



Comment améliorer l'utilité et l'efficacité des évaluations ?

Le rôle des normes de qualité

Les deux objectifs principaux de l'évaluation sont connus : d'une part rendre compte des actions publiques à l'égard du contribuable et du citoyen permettant d'assurer la « redevabilité » et d'autre part tirer les enseignements du passé pour améliorer les actions futures. Améliorer l'utilité et l'efficacité des évaluations participe donc à l'amélioration de l'efficacité d'une politique publique. Dans ces conditions, améliorer l'utilité et l'efficacité de l'évaluation, c'est améliorer l'efficacité d'une action ou d'une politique publique.

Conscient de cet enjeu, les sociétés nationales d'évaluation, les réseaux des évaluateurs, les organisations internationales à caractère normatif, au premier rang desquels le Comité d'Aide au Développement de l'OCDE en ce qui concerne les évaluations des actions de développement ont cherché le meilleur moyen d'atteindre ce résultat en érigeant des normes de qualité ou des guides de bonnes pratiques.

L'édiction de ces normes fut le résultat d'un long processus entamé dans les années 60 notamment dans le domaine de l'éducation à travers les « Standards for Educational and Psychological Tests and Manuals ; American Psychological Association ».

Puis, il se développa dans les années 70 pouvant aboutir à une identification particulièrement détaillée et rigoureuse, à l'instar des normes que la société suisse d'évaluation a développées quelques années plus tard.

Au total, nous pouvons estimer que l'efficacité d'une évaluation repose essentiellement sur quatre piliers :

Toute évaluation doit être utile ; en ce sens, il est vain de vouloir tout évaluer. Mener une évaluation coûte chère et est chronophage. Dépenser par exemple 30 000€ pour mener l'évaluation d'un projet qui n'en a coûté que 20 000€ est évidemment absurde. Il faut donc éviter le syndrome du celui qui veut être plus royaliste que le roi. Deuxièmement, la singularité d'une évaluation repose sur le temps nécessaire à sa bonne conduite. Ce temps long s'oppose au temps court de la décision publique ou politique. Il est donc essentiel que l'évaluation reste connectée au temps court du politique en liant son exercice au processus ou au circuit de décision. Par exemple, mener une évaluation en amont d'un processus décisionnel est souvent pertinent ; mener une évaluation sans orientation temporelle risque de souffrir de ce décalage et in fine être inutile.

Le second pilier consiste à s'assurer de la faisabilité d'une évaluation. Les objectifs de l'évaluation doivent être précis, ciblés et au champ délimité. Le commanditaire peut avoir la tentation gargantuesque, c'est-à-dire celle de vouloir évaluer le plus de choses possibles avec l'antériorité la plus extensive. L'expérience montre qu'une évaluation doit être ciblée dans le temps et dans l'espace avec l'ambition de définir des objectifs mesurables et limités. En ce sens, les évaluations de politique sont les plus complexes dans la mesure où elle relève le plus souvent d'une forte transversalité. Identifier l'objectif, le donneur d'ordre et la mise en œuvre permet de s'assurer de la bonne faisabilité d'une évaluation.

Troisièmement, la conduite de l'évaluation doit être menée dans le respect des règles d'éthique et de déontologie. Il est bien sûr évident que l'évaluateur ne peut avoir été impliqué de manière directe ou indirecte à la gestion d'un projet, d'un programme ou d'une politique. Mais, l'équipe d'évaluation doit également suivre un ensemble de règles permettant s'assurer sa crédibilité et in fine son professionnalisme. Dans ce sens, l'exactitude des jugements et des informations sont les témoins de la validité et de la rigueur des données et de l'analyse. Ils constituent le dernier pilier. Les constats sont fondés sur faits et des données vérifiés et croisés ; leur traitement permet une analyse et donc un jugement objectif permettant de formuler des conclusions impartiales.

Dans ce contexte, en matière d'évaluation des activités de développement, les normes du Comité d'Aide au Développement de l'OCDE pour une évaluation de qualité furent adoptées en 2006 et résuma ces principes en dix critères principaux :

- justification, finalité et objectifs de l'évaluation
- paramètres de l'évaluation
- contexte
- méthodes d'évaluation
- sources d'information
- indépendance
- déontologie de l'évaluation
- assurance de la qualité
- pertinence des résultats de l'évaluation
- exhaustivité

Suite à l'application de ces normes dans la plupart des Etats membres, le groupe de travail sur l'évaluation du CAD a souhaité tirer les enseignements de cette première application des normes. Cet exercice fut lancé en 2009 et a pour mission d'aboutir le cas échéant à leurs révisions.

La mise en œuvre de telles normes répond à deux objectifs distincts :

D'une part l'existence de ces normes et leur caractère formalisé permet aux services opérationnels de disposer d'un *modus operandi* développant leur savoir et leur savoir-faire en matière d'évaluation. L'évaluation n'est pas processus aléatoire, subjectif et visant à mener une forme de contrôle ou d'inspection qui n'en porte pas le nom. La formation de ces normes permet donc de répondre aux principales interrogations en apportant des réponses concrètes, objectives, étayées et donc professionnelles.

D'autre part, une des faiblesses de l'efficacité des évaluations portait sur la mesure de la qualité du travail des consultants. Les services opérationnels, c'est-à-dire les services gestionnaires d'un projet, d'un programme, d'un instrument ou d'une politique pouvaient avoir la tentation de décrédibiliser une étude si cette dernière mettait en évidence l'inefficacité d'une action évaluée ; à l'inverse, une évaluation objectivement de mauvaise qualité (absence de rigueur méthodologique, fiabilité des données non vérifiées, autonomie aboutissant à un rapport de force, etc.) ne pouvait pas être sanctionnée. Enfin, les deux parties (le gestionnaire et l'évaluateur) pouvaient être frustrées dans la mesure où leur interaction s'arrêtait à la remise du rapport sans appréciation de l'utilité de leur collaboration ni des efforts fournis.

Pour répondre à toutes ces interrogations, sur la base de ces normes, une grille d'appréciation du travail des consultants est élaborée. Cette grille fondée sur quatre ou cinq

critères de notation¹ (de très bon à très insuffisant) permet d'apprécier pour chaque critère la performance du travail des évaluateurs.

La notation doit être réalisée de manière collégiale. La notation du seul commanditaire est évidemment à proscrire puisqu'elle réintroduit un rapport de force (l'évalué évalue l'évaluateur) ; de la même manière, on peut estimer que le tête-à-tête unité ou département d'évaluation vs. cabinet de consultants n'est pas idéal dans la mesure où il met la structure d'évaluation dans une position d'arbitrage. La solution la plus satisfaisante semble être du côté du comité de pilotage à partir du moment où sa composition est diversifiée et assure l'indépendance de l'évaluation. Dans ces conditions, chaque membre du comité de pilotage apprécie de manière personnelle et la plus impartiale possible la qualité du travail fourni.

La mise en œuvre de cette grille permet d'atteindre trois objectifs dont les retombées sont particulièrement intéressantes :

Tout d'abord, sa première fonction est de permettre une appréciation du travail des consultants. Ce faisant, les consultants ont connaissance de l'avis des membres du pilotage et peuvent donc avoir un retour du travail qui a été effectué. Ce procédé constitue une valorisation des efforts fournis et permet aux consultants eux-mêmes d'en tirer des enseignements pour améliorer leurs travaux futurs.

Ensuite, la collégialité est une assurance contre la mauvaise foi ou l'instrumentalisation dans un sens positif ou négatif du résultat d'une évaluation. Le service gestionnaire ne pourra pas instrumentaliser l'évaluation si les conclusions sont favorables uniquement par complaisance de la part des consultants ; à l'inverse, si les conclusions d'une étude sont critiques à l'endroit d'un service gestionnaire, ce dernier ne pourra pas s'en exonérer simplement en voulant discréditer la qualité de l'évaluation. Dans ces conditions, la notation collégiale permet de mettre en valeur les points focaux, c'est-à-dire les éventuels points de clivage de l'action évaluée. Ce faisant, cette grille permet d'apporter une démarche objective face à une perception subjective de la qualité d'une évaluation. La collégialité des appréciations est ici particulièrement importante. La participation des représentants d'autres départements ministériels, des représentants de la société civile, du secteur privé, des experts universitaires ou d'instituts de recherche permet d'optimiser l'indépendance et l'impartialité des jugements.

Enfin, l'établissement d'une telle grille est d'une grande utilité pour les unités ou départements d'évaluation. En son absence, il est juridiquement fragile d'établir des listes de « bons cabinets » et de « mauvais cabinets », faute de critères de notation objectifs et transparents. Les appréciations sont donc orales, nécessairement plus subjectives et reposaient sur la mémoire des individus, par définition faillible notamment en cas de rotation élevée du personnel. Cette méthode ne peut être satisfaisante. Le recours aux grilles de notation, publiques et publiées, permet ainsi d'apprécier dans la durée la qualité des travaux des cabinets de consultants et de fournir une base objective, collégiale et transparente des performances.

L'évaluation est un processus d'apprentissage. L'utilité et l'efficacité des actions ou politiques publiques reposent une sur meilleure formalisation et professionnalisation des évaluations. La mise en œuvre des normes de qualité semble fournir des résultats encourageants quant à la formalisation des évaluations.

Toutefois, les défis restent encore nombreux. Tous les pays, tous les ministères, toutes les agences ne les ont pas encore appliquées ou mises en œuvre dans leur intégralité. Or, seule l'intégralité de leur mise en œuvre permet d'atteindre les résultats escomptés. Par exemple, l'absence d'un comité de pilotage diversifié et compétent ne permet pas d'assurer l'indépendance et donc la pertinence de ses jugements. S'il n'existe pas une seule méthode

¹ Deux méthodes sont possibles : le recours à une notation sur une échelle impair de 5 (--/=+/++), elle a le souci d'être la plus précise possible ; le recours à une notation sur échelle pair de 4 (--/=+/++) ; elle supprime la notation moyenne permettant de trancher entre une performance plutôt bonne ou plutôt moyenne dans l'atteinte des objectifs.

permettant d'atteindre ces objectifs – le contexte socioculturel doit être pris en compte- leur déclinaison doit constituer un ensemble cohérent.

Auteur : Benoît Chervalier

Bibliographie : Benoît Chervalier est le chef de l'unité d'évaluation des activités de développement à la Direction générale du Trésor et de l'économie (Ministère de l'économie, des finances et de l'économie) et vice-président de l'organe subsidiaire sur l'Evaluation au Comité d'Aide au Développement de l'OCDE.