

**Conférence « The Future of Evaluation in Modern Societies »
Compte-rendu de la conférence, 14 – 15 juin 2012, Saarbrücken**

Laurent Crémieux, EFK

Organisation : Centrum für Evaluation (CEVAL)
Lieu : Universität des Saarlandes, Saarbrücken

Cette conférence était organisée pour les 10 ans d'activité du CEVAL. Ce centre d'évaluation a été créé et est toujours dirigé par Reinhard Stockmann, une des personnes de référence en Allemagne dans le domaine de l'évaluation. Stockmann a publié de nombreux ouvrages de référence et est actif dans la DeGEval, la Société allemande d'évaluation. Le CEVAL est un bureau d'évaluation travaillant sur mandat. Il réalise des projets principalement dans le domaine de l'aide au développement mais reçoit aussi des mandats sur des questions économiques et sociales de la part du Land de Saare ou d'autres régions allemandes. Une vingtaine de personnes travaillent dans ce centre. Il couvre ses frais de fonctionnement à 90%. En 10 ans, il a réalisé 82 évaluations et expertises. De par son ancrage universitaire, le centre bénéficie des activités d'enseignement et de recherche. L'Université de Saare propose une formation en évaluation (Master). Reinhard Stockmann a de plus développé des partenariats avec différentes universités (Ouganda, Costa Rica,...) afin de contribuer à renforcer la culture d'évaluation dans les pays en développement, notamment par le soutien à des formations type Master. Le rapport annuel 2011 donne une idée des activités et ressources du CEVAL, de même que des évaluations réalisées.

La conférence était articulée autour de différents axes :

- Le rôle de l'évaluation et les attentes par rapport au changement politique (bilan)
- L'avenir avec les défis de la professionnalisation, les perspectives dans les pays en développement ou les thèmes émergents
- Présentation de projets particuliers (« Market Square of Evaluation »)

Voici les principaux points que je retiens de cette conférence :

L'évaluation est une construction sociale basée sur des normes et des valeurs. Elle sert à définir ce qui est bien et ce qui est moins bien. On peut distinguer 5 niveaux par rapport au développement de l'évaluation depuis 40 ans (*présentation de Peter Dahler Larsen, University of Southern Denmark*) :

1. « Popularisation » :

Sous cet aspect se distinguent l'institutionnalisation selon laquelle l'évaluation doit être reconnue (standards, normes, valeurs, soutien,...) et la diffusion auprès de la population pour qu'elle puisse s'approprier elle-même les outils d'évaluation (auto-évaluation)
→ risque : il se développe une « fatigue » ou une lassitude vis-à-vis de l'évaluation.

2. « Outcome evaluation » :

Initialement développé comme instrument antibureaucratique et pour améliorer le système d'informations, il y a un risque de se perdre dans des débats conceptuels : que peut-on

exactement comparer ? On cherche à isoler des interventions et on aboutit à des résultats très spécifiques, voire individuels, et il est impossible de généraliser.

→ risque : l'évaluation se perd dans des débats conceptuels.

3. « Systématisation de l'évaluation » :

Sur le modèle des enquêtes PISA pour mesurer la qualité de l'enseignement scolaire, on crée des « machines » à évaluer. On mesure les écoles, mais pas la politique scolaire.

→ risque : trop grande standardisation des évaluations ; l'évaluation se résume à l'utilisation d'un logiciel ; pas d'influence sur la politique.

4. « Evaluation et recherche » :

L'évaluation n'est pas reconnue comme discipline académique. La rationalité entre l'évaluation et la recherche n'est pas identique. Les prémices pour la recherche en évaluation ne sont pas les mêmes que ceux pour réaliser des mandats d'évaluation. Les mandats d'évaluation sont régis par des règles du marché puis sur l'intégration des résultats par les mandants (relation mandant / mandataire). Dans la sphère de la recherche, les professeurs définissent les domaines à explorer. Il serait nécessaire d'avoir plus de recherches sur l'évaluation.

→ risque : le monde académique se détourne de l'évaluation.

5. « Utilisation des évaluations » :

C'est le « trend » actuel au niveau de la recherche. Comment démontrer l'effet des évaluations ? Faut-il « imposer » l'utilisation des résultats ? Comment faire pour que les résultats des évaluations restent en mémoire ? Des modèles de développement intégrant le niveau de conflit (parmi les stakeholders) et la pression existante pour le changement.

→ risque : on ne progresse pas dans les explications relatives à l'utilisation qui dépend avant tout du contexte (situative view). L'évaluation n'est pas perçue comme instrument du changement.

L'évaluation s'est beaucoup développée dans le secteur de l'aide au développement. Ce n'est pas un hasard, car il y avait l'espoir que l'évaluation pouvait être vecteur du changement politique (amélioration des conditions socio-économiques grâce aux programmes d'aide au développement ; baisse de la pauvreté,...). C'est un des échecs de l'évaluation qui non seulement n'a pas été un vecteur du changement politique, mais a même eu une tendance inverse à pérenniser des programmes existants. *Hans-Martin Boehmer (Banque Mondiale)* a présenté quelques exemples actuels sur les difficultés face à l'institutionnalisation de l'évaluation ou l'exploitation des résultats des évaluations.

L'Afrique du Sud est un exemple où l'évaluation a été institutionnalisée au niveau de la Présidence, au niveau de l'Audit Office qui rend ses rapport au Parlement et au niveau du ministère des finances. La transparence existe. Cependant, il n'existe pas d'attentes vis-à-vis de l'utilisation des résultats pour l'action politique. Autrement dit, l'existence d'un rapport démontrant une faible performance des programmes politiques n'est pas suffisant pour générer une action politique. Car en Afrique du Sud, la loyauté politique (appartenance au même parti politique ; appartenance à la même région d'origine ; etc.) prime sur tout autre facteur pour l'action politique. Les décisions ne se prennent que sur la base de ce critère.

Autre limite, il peut exister des perceptions différentes quant à l'urgence de mettre en œuvre des recommandations. En effet, comme les résultats d'évaluation ne sont en principe ni tout blancs, ni tout noirs, le management ou les responsables politiques peuvent avoir une perception différente suivant s'ils interprètent les résultats de manière optimiste (pas de besoin d'action) ou pessimiste (besoin d'action). Le statu quo va souvent l'emporter.

De manière générale, la culture d'évaluation demeure très variable selon les pays. Cependant on note dans les pays développés un « tsunami » d'évaluations dans le secteur social, le domaine de la santé et celui de l'école (éducation et formation). A l'inverse, il n'existe quasiment pas d'évaluations dans le domaine de la défense ou de la politique fiscale.

Au niveau des pays en développement, il est important que les pays planifient et réalisent les évaluations eux-mêmes, ce qui suppose le développement d'une culture d'évaluation. Il faut développer les compétences (capacity building) en offrant des possibilités de formation dans les universités locales. Il faut aussi que la population locale s'approprie les outils et les résultats d'évaluation. Le changement au niveau de l'action politique ne passe seulement pas les décisions des managers, mais aussi par la diffusion des idées et l'évolution des mentalités au niveau local (ce constat est valable pour tous les pays).

Pour la formation au niveau universitaire, il est important de créer des cursus avec Master en évaluation selon les mêmes critères de formation qu'en Europe. Ceci a ainsi été développé, par exemple, au Costa Rica et en Ouganda, grâce à un partenariat développé avec l'Université de Sarrebruck (Stockmann). Mais une fois l'existence d'une formation de qualité et d'un processus d'évaluation institutionnalisé, se pose la question de l'utilisation et de la diffusion des résultats. En Ouganda (James Muwanga, Makerere University), le développement est récent mais rapide. A partir de la moitié des années 1990, des principes de prévisions budgétaires, de définitions d'objectifs et de monitoring ont été adoptés. A partir de 2005, il y a eu une réforme du secteur public, qui doit désormais travailler selon une approche orientée résultats. Il existe une stratégie d'évaluation (National Integrated Monitoring and Evaluation Strategies) définie au niveau du Premier Ministre à l'aide de collaborateurs spécialisés. Dès 1997, l'Ouganda a défini ses propres Plans d'action pour lutter contre la pauvreté (sans se baser sur les programmes d'aide extérieur). Il a fallu ensuite créer des bases de données pour apprécier l'évolution : c'est le plus difficile ! Sans données, il est impossible de mener un monitoring (et encore moins une analyse sur les impacts). Pour ce faire et sous la pression de la société civile pour plus de liberté et de transparence, les communautés locales ont été associées. La récolte des données de base se fait grâce à un réseau de 6'000 « monitors » formés et agréés, répartis dans les 10 districts du pays (Community based monitoring and evaluation system). Ils transmettent leurs données aux districts puis au gouvernement, le tout sous la supervision de représentants de la société civile. Ces données de monitoring font l'objet de rapports trimestriels. Les analyses se limitent pour le moment au niveau output / résultats.

Sur le thème de la professionnalisation *Donna Martens (Gallaudet University, Washington)* rappelle l'évolution aux Etats-Unis et au sein des organes des Nations-Unies. Le développement de l'évaluation est intimement lié à son rôle dans le changement social. Puis dans les années 1990, l'évaluation s'est développée autour du thème de la justice sociale.

Il a fallu attendre 2003 pour que l'Association américaine d'évaluation définisse nécessaire d'avoir

une base scientifique pour montrer l'évidence. Plus tard, l'administration Obama a précisé que les programmes doivent être évalués selon des « standards clairs et élevés ». Ceci laisse l'impression que l'évaluation se trouve toujours confrontée aux mêmes questions de base (c'est un éternel recommencement) ! Il existe beaucoup d'attentes : l'objectivité au niveau des méthodes ; l'utilité de l'évaluation pour les managers ; le positionnement autour des valeurs (bien / mal) relevant de la justice sociale ou d'autres dimensions ; l'utilisation des résultats dans une perspective de changement politique.

Mais il existe souvent une tension : d'une part l'évaluation est jugée trop longue et trop chère ; d'autre part l'évaluation n'a pas assez de crédibilité. Comment trouver le juste équilibre entre rapidité et crédibilité ?

Or on ajoute toujours plus de dimensions en lien avec les valeurs : les compétences culturelles ont pris une grande importance ces dernières années au sein des organes des Nations-Unies avec la promotion de la diversité culturelle, la prise en compte de la dimension du genre dans l'évaluation. L'Unicef ou l'UNDP ont intégré ces dimensions de manière systématique dans leurs critères d'évaluation et recherchent des évaluateurs avec ce type de profil.

Le principal défi demeure dans l'introduction de nouvelles dimensions qualitatives. Il faut de plus surmonter les problèmes liés à l'interprétation et aux jugements de valeurs. Pourquoi interprète-t-on les résultats de cette manière (bien) plutôt que de cette manière (pas bien) - analogie avec le verre à moitié plein ou à moitié vide ? Le référentiel sur lequel se base le jugement est-il juste ?

Jan-Erik Furubo, Swedish National Audit Office évoque le développement de l'évaluation en Suède où dominant une certaine perception de l'harmonie et du consensus pour la prise de décision, de même qu'une croyance en la science et dans l'action du gouvernement pour influencer le changement. Aujourd'hui, l'évaluation est bien institutionnalisée en Suède, tant au niveau des agences gouvernementales, de l'Audit office (performance audit) ou des commissions parlementaires et extra-parlementaires. Mais l'évaluation est en concurrence avec d'autres « producteurs de connaissances ». En effet, il n'y a pas que ce qui porte le label « évaluation » qui permet d'améliorer ou d'encourager des actions. De plus l'évaluation est trop liée à deux croyances : 1) le développement de l'évaluation est intimement lié au New Public Management et à ses valeurs, la culture administrative actuellement dominante ; or il est probable qu'à moyen ou long terme, une nouvelle culture administrative émergera ; 2) la plupart du temps, il existe une croyance selon laquelle il est toujours possible d'améliorer un programme politique grâce à l'évaluation (modèle d'impacts avec feedback). Or parfois, il faudrait mieux laisser tomber un programme, si les résultats sont mauvais (non atteinte des objectifs ; pas d'effets pour les bénéficiaires ; fausses incitations ;...). Il est plus facile de proposer des améliorations dans un cadre existant que d'abandonner purement et simplement un programme. Au lieu d'inciter au changement politique, l'évaluation sert à légitimer les politiques existantes et pousse au conservatisme. De plus, en suggérant des améliorations, les résultats des évaluations impliquent très souvent l'allocation de ressources, donc de dépenses supplémentaires.

Verena Friedrich (Université de Berne) a effectué une recherche sur les universités offrant des masters en évaluation (master of advanced studies in evaluation). Cette recherche se limite aux universités de l'Europe de l'ouest. Ainsi 9 universités proposent ce type de Master (Madrid, Lyon, Bern, Amsterdam, Anvers, Odense, Rennes, Saarbrück et Gênes). 3 universités (Bonn, Londres, Séville) ont abandonné cette formation, car il n'y avait pas assez de participants. Les deux masters

français (Rennes et Lyon) produisent le plus de diplômés. Le Master en évaluation de Rennes existe depuis 1996 (c'est le plus ancien) et a le plus d'étudiants. Depuis 1996, 236 étudiants ont reçu le Master à Rennes.

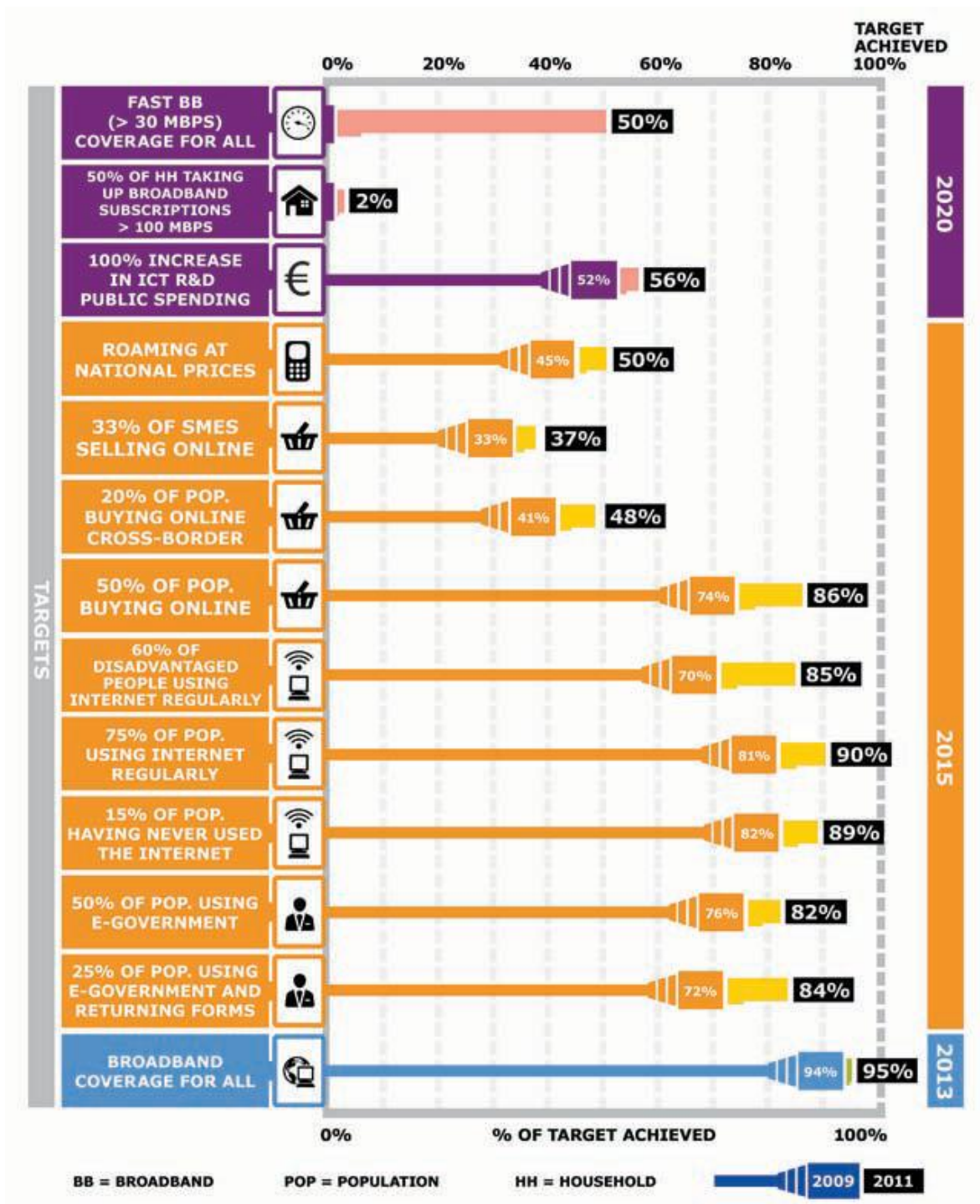
Le rattachement académique des Masters est très variable (sciences de l'éducation, sciences politiques, ou business administration). L'évaluation n'est pas une discipline et il est très difficile de faire de la recherche. L'argent est principalement généré par des mandats. La professionnalisation de l'évaluation dépend d'une plus grande reconnaissance par le milieu académique, voire par le développement de l'évaluation comme discipline universitaire.

Pour *Frans Leeuw (Research and Documentation Center, Maastricht)*, l'évaluation a tendance à rendre l'action politique encore plus complexe alors qu'elle devrait contribuer à réduire la complexité. On parle toujours plus d'interdépendances, de réseaux globaux, de politiques globales, de régulations transnationales, de soft factors, etc. Peut-être faudrait-il revenir à un niveau local. L'évaluation peut apporter davantage de réponses aux niveaux local et régional. L'évaluation n'apporte pas de réponses face à des politiques globales et à leur échec, par exemple Kyoto. Autre exemple, il n'y a jamais eu tant de campagnes de prévention et d'actions dans le domaine la sécurité alimentaire alors que le nombre de personnes obèses ne cessent d'augmenter ! A quoi servent ces programmes et la multiplication des agences de contrôle sanitaire, tant au niveau national, européen que mondial ?

Autre problème fondamental, si le politique cherche à légitimer son action, l'évaluation tombe aussi dans ce piège. Or le manager ou le politique ne veut plus avoir de rapports qui lui disent que les objectifs n'ont pas été atteints. Il y a dès lors un grand risque de sélectionner des buts faciles à atteindre, le tout confirmé par un rapport final attestant que les objectifs ont effectivement été atteints !

Frans Leeuw estime que les évaluateurs sont absents du débat sur les effets des programmes pour la promotion ou la restriction des nouvelles technologies de l'information. Il a écrit un article à ce propos dans la revue *Evaluation* (vol. 18, n°1, pp. 111-127) : « Cyber society and digital policies : challenges to evaluation ». Dans de nombreux pays, il existe une multitude de programmes et d'actions visant à promouvoir les nouvelles technologies de l'information ou à limiter leur usage ! Par exemple l'Union européenne a adopté un programme comprenant plus d'une centaine d'actions (EC Digital Agenda) http://ec.europa.eu/information_society/digital-agenda/index_en.htm Cela va de la promotion du e-learning à la promotion de la vente par interne ou encore le vote électronique et la régulation des télécommunications. L'UE a même défini des objectifs chiffrés (Digital Agenda Targets – voir page suivante) !

Laurent Crémieux, 5 septembre 2012



EU Digital Agenda Targets