The 9th European Evaluation Society International (EES) Conference
Bref compte-rendu de la conférence, 6 - 8 octobre 2010, Prague

Thème: évaluation et intérêt public: participation, politiciens et politiques publiques

La « grande messe » européenne de l'évaluation a lieu tous les deux ans. Cette année, elle a enregistré 688 participants, 461 propositions soumises, 39 panels et symposiums, 97 "paper sessions", 44 posters. Il y a eu près de 140 sessions portant sur cinq axes principaux :

1. Ethique, capacités et transparence
2. Evaluation et politique
3. Evaluateurs, bénéficiaires, usagers et décideurs
4. Evaluation de politiques sectorielles
5. Evaluation dans les pays en développement et en transition

Trois sessions plénières et douze sessions parallèles ont été organisées tout au long des trois jours de la conférence, entre 8h30 et 19h00. Lors des sessions parallèles, les participants avaient chaque fois le choix entre 12 sessions différentes (panels, symposiums et "paper sessions"). Le programme, ainsi que les slides et les textes présentés peuvent être trouvés dans le website EES. Vu la thématique très générale du congrès, il y avait à boire et à manger quant aux sujets abordés et à la qualité des présentations. Il était important de bien choisir préalablement les sessions suivies. Ci-dessous quelques enseignements concernant des thèmes intéressant plus particulièrement l'auteur de la note.

1. Méthodes d'évaluation – évaluation contrefactuelle…

Les exposés montrent que l'utilisation de méthodes expérimentales (très populaire avant les années 1980) reviennent à la mode sous le nom « **évaluation d'impact contrefactuelle** » (« counterfactual impact evaluation »). L'idée essentielle des méthodes contrefactuelles est de reconstruire ce qui se serait passé en l'absence de l'intervention, en établissant des groupes de contrôle. Bénéficiaires et non bénéficiaires d'une mesure étatique.

Pour les besoins de l'étude d'impact, la sélection est réalisée le plus souvent de manière aléatoire (*Randomised Control Trial RCT*). La sélection devrait intervenir avant la sélection des bénéficiaires du projet. Cette méthode, proche des expérimentations à l'aveugle couramment utilisées dans le domaine médical permet d'attribuer sans ambiguïté à l'intervention évaluée les différences entre groupe de traitement et groupe témoin.

La construction statistique d'un groupe de comparaison peut aussi être réalisée ex post. Différentes méthodes de constitution de groupe de contrôles ont été discutées comme les techniques d'appariement (*matching*) de bénéficiaires à des individus non-bénéficiaires partageant des caractéristiques socio-économiques similaires, la séparation des individus entre groupes de traitement et groupes témoin par des évènements non liés au projet lui-même ni aux caractéristiques des bénéficiaires (*discontinuity design)*, l'élimination des différences initiales entre bénéficiaires et non-bénéficiaires par des méthodes comparant les différences dans l'évolution (plutôt que dans le niveau final) (*double différence).*

Ces méthodes supposent un grand nombre de bénéficiaires. Elles présentent les avantages et les risques liés au caractère rigoureusement scientifique de la méthode utilisée : coûts en général élevé,

peu ou pas d'explication sur les raisons pour lesquels un programme fonctionne ou non. Ce sont des méthodes qu'il ne faut cependant pas négliger dans certains domaines sensibles politiquement : comparaison du développement des entreprises qui ont bénéficié de fonds publics avec des entreprises n'ayant pas reçu d'aide, analyse de la taille de la classe et des performances des élèves, gros projets de coopération au développement, etc.

Le tableau ci-dessous tiré d'une brochure de l'Union européenne (Panorama inforegio 33), donne une bonne vue d'ensemble des méthodes que l'on peut utiliser dans le domaine de l'évaluation.

Evaluation d'impact «contrefactuelle» *(counterfactual impact invaluation)* Un programme fonctionne-t-il?	**Evaluation à base théorique** *(theory based evaluation)* Pourquoi un programme fonctionne-t-il?
• groupe témoin composé aléatoirement (*Randomised Control Trial RCT*) • Double différence *(difference in difference)* • *Discontinuity design* • Méthodes d'appariement (*matching methods*) • Variables instrumentales	• Enquêtes auprès des bénéficiaires • Etudes de cas, interviews • Evaluation réaliste • Evaluation participative

Un autres thème méthodologique très discuté a été la méthode «**Contribution Analysis** ». John Mayne a développé il y a une dizaine d'années une méthode pragmatique d'analyse de liens de causalité en opposant « Contribution Analysis» à « Attribution Analysis » qui suppose un scénario contrefactuel. L'analyse de contribution vise à démontrer si l'intervention évaluée est *une* des causes des changements observés ou non. Elle peut également classer l'intervention évaluée parmi les différents facteurs explicatifs des changements observés. L'analyse de contribution repose sur des chaînes d'arguments logiques qui sont vérifiés systématiquement, en général de manière qualitative. Jacques Toulemonde a utilisé cette méthode dans le cadre de plusieurs évaluations importantes au niveau de l'Union européenne. John Mayne présentera sa méthode lors d'un workshop SEVAL à Berne les 19 et 20 mai 2011.

2. Coopération au développement

Un grand de nombre de présentations a concerné la coopération au développement avec de nombreux intervenants provenant d'organisations internationales. Il y a plus de 3000 évaluations produites par année par les agences de coopération nationales ou multilatérales.

On parle actuellement beaucoup des méthodes d'évaluation d'impact contrefactuelle dans le domaine de la coopération au développement. Exemple d'un projet très perfectionné : Etude d'impact de programmes de micro-finance en milieu rural : le cas d'Al Amana au Maroc**.** L'objectif final de l'étude est l'évaluation de l'impact de l'offre de microcrédit (400 millions d'euros sur 20 ans) par Al Amana sur les activités agricoles et non-agricoles, les revenus des ménages et leurs actifs, le niveau de pauvreté et de sécurité alimentaire, la place des femmes dans le ménage.

A la suite de la Déclaration de Paris, la volonté est d'essayer d'impliquer au maximum le pays, les autorités, les ONG (NGO) et la population locale dans les évaluations et de favoriser les évaluations conjointes. Une grande difficulté provient du fait que les critères d'évaluations ne sont pas forcément les mêmes entre le pays bénéficiaire et le pays donataire. Ce dernier doit rendre compte de

l'utilisation de ses moyens et a développé son style et ses critères d'évaluation depuis de nombreuses années. Ces différences de reporting posent également des problèmes lorsque deux organisations d'aide cherchent à faire des évaluations conjointes. Différents exposés ont montré les difficultés et l'énergie nécessaire pour les surmonter. L'Unicef semble être à la pointe dans l'intégration du pays bénéficiaire. Il s'agit de renforcer son leadership et ainsi que l'appropriation des résultats, de sélectionner des thèmes d'intérêt commun. Il faut recourir aux experts locaux et cesser de présupposer qu'ils ont de faibles capacités d'évaluation. Le cas échéant, il faut contribuer à leur formation. Il est nécessaire d'investir dans le processus pilotés par les pays, même si cela exige des efforts supplémentaires et moins de contrôle… (Voir Marc Segone, Country-led M&E Systems. Better evidence, better policies, better development results").

3. Compétences des évaluateurs

Cela fait des années que différents pays s'intéressent à ce thème : Etats-Unis (Russ-Eft et al. 2008; King, Stevahn, Ghere et Minnema, 2001, 2005; Mertens, 1994; Torres, 1996), Canada (Kishchuck & Long, 1997; McGuire, Zorzi & Perrin, 2002), Australie (English, 2002), DeGEval (A. Caspari, M. Hennen, D. Scheffler, U. Schmidt & O. Schwab 2004), actuellement l'EES.

Pourquoi s'intéresser aux compétences des évaluateurs ? Comment pouvons-nous aider les débutants à développer leur propre expertise, leurs connaissances pratiques. La présentation de Claire Tourmen montre que les praticiens expérimentés...

- font un diagnostic plus large des situations d'évaluation – prennent davantage d'éléments en compte,
- pensent davantage en termes d'évaluabilité et de légitimité pour interpréter ces situations,
- ont des connaissances pragmatiques sur
  - comment un programme fonctionne,
  - comment les personnes agissent,
  - les différents types de résultats que peuvent produire les méthodes et les moyens utilisés.

Comment définir les compétences ? Sur la base d'une longue liste de compétences (104) soumise à une série d'experts dans quatre universités, Jacques Toulemonde essaie de montrer
1) quelles sont les compétences effectives pour le métier d'évaluateur,
2) si la compétence est spécifique à l'évaluation
3) si elles sont essentielles pour l'évaluateur et
4) si elles sont essentielles et spécifiques à l'évaluation comme par exemple la compétence :

- d'intégrer l'intérêt public dans le design de l'évaluation
- de définir et d'intégrer les stakeholders (parties prenantes)
- de rendre explicite la théorie du programme (modèle d'impact) ou de la reconstruire
- de focaliser l'évaluation et les questions auxquelles il faut répondre
- d'estimer l'évaluabilité du thème.

Voir les différentes listes qui se trouvent sur les slides de Jaques Toulemonde.

La **DEGEVAL** (A. Caspari, M. Hennen, D. Scheffler, U. Schmidt & O. Schwab 2004 – Empfehlungen für die Aus und Weiterbildung in der Evaluation) a développé des standards pour 1) aider ceux qui développent des programmes et pour les enseignants afin d'établir des programmes de formation à l'évaluation et pour pouvoir évaluer ces programmes et 2) évaluer les évaluateurs par exemple pour

aider les clients à choisir un évaluateur et pour permettre aux évaluateurs de s'auto-évaluer. Voici les prochaines étapes prévues par le groupe de travail de la DEGEVAL "Education et formation à l'évaluation":

- Enquête en ligne notamment auprès des membres DEGEVAL, d'experts pertinents, des diplômés de programmes d'études avancées et autres parties intéressées afin de mieux connaître la diffusion du document parmi les principaux utilisateurs visés, d'identifier des mesures visant à améliorer la diffusion et les nécessités de révision du point de vue des utilisateurs actuels.
- Révision du texte : notion de compétences (KSA), différenciations, élaboration des compétences méthodologiques, compatibilité avec les modèles internationaux, etc.
- Vue d'ensemble des cours de formation et séminaires offerts
- Renforcement des capacités
- Accréditation des évaluateurs

Personne de contact: Dr. Jan Hense, Ludwig-Maximilians-Universität München, DeGEval working group "Education and Training in Evaluation" (jan.hense@psy.lmu.de)

La **Société canadienne d'évaluation** est actuellement la plus en avance dans ce domaine avec son programme d'accréditation des évaluateurs : voir le référentiel des compétences professionnelles requises pour l'évaluation au Canada, version 12 du 29 avril 2010 qui comporte actuellement 49 indicateurs répartis dans cinq chapitres différents et le « Guide du postulant » de mai 2010. Quatre exigences sont formulées : 1) Diplôme ou certificat d'études supérieures, 2) Deux ans (équivalent temps plein) d'expérience professionnelle liée à l'évaluation acquise au cours des dix années précédentes. 3) Attestation d'études et/ou d'expérience pour 70 % des compétences énumérées sous chacun des cinq volets des Compétences requises à l'exercice de l'évaluation au Canada. 4) 40 heures de formation professionnelle par période de trois ans. Le système d'accréditation a débuté en juillet de cette année.

Le **débat** a porté sur la nécessité d'accréditer des personnes ou plutôt les cours donnés ou sur la nécessité de prendre en compte l'ensemble des compétences de l'équipe chargée d'une évaluation. L'EES s'est aussi attelée à définir ces compétences. Linda Morra Immas juge en effet que les compétences nécessaires devraient être déterminées au niveau international pour éviter des certifications multiples non reconnues. Les pays devraient ne devraient compléter cette liste que par rapport à leurs spécificités.

Note de l'auteur: pour la Suisse, il est important de choisir un modèle et de définir ses spécificités, tâche du groupe de travail « recherche et formation ». Le processus pour aboutir à une solution est aussi important que la solution elle-même. C'est important pour la reconnaissance internationale de la formation en évaluation prodiguée en Suisse et pour les bénéficiaires de cette formation. Ensuite, il s'agit de voir comment se rattacher à une solution internationale

4. Autres

La responsable des statistiques du Royaume-Uni, Denise Lievesley, a donné un brillant exposé sur « Counting with confidence – the importance of trustworthy statistics for evaluation ». Il y a très grandes analogies avec l'évaluation quant au rôle de la statistique par rapport à la politique et à l'importance de l'indépendance. La crédibilité de statistiques provenant d'unités décentralisées (cf.

USA) semble moins élevée que lorsqu'elles proviennent d'une seule unité administrative nationale. Denise Lievesley insiste sur l'importance de renforcer les liens entre statisticiens, évaluateurs, auditeurs et chercheurs.

Présentation intéressante concernant l'analyse (descriptive) du bon fonctionnement d'un réseau d'acteurs (coalition capacity), notamment les différents aspects à prendre en compte (capacité des membres du réseau, capacité du réseau lui-même) (Jared Raynor)

Présentation intéressante des difficultés liée à une législation néerlandaise d'imposer un service d'audit obligatoire au niveau des communes (Dick Hanemaayer).

Berne, 15 novembre 2010, Emmanuel Sangra