations ont produit des connaissances nouvelles et crédibles sur la différence d'impact entre hommes et femmes mais ces connaissances n'ont pas contribué directement à concevoir les réformes des programmes intervenues sur ce thème dans les années suivantes.

La « théorie » doit donc être repensée et la méta-évaluation apporte des éléments pour le faire. Il est en effet apparu que les connaissances sur la plus ou moins grande ampleur des impacts ont été finalement utilisées à partir du moment où elles ont été complétées par d'autres connaissances sur l'explication des différences d'impact constatées.

En résumé, le travail présenté est une évaluation d'un ensemble d'évaluations ayant réalisé des estimations d'impact particulièrement rigoureuses. On s'attendait à ce qu'une plus grande rigueur conduise à une plus grande utilisation décisionnelle. Or ce n'est pas ce que montre l'expérience. Malgré leur rigueur, les estimations d'impact n'ont été utilisées qu'à partir du moment où elles ont été accompagnées d'une explication. Cette leçon a au moins une portée européenne dans le champ des politiques de l'emploi. N'étant pas contredite par la littérature sur l'utilisation des connaissances, elle a probablement une portée plus générale.

3.2 Vers une méta-évaluation de l'évaluation des expérimentations du RSA : une méthode scientifique ou une stratégie politique ?

En s'appuyant sur l'évaluation des expérimentations du Revenu de Solidarité Active, menée selon la méthode d'évaluation expérimentale, l'étude de cas présentée par Nadia Okbani (2013) a permis de revenir sur les critères propres à la méta-évaluation et d'analyser la structuration (validité) et l'utilité (voire utilisation) de cette évaluation dans l'optique de la généralisation du dispositif. Les conditions de lancement de cette expérimentation et de son évaluation sont conçues comme une étape de la généralisation du RSA. Par-là, la place de l'évaluation est interrogée comme instrument d'action publique (Lascoumes et Le Galès, 2005) dans la conduite de l'action publique. Il est mis en avant une tension entre les objectifs scientifiques des évaluateurs et les objectifs politiques des décideurs (Conseils généraux et Haut-commissariat aux solidarités actives) dans le cadrage de l'évaluation comme dans l'instrumentalisation de ses résultats, particulièrement discutables. Au final cette étude cas analyse les enjeux et les usages politiques de ce type d'évaluation, entre une volonté de rationalisation, de dépolitisation et de légitimation d'une réforme controversée orientée vers l'activation de la protection sociale.

Les principaux éléments de cette présentation sont disponibles sous HAL dans le working paper (Okbani, 2013).

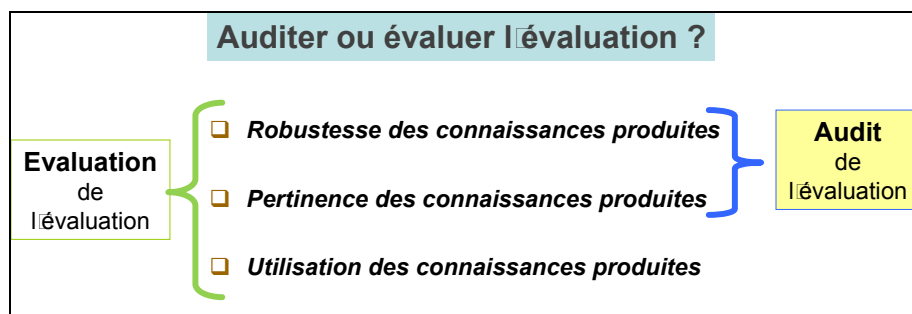
4. Quelques pistes pour la recherche

- Quelles sont aujourd'hui l'importance et la place accordée au processus de méta-évaluation dans la décision publique ?
- Ce processus de méta-évaluation a-t-il le même degré d'utilité pour toutes les politiques ?
- Ce processus est-il plus important à conduire pour certaines politiques (politiques constitutives *versus* substantielles) ou domaines thématiques ?
- Les types de preuves sont-elles de nature très différente par domaine d'intervention (politiques sociales, environnementales, d'éducation, etc.) ?

- Serait-il possible que les évaluateurs donnent à voir comment ils hiérarchisent les preuves dans leur réponse ?
- Est-il possible et utile de proposer à la décision publique des principes d'établissement de niveau de niveaux de preuves selon qu'on se situe dans un processus ex post ou ex ante d'évaluation ?

5. Quelques recommandations pratiques lors de la conduite d'une évaluation

- Il convient de ne pas utiliser le concept de méta-évaluation lorsque le travail réalisé conduit à une "seule" synthèse de connaissances de plusieurs évaluations.
- Lors d'une réponse à un cahier des charges par un bureau d'étude, il serait important que ce dernier explicite le principe de hiérarchie des preuves qu'il retient (places de données issues d'experts, de simulation, de mesures, d'observations systématiques ou d'enquêtes, etc.)
- Lors de la réalisation d'une méta-évaluation, il importe de rendre explicite et transparent les critères de jugement sur lequel se fonde le jugement de l'évaluateur. Les critères tels que validité, clarté, crédibilité, éthique, ratio coût-utilité et capacité à généraliser les résultats (Scriven, 2010) ou Utilité-pertinence, Fiabilité, Objectivité, Possibilité de généralisation et Transparence (CSE, 1996) sont aujourd'hui reconnus.
- Il convient de bien distinguer le processus d'audit d'une évaluation de la conduite d'une méta-évaluation.



Source : Toulemonde, 2013

- Afin de choisir **la méthode d'évaluation la plus appropriée**, les décideurs publics accordent désormais une grande importance à ce que signifie donnée « probante ».
- Il est nécessaire de clairement différencier dans l'analyse des connaissances mobilisées **le type de preuves** de celui sur les **niveaux de preuves**.

Références citées :

- CE, 2010, Study on the Return on ESF Investment in Human Capital, European Commission □ DG Employment, Eureval / Ecorys / Ramboll Management
- Cook T. D., Gruder C. L., 1978, Metaevaluative research, *Evaluation Quarterly*, 2, pp. 5-51
- CSE, 1996, Petit guide de l'évaluation du Conseil Scientifique de l'Evaluation, France, Paris
- Illari P. 2011. Mechanistic evidence: disambiguating the Russo-Williamson Thesis. *International Studies in the Philosophy of science*, 25:2, 139-157.
- Laurent C., Baudry J. Berriet-Sollicec, Kirsch M., M., Perraud D., Tinel B., Trouvé, A., 2009, Pourquoi s'intéresser à la notion □Evidence Based Policy□ *Revue Tiers-Monde*, N° 200, pp. 853-873
- Okbani N., 2013, L'évaluation des expérimentations du RSA : entre légitimation de l'efficacité et usages politiques, Working Paper AFS RT6, disponible sous HAL
- Scriven M., 1969, An introduction to meta-evaluation, Educational Product Report, 2, pp. 36-38
- Scriven M., 2010, Evaluating evaluation : a meta-evaluation checklist
- Shadish WR, Cook TD, Campbell DT, 2002. *Experimental and quasi experimental Design for generalized inference*. Boston, New-York:Houghton Mifflin Company.
- SFE, 2010, Evaluation des impacts des programmes et services publics, *Les cahiers de la SFE*, 72 p.
- Stufflebeam D. L., 2001, The Metaevaluation Imperative, *American Journal of Evaluation*, Vol. 22, No. 2, pp. 183-209

Pour en savoir plus :

- Cartwright N, Hardie J, 2012. Evidence-Based Policy: A Practical Guide to Doing It Better.Oxford University Press.
- Nutley S. Walter I., Davies H. 2007. Using evidence. How research can inform public services. Policy press. University of Bristol. 363 p.
- Donaldson, S.I. Christie C.A. Mark, H.H. 2008. What counts as credible evidence in evaluation and evidence-based practice ? Thousand oaks, CA: Sage
- Laurent, C., Berriet-Sollicec, M. Labarthe, P. et Trouvé, A. 2012. Evidence-based policy : de la médecine aux politiques agricoles ? Les enjeux d'une approche méconnue en France. Notes et Etudes Economiques, N° 36, Juin 2012, pp. 79-101
- Laurent C., Berriet Sollicec M., Kirsch M., Labarthe P., Trouvé A. 2010. Multifunctionality of agriculture, public policies and scientific evidences: some critical issues of contemporary controversies. *Applied studies in Agribusiness and Commerce* :1-2, 53-58.
- Pawson R., 2002. □Evidence-based Policy: In Search of a Method□ *Evaluation*, Apr 2002; vol. 8: pp. 157 - 181.
- Pawson R. 2006. Evidence-based policy: a realistic perspective. Sage, London, UK
- Stufflebeam D.L., 1999, Programm evaluation models metaevaluation checklist (Based on The Program Evaluation Standards), Evaluation Checklists Project voir www.wmich.edu/evalctr/checklists, 11 p.
- Shadish W.-R., 1998, Evaluation Theory is Who We Are, *American Journal of Evaluation*, Vol. 19, N°1, pp. 1-10
- Toulemonde J., Jacob J.L., Bern S., Lang, Genard P., 2006, Appropriation des résultats de l'évaluation : leçons de la pratique en Région Limousin, dans : *L'évaluation au niveau régional*, Bruxelles, P.I.E Peter Lang, pp. 131-142

Consultation possible de sites internet

- http://www.wmich.edu/evalctr/archive_checklists/checklistmenu.htm
- www.wmich.edu/evalctr/checklists
- <http://www.wmich.edu/evalphd/wp-content/uploads/>

Une revue américaine

Journal of Multidisciplinary Evaluation voir : http://survey.ate.wmich.edu/jmde/index.php/jmde_1