



Société Française de l'Évaluation

LES CAHIERS DE LA SFE

*Charte de l'évaluation :
Témoignages, débats et mise en œuvre des standards*

*N°1
Septembre 2005*

1 avenue de la Porte de Montrouge - 75014 PARIS

☎ : 01-45-41-58-40

Site Internet : www.sfe.asso.fr 📧 : evaluation.sfe@wanadoo.fr

Ont contribué à la rédaction de ce cahier :

- *Jean-Claude BARBIER*
- *Guy CAUQUIL*
- *Dominique GAGEY*
- *Claire GUIGNARD-HAMON*
- *Frédéric LEFEBVRE-NARE*
- *Manuela LOUE*
- *René PADIEU*
- *Bernard PERRET*
- *Jacques TOULEMONDE*
- *Thomas WIDMER*

Comité de lecture :

- *Claire GUIGNARD-HAMON*
- *Marie-Pierre ARLOT*
- *Jean-Claude BARBIER*
- *Béatrice BLANCHET-LACHENY*
- *Jean-Louis DETHIER*

Directeur de publication :

- *Emmanuel JAECK*

Responsable d'édition :

- *Stéphanie BRETON*

© Société Française de l'Évaluation, 2005.

La Société Française de l'Évaluation souhaite la meilleure utilisation et la plus large diffusion des travaux et analyses qu'elle publie. Toute citation ou reproduction partielle est donc autorisée. Toutefois, dans l'intérêt même des lecteurs, la SFE demande que la mention précise de la publication et la référence exacte (auteur, chapitre ou page) soient explicitement indiquées.

SFE

1 avenue de la Porte de Montrouge - 75014 PARIS

☎ : 01-45-41-58-40

Site Internet : www.sfe.asso.fr ✉ : evaluation.sfe@wanadoo.fr

SOMMAIRE

- *Introduction* p. 5
- *De la création de la SFE à l'adoption de la charte* p. 7
- *Charte de l'évaluation : cinq années de débats* p. 17
- *Une perspective européenne et une vue suisse sur des standards d'évaluation* p. 25
- *Charte de l'évaluation : les principes à l'épreuve des pratiques* p. 35
- *La charte des politiques publiques et des programmes publics* p. 51
- *Bibliographie* p. 53

INTRODUCTION

Ce numéro 1 des cahiers de la SFE ouvre une ligne éditoriale complémentaire à celle de l'édition annuelle (ou bi-annuelle) des actes des Journées Françaises de l'Évaluation. Ces cahiers se veulent avant tout un support de valorisation des productions des groupes de travail et clubs régionaux de la SFE et viennent donc compléter les réflexions - parfois thématiques, parfois stratégiques- de ces colloques nationaux qui rassemblent bien plus largement que les seuls adhérents de la SFE, praticiens ou théoriciens de l'évaluation.

Il nous a en effet semblé nécessaire de disposer d'un cadre rédactionnel qui structure la discussion sur le positionnement spécifique de l'évaluation des politiques publiques face à d'autres disciplines, et qui permette de faire valoir la situation française par rapport à celle de nos voisins les plus proches comme à celle d'autres espaces culturels ou institutionnels dont la tradition et l'histoire sont plus éloignés. Ces cahiers seront donc à la fois un instrument de débat interne et de confrontation avec d'autres acteurs français et étrangers du champ de l'évaluation des politiques publiques.

La première livraison a pour ambition de présenter et approfondir les débats suscités par notre « Charte de l'Évaluation », qui sont récurrents depuis la création de la SFE. Elle prend comme point d'appui la préparation de la Charte, pour ensuite étudier les pratiques qui en découlent, préciser l'usage qui est fait des standards en France, et confronter cet usage aux discussions qui ont lieu dans les principaux pays voisins.

Il s'agit donc d'un ensemble de contributions collectives, aussi denses que possible, qui ont l'ambition d'illustrer les « bonnes pratiques » et de retracer le déroulé des questions rencontrées. Celles-ci traversent des acteurs dont la définition, l'allure et l'envergure diffèrent ; reste que la réflexion sur l'évaluation leur est commune et semble résister aux aléas comme à la conjoncture. Une espèce de musique plus ou moins entendue, qui peut faire écho et revenir...

Assez spontanément, puisque le propos est celui d'une « pensée collective », c'est au groupe « standards de qualité et déontologie » qu'il a été proposé de faire l'ouverture et de rassembler les matériaux constituant ce premier cahier. Celui-ci, s'il n'est encore périodique, devrait être suivi d'autres, étalonnés aux besoins de leurs destinataires et à ce que la SFE peut être en mesure de produire et publier.

Le contenu de ce premier cahier court le risque de paraître parfois un peu trop "nombriliste" ou procédural. Il est cependant essentiel pour marquer deux bornes.

D'une part, il devrait permettre d'indiquer sur quels fronts se déplacent aujourd'hui les méfiances à l'égard de l'évaluation : situées il y a quelques années du côté des élus, elles se manifestent désormais par une absence du côté de l'exécutif qui -procédure budgétaire aidant- fait courir le risque d'une instrumentalisation de l'évaluation à des fins purement gestionnaires et se satisferait parfois d'une reconfiguration procédurale visant à concrétiser la modernisation de l'État. Dans ce cas, la défiance à l'égard de l'évaluation des politiques publiques n'est plus déclarée mais agit sur le plan instrumental, au contraire du propos

évaluatif. Tandis que par un phénomène de bascule, des élus souvent initialement défiants s'y entraînent progressivement, comme s'il était nécessaire de tomber dans la marmite avant d'y prendre intérêt et goût.

Cette analyse critique pourrait s'inscrire dans une réflexion sur les impasses du pilotage volontariste et sur la difficulté éprouvée à renouveler les formes de gouvernabilité.

Seconde grille d'analyse des débats conduits par le « groupe standard », la reprise des questions brillamment illustrées par la tradition anglo-saxonne du « value for money » : produire de l'action publique selon les coûts les plus faibles et au meilleur rendement, atteindre les objectifs que l'on s'est fixé, bref, montrer que les agencements choisis fonctionnent ou sinon sont révisables et adaptables. On serait là dans une position de surveillance et de contrôle; la perversion serait alors d'y faire procéder par des structures dites indépendantes et formées à bon escient. En ce cas, le discours sur les standards, très rationalisé, peut sembler parfois extrêmement violent. Il témoignerait, au contraire de la borne précédente, d'une défiance à l'égard des acteurs, des groupes sociaux et finalement de la démocratie.

Selon cette grille, on se placerait là, à force de recherche, de clarté et de mesures, sous l'autorité d'un bureaucrate apprenti sorcier.

Ces deux bornes virtuelles signalent les positions extrêmes qui ont marqué le débat sur « standards et déontologie » en France, et expliquent pour partie la lenteur qui pourrait paraître byzantine du débat sur la charte de l'évaluation : le consensus ne s'est bâti que très progressivement et a dû garder une certaine élasticité dans l'énoncé des principes, pour que chacun à la fois puisse y trouver son ou ses comptes, admettre que des variances étaient possibles et que cette charte était progressivement amendable. Elle est d'ailleurs promise à évolutions successives.

Ce premier cahier risque donc de rester atypique, par son objet et de par sa dynamique : il s'agit en effet d'un récit et d'une certaine explicitation des positions prises par différents groupes d'évaluateurs : chacun devrait ainsi être en mesure de s'appropriier les termes du débat, de s'y situer et de réfléchir aux divers possibles.

Bonne lecture donc.

Mais avant de passer à la page suivante, permettez-nous de remercier chaleureusement en votre nom et les membres du « groupe standards »¹ qui ont accepté de formaliser leurs débats à notre intention et chacune des petites mains² qui ont, avec assiduité, contribué à l'ouvrage.

Guy CAUQUIL
Président de la SFE

Claire GUIGNARD-HAMON
Secrétaire de la SFE

¹ ... dont les principaux acteurs sont cités en seconde de couverture.

² et plus particulièrement Stéphanie Breton qui a assuré la mise en page et Manuela Loué, stagiaire du cabinet Evalua, qui a contribué à sélectionner les premières références bibliographiques.

De la création de la SFE à l'adoption de la charte

Dès sa création, la question des « standards de qualité » et de la « déontologie » s'est imposée à la SFE comme l'un des enjeux majeur de la structuration d'un milieu professionnel de l'évaluation. Les statuts de la société prévoyaient expressément, dans leur article 3 (alinéa b), que la société « se fixe notamment pour objectifs de faire progresser les techniques et méthodes et de favoriser le respect des règles éthiques et procédurales propres à garantir la qualité des évaluations ainsi qu'un usage approprié des résultats ». Cette base statutaire inspira la création d'un premier groupe de travail³.

Le groupe « standards de qualité et déontologie »⁴ : rassemblement et interprétation des normes possibles (1999-2001)

La première réunion du groupe a été organisée à la fin septembre 1999, peu après le colloque fondateur de Marseille. Elle s'est tenue au siège de la SFE, à l'époque hébergée par le Centre d'études de l'emploi. Son ordre du jour comprenait la réflexion autour d'un nécessaire renforcement de la qualité des évaluations et du respect de règles éthiques et procédurales. Le premier texte d'invitation était ainsi libellé :

« Dans un premier temps, la question dont s'occupera le groupe de travail peut être située ainsi :

- *d'un côté, la possibilité de l'évaluation (la « zone d'autonomie » du politique et de l'administrativo-politique dont parle Jean Leca) suppose l'existence de normes collectives reconnues (qui dépassent la déontologie personnelle); celles-ci peuvent aider à construire l'indépendance de l'activité évaluative, face à l'utilisation stratégique par le politique et l'administratif et, en particulier, la légitimation cynique liée à ce qu'il est actuellement convenu d'appeler « gouvernement de l'opinion » ;*
- *de l'autre, symétriquement, les chargés d'évaluation ne peuvent prétendre à ce respect de l'autonomie que si leurs produits répondent à des critères justifiables de qualité sur lesquels producteurs et utilisateurs des résultats de l'évaluation peuvent tomber d'accord (sinon, c'est uniquement l'éthique individuelle de la conviction de l'évaluateur qui est opposable, comme, par exemple, dans le cas des travaux de recherche qui peuvent s'affranchir largement de critères d'utilité sociale, uniquement parce qu'ils sont soumis à l'évaluation par leurs pairs ; ou encore dans le cas de la commande d'étude où des standards professionnels sont portés par les regroupements représentatifs des cabinets d'étude). »*

La première formation du groupe, le 19 novembre 1999, comptait quatre participants : Jean-Claude Barbier, Pierre Dubois, Jean-Pierre Nioche et Bernard Perret ; plusieurs autres adhérents ayant manifesté leur intérêt mais avaient dû s'excuser : Françoise Chevallier, Jean-Louis Dayan, Jean Marc Dutrénit, Eric Monnier et Bernard Simonin. La réunion du 19 novembre procéda à une première classification de références normatives : des principes généraux de conduite ; des standards de qualité des produits de l'évaluation ; des indicateurs permettant l'appréciation de la qualité ou de la non qualité des travaux évaluatifs ; des références concernant les processus, le *design*, les dispositifs d'évaluation.

³ Lors de l'assemblée générale fondatrice, en 1999, à Marseille, les candidats à l'élection au bureau ont présenté les orientations qu'ils défendraient s'ils étaient élus ; Jean-Claude Barbier cita la question des standards comme l'une de ses deux orientations prioritaires, la seconde étant de faire mieux reconnaître l'activité d'évaluation dans la communauté académique.

⁴ Par simplification, on l'appelle, le plus souvent, le groupe « standards » dans le présent texte.

Le groupe adopta alors le projet de présenter un texte à débattre au congrès prévu à Rennes en 2000. Des contacts furent établis avec la société australasienne (C. Milne), la société suisse (T. Widmer), la société allemande (W. Beywl), la société britannique (E. Stern). Le groupe se répartit le travail de recensement des normes existant à l'étranger et en France (dont ce qui pouvait être considéré comme « l'héritage » du Conseil scientifique de l'évaluation, particulièrement porté dans le groupe par B. Perret), afin de constituer une base documentaire commune.

Un atelier fut finalement organisé pour le colloque de Rennes, avec des présentations de Jacques Toulemonde, Jean-Louis Dayan, Jean Marc Dutrénit, Bernard Perret et Jean-Claude Barbier. Une information fut également faite en assemblée générale sur le processus de réflexion sur les standards de qualité et sur la déontologie. Le bref débat qui suivit fit surtout ressortir des positions d'adhérents critiquant la nécessité même de standards pour la SFE, et doutant en particulier de la possibilité de définir ce qu'était « l'intérêt général ».

Le groupe « standards » continua son travail de rassemblement de normes, la première typologie évoluant au fur et à mesure des discussions. Jean-Claude Barbier et Bernard Perret⁵ firent une communication au colloque de la société européenne de l'évaluation, à Lausanne en octobre 2000, sur la comparaison des normes internationales. Ils y proposèrent d'apprendre, de façon privilégiée, de l'exemple québécois et canadien. Ils mirent notamment en évidence les larges plages de recoupement entre standards et principes d'éthique, tout en soulignant que les deux types de normes n'avaient pas la même fonction : références de qualité d'un produit ou d'un processus pour la première, versus principe de conduite des professionnels pour la seconde. En conclusion, ils soulignèrent, d'une part, la nécessité d'adapter les standards et principes au contexte français et, d'autre part, l'intérêt de lier la rédaction d'une charte à une réflexion sur ses usages.

La charte fut d'abord proposée pour :

- inviter les évaluateurs à souscrire à la charte, en s'engageant à s'y conformer dans leurs réponses aux appels d'offre,
- faire reconnaître la charte comme document de référence en cas de litiges entre commanditaires et évaluateurs.

Parallèlement à ce rôle direct dans la régulation du marché de l'évaluation, les auteurs mentionnaient plusieurs autres usages possibles d'une charte :

- servir de point d'appui pour contester l'utilisation abusive du terme évaluation dans le débat public,
- étayer les avis rendus par les divers comités scientifiques chargés de valider les évaluations,
- fournir un référentiel aux méta-évaluations.

Les deux auteurs étaient, à l'époque, membres du bureau de la SFE et rendaient compte de l'avancement des travaux à celui-ci⁶.

⁵ Cette communication n'a jamais été traduite en français, ni publiée ; elle est datée, du point de vue des références étrangères examinées (d'autres sociétés d'évaluation se sont créées depuis, comme la danoise et l'espagnole, pour n'en citer que deux) ; mais elle présente une approche toujours actuelle (Barbier J.-C., et Perret B., 2000, "Ethical Guidelines, Process and Product Quality Standards, What For An SFE (French Evaluation Society) Perspective", paper presented at the European Evaluation Society Conference, Lausanne, October, 12th -14th. ; le texte est accessible sur le site des auteurs, voir http://www.cce-recherche.fr/uk/fiches_chercheurs/barbier.htm; <http://perso.wanadoo.fr/bernard.perret/standards-SFE.htm>). Thomas Widmer, de la société suisse d'évaluation, a présenté une analyse plus récente et complémentaire (voir Widmer, Thomas, 2004, « The development and status of evaluation standards in Western Europe », *New directions for evaluation*, n° 104, Winter, p. 31-42).

⁶ Un troisième membre du groupe, Jean-Pierre Nioche, était également membre du bureau de la SFE.

A la suite de ce colloque international et du colloque de Rennes, le groupe « standards » s'est réuni plusieurs fois, dans l'intention de présenter un document à discuter à l'assemblée générale de 2001 ; il a associé de nouveaux membres, comme Françoise Bouygar, qui mit notamment l'accent sur l'utilité sociale de l'évaluation, René Padieu, Michael Ruleta et Dominique Gagey. Lors du séminaire du 13 mars 2001, organisé en prévision de l'assemblée générale de 2001, le groupe invita les membres du bureau de la SFE à débattre, avec lui et d'autres invités, à propos des résultats de ses travaux. Une partie des membres du bureau, dont le président, Maurice Baslé, et des membres parisiens de la SFE intéressés, débattirent avec un invité de la Commission européenne, sur une section du texte alors en préparation : Jacques Toulemonde organisa le débat sur la méthode, afin d'illustrer la façon dont le document pourrait être discuté au sein de la société.

Le colloque d'Issy les Moulineaux de juin 2001 approchant, le groupe décida de rassembler dans un document l'ensemble des principes et standards qu'il avait collectés et sélectionnés après plusieurs débats de classification des normes. Les normes furent présentées selon une grille de quatre types de normes : (a) les principes généraux et l'utilité sociale ; (b) les standards de qualité des produits de l'évaluation ; (c) les standards de qualité du processus ; (d) les standards de qualité du système d'évaluation. Un premier cycle de réflexion du groupe de travail débouchait ainsi, en avril 2001, sur un document relativement volumineux (d'une douzaine de pages structuré en quatre parties correspondant aux quatre types de normes identifiés ci-dessus).

Ce document⁷ fut donc diffusé lors du colloque de 2001. Présenté en assemblée générale par Jean-Claude Barbier et Bernard Perret, il donna lieu à une première discussion d'assemblée générale où les avis furent partagés. On fit référence à l'expérience des sociétés étrangères, notamment allemande et britannique : la société britannique refusait à l'époque d'adopter des « standards » alors que la société allemande en avait accepté le principe, mais connaissait cependant des difficultés de mise en œuvre⁸. La réflexion était ainsi ouverte mais non soumise à un vote, repoussé à l'assemblée générale de l'année 2002 devant se tenir à Lille. Le préambule du document, donné aux membres de la SFE, indiquait explicitement le processus envisagé à l'époque (voir encadré) ; les membres du groupe standards comme ceux du bureau étaient, en effet, conscients de la relative faiblesse du débat interne à l'association sur cette question.

⁷ Ce document est consultable sur le site de la SFE, <http://www.sfe.asso.fr/>.

⁸ Depuis, ces deux sociétés ont adopté des standards ; voir, sur ce point, l'article de Thomas Widmer déjà cité et le site de la société européenne d'évaluation, <http://www.europeanevaluation.org/>.

Encadré : Extrait du préambule du projet de charte

I - Statut du projet de charte

L'évaluation en France, en 2001, n'est ni un secteur professionnel ni un champ strictement délimité, ni même une activité dont la spécificité est bien établie, comme cela est le cas dans d'autres pays. Il convient donc d'adopter, en matière de diffusion de standards et de principes de déontologie, une démarche progressive et bien adaptée aux besoins des protagonistes de l'évaluation. C'est la raison pour laquelle le bureau de la SFE a choisi de mettre en débat, parmi les adhérents, l'adoption d'une charte « Standards de qualité et principes de déontologie », à adhésion volontaire.

Le présent texte⁹ est issu des travaux du groupe de travail « Standards et déontologie », d'une des contributions du groupe « Qualité de la commande », et du séminaire de travail, du 13 mars 2001, sur les standards et la déontologie.

Il prend la forme d'un document de travail, comportant aussi bien des points qui ont fait consensus dans les participants à son élaboration que des points controversés.

Diffusé à l'ensemble des adhérents, en prévision du colloque de la SFE de juin 2001, il sera discuté lors de la pré-session consacrée à ce thème, le 13 juin.

Lors du colloque de 2002, l'assemblée plénière de la SFE sera invitée à valider les points sur lesquels un consensus se sera dégagé.

II - Finalité de la charte

La Société française de l'évaluation (SFE), conformément à ses statuts, œuvre pour la diffusion des bonnes pratiques de l'évaluation, la professionnalisation des chargés d'évaluation, des membres des instances d'évaluation et des commanditaires, ainsi que le respect, par tous ses participants, de normes éthiques de conduite dans l'évaluation des politiques et des programmes.

La présente charte a vocation à s'appliquer à l'ensemble des participants à des opérations d'évaluation de politiques ou de programmes. Les chargés d'évaluation peuvent librement décider d'y adhérer. Dans ce cas, ils joignent la charte à leurs réponses aux appels d'offre d'évaluation. De leur côté, les commanditaires peuvent faire référence à la charte dans leurs appels à prestations d'évaluation. La charte peut également être jointe à la lettre nommant les membres des instances d'évaluation.

Parallèlement, la charte a vocation à servir de référentiel lors de la validation de la qualité des évaluations (par exemple, par les conseils scientifiques régionaux). Elle peut aussi servir de référence dans la réalisation de méta-évaluations portant sur un ensemble d'opérations.

Les normes de la présente charte envisagent, d'une part, les principes qui relèvent de la *conduite éthique des personnes* engagées dans ces opérations d'évaluation, et, de l'autre, les *critères de la qualité* des évaluations. Ces deux ordres de normes sont toutefois profondément interdépendants l'un de l'autre.

⁹ Une réflexion spécifique reste à conduire pour l'application de la Charte aux « évaluateurs internes » des administrations.

Maturation et controverses internes (2001-2002)

A partir de ce moment, s'ouvre, au sein du bureau et de la société, un débat interne qui peut être analysé de bien des façons¹⁰ ; nous en retracerons les grandes lignes sans insister sur les porteurs de telle ou telle position.

Fondamentalement, parmi ceux qui s'étaient exprimés, coexistaient trois positions dans la société : celle de ceux et celles qui pensaient qu'on pouvait avancer et faire adopter rapidement des standards de qualité et des principes de déontologie par les adhérents ; celle de ceux (celles) qui y étaient réticents par principe ; celle, enfin de ceux (celles) qui pensaient que le choix était prématuré.

Cependant, le débat ne se situa pas sur ce plan explicite : il donna lieu à des controverses vives à propos de deux points précis. Le premier était la question, elle aussi controversée dans la littérature internationale (voir, sur ce point, M.Q. Patton, par exemple¹¹), de savoir à qui l'évaluateur ou le chargé d'évaluation devait rendre des comptes ; pour les uns, l'intérêt général était engagé chaque fois qu'il y avait évaluation, par le fait qu'ils pensaient la question en termes de politiques publiques ; pour les autres, l'intérêt général n'était pas toujours pertinent à considérer.

Ce premier point de divergence s'articulait avec un second portant sur ce que l'on pouvait appeler « évaluation » ; les uns pensaient qu'un face à face, entre un commanditaire et un chargé d'évaluation, dans une relation classique de consultant ou d'expert, pouvait être rangé dans la catégorie des études d'évaluation ; à l'inverse, les autres soutenaient, pour qu'il y ait vraiment évaluation, qu'il fallait sortir de cette relation. Dans le « codage » de la controverse, cette question apparaissait sous la rubrique du « pluralisme des méthodes ». Le « pluralisme », ainsi défini, permettait, par ailleurs, une critique implicite de l'expérience du Conseil scientifique de l'évaluation.

Entre le colloque d'Issy les Moulineaux de juin 2001 et celui de Lille de juin 2002, le débat ne put être tranché. Avant le colloque de Lille, le bureau avait demandé au groupe « standards » de lui faire tenir un « exposé des motifs » pour une charte, texte qui fut rédigé par Jean-Claude Barbier et Bernard Perret au nom du groupe (voir annexe 1), le 9 octobre 2001. En février 2002, le groupe écrivit au bureau en s'interrogeant sur les intentions de ce dernier et en lui proposant de prendre la première section du document (principes généraux de conduite et utilité sociale de l'évaluation), discuté à Issy les Moulineaux, pour en faire la base d'une présentation au vote de la future assemblée de Lille.

Ayant délibéré, le bureau décida de présenter au colloque un texte intermédiaire, signé de son président (« Note d'étape sur la charte SFE et les standards et la déontologie dans les pratiques évaluatives et les évaluations de politique publique »¹²) qui le présenta au colloque.

¹⁰ Pour une analyse du contexte institutionnel français, voir Barbier J-C., 2004, "Devising and using evaluation standards, the French paradox", in Schwartz R., Mayne J., (eds), *Quality Matters, Seeking confidence in evaluation, auditing and performance reporting*, Transaction Books, Rutgers, N Jersey, p. 26-57.

¹¹ Dans son ouvrage de 1997, Patton discute ce point à plusieurs reprises (voir page 16, page 341, notamment, dans Patton M-Q., 1997, *Utilization Focused Evaluation*, Sage, Thousand Oaks.

¹² Le préambule de cette note, adoptée par le bureau le 13.5.2002, était le suivant : « Le besoin de principes spécifiquement ré-affirmés pour guider les actes des commanditaires et des prestataires dans l'évaluation des politiques publiques s'impose. Certes, les principes de la commande et de la prestation intellectuelle obéissent déjà à des codes qui sont ceux de la communauté d'appartenance, du milieu scientifique (regard des pairs, construction des réputations d'experts...), ou du contrat. Mais le problème est que la convention ou le contrat peuvent donner l'apparence qu'il s'agit d'un travail sans enjeu collectif : par définition, dans les contrats, les parties prenantes contractent parce qu'ayant d'abord un intérêt privé à la prestation. Mais, dans l'évaluation des politiques publiques, le contrat lui-même et la qualité de la prestation fournie dépassent les intérêts des parties prenantes. La raison de ce débordement d'intérêt est très simple : il s'agit d'aboutir à la production d'une connaissance publique de l'action publique. Cette connaissance va orienter les politiques publiques et donc produire des effets sur de larges communautés. Les politiques publiques, ayant pour objectif de réguler des conflits d'intérêt

L'animateur de groupe "standards et déontologie", absent du colloque, fit lire une courte mise au point, dans laquelle il indiquait que la lettre du bureau, qui faisait référence aux travaux du groupe « standards », s'écartait des choix préconisés par ce groupe, notamment, du texte des « principes généraux de conduite et utilité sociale de l'évaluation ». La controverse était latente au sein de la SFE, mais ne mobilisait vraisemblablement qu'une partie des adhérents.

Après le débat animé du congrès d'Issy les Moulineaux, l'absence de débat, lors du colloque de Lille, et du fait du contexte de controverse latente, le groupe de travail et le bureau firent évoluer le texte d'Issy vers un texte plus court, visant à formuler des principes plutôt que des recommandations détaillées de normes de qualité, qui serait probablement plus aisé à faire admettre par les adhérents. Ceux-ci n'étaient d'ailleurs pas très mobilisés, à la différence des divers groupes de travail et clubs de la SFE, qui organisèrent, dans plusieurs régions, des débats autour du projet de charte.

L'initiative du bureau en amont du colloque de Limoges (2003)

Finalement, le nouveau bureau, élu à Lille en 2002, se saisit de la documentation existante pour en tirer un document très court, devenu le texte actuel de la charte adoptée le 28 octobre 2003, lors de l'assemblée générale de Limoges ; ce texte fut voté à une confortable majorité, mais non sans débats. Pour faire droit aux questions posées, il fut précisé que cette charte serait périodiquement révisée et soumise à un processus élargi de confrontation avec les praticiens de l'évaluation, au-delà même des adhérents de la SFE. Ce texte, tout en s'inspirant de la partie « principes » du document d'Issy, innove et synthétise, d'une façon claire et succincte, les six grands principes adoptés.

Quelques enseignements

A ce stade, il est possible de revenir brièvement sur le processus qui a conduit à l'adoption de la charte et d'en tirer quelques commentaires.

Un trait caractérise le débat interne à la SFE : un mélange constant d'intérêt (cf. les invitations à discuter dans plusieurs régions) et de faible mobilisation des adhérents pour participer individuellement au débat ; à plusieurs reprises, les courriers du bureau, adressés aux adhérents, pour les consulter sur cette question des standards, restèrent sans réponse.

La réflexion sur les standards de qualité et la déontologie a soulevé, de façon récurrente, la question de la définition de l'évaluation. A l'origine de la société, le groupe fondateur, discutant et rédigeant les premiers statuts, fit le choix explicite de ne pas prendre cette question de front, optant, au contraire, pour la palette la plus large de qualités possibles pour les adhérents, incluant les contrôleurs comme les gestionnaires, les professionnels de l'audit, etc. (voir le dernier alinéa du point b de l'article 3 des statuts).

autrement que par des arrangements contractuels, les participants à l'évaluation comprennent que les enjeux les dépassent en partie. Ils acceptent donc que les conventions ou les contrats qui les lient reflètent, comme dans la concession de service public ou l'achat public, la dimension "publique" de l'évaluation.

Les évaluateurs et leurs commanditaires sont avertis des conflits potentiels et des enjeux de l'évaluation : de ces enjeux et conflits d'intérêt naît le besoin, au-delà de la commande habituelle d'études, de repères et de quelques lignes indicatrices, ou de guides, précisant la procédure tant au niveau des commanditaires, des prestataires et des citoyens. »

Ce débat sur la nature de l'évaluation, qui n'a jamais eu lieu explicitement au sein de la SFE, resurgit d'une autre façon au moment où ces lignes sont écrites, en raison, notamment, de l'absence de renouvellement d'une instance compétente au niveau de l'Etat central en matière d'évaluation, mais aussi au moment où la loi d'orientation (LOLF) apporte un bouleversement des pratiques budgétaires, explicitement référé à la mesure de l'efficacité et de l'efficience des programmes de l'Etat.

Si, à Limoges, l'on décida, sous la présidence de Jacques Toulemonde, de ne pas mettre en discussion la lettre de la charte, mais de recenser, dès ce moment, les débats, deux points ressortirent : premièrement des participants à l'assemblée générale s'interrogeaient sur l'opportunité de garder la formule du préambule faisant référence au jugement sur la « valeur » des politiques, surtout en s'appuyant sur les inconvénients présumés de cette formule vis-à-vis des hommes et des femmes politiques. Deuxièmement, si la majorité était acquise au principe de la charte, des positions contraires explicites se manifestèrent de nouveau, comme à Issy les Moulineaux en 2001. Enfin, le principe de pluralité exposé par la charte de Limoges fut également contesté vivement. Ces points sont étudiés dans l'article suivant. La charte de Limoges se présente ainsi comme une étape pour une discussion toujours en cours.

ANNEXE - EXPOSE DES MOTIFS POUR UNE CHARTE¹³*Une charte de l'évaluation, pour quoi faire ?*

De nombreux groupes professionnels sont amenés à rédiger des codes de déontologie ou chartes de qualité (on pourrait montrer que cette préoccupation, qui a toujours existé, est renforcée par un contexte où l'éthique fait problème et où l'on ressent davantage le besoin d'explicitier les normes régissant les pratiques sociales).

Le besoin de définir des critères de bonnes pratiques et de déontologie est particulièrement fort en matière d'évaluation, compte tenu du caractère politique de cette activité et de la vigueur des débats concernant le rôle et le positionnement des évaluateurs.

Dans les principaux pays où l'évaluation est pratiquée à grande échelle, les associations d'évaluateurs ont élaboré des systèmes de normes ou standards, relatifs, d'une part, à la déontologie (des évaluateurs, ou, plus largement, des commanditaires et des acteurs qui participent au pilotage de l'évaluation, comme les membres des instances d'évaluation) et, d'autre part, à la qualité des évaluations (qualité de ses produits, qualité de ses processus, qualité des systèmes d'évaluation).

Les États-Unis disposent de telles normes, le Canada et l'Australasie également : dans ces trois univers « anglo-saxons », un système de normes développé sert d'outil à des professions plus organisées qu'en France.

En Europe, la société suisse de l'évaluation a adapté le principe des normes américaines selon le principe de commencer à s'en servir. La société allemande travaille, depuis plusieurs années, sur un projet analogue. La société italienne a adopté, dès son origine, une charte déontologique plus générale, pour les chargés d'évaluation. La société britannique est la seule société d'évaluateurs d'un grand pays développé qui, pour l'instant, ait écarté le projet d'investir dans des normes particulières de l'évaluation.

Dans des contextes où l'évaluation constitue un marché très concurrentiel, l'objectif est d'abord de protéger le client et d'inciter les fournisseurs d'études évaluatives à respecter des exigences de rigueur, d'impartialité et de respect des personnes.

La situation française est certes différente. L'évaluation se développe dans des contextes institutionnels variés qui posent, dans chaque cas, des problèmes spécifiques. La responsabilité personnelle de l'évaluateur est parfois partiellement transférée à une instance d'évaluation.

Une réflexion collective, visant à formuler des critères de bonnes pratiques et de déontologie, peut, dans ce cas, contribuer à la structuration du milieu professionnel de l'évaluation, au progrès des méthodes et à la clarification des rapports entre commanditaires et chargés d'évaluation.

¹³ J-C. Barbier et B. Perret, 9 Octobre 2001, Exposé des motifs pour la charte « standards de qualité et déontologie ».

Dès son origine, la SFE a inscrit, dans ses statuts (article 3b), l'objectif de « favoriser le respect des règles éthiques et procédurales propres à garantir la qualité des évaluations ainsi qu'un usage approprié de leurs résultats ».

C'est dans cet esprit que le projet de charte « Standards de qualité et déontologie » d'avril 2001, émanant du groupe de travail consacré à ce sujet, a été soumis au débat à l'assemblée générale de la SFE, en juin dernier.

Comme cela a été évoqué à cette occasion, conformément aux objectifs proposés par le groupe de travail, la démarche d'adoption de références communes n'a de sens, en France, que si elle est « pédagogique », c'est-à-dire si elle vise à enrichir les pratiques et élever les compétences de tous les acteurs engagés dans l'évaluation.

Sa faisabilité suppose une progressivité. C'est d'un travail d'apprentissage collectif qu'il s'agit.

Elle prend tout son sens également dans le projet de rassembler, dans un texte commun, un répertoire de normes qui puissent être adaptées à divers environnements de l'évaluation (évaluation dans un secteur particulier, évaluation interne, etc.).

Le document, qui est soumis au débat, ne doit donc pas être considéré comme la totalité des normes qui devraient, dans un idéal sans doute impossible à définir, être respectées par tous les acteurs de l'évaluation à un moment donné. Seules certaines de ces normes peuvent être considérées comme suffisamment prééminentes pour être retenues comme un corpus commun.

Dans son état actuel, tel qu'il a été diffusé à tous les membres de la SFE, le projet de charte est donc susceptible de beaucoup d'améliorations, auxquelles les adhérents et les groupes de travail, les clubs régionaux, sont invités à contribuer, soit via le groupe de travail « standards et déontologie », soit directement via le bureau.

Cinq années de débats autour des principes de la charte

Introduction

Comme l'indique le rappel historique introduisant ce cahier, le texte actuel de la charte fut adopté fin 2003 à Limoges, à une majorité confortable, mais non sans débats. Pour faire droit aux questions restant posées, il fut précisé que cette charte devrait être périodiquement révisée et soumise à un processus élargi de confrontation avec les praticiens de l'évaluation, au-delà même des adhérents de la SFE.

C'est dans ce but que le groupe de travail « standards et déontologie » de la SFE a organisé plusieurs débats et tables rondes au cours de l'année 2004 (deux à Paris en juin et décembre et un à l'occasion du congrès de Bordeaux en septembre). En mai 2005, au moment où ce texte est écrit, le groupe vient de tenir une réunion publique sur les rapports entre la charte de l'évaluation et la LOLF.

Le travail a été organisé autour de tables rondes auxquelles le groupe Standards et Déontologie a invité des personnes appartenant à diverses institutions : commanditaires, universitaires, praticiens, etc. Ces tables rondes visaient à préparer d'éventuelles évolutions de la charte, mais surtout à mieux expliciter les questions en débat sur la pratique de l'évaluation, à partir de l'utilité de la première version de la charte. Il s'agissait notamment de distinguer les simples commentaires, ou modifications, d'éventuelles évolutions substantielles des principes eux-mêmes.

Les participants ont marqué leur intérêt, du point de vue de leur pratique professionnelle, pour la première version de la charte ; ils ont également relevé les questions qui restent en suspens à partir de la présentation de cas illustrant des problèmes d'interprétation ou d'application concrète de la charte (voir synthèse ci-après).

Le bilan synthétique de ces tables rondes fait l'objet des parties 2 (débat sur les attendus) et 3 (débat sur les principes) du présent texte.

Bien que de manière limitée, ces discussions ont un peu débordé le cadre de la SFE. Elles prouvent que la charte peut aider les praticiens à réfléchir aux enjeux déontologiques des choix méthodologiques auxquels ils sont confrontés dans leur pratique professionnelle. Nombre d'entre eux perçoivent l'intérêt d'une charte pour la structuration du champ de l'évaluation.

Débat sur les attendus de la Charte

Dans la phase initiale du débat (notamment lors du congrès de Limoges), les oppositions à la Charte se sont cristallisées autour des mots « jugement » et « valeur ». Or, dans le texte actuel de la Charte, il n'est pas écrit que l'évaluation est, en elle-même, un jugement de valeur, mais seulement qu'elle « cherche à fonder » un jugement « sur la valeur ». Il y a là plus qu'une nuance : reconnaître que la visée ultime de l'évaluation est de permettre l'exercice d'un jugement n'implique pas qu'elle constitue, par elle-même, un acte de jugement. Cet effort de précision n'a cependant pas suffi à lever les oppositions. Celles-ci renvoient, pour une part, à des malentendus sur le sens des mots « jugement » et « valeur », mais également, à des clivages plus fondamentaux sur la définition de l'évaluation.

Le débat sur la portée du jugement évaluatif n'est pas sans rapport avec certaines des questions posées à propos des principes (notamment ceux de pluralité, distanciation et transparence). L'enjeu fondamental est la reconnaissance du caractère « public » de l'évaluation¹⁴, comme partie intégrante d'un processus politique (au sens de *Policy*) lui-même public, c'est à dire conduit pour l'ensemble des citoyens et sous leur regard. La différence de définitions proposées par certains peut aider à expliciter la nature du débat ; on rappelle la seule définition officielle française est celle inscrite dans un décret de 1990 : « *évaluer une politique, c'est rechercher si les moyens juridiques, administratifs et financiers mis en œuvre permettent de produire les effets attendus de cette politique et d'atteindre les objectifs qui lui sont fixés* » ; une autre définition proposée insiste plus sur le cadre institutionnel formalisé à utiliser pour mieux satisfaire aux principes, notamment de pluralité : « *évaluation, activité d'étude et d'analyse portant sur la mise en œuvre et les résultats d'une action, menée dans un cadre méthodologique et institutionnel formalisé dans le but de former des jugements empiriquement et normativement fondés sur la pertinence, la cohérence, l'efficacité et/ou l'efficience de cette action* ». Mais certains, sans méconnaître l'importance du principe de pluralité, s'interrogent sur l'intérêt de prôner quasi-systématiquement des formes uniques d'évaluation avec des instances, pensant qu'il peut être plus efficace d'utiliser des modes d'évaluation plus divers.

Non sans rapport avec le point précédent, la question a été posée, à plusieurs reprises, de l'insertion dans le préambule d'une définition plus classique de l'évaluation, rappelant qu'elle cherche à mesurer les effets d'une action. Mettre l'accent sur les finalités informationnelles et cognitives de l'évaluation permettrait peut-être de lever une partie des malentendus sur les mots « jugement » et « valeur ». Dans le même ordre d'idée, il convient de mentionner la suggestion de Jean-Pierre Nioche d'ajouter au préambule la phrase suivante : « l'évaluation s'inscrit dans la double logique de la démocratie pluraliste et du débat scientifique ; elle s'exerce donc dans une pluralité de commandes, de dispositifs et de méthodes... ».

Reste cependant ici un point délicat pour la SFE, qui doit tenir compte de plusieurs éléments :

- d'un côté, les choix normatifs faits dans la charte privilégient parmi d'autres une conception particulière de l'évaluation, sans pour autant s'y résumer ; cette conception est marquée par la première institutionnalisation de l'évaluation des années 1990 en France, cette « évaluation à la française », dont parlaient Eric Monnier et Patrice Duran, dans leur article de la revue *Evaluation*, en 1995¹⁵ ;
- de l'autre côté, la société est ouverte à de nombreuses catégories d'adhérents, dont le travail ou les intérêts concourent à l'évaluation au sens le plus large ;
- en troisième lieu, comme il fut de nouveau constaté dans la réunion publique animée par J. Toulemonde à propos d'études de cas, le terme même d'évaluation continue en France à bénéficier d'une connotation positive que les usages divers, plus ou moins stratégiques ou politiques cherchent à s'approprier ; se référant à un « état de l'art » international de l'évaluation, la SFE ne peut donc manquer de prendre position vis-à-vis de possibilités d'usages relativistes, qui incluraient, à la limite, le conseil, le contrôle de gestion, celui de conformité, etc.

¹⁴ C'est, au fond, le même débat que celui indiqué dans l'introduction historique à ce cahier, à propos de la présence des termes « intérêt général » dans le préambule de la charte.

¹⁵ Duran P., Monnier E., Smith A. 1995. "Evaluation a la française, Towards a New Relationship between Social Science and Public Action." *Evaluation*, 1, n°1: 45-63.

- enfin, la situation française continue de manifester un paradoxe vis-à-vis du contexte européen, qui, certes peut s'expliquer sur le plan sociologique et institutionnel, mais ne semble nullement en voie de se résorber : la réticence à normaliser ou institutionnaliser l'évaluation dans le pays, et l'usage constant de la notion même d'évaluation à des fins stratégiques, légitimantes ou commerciales, à côté de fins tout à fait conformes [précisément à cet état de l'art international], alors même que ces derniers sont bien représentés par les critères et normes maintenant anciens établis par la Commission européenne.

Au total, il n'est pas aisé de concevoir, pour la SFE et plus généralement pour l'évaluation en France, une doctrine de référence, des standards et des principes. Ce débat reste très largement ouvert aujourd'hui.

Débat sur les principes

Dans son état actuel, la charte ne traite que de certains des principes, normes et critères qui peuvent être mis en avant en matière d'évaluation. En particulier, si l'on excepte le principe de transparence, elle n'inclut pas de critères de qualité des « produits » de l'évaluation, et ce contrairement au texte construit et discuté en 2001 à Issy les Moulineaux¹⁶.

➤ **Principe de pluralité**

1. Comme on vient de le rappeler, le débat se poursuit sur le point de savoir si l'évaluation pluraliste est, ou non, la seule forme d'évaluation. Il n'en demeure pas moins que, dans les pratiques françaises contemporaines, l'évaluation semble de plus en plus conçue comme un exercice pluraliste.
2. Ce relatif consensus - qui repose sur des raisons au moins autant méthodologiques qu'éthiques (intérêt de connaître l'ensemble des points de vue, notamment pour la pertinence et l'utilité sociale de l'évaluation) - ne résout pas toutes les questions. Ainsi, celle du mode de participation des usagers à l'évaluation se pose de manière récurrente, notamment pour les politiques sociales, sans que puisse se dégager une solution univoque et claire.
3. La formulation actuelle du principe de pluralité ne précise pas à quel stade de l'évaluation il est censé s'appliquer : construction du projet, pilotage de l'évaluation, enquêtes, etc. Le principe de pluralité est d'autant plus exigeant qu'il s'applique aux phases amont de l'évaluation. Tout le monde ne peut être associé à l'origine de l'évaluation. Or, la charte ne distingue pas les différents stades. Dans certains cas, l'application du principe de pluralité au stade de la formulation des questions d'évaluation peut se heurter à la légitimité du commanditaire politique. Enfin, la question du poids relatif des *insiders* et des *outsiders* dans les comités de pilotage se pose très différemment selon le contexte et la nature de l'action.
4. Les limites rencontrées dans l'application du principe de pluralité ne sont pas seulement politiques, elles sont aussi pratiques au regard du fonctionnement des instances. Pour tenter de concilier l'exigence de pluralisme et les contraintes pratiques, l'idée a été lancée d'instances de pilotage à géométrie variable, ne réunissant tous leurs membres que deux ou trois fois au cours d'une évaluation et travaillant, le reste du temps, en formation plus réduite.

¹⁶ Se reporter à l'historique chapitre 1. Ce texte est accessible sur le site de la SFE.

➤ Principe de distanciation

La formulation de ce principe fait référence au fait que l'évaluation doit être conduite de manière autonome par rapport au processus de gestion et de décision.

1. On passera rapidement sur le cas des évaluations « sous influence », dont les résultats sont manipulés par la communication gouvernementale (cf. l'exemple de l'évaluation des 35 heures, présenté par Jean-Louis Dayan lors du colloque de Rennes, et sa défense de la possibilité d'une distanciation au sein même des services de l'Etat). La plupart des cas mentionnés dans les discussions ne relèvent heureusement pas de cette situation : l'évaluation est le plus souvent reconnue comme un processus autonome dont le déroulement et les conclusions ne peuvent être manipulés. Mais un certain nombre de questions restent entières.
2. L'une des présentations faites au colloque de Bordeaux, en 2004, montre combien la distanciation peut être limitée par des impératifs managériaux et, plus précisément, par l'inscription de l'exercice d'évaluation dans un mécanisme de type contractuel, en l'occurrence les « conventions d'objectifs et de gestion » avec les organismes de protection sociale. Dans un tel cas de figure, qui n'a rien d'exceptionnel et pourrait même se généraliser dans la logique de la LOLF, l'évaluation est inscrite et utilisée dans une relation « tutelle - agence ». Réalisée à l'initiative de l'un des deux co-contractants, elle s'apparente parfois à un audit externe de performance auquel les standards de l'évaluation des politiques publiques ne peuvent totalement s'appliquer. Utile et légitime, cette forme d'évaluation n'est manifestement pas « conduite de manière autonome par rapport au processus de gestion et de décision ». Dans l'exemple cité, le principe de distanciation, basé sur une instance d'évaluation pluraliste et autonome, est écarté au profit d'un « principe de contradiction » qui se concrétise par la confrontation de deux évaluations distinctes. Cela peut permettre d'aboutir à des conclusions partagées, mais on ne peut, malgré tout, parler de distanciation au sens de la charte. Au-delà de l'évaluation elle-même, le cas que l'on vient d'évoquer révèle la tension qui peut exister entre le caractère politique d'une action (et l'exigence de démocratie qui en découle pour son évaluation) et la forme contractuelle d'une relation entre acteurs publics¹⁷. La question « évaluation et contractualisation » mériterait, sans doute, une réflexion spécifique. La tension contrat/ intérêt général se pose à deux niveaux emboîtés : celui de l'objet sur lequel porte l'évaluation (action publique ou contrat entre deux ou plusieurs acteurs ?) et celui de la prestation évaluative (prestation intellectuelle achetée librement par un commanditaire, ou activité inscrite dans l'action publique et participant structurellement de son caractère public ?).
3. Dans le cas d'évaluations qui revendiquent leur caractère « pluraliste », les interactions entre l'instance d'évaluation et l'évaluateur ou le chargé d'évaluation posent un certain nombre de questions à approfondir. Il convient tout d'abord de distinguer les cas où l'instance d'évaluation est l'auteur collectif du rapport, des cas où le rapport d'évaluation est assumé par un évaluateur professionnel¹⁸. Dans le premier cas, l'instance assume clairement ses responsabilités vis à vis du commanditaire. Dans le second cas se pose la question du degré d'autonomie de

¹⁷ De nombreux exemples de cette situation existent en France, relatés parfois dans les rapports officiels : c'est le cas du rapport Marimbert (janvier 2004) à propos de la réforme du service public de l'emploi.

¹⁸ La tendance dominante de la référence de la Commission européenne est plutôt la première, qui exige de l'évaluateur un jugement. On peut penser que cette tendance est également dominante dans l'état de l'art international de l'évaluation.

l'évaluateur par rapport à l'instance : faut-il élever au rang de norme la distinction suggérée par l'un des contributeurs à l'atelier de Bordeaux , à savoir la distinction entre les remarques factuelles de l'instance (corrections d'erreurs manifestes, informations complémentaires), qui doivent être intégrées au rapport, et les observations concernant le fond de l'analyse qui doivent être prises en compte en tant que commentaires distincts du rapport ?

4. A partir de quelle étape du processus l'évaluation doit-elle s'autonomiser par rapport au commanditaire ? En particulier, le moment de la formulation des questions doit-il être assumé par une instance d'évaluation libre de jugement par rapport au commanditaire ? Dans le cas du Conseil régional de Rhône-Alpes, le projet d'évaluation est discuté directement par l'assemblée délibérante. N'est-ce pas une confusion des rôles ? Il est certes de la responsabilité du commanditaire de fixer les objectifs généraux de l'évaluation, mais il semble souhaitable que le questionnement soit complété et précisé par une instance autonome.
5. Les corps d'inspection sont-ils légitimes pour l'évaluation ? L'organisation du processus est, certes, plus importante que le profil des personnes qui participent à l'évaluation. Le plus important est que des acteurs, ayant un intérêt ou un point de vue particulier à défendre, ne soient pas en mesure de peser de manière excessive sur le déroulement de l'évaluation. A l'inverse, l'intervention d'un consultant externe ne garantit pas une distanciation *ipso facto*¹⁹. Dans le cadre du modèle pluraliste, le choix du président de l'instance est crucial pour garantir une posture adéquate vis à vis du commanditaire. Confier la présidence de l'instance à une personnalité indépendante peut être un bon moyen pour concilier la distanciation et la prise en compte des préoccupations du commanditaire.
6. Enfin, par définition, le principe de distanciation ne s'applique pas aux auto-évaluations. Toutefois, une auto-évaluation peut chercher par divers moyens à construire un point de vue objectif sur son objet. Elle peut même constituer un élément indispensable s'il est intégré dans une perspective plus large²⁰.

➤ **Principe de compétence**

1. La question de la compétence est à mettre en relation avec les enjeux de professionnalisation et de progrès méthodologique. Une question se pose de manière récurrente : faut-il privilégier la compétence sectorielle (connaissance minimale de la politique évaluée et de son contexte), les compétences techniques (techniques d'enquête) ou les compétences méthodologiques relatives à la conduite d'un projet d'évaluation ? Etant entendu que les compétences scientifiques, de méthode et de pilotage sont, en tous cas, indispensables, la réponse ne peut être que contextuelle. Dans certains cas, le « coût d'entrée » serait trop élevé pour un évaluateur dénué de culture « sectorielle ». Dans d'autres, l'extériorité du regard d'un non spécialiste peut être un atout.
2. Une autre question abordée dans les débats est celle de la diversité des approches méthodologiques et du rôle de l'innovation. Il est parfois reproché à la formulation

¹⁹ Ce problème est également soulevé par la pratique des évaluations commanditées par la Commission européenne, qui, en outre, tendaient encore récemment, de façon majoritaire, à établir une relation indue entre l'indépendance et le caractère privé du consultant, alors que la concurrence sur le marché ne répond pas à cette question de façon satisfaisante.

²⁰ Ainsi, les auto-évaluations, produites par les plans locaux d'insertion (PLIE) dans les années 90, furent-elles utilisées par l'évaluation d'ensemble commanditée conjointement par le ministère du travail et la Commission européenne (DG emploi et affaires sociales).

actuelle du principe de compétence de privilégier les standards méthodologiques internationalement reconnus. Toutefois, comme l'a fait remarquer une intervenante à la table-ronde de juin 2004, s'y référer n'interdit pas de s'en écarter. Ainsi, dans le cas des politiques ciblées sur des personnes (éducation, emploi...) ou des entreprises, il est pertinent de se poser systématiquement la question de l'expérimentation avec groupe de contrôle, même si les contraintes pratiques conduisent, le plus souvent, à écarter cette méthode.

3. Reste enfin la question : qui apprécie la compétence ? Elle renvoie à la création de diplômes et de qualifications reconnues en évaluation, voire à la certification des évaluateurs. Il ne semble pas que l'on puisse avancer beaucoup sur ces sujets actuellement compte tenu du faible développement des formations spécialisées. L'expérience montre cependant que des évaluateurs de profils variés peuvent se montrer compétents et s'approprier par l'usage les principes de la démarche évaluative.

➤ Principe du respect des personnes

1. Ce principe, qui s'applique d'ailleurs à toute recherche en sciences sociales, revêt un caractère méthodologique autant qu'éthique : pour ne pas se mettre en défaut et compromettre l'idée même d'évaluation, l'évaluateur doit respecter et protéger ses informateurs. D'où la nécessité de rendre anonymes les citations, nécessité rappelée dans une contribution de l'atelier de Bordeaux, sur la base d'un principe largement admis qui ne va pas, parfois, sans conflit avec le principe de transparence.
2. Il n'en demeure pas moins qu'un évaluateur peut se trouver confronté à des dysfonctionnements administratifs difficiles à passer sous silence, surtout s'il est membre d'un corps d'inspection.
3. La question du respect de l'intégrité du sens que les personnes interrogées donnent à leurs propos a quant à elle été soulevée par un autre contributeur à l'atelier de Bordeaux. Cette question est à rapprocher d'une recommandation issue d'un guide méthodologique australien : « *l'information fournie par les personnes interrogées est utilisée en cohérence avec les représentations (understandings) sur lesquelles elle s'appuie* ». Or, le problème est d'autant plus aigu que l'on paraphrase les personnes interrogées au lieu de les citer. Ce point peut cependant relever davantage de la problématique de la qualité du rapport d'évaluation que de celle du respect des personnes.

➤ Principe de transparence

1. Deux aspects peuvent être distingués : la transparence dans la diffusion des rapports et dans l'explicitation de la méthode. La question est le plus souvent abordée sous l'angle de la publicité des rapports. L'impression qui se dégage est que les cas de censure sont plutôt rares et que la diffusion gagne du terrain. En revanche, les problèmes de communication, d'accessibilité et de médiatisation demeurent importants²¹.

²¹ On notera que la généralisation de l'usage des sites web va dans le sens d'une transparence de diffusion qui s'accroît ; un exemple parmi d'autres : alors que l'évaluation du contrat de progrès de l'ANPE existe depuis son premier exercice en 1990, le premier rapport publié sur le site du ministère du travail ne l'a été qu'en décembre 2004.

2. En matière de communication, des principes doivent être fixés au départ : une fois l'évaluation lancée, l'initiative de la communication appartient au commanditaire. Le cadre fixé à l'évaluation doit garantir son caractère public : l'évaluateur ne peut porter sur ses seules épaules l'exigence de transparence. Il est là pour répondre à une question et pour aider le commanditaire à former son jugement. Force est de constater que la transmission des rapports se fait généralement davantage au profit des membres de l'administration qu'à celui des élus. Si l'on veut que l'évaluation serve au débat démocratique, une plus ample diffusion des rapports d'évaluation doit être recherchée.
3. La question de la transparence inclut également celle de la clarté et de l'intégrité des rapports d'évaluation : certaines des informations portées à la connaissance des évaluateurs peuvent être occultées, délibérément ou non. Cette occultation est parfois justifiée par le principe de respect des personnes évoqué précédemment.
4. L'exigence de transparence se heurte au caractère stratégique du jeu des acteurs. On ne peut exiger des acteurs qu'ils explicitent totalement leurs objectifs et leurs critères de jugement (agenda caché...). En revanche, on peut exiger d'une évaluation qu'elle clarifie ses propres limites en explicitant les règles du jeu dans lesquelles elle se trouve cadrée. Il est donc important de rappeler que la transparence implique l'explicitation des méthodes et des critères d'évaluation.

➤ **Principe de responsabilité**

Ce principe est très général et difficile à confronter à des observations précises. Il renvoie, pour partie, à l'agencement du dispositif institutionnel d'évaluation et, au-delà, à l'ensemble du processus de *policy making*. Il est de la responsabilité du « management » de la fonction évaluative de créer les conditions d'une évaluation utile, rigoureuse et éthique. Comme l'a noté l'un des participants à l'atelier de Bordeaux, la réflexion évaluative, en amont de la mise en œuvre des politiques publiques, est de nature à favoriser la clarification de la commande, la reconnaissance de l'autonomie du processus évaluatif et une bonne articulation entre son pilotage politique et son pilotage technique.

➤ **Principe d'utilité et d'opportunité**

1. La discussion de l'atelier de Bordeaux a fait émerger la proposition d'ajouter, à la Charte, un principe supplémentaire, qui pourrait être formulé comme « principe d'opportunité ». L'évaluation doit être justifiée par le fait qu'elle est susceptible de produire du changement. On observera, toutefois, que ce principe est implicitement contenu dans le préambule de la Charte.
2. La formulation d'un principe d'opportunité devrait tenir compte d'une tension fondamentale entre, d'une part, l'intérêt de systématiser l'évaluation dans une perspective de « redevabilité » (*accountability*) démocratique et, d'autre part, le risque de voir l'évaluation centrée sur la mise en œuvre et la performance au détriment d'une réflexion plus approfondie sur la pertinence des politiques.

➤ **Un débat à poursuivre**

Comme on peut le constater, le débat est loin d'être clos. Toutefois, il a déjà permis de cerner certaines difficultés et de suggérer certaines améliorations. Ainsi, le préambule de la Charte devra mieux préciser que le fait de promouvoir certains principes ne signifie pas qu'ils doivent être appliqués de manière dogmatique dans tous les contextes. D'une manière générale, la réflexion à venir devra chercher à lever les ambiguïtés sur l'interprétation des principes et apporter davantage d'éléments pratiques sur leur mise en œuvre dans différentes situations. Ces points feront l'objet de prochaines discussions, notamment, à l'occasion des Assemblées Générales de la SFE de 2005 et 2006.

Une perspective européenne et une vue suisse sur des standards d'évaluation

Dr Thomas Widmer, Département de science politique, Université de Zurich

Communication présentée à l'Université Victor Segalen de Bordeaux 2, les 20 et 21 septembre 2004, lors des 6èmes Journées Françaises de la Société Française de l'Évaluation²².

Avertissement

Je dois d'abord mentionner que, lorsque je parle de standards, sauf mention contraire, ce terme correspond à des références telles que principes, guide, normes, lignes directrices ou charte.

Je précise aussi que je ne suis mandaté ni par la Société Européenne de l'Évaluation (SEE) ni par la Société Suisse d'évaluation (SEVAL). J'exprime ici mon opinion personnelle.

1. Standards, guides, principes et chartes en Europe

En Europe, pendant longtemps, les standards d'évaluation développés aux Etats-Unis ne furent souvent pas pris en considération. Durant les années 1970 et 1980, les références aux standards d'évaluation étaient très rares. Vers la fin des années 1980, et davantage pendant les années 1990, cette situation a radicalement changé.

Pour illustrer ce développement en Europe, je vais présenter la situation dans trois pays européens - la Suisse, le Royaume-Uni et l'Angleterre - et pour l'Union Européenne. J'ai exclu la France, car je pense ne pas être dans la meilleure position pour en parler.

➤ La Suisse

Je vais commencer par la Suisse. Ma nationalité n'en est pas la raison. La communauté des évaluateurs en Suisse a été l'une des premières, en Europe, à avoir ses propres standards.

La pratique de l'évaluation en Suisse n'est pas très ancienne. A la fin des années 1980, une commission fédérale avait travaillé sur ce thème. Ses travaux et un Programme National de Recherche, consacré à l'effectivité des mesures étatiques, incitent à un développement rapide de l'évaluation en Suisse. En résulte la création de la Société suisse d'évaluation (SEVAL) en 1996.

Dans le cadre du Programme de Recherche mentionné ci-dessus, deux projets ont été dédiés à la méta-évaluation, c'est-à-dire à l'évaluation des évaluations. Dans le cadre de ces études mesurant la qualité des évaluations, les standards des Etats-Unis ont été appliqués avec un bon succès.

Après son établissement, le comité directeur de la SEVAL a décidé d'instaurer un groupe de travail pour discuter des standards et, éventuellement, formuler des standards suisses. Comme j'avais été responsable de ces études méta-évaluatives, le comité m'a délégué cette tâche. Le groupe de travail comprenait à peu près quinze membres, représentant des bureaux privés, l'administration publique cantonale et fédérale et des universités.

²² Version française revue par René Padieu.

Tous avaient une expérience d'évaluation, soit comme évaluateurs soit comme commanditaires. Après une discussion approfondie des standards, guides, principes et codes existants, le groupe a décidé de développer des standards propres à la SEVAL, considérant les spécificités de la situation en Suisse. Au lieu de repartir de zéro, nous avons décidé de partir des standards américains. Un premier projet a été discuté avec les membres de la SEVAL, à son assemblée générale au printemps 1999. Pour susciter la discussion sur les standards, un congrès additionnel a été organisé, au printemps 2000, dédié exclusivement à la qualité de l'évaluation et aux standards. Avec des petites adaptations, le projet a été adopté comme « Standards de la SEVAL » par l'assemblée générale du printemps 2001. Ces standards SEVAL existent en français et en allemand. Une traduction en anglais est également disponible.²³

Les standards SEVAL s'inspirent des «Program Evaluation Standards» du «Joint Committee on Standards for Educational Evaluation». Ils comportent 27 principes, répartis en quatre groupes : utilité, faisabilité, déontologie et précision. Parmi les changements effectués, on a réuni certains standards, reformulé plusieurs, supprimé deux sans remplacement et introduit un standard tout nouveau. Ces différents changements sont expliqués et motivés dans les commentaires annexés aux standards SEVAL. Mais de façon générale, le document est très proche des standards américains.

➤ *L'Allemagne*

Inspirée par les discussions sur les standards suisses et par la parution d'une traduction allemande des standards américains en 1999, la Société allemande d'évaluation (DeGEval), fondée en 1997, a entrepris en 2000 d'établir des standards d'évaluation. Après une enquête auprès des membres de DeGEval, l'assemblée générale a établi une commission pour formuler des standards propres sur la base des standards américains et de ceux de SEVAL. La commission comprenait sept évaluateurs et deux commanditaires. Le projet était soumis à treize commentateurs. En fin d'automne 2001, l'assemblée générale a formellement adopté la proposition de la commission comme "Standards DeGEval".

Les standards DeGEval se composent de 26 principes. Le document (50 pages) est disponible en allemand. Il comprend aussi un sommaire en anglais, mais, à ma connaissance, pas de version française. Comme ceux de la SEVAL, les standards DeGEval ressemblent fortement aux standards américains. Par rapport aux standards SEVAL, deux principes sont fondus, une grande part sont reformulés et regroupés autrement.

➤ *Le Royaume-Uni*

En Angleterre, la situation est très différente. La société d'évaluation du Royaume-Uni (UKES), fondée en 1994, était la première en Europe. Malgré une tradition d'évaluation relativement longue, UKES n'a pas, durant longtemps, établi de standards. Enfin, en 2003, il a été décidé, sans prendre les standards américains comme point de départ, de formuler des lignes directrices pour une bonne pratique en évaluation. Cela, bien sûr, en connaissant les autres documents existants et dans une volonté explicite de ne pas suivre les exemples américain, suisse ou allemand.

Au contraire des standards mentionnés, ces lignes directrices sont divisées en quatre groupes s'adressant à un groupe de personnes : séparément pour les évaluateurs, les commanditaires et pour les participants à une évaluation, ainsi qu'un groupe spécifique pour les autoévaluations.

²³ L'auteur a distribué la version française. Des copies additionnelles sont disponibles sur le site www.seval.ch.

➤ *L'Union Européenne*

Après plusieurs documents en matière d'évaluation, la DG Budget de la Commission Européenne a élaboré un texte, en collaboration avec plusieurs agences de standards et bonnes pratiques, afin d'uniformiser les activités d'évaluation au sein des différents services. Il a été décidé pour ce faire de ne pas suivre les publications existantes, et de rechercher un document qui soit très proche des pratiques d'évaluation à la Commission.

Ce document établit des standards (contraignants) et propose des bonnes pratiques (recommandées) couvrant les points suivants :

- les profils et les rôles, les tâches et les ressources de l'évaluation ;
- la gestion de l'évaluation, y compris sa planification, la rétroaction et la communication externe des résultats ;
- le processus d'évaluation, incluant le projet d'évaluation, le groupe de pilotage et la conduite de l'évaluation ;
- la qualité des rapports.

2. Politique de la Société Européenne d'Évaluation concernant les standards

La Société Européenne d'Évaluation (SEE) a décidé de contribuer aux discussions des sociétés d'évaluation nationales. Comme première étape, elle a analysé les pratiques existantes chez les sociétés d'évaluation et autres organisations intéressées. La situation comparée est présentée au tableau 1.

Tableau 1 : comparaison des standards d'évaluation en Europe

Caractéristiques	Standards SEVAL	Standards DeGEval	Charte SFE	Lignes directrices RU	EC Standards & Bonnes Pratiques
Champ d'application	Suisse	Allemagne et Autriche	France	Royaume-Uni	Europe
Organisation	société nationale d'évaluation	société nationale d'évaluation	société nationale d'évaluation	société nationale d'évaluation	Commission Européenne
Objet	évaluation (processus et résultat)	évaluation (processus et résultat)	évaluation (processus et résultat)	acteurs participants à une évaluation	agences de la CE
Destinataires	acteurs participants à une évaluation	acteurs participants à une évaluation	membres de la SFE	acteurs participants à une évaluation	agences de la CE
Date d'adoption	2000 (1999)	2001	2003	2003	2002
Types d'évaluation	tous, sans évaluations des personnes	tous, sans évaluation des personnes ni autoévaluations	tous, sans évaluation des personnes	tous	sans évaluation des projets individuels dans programmes
Fonction	assurer / promouvoir qualité en évaluation	protection / développement qualité d'évaluation et incitation du dialogue	expression des valeurs partagées par membres SFE	incitation / référence pour discussions déontologie et pratiques.	améliorer la cohérence et la qualité de l'évaluation chez la CE
Nature	décrire meilleures pratiques	décrire meilleures pratiques	établir des valeurs communes	informer sur des bonnes pratiques	standards minimaux et bonnes pratiques
Obligation	pas obligatoire	pas obligatoire	pas obligatoire	pas obligatoire	partiellement obligatoire
Sujets traités	utilité (8)	utilité (8)	pluralité	lignes directrices pour:	profil, rôle, tâches et ressources (8)
(nombre d'articles)	faisabilité (3)	faisabilité (3)	distanciation	- évaluateurs (19)	gestion d'évaluation (13)
	déontologie (6)	fair-play (5)	compétence	- mandants (18)	
	précision (10)	précision (9)	respect des personnes	- participants (7)	processus d'évaluation (12)
			transparence	- pour auto-évaluations (17)	qualité des rapports (6)
			responsabilité		*
Nombre de standards	27	25	6	61	39*
Structuration	Par sujets	Par sujets	Par sujets	Par rôles	Par processus
Accent de l'orientation	recherche	recherche	déontologie	délibération	gestion

* Standards seuls, sans les bonnes pratiques.

La SEE a discuté des standards, lignes directrices, charte, etc., une première fois, au congrès de Séville en 2002. Puis, le comité de la SEE a démarré un processus de dialogue entre les parties intéressées en Europe.

En novembre 2003, une discussion a eu lieu entre des experts du domaine et des membres du comité. Il est alors apparu qu'il n'était pas utile que la SEE établisse aussi ses propres standards, mais plutôt qu'elle formule une politique en ce domaine. En février 2004, le comité de la SEE a adopté cette politique.²⁴

De façon générale, cette politique se veut modeste. La SEE soutient les actions d'autres parties en matière de standards d'évaluation. Elle veut inciter à la discussion les personnes actives dans ce domaine dans les pays européens. La SEE ne prend pas position sur un ensemble de standards. Elle veut au contraire laisser faire les sociétés nationales et autres organisations. Elle affirme aussi son intention de continuer le débat, la prochaine occasion étant son congrès à Berlin en septembre/octobre 2004. Là, trois sessions ont été exclusivement réservées aux standards.

3. Application des standards SEVAL en Suisse

Les standards SEVAL ont été appliqués dans différents contextes et pour diverses raisons. Quelques exemples vont être présentés, avec, pour commencer, l'usage bien connu des standards pour la méta-évaluation.

➤ Standards et méta-évaluation

Une méta-évaluation est une évaluation qui a comme objet une évaluation. Le méta-évaluateur a pour tâche de déterminer la valeur, ou la qualité, d'une évaluation. Pour ce faire, il est nécessaire d'avoir des critères de qualité.

Tableau 2 : Les bases pour des critères à disposition

<i>Niveau</i>	<i>Objet</i>	<i>Base pour des critères internes</i>	<i>Base pour des critères externes</i>
Construction de premier niveau	programme, projet, mesure, etc.	objectifs du programme, du projet etc..	importance sociale du programme
Construction de deuxième niveau	<i>évaluation</i>	objectifs de l'évaluation	théorie et méthodologie d'évaluation
Construction de troisième niveau	méta-évaluation	objectifs de la méta-évaluation

Les standards – tels ceux de la SEVAL – peuvent être pris comme source pour la théorie et la méthodologie d'évaluation.

En Suisse, la pratique de méta-évaluation a commencé au milieu des années 1990. À cette époque, les standards SEVAL n'étaient pas encore parus, aussi a-t-on appliqué les standards américains. Aujourd'hui, on utilise normalement les standards SEVAL. Le tableau 3 présente des exemples de méta-évaluations.

²⁴ Le document, exprimant cette politique de la SEÉ, a été distribué lors de la conférence de Bordeaux. Il n'existe toutefois qu'en anglais.

Tableau 3 : Des exemples des méta-évaluations

Etude	Widmer 1996a-h	Widmer et al. 1996	Kuffner 2000	Nideröst 2001	Widmer et al. 2001	Lulofs et Arentsen 2001
Promoteur	Intérêt personnel	Fonds national suisse de la recherche	Intérêt personnel	Intérêt personnel	Chancellerie fédérale	Office fédéral de l'énergie
Contexte	Promotion de la recherche	Promotion de la recherche	Recherche académique	Recherche académique	Recherche contractuelle	Recherche contractuelle
Mandant	Fonds national suisse de la recherche	Fonds national suisse de la recherche	Auto-financement	Auto-financement	Chancellerie fédérale	Office fédéral de l'énergie
Nombre d'études évaluées	10	15	43	1	3	5
Critères de sélection	Variation maximale	Variation en ressources	Thème	Étude de cas	Variation maximale	Programme
Critères d'évaluation	Standards JC	Standards JC	Standards JC	Standards SEVAL	Standards SEVAL	Standards SEVAL
Méthode	Étude de cas comparé	Étude de cas comparé	Enquête	Étude de cas	Étude de cas comparé	Étude de cas comparé

Plusieurs autres méta-évaluations ont eu lieu ces dernières années. On peut trouver d'autres exemples sur le site de la SEVAL.

➤ *Conseils pour la pratique d'évaluations*

Tout évaluateur connaît des situations difficiles lors d'une évaluation. Le mandant, par exemple, a l'intention d'influencer les résultats ; ou bien il a l'impression que l'évaluateur n'est pas compétent. Sans prétendre à l'exhaustivité, les standards SEVAL offrent des réponses pour quelques unes des questions qui se posent alors.

La liste des standards, 27 au total, est un peu longue pour une vue d'ensemble. C'est pourquoi les standards SEVAL offrent aussi un aperçu fonctionnel. Cette liste pose quels sont les principes auxquels il convient de prêter une attention particulière lors des principales activités d'évaluation. Grâce à cet aperçu, il est possible de construire un instrument de travail très utile, une liste de contrôle (ou, en allemand, Checklist).

Le tableau 4 montre quels sont les standards importants selon les différentes phases de l'évaluation.

Tableau 4 : Les standards correspondant aux activités évaluatives

Standards	N 1	N 2	N 3	N 4	N 5	N 6	N 7	N 8	D 1	D 2	D 3	K 1	K 2	K 3	K 4	K 5	K 6	G 1	G 2	G 3	G 4	G 5	G 6	G 7	G 8	G 9	G 10
Action																											
Décision d'évaluer																											
Définition du problème																											
Planification de l'évaluation																											
Collecte d'information																											
Exploitation des informations																											
Rapport et communication																											
Budget de l'évaluation																											
Contrat de l'évaluation																											
Gestion de l'évaluation																											
Ressource personnelle																											

Tableau 5 : Checklist méta-évaluation 1: Décision sur l'exécution d'une évaluation

Standard	pertinent	utilisé	Commentaires
N1 Identification des parties prenantes et groupes concernés			
N2 Clarification des buts d'évaluation			
N3 Crédibilité			
N8 Effet de l'évaluation			
D2 Acceptabilité politique			
D3 Rapport couts//avantages			
K1 Convention formelle			
K6 Identification des conflits d'intérêt			
G1 Documentation de l'objet d'évaluation			
G2 Analyse du contexte			
G10 Méta-évaluation			

Mentionnons encore deux autres formes d'application des standards SEVAL : l'assurance de qualité et l'éducation en évaluation.

➤ *Assurance de qualité*

Pour l'assurance de qualité de l'évaluation, voici deux exemples :

1. L'Office fédéral de la santé publique (OFSP) met régulièrement au concours des mandats d'évaluation. Les réponses à l'appel d'offres sont jugées par un groupe d'experts, spécialistes du contenu de l'objet de l'évaluation, constitué par le Centre de compétences en évaluation (CCE) de l'OFSP. Les critères de jugement sont les standards de la SEVAL ainsi que les informations présentées dans le «Guide pour la planification de l'évaluation de projets ou de programmes de santé » de l'OFSP (Berne, 1997).
2. Le centre de compétences « audit de rentabilité et évaluation » (CC-EVAL), du Contrôle Fédéral des Finances, écrit dans une brève présentation du service : « Les analyses menées par le CC-EVAL respectent les standards de la SEVAL (Société suisse d'évaluation) ». Il est difficile de dire s'il s'agit seulement d'une déclaration symbolique.

➤ *Éducation en évaluation*

Comme les standards de la SEVAL représentent aussi une définition détaillée de l'évaluation en général, ils constituent un outil didactique de base. Avec eux, il est possible de communiquer les particularités de l'approche évaluative aux personnes ne connaissant pas le domaine.

Mais les standards peuvent être utilisés comme outil pédagogique, par exemple pour guider des groupes de travail dans leurs projets de formation mais aussi pour des personnes ayant des connaissances de base de l'évaluation ou des personnes confrontées régulièrement au concept d'évaluation dans leur domaine de travail. Les standards de la SEVAL sont aussi appropriés pour juger du bien fondé des études d'évaluation préexistantes, afin de développer chez les étudiants l'aptitude à une réflexion critique.

Les standards de la SEVAL sont régulièrement utilisés dans plusieurs institutions d'enseignement supérieur en Suisse – par exemple à l'Université de Berne, à l'Institut de hautes études en administration publique (IDHEAP) à Lausanne ou à l'Université de Zurich – mais aussi dans l'éducation continue dans différents contextes comme l'administration, l'école, le travail social, etc.

Conclusion

En général, les standards de la SEVAL apparaissent comme un outillage pratique ayant plusieurs fonctions dans des contextes différents. Cependant, par delà l'impression générale très favorable, des perspectives restent ouvertes pour améliorer la situation actuelle en Suisse. Le cadre d'application des standards de la SEVAL n'est pas encore « global ». Dans beaucoup de cas encore, on observe des actions évaluatives qui ignorent les standards. En outre, les matériaux accompagnant les standards peuvent être étendus et perfectionnés pour permettre encore plus de succès, et établir une culture de l'évaluation fondée sur une compréhension en commun du rôle de l'évaluation dans la société contemporaine.

Charte de l'évaluation : Les principes à l'épreuve des pratiques

Réunion débat

7 décembre 2004 à 16 h 30 à l'Hôtel de Ville de Paris

➤ *Une étape dans le débat sur la charte*

Après avoir adopté la « Charte » en octobre 2003, la Société Française de l'Évaluation a choisi de la mettre en débat en la soumettant à l'avis des praticiens, évaluateurs et commanditaires d'évaluations, qu'ils soient adhérents ou non de cette association. En 2004, le débat s'est déroulé en plusieurs étapes : table-ronde du 8 juin 2004 à l'université de Paris Dauphine, atelier des 20 et 21 septembre au Colloque de la SFE à Bordeaux et réunion-débat du 7 décembre à l'Hôtel de Ville de Paris. Les pages qui suivent rendent compte de cette dernière étape.

➤ *Etudes de cas et cas de conscience*

Après une introduction de Madame Stievenin, Maire-adjoint chargée de la solidarité, et de Jean-Claude Barbier, président du groupe de travail, la réunion a permis de débattre successivement de trois cas concrets mettant en jeu les principes de la charte. Ces cas ont été proposés par des membres de la SFE à partir d'expériences vécues²⁵. La réunion a été préparée et animée par Jacques Toulemonde.

Chaque cas a été formaté en vue d'une compréhension rapide par des non spécialistes, puis soumis à l'avance à plusieurs discutants, tous professionnels membres de la SFE. Lors de la réunion, ces cas ont été présentés et débattus selon la séquence suivante : lecture du cas, commentaires des discutants, commentaires de la personne ayant proposé le cas, commentaires de la salle et débat, brève synthèse par l'animateur.

Cette forme de discussion s'inspire d'une pratique de l'Association Américaine d'Évaluation²⁶ qui procède également par études de cas commentés par des professionnels, mais sans réunion ni débat.

1 - Évaluation ou expertise ?

➤ *Exposé*

La politique forestière met en œuvre des aides financières et fiscales aux propriétaires forestiers pour les inciter à mieux mettre en valeur leurs forêts, notamment par une réduction du morcellement de la propriété forestière et par un accroissement du prélèvement de bois, ressource renouvelable insuffisamment exploitée.

²⁵ Plusieurs autres cas ont été élaborés ou suggérés. L'exercice pourra donc se renouveler facilement à l'avenir.

²⁶ American Journal of Evaluation, 'Ethical challenges'

Le directeur du ministère de l'Agriculture en charge de cette politique, souhaite savoir si on pourrait faire d'avantage appel à l'outil fiscal dans le cadre de cette politique. Il considère qu'il manque d'analyse dans ce domaine et demande au Conseil général du génie rural, des Eaux et des Forêts (CGGREF), de réaliser un travail succinct.

En liaison avec la direction gestionnaire, le directeur concerné précise, dans une courte note, l'objet du travail demandé, à savoir : faire le point sur l'atteinte des objectifs visés et faire des propositions pour la suite. Une question clé est : les aides n'incitent-elles pas, en priorité, à de nouvelles plantations au détriment de l'objectif d'un prélèvement de bois suffisant ?

Les questions posées n'ont pas un caractère d'urgence. Un délai de 12 à 18 mois est considéré comme satisfaisant. La demande est acceptée et deux ingénieurs généraux prennent en charge le travail.

L'évaluation est conçue, dans un premier temps, comme relativement légère, avec une enquête auprès des principaux acteurs au niveau national et le recueil des principales statistiques disponibles (notamment à partir d'informations non exploitées des services fiscaux sur les abattements fiscaux dans ce domaine).

Au cours de leur travail, les ingénieurs généraux relèvent :

- un écart très important entre les objectifs affichés et les résultats,
- une faiblesse notable des moyens affectés à l'objectif prioritaire d'augmentation du prélèvement de la ressource de bois.

Les ingénieurs généraux répondent aux questions qui leur ont été posées. Ils n'identifient pas d'autres points de vue qui créeraient des difficultés significatives pour la réalisation de leur travail. Si tel avait été le cas, les règles déontologiques des ingénieurs généraux leur auraient imposé de prendre en compte tous les points de vue pertinents.

Au moment de finaliser leur rapport, les ingénieurs généraux hésitent sur le titre à donner : évaluation ou expertise ?

➤ *Introduction au débat*

Ce cas met en jeu plusieurs parties de la charte et de ses attendus, et notamment le principe de pluralité : « *L'évaluation prend en compte de façon équilibrée les différents points de vue légitimes qui ont été exprimés sur l'action évaluée. Cette prise en compte de la pluralité des points de vue se traduit, chaque fois que possible, par l'association des différentes parties prenantes au processus d'évaluation, ou par tout autre moyen approprié* ».

➤ *Commentaire de M. RULETA (Bureau de l'Évaluation, Ministère des Affaires Étrangères)*

L'interrogation de savoir s'il s'agit d'une évaluation ou d'une expertise serait ainsi apparue à la fin de l'exercice. Ce qui laisse entendre, déjà, qu'une expertise ne serait pas une évaluation. Une évaluation est réalisée par des évaluateurs (non experts) et une expertise est du ressort des experts (qui seraient ou non des évaluateurs ?). Mais, sans entrer dans les controverses sémantiques, il est possible de s'en tenir au débat sur les méthodes. La question posée, « évaluation ou expertise », telle qu'elle est posée à l'issue d'une étude, analyse, enquête, pourrait laisser croire, d'abord, qu'il est possible de réaliser une évaluation sans le savoir. Sans l'avoir décidé préalablement.

A ce stade, de nouvelles sous-questions apparaissent.

1. Qui décide du statut d'une évaluation ? Est-ce le commanditaire ou l'expert ? Est-ce les parties prenantes ou une instance indépendante (la SFE par exemple) ? On voit

bien que, dans la situation actuelle, on arrive à une sorte de consensus avec un plateau de la balance qui penche nettement du côté de l'évaluation car elle exerce une certaine fascination. Mais est-ce une « vraie » évaluation ?

2. Quels sont les éléments nécessaires pour déterminer une évaluation ? Plusieurs caractéristiques sont souvent avancées (indépendance, compétence, rigueur, etc.), mais retenons celle de la pluralité qui est proposée dans l'étude de cas présentée. Effectivement, l'évaluation de politiques publiques relève d'une analyse sociétale composite justifiant la prise en compte de points de vue pluralistes. La pluralité est donc un élément indispensable d'une évaluation.
3. Mais comment s'exerce la pluralité ? S'agit-il, seulement, de prendre en compte les points de vue qui peuvent s'exprimer par les divers acteurs concernés (en laissant aux évaluateurs la possibilité d'effectuer la sélection des opinions, leur pondération, la hiérarchie des problématiques, sans aucun contrôle), ou s'agit-il de faire participer les protagonistes tout au long du déroulement de l'exercice, notamment au sein d'un comité de pilotage qui oriente, valide, décide, etc. La pluralité passive ou active ne peut donc se concrétiser toute seule et se déterminer ex post.
4. Par conséquent, poser la question, après sa réalisation, de savoir s'il s'est agi d'une évaluation ou d'une expertise ne paraît pas rationnel : une expertise ne peut pas devenir une évaluation si elle n'a pas été conçue comme telle dès son lancement, mais, sans doute, une évaluation peut devenir une simple expertise si elle perd en cours de route ses fondements princeps.

La juxtaposition entre expertise et évaluation conduit à s'interroger sur certaines dérives. Il est en effet essentiel de prendre en compte les attentes d'un groupe de population pour lequel, par exemple, une politique de dégrèvement fiscal est envisagée afin d'anticiper sa réaction et s'assurer de l'efficacité de la politique prévue. Il reste que les demandes, voire les exigences de ce groupe ne doivent pas guider les décisions des évaluateurs qui ne prévoient pas une politique, dans le cas proposé, en faveur des forestiers, mais en faveur de la forêt. Les intérêts particuliers des uns, à court terme, ne correspondent pas toujours à l'intérêt général de tous, à plus long terme.

Pour conclure sur ce point, et montrer que des évolutions contradictoires sont parfois en cours, le ministère des Affaires Étrangères, estimant que les évaluations "traditionnelles" (avec la panoplie habituelle : termes de référence, comité de pilotage, mise en concurrence, expert indépendant) sont trop longues, lourdes et parfois peu efficaces, a mis en place des « expertises évaluatives » en confiant des évaluations, dites "légères", à des experts publics en vue d'obtenir un rapport dans un délai réduit et pour un coût plus modeste. Il s'agit donc d'un PACS *ad hoc* entre expertise et évaluation.

➤ *Commentaires de M. BAUCHET (École Nationale de Santé Publique)*

La question qui se pose est, à partir des principes tels qu'ils sont exposés dans la charte établie par la SFE, de vérifier si le cas recouvre davantage le champ de l'évaluation ou celui de l'expertise ; l'idée n'étant, évidemment pas, d'apporter un jugement de valeur sur l'intérêt de l'une ou l'autre de ces deux démarches.

Pour le cas présent, deux options ont été discutées : l'une en faveur de l'évaluation, l'autre en faveur de l'expertise.

La première, en faveur de l'évaluation, tient à la nature du questionnement soumis aux ingénieurs généraux : « les aides n'incitent-elles pas en priorité à de nouvelles plantations au

détriment de l'objectif d'un prélèvement de bois suffisant ? ». Il s'agit alors de faire le point sur l'atteinte des objectifs et de faire des propositions. Ce questionnement renvoie bien aux trois dimensions dont la présence conjointe caractérise la démarche d'évaluation :

- une dimension cognitive : que s'est-il passé ?
- une dimension normative : la politique engagée est-elle efficace au regard des objectifs fixés ?
- une dimension instrumentale : quels ajustements doivent être opérés ?

La deuxième option penche en faveur de l'expertise, et ce, au regard du principe de pluralisme de la charte : « *L'évaluation prend en compte de façon équilibrée les différents points de vue légitimes qui ont été exprimés sur l'action évaluée. Cette prise en compte de la pluralité des points de vue se traduit –chaque fois que possible– par l'association des différentes parties prenantes au processus d'évaluation, ou par tout autre moyen approprié* ». En effet, dans le dispositif d'évaluation mis en place, il n'existe pas de structure (comité de pilotage) permettant l'expression des points de vue des différentes parties prenantes à cette politique. Seuls les gestionnaires et principaux acteurs au niveau national sont directement impliqués dans le processus. La présence d'autres acteurs concernés par cette politique (bénéficiaires, lobby...) aurait pu élargir le questionnement évaluatif et/ou réinterroger la méthodologie employée, voire ouvrir vers de nouvelles perspectives (recommandations...).

En conséquence, et au regard du non-respect de ce principe, ce travail relève davantage de l'expertise.

➤ *Commentaires de B. PERRET (Conseil Général des Ponts et Chaussées)*

L'intérêt du cas vient de ce qu'il présente une situation fréquente, notamment pour les corps d'inspection à qui on pose des questions d'évaluation, en l'occurrence : les objectifs ont-ils été atteints et a-t-on choisi les bons moyens pour les atteindre ?

Quid des principes de la charte ? Les membres de l'Inspection Générale ont un principe de neutralité et ne sont pas liés aux décideurs. De même, le principe de compétence ne pose pas trop de problèmes : qu'il y ait des limites techniques ne suffit pas à disqualifier l'exercice. De même encore, les inspecteurs sont tenus au principe de respect des personnes et de responsabilité.

Le principe de pluralité suppose la capacité à prendre en compte les différents points de vue. Il implique aussi un dispositif de pilotage pour que la pluralité contribue aux choix orientant l'opération. Ainsi, au Conseil Général des Ponts, on n'appelle « évaluation » que ce qui a un comité de pilotage comprenant des personnes extérieures à l'administration. Dans ces conditions, le cas examiné ici n'est pas une évaluation.

➤ *Commentaires de D. GAGEY (Conseil Général du Génie Rural, des Eaux et des Forêts)*

Le respect du pluralisme se constate a posteriori. Les gens ont-ils le sentiment que leur point de vue a été pris en compte ?

L'exercice était bien une évaluation si l'on se réfère à la définition des années 90 : « rechercher si les moyens mis en œuvre permettent d'atteindre les effets visés ».

Le principe de distanciation a été respecté dans la mesure où l'un des deux inspecteurs n'avait pas été impliqué dans la politique concernée, ce qui aidait à prendre de la distance.

➤ *Débat général*

- La réponse était dans la question : « les aides n'incitent-elles pas ... ? ». L'orientation prise ne poussait pas à aller jusqu'au pluralisme et à s'intéresser au point de vue des bénéficiaires. D'accord, cependant, pour considérer l'exercice comme une évaluation.
- D'accord, au contraire, avec B. Perret : c'est un rapport d'expertise et non une évaluation, car il n'y a pas d'expression des parties prenantes. Les bénéficiaires ont peut-être été consultés mais n'ont pas participé au processus. Il existe une suspicion de manipulation envers les enquêtes, comme envers les statistiques. La confrontation des points de vue permet de lever cette suspicion, ce qui n'est pas le cas pour un rapport d'expertise. Personne ne détient la vérité, chacun en détient un bout.
- Il faut retenir l'élaboration conjointe du jugement évaluatif par toutes les parties prenantes, en particulier les destinataires de la politique mais aussi ses concepteurs, lesquels souvent se retirent de l'évaluation en attendant le rapport. Il y a délégation du regard à l'expert. Les "politiques" préfèrent déléguer les évaluations à des tiers, mais il faudrait les impliquer dans l'élaboration du jugement.
- Il est important de savoir qui sont les parties prenantes, y compris au-delà des destinataires de la politique, par exemple les lobbies... Associer toutes les parties prenantes pose un problème de faisabilité : une autre voie est de commencer avec un dispositif plus fermé, mais d'avoir ensuite une mise en débat et de réfléchir ensemble au jugement.
- La question « expertise ou évaluation » peut varier selon le moment au cours de l'opération. Ainsi, on peut démarrer, en essayant d'associer, pour finir par ne pas arriver à dépasser la simple expertise. Par exemple, lors des travaux de la Région Ile-de-France sur la vie étudiante, des enquêtes ont été conduites auprès des bénéficiaires directs et ont fourni des statistiques et des commentaires explicatifs très utiles, mais on n'a pas pu passer au stade de l'évaluation par-delà le diagnostic partagé, car la volonté politique des deux commanditaires a fait défaut au moment où il aurait fallu aller du diagnostic au questionnement évaluatif.
- Il peut y avoir un pluralisme du questionnement et/ou du jugement.
- Le pluralisme est toujours relatif. On n'a jamais tous les acteurs. On ne pose pas un principe de représentativité : seulement avoir la garantie de ne pas rester confiné à un point de vue unique. D'autre part, privilégier l'évaluation avec comité pluraliste ne signifie pas qu'il faille toujours faire comme cela. Nombreux sont les cas où il est sage de se limiter, de rester en deçà de l'évaluation pluraliste (exemple d'expertise sur la politique d'accueil des gens du voyage.)
- D'accord avec cela, mais subsiste la question du choix du terme "évaluation" plutôt que "expertise" ou "contrôle". La loi de 2002 sur l'action médico-sociale évoque 4 ou 5 versions de l'évaluation, incluant l'expertise et le suivi.
- Dans le cas présenté, on parle « d'enquête auprès des principaux acteurs nationaux ». Il n'y en a donc pas de locaux : destinataires, administrations, voisins... Le texte précise aussi que les inspecteurs « n'identifient pas d'autres points de vue », ce qui signifie qu'ils ont vu le problème à un moment donné. L'important, avec le principe de pluralité, c'est d'avoir une démarche de travail dans la durée et de ne pas se limiter à un seul diagnostic.

- L'évaluation pose la question de l'intérieur versus l'extérieur. L'expertise et le contrôle appliquent des critères imposés. L'évaluation implique de construire ensemble les critères.
- L'évaluation se définit-elle selon la nature des questions posées ? Ou, selon le processus pour former le jugement ?
- Evaluation ou expertise ? L'important est de contribuer à éclairer la décision. Rappelons que le principe de pluralité a fait controverse au moment du vote de la charte. L'existence de la Charte est importante pour l'expert !

➤ *Synthèse*

Le principe de pluralité suppose d'introduire plusieurs points de vue dans le processus d'évaluation, par exemple les bénéficiaires, les lobbies, les concepteurs de la politique concernée, etc. Qui, alors, a la responsabilité de ne pas oublier un point de vue important ?

A quelle étape du processus d'évaluation le principe s'applique-t-il ? Pluralité du questionnement, de la supervision du travail et/ou du jugement ?

Le principe de pluralité doit-il toujours être mis en œuvre de la même façon, sachant que se posent des problèmes de faisabilité ? Par exemple, comment faire fonctionner un comité de pilotage ouvert ?

La charte définit l'évaluation comme un exercice qui « cherche à fonder des jugements sur la valeur des actions publiques », ce qui « met en jeu l'intérêt général » et motive le principe de pluralité. Celui qui choisit de dénommer un travail « évaluation » prend une responsabilité au nom de l'intérêt général et s'oblige à prendre en compte les principes qui en découlent. Au contraire, les principes de la charte ne s'appliquent pas à une expertise.

2 - Conflit entre évaluation et communication politique

➤ *Exposé*

Un grand ministère social lance un programme fortement « médiatisé » avec une échéance relativement courte qui coïncide avec les débuts d'une période électorale. Pour en accroître la crédibilité, le ministère annonce à la presse que le programme sera évalué.

Au moment prévu, l'évaluation est lancée et l'administration est chargée de la mettre en œuvre. Le service pilotant l'évaluation agit dans un cadre fortement contrôlé par le cabinet du ministre, qui intervient directement dans le questionnement évaluatif. La méthode choisie est rudimentaire mais robuste.

Les résultats de l'évaluation montrent que le programme a connu une efficacité limitée. Cette conclusion figure dans le rapport remis par le service responsable de l'évaluation à sa direction générale et, par celle-ci, au cabinet du ministre.

Au moment de présenter les résultats à la presse, ces autorités utilisent le contenu du rapport en le modifiant afin de présenter les effets du programme sous un jour plus favorable. La presse reprend alors les données qui lui sont communiquées.

Le responsable de l'évaluation pouvait-il prendre des précautions pour éviter d'en arriver là ? Si oui, lesquelles ?

➤ *Introduction au débat*

L'histoire date du début des années 90, à un moment où il n'existait ni société française de l'évaluation, ni charte. Déplaçons cette histoire à aujourd'hui et réfléchissons à l'application de la charte.

C'est le principe de transparence qui est principalement concerné : «*La présentation des résultats d'une évaluation s'accompagne d'un exposé clair de son objet, de ses finalités, de ses destinataires, des questions posées, des méthodes employées et de leurs limites, ainsi que des arguments et critères qui conduisent à ces résultats. La diffusion publique des résultats d'une évaluation est souhaitable. Les règles de diffusion des résultats sont établies dès le départ. L'intégrité des résultats doit être respectée, quels que soient les modalités ou les supports de diffusion retenus.*», mais d'autres principes pourraient aussi être convoqués.

➤ *Commentaires de L. TOUGARD (FASILD)*

Il s'agit d'une situation fréquente en matière d'évaluation des politiques publiques : les responsables de la politique font dire à l'évaluation ce qu'ils en attendaient ou en isolent les seuls éléments pouvant servir leur propos.

De mon point de vue, les principes de la charte sont suffisamment clairs et explicites pour éviter ces dérives si on se situe dans un exercice formel ou si on traite d'un cas d'école.

Je propose de les décliner les uns après les autres et d'interroger leur traduction en points de méthode opérationnelle dans la conduite de l'évaluation, de l'amont (la définition du mandat d'évaluation) à l'aval (la diffusion des résultats) :

- le principe de pluralité : « l'évaluation prend en compte, de façon équilibrée, les différents points de vue légitimes qui ont été exprimés sur l'action évaluée » ; les évaluateurs ont-ils suffisamment pris en compte le point de vue politique (dans son contexte) et suffisamment associé les politiques à l'élaboration du cahier des charges et aux instances d'évaluation ?
- le principe de transparence : « la présentation des résultats d'une évaluation s'accompagne d'un exposé clair... les supports de diffusion retenus » ; les politiques ont-ils été informés, dès le lancement de l'évaluation, de ces exigences de transparence ? Les règles de diffusion des résultats ont-elles été proposées par les évaluateurs avant même de commencer (principe essentiel à mes yeux) ?
- le principe de responsabilité : « la répartition des rôles entre les différents acteurs est établie dès le départ », notamment « la diffusion des résultats ». Les évaluateurs ont-ils été vigilants sur ce point ?

En fait, le problème posé, au-delà des exigences méthodologiques pour faire vivre les principes de la charte, est celui de la légitimité de cette charte : comment l'opposer à des ministres, à des élus ? Comment l'utiliser pour faire vivre le principe de responsabilité qui précise que « les personnes et institutions impliquées... sont conjointement responsables de la bonne application des principes énoncés dans la charte ».

Sans doute la réponse se situe-t-elle (comme précisé dans le préambule de la charte) dans l'élaboration de normes de qualité ; mais peut-être serait-il intéressant également que la SFE agisse pour enrichir le décret de janvier 1990 proposant une définition de l'évaluation des politiques publiques de façon à y intégrer les principes de la charte.

Au-delà de l'évocation des principes concernés, rappelons l'un des attendus « l'évaluation participe au débat démocratique » ; dans le cas présent, comme souvent, l'évaluation ré-interroge la conception que l'on a de la démocratie...

➤ *Commentaires de F. LEFEBVRE-NARE (EVALUA)*

Sachant qu'il s'agit ici d'une évaluation externe, réalisée par un consultant dont le rapport n'a pas été publié : à qui revient-il d'appliquer ou de faire appliquer la Charte ?

Pour ce qui est du rôle du consultant, il n'a, en première analyse, aucun pouvoir de rendre transparente l'évaluation, car son travail appartient entièrement au commanditaire et le consultant est tenu à la confidentialité.

Il peut toutefois limiter les risques de déformation des informations issues de l'évaluation, par exemple :

- en faisant lui-même la synthèse,
- en écrivant en langage brut (rien que les faits) pour laisser peu de place à la distorsion,
- en proposant d'être présent lors des présentations de résultats.

À l'image des cadres de l'industrie argentine, évoqués par d'Iribarne et al. dans "Le Tiers-Monde qui réussit", toutes les parties prenantes de l'évaluation sont attachées, en principe, à se conduire de façon éthique. L'intérêt d'une Charte est alors de dire *comment* faire.

Un exemple d'application, certes problématique, est la double disposition de la Charte selon laquelle " *La diffusion publique des résultats d'une évaluation est souhaitable. Les règles de diffusion des résultats sont établies dès le départ* ". Une application rigoureuse de cette double disposition conduit, comme on l'observe en mission d'assistance au maître d'ouvrage (AMO), à la décision, dès le départ, de ne pas diffuser ; tout en se laissant, à vrai dire ... la possibilité de le faire ultérieurement ! Solution peu satisfaisante au regard de la Charte elle-même.

➤ *Commentaires de M. RULETA (Bureau de l'Evaluation, Ministère des Affaires Etrangères)*

Le cas présenté permet d'abord de dresser le profil du commanditaire idéal. Le commanditaire est celui qui rendra l'évaluation possible, efficace et pertinente.

Le commanditaire idéal passe par trois phases : ex ante, in itinere et ex post. Il faut noter que ses mérites évoluent à chaque étape et peuvent paraître contradictoires. (1) Ex ante, avant la décision de lancement, il est convaincu de l'intérêt de réaliser des évaluations. Il connaît leurs qualités et se déclare prêt à prendre en compte les conclusions et les recommandations. Il est demandeur, motivé, proactif. (2) En cours de réalisation, il comprend les impératifs des évaluations, adapte son agenda en conséquence, met en place les crédits nécessaires, apporte les informations nécessaires et accepte les méthodes proposées respectant l'indépendance. (3) Enfin, dès que l'évaluation est terminée, le commanditaire reprend ses prérogatives, décide, donne des instructions, assure le suivi.

Personne ne connaît le commanditaire idéal. De fait, le commanditaire idéal ne peut exister si notre perception du commanditaire se limite aux seuls décideurs, avec un pouvoir sans contre-pouvoirs. Pour l'instant, les évaluations de politiques publiques sont commanditées par les seules instances de l'exécutif. Les pouvoirs législatifs (au niveau central ou décentralisé), et moins encore les corps sociaux, ne sont commanditaires d'évaluations de politiques publiques. Le Parlement s'est intéressé, par éclipse, aux évaluations, et le Conseil économique et social a produit un rapport.

On devrait plutôt s'étonner que des évaluations soient actuellement conduites sous la responsabilité d'un pouvoir exécutif qui, finalement, ne rend compte qu'à lui-même. Il existe bien sûr des effets croisés par la mise en place des structures de l'Union européenne, par la décentralisation, la privatisation de secteurs financiers et économiques (et même sociaux), de comparaisons internationales plus vives dans un contexte de mondialisation, etc.

L'évaluation a pris pied dans les secteurs « secondaires » sans remettre en cause les domaines régaliens de la puissance publique. Il pourrait en être autrement avec l'introduction de la LOLF, encore que l'évaluation puisse aussi se trouver instrumentalisée, manipulée, réduite aux acquêts.

Pour ces motifs, l'évaluation reste généralement la propriété des payeurs-commanditaires, sauf dans les marges de la puissance publique, dans les zones où les enjeux sont internationalisés. Notamment, dans le cas de l'aide publique au développement (et encore, car la Direction du Trésor ou l'Agence française de développement gardent précieusement secrets les résultats, aucun rapport n'étant publié).

On peut observer que dans le cas n° 2 comme le cas n° 1, pluralisme et transparence sont deux principes complémentaires. Un pluralisme ouvert à des parties représentant la société civile conduit à la transparence des procédures, à la publication des résultats, à la mise sur la place publique des recommandations suivies. Il faudrait, pour cela, que l'évaluation soit exigée. Mais peut-on faire boire un âne qui n'a pas soif ? (l'âne étant le citoyen, vous et moi).

Il faudrait faire de l'évaluation un « objet de désir ». On n'ose pas faire appel aux publicitaires, mais on évoque généralement les médias, en espérant qu'ils empoignent le sujet de l'évaluation pour en faire un enjeu de politique citoyenne. C'est oublier, un peu vite, que les médias répondent aux attentes du lecteur-auditeur-télespectateur (ou sont aux ordres des lobbies économiques) et ne sont certainement pas en charge de la seule éducation des citoyens. Sauf, peut être, les médias des structures non gouvernementales privées, associations de consommateurs et syndicats.

Comment donner de la publicité aux évaluations ? En impliquant les organisations de citoyens dans les exercices car espérer leur implication, en leur communiquant des rapports indigestes, reste encore illusoire. C'est donc un changement de culture administrative, politique, de société. La culture de l'évaluation ne peut germer et se reproduire comme les OGM (organismes génétiquement modifiés).

➤ *Commentaires de J-C. BARBIER (Centre d'Etudes de l'Emploi)*

Un point n'a pas encore été souligné : la relation avec la presse et l'utilisation du mot « évaluation » dans le débat politique. Le ministère avait refusé la procédure passant par le Conseil Scientifique de l'Évaluation et le rapport n'était pas publié mais le mot « évaluation » a été employé dans la présentation à la presse. Le cabinet du Ministre tenait au mot et utilisait même, avec la presse, le terme « évaluation scientifique ». Pour convaincre du sérieux, il a même fait augmenter l'échantillon plus que nécessaire. La mise en débat, via la presse, suffit-elle pour que ce soit une évaluation ?

➤ *Débat général*

L'intérêt de la Charte, même si elle ne résout pas tous les problèmes, est qu'on peut demander au commanditaire s'il accepte de la respecter. Il est possible de faire référence à la Charte dans le contrat.

La presse a un rôle important à jouer, si elle est critique. Elle a une capacité de curiosité. On peut aussi convaincre le politique que dire la vérité n'est pas forcément mauvais pour lui. Enfin, écrire, sur un problème délicat, implique une rédaction acceptable.

On peut faire l'analogie avec la statistique, domaine dans lequel les ministres ont souvent été tentés de trafiquer ou de ne pas publier. Pour parer à cette tendance, il existe un positionnement du système statistique de l'administration et de l'INSEE. Les statisticiens ont une charte, qu'ils ne mettent pas assez en avant. On peut aussi évoquer le rôle de Pénombre²⁷, association qui pointe les usages abusifs des chiffres.

Si j'ai à évaluer, j'ai un commanditaire. J'ai aussi un devoir de transparence envers les sujets associés à l'évaluation, ce qui implique restitution. Le rapport est remis au Comité de pilotage, au commanditaire ou ministre. Que le rapport soit publié, ou pas, est le problème de ce dernier. Cela m'échappe totalement. Le seul problème, pour moi, est lorsque le commanditaire vient me chercher pour accompagner sa communication. Jusqu'où dois-je l'aider, avec des "éléments de langage" ? Je comprends son désir, et je défends mon rapport.

D'accord pour la Charte, qui est une « colonne vertébrale ». Mais peut-on aller plus loin et délivrer des labels ? Nul n'est légitime et notamment pas la SFE. Il y aurait un risque d'étouffement si l'exigence était trop contraignante. Par exemple, pour l'analyse des cycles de vie en environnement, il existe des normes, notamment pour la communication. Si l'on vise à une communication publique, la transparence s'impose, mais pas si l'usage est interne.

Il advient aussi qu'on ait la visite d'une Inspection, de la Cour des Comptes et qu'elle demande : avez-vous fait une évaluation ? Alors, on court un risque si l'on a biaisé.

Est-il possible que les commanditaires choisissent une autre dénomination plutôt que de devoir appliquer les principes associés à l'évaluation ? Il y a peu de risque car le mot « évaluation » a, en lui-même, de la valeur ! Il a son histoire, son étymologie, son usage international. Notre raison d'être est de veiller à ce que cette valeur (qui va croissant) ne soit pas invoquée sans conditions. Les conditions d'usage du mot « évaluation » sont un enjeu considérable.

Il faudrait invoquer ici le principe d'indépendance, lequel n'est pas dans la Charte ! Mais il est dans celles de l'OCDE, du FMI, de la Banque Mondiale. Pour cette dernière, l'indépendance s'impose à tous les stades : programmation, termes de référence, "staffing", publication des rapports et discussion du rapport avec les services de la Banque. De plus, le patron du département d'évaluation n'est jamais réintégré dans les services de la Banque. Cela ressemble au statut de la Cour des Comptes ou de l'IGF, qu'on n'a jamais vu se censurer. Indépendance entre en conflit avec transparence ; combien de fois entend-on dire "on peut tout se dire entre nous".

Si la Charte avait existé, à l'époque du cas en discussion, il aurait été plus facile de refuser de signer le rapport biaisé.

D'accord sur l'usage du mot « évaluation », même s'il n'y a personne pour labelliser. La solution est dans la construction du système pour arriver à une sorte de "label collectif", mais nous en sommes encore loin.

²⁷ L'association Pénombre, créée en 1993, se donne pour objet "l'usage du nombre dans le débat public".

Elle compte quelque 500 membres : statisticiens, chercheurs, enseignants, professionnels divers (magistrats, journalistes, ...) mais aussi citoyens concernés par le sujet. Elle publie une "Lettre blanche" trimestrielle, critique et humoristique, et une "Lettre grise" thématique aperiodique. Elle organise aussi des "nocturnes" : débats publics sur des thèmes d'actualité (immigration, évaluation des lycées, réforme des retraites, élections, LOLF, etc.). <http://www.penombre.org>

On débouche aussi sur l'idée (accompagnant la LOLF) que tout doit être évalué. Il y a donc intérêt à croiser le travail sur les standards et la réflexion sur la LOLF.

Pour éviter les biais dans la presse, nous avons Internet à notre disposition. Il faut aussi que les évaluateurs rédigent des synthèses accessibles à tous, très communicantes !

➤ *Synthèse*

Le mot « évaluation » ayant de la force, il est possible d'utiliser cette force pour faire de la Charte un instrument dans les jeux de pouvoirs et contre-pouvoirs. Il s'agit d'un jeu qui dépasse le couple commanditaire-évaluateur. La presse a aussi un rôle ainsi que les institutions de contrôle.

Diverses recommandations pratiques ressortent du débat : s'obliger à annoncer, dans un mandat, si l'évaluation sera publiée ou pas, assurer la diffusion par un service indépendant, assurer la présence de l'évaluateur aux moments forts de la diffusion, s'efforcer à une rédaction factuelle des rapport.

D'autres professions arrivent à faire prévaloir leurs principes (statisticiens, environnement...). La SFE devrait aller les voir.

3 - L'évaluateur découvre une faute professionnelle

➤ *Exposé*

L'inspection des services d'une direction, au sein d'un ministère, est chargée par le directeur d'administration centrale de conduire l'évaluation d'un programme expérimental concernant les établissements accueillant des publics en grande difficulté.

Le chef de l'inspection constitue une équipe d'évaluation pour réaliser le travail. Il s'agit de plusieurs inspecteurs et d'autres professionnels du ministère, choisis sur la base de leurs compétences et non de la représentativité des services et bureaux. Il est convenu que le chef d'inspection rendra compte des conclusions de l'équipe d'évaluation et communiquera son rapport au comité de pilotage du programme évalué.

Les inspecteurs en charge de l'évaluation préparent leur démarche méthodologique et les outils spécifiques à mettre en œuvre. Ils réfléchissent aux risques (internes et externes) de confusion possible entre évaluation et contrôle, et travaillent à les limiter en choisissant une posture et des concepts différents par rapport aux missions habituelles de l'inspection.

Après validation de la démarche et des outils par le comité de pilotage du programme, l'équipe d'évaluation entre en contact avec les opérateurs engagés dans la réalisation concrète du programme, notamment en effectuant plusieurs déplacements sur site. L'équipe informe largement sur l'évaluation qui commence, sa démarche et ses outils, les principes déontologiques applicables, les modes de recueil et de traitement de l'information, la destination du rapport et le fait que le commanditaire aura la maîtrise de la communication du rapport. A ce stade, les opérateurs chargés de l'exécution du programme ont donc eu connaissance de ce que plusieurs membres de l'équipe d'évaluation appartiennent à l'inspection des services.

A mi-parcours, lors d'un déplacement sur l'un des sites concernés, deux membres de l'équipe d'évaluation, dont un inspecteur, constatent des dysfonctionnements sérieux et une volonté délibérée de s'affranchir des normes établies par l'administration. Les écarts

constatés sont dus à la libre interprétation du cahier des charges, à l'absence de travail méthodologique dans la mise en place du projet. De nombreux incidents se sont produits, mettant les professionnels en grande difficulté.

Bien qu'ayant eu connaissance des dysfonctionnements, à l'occasion de l'évaluation et non pas d'un contrôle, l'inspecteur doit-il alerter le service gestionnaire du programme ?

➤ *Introduction au débat*

Un des principes de la charte susceptible d'être mobilisé dans ce cas est celui du respect des personnes : «*Les participants au processus d'évaluation respectent les droits, l'intégrité et la sécurité de toutes les parties concernées. Ils s'interdisent de révéler l'origine nominative des informations ou opinions recueillies, sauf accord des personnes concernées*». Cependant, on peut aussi invoquer le dernier des « attendus » de la charte, à savoir : «*Les enjeux de l'évaluation des politiques et programmes publics dépassent ceux de ses protagonistes directs et concernent l'ensemble des citoyens ; ces évaluations engagent l'intérêt général ; c'est pourquoi elle sont régies par quelques principes spécifiques* ».

➤ *Commentaires de F. LEFEBVRE-NARE (EVALUA)*

On a là un exemple de dysfonctionnement. Entre-t-il dans le champ de l'évaluation ? S'il s'agissait d'une évaluation externe, tout dépend du cahier des charges : soit le sujet est dans le champ, et on dit ce qu'on a vu en rendant les informations anonymes, soit il est hors champ.

Certes, "l'efficience" est presque toujours dans le champ. Mais à quelle aune le cahier des charges de l'évaluation demande-t-il de l'apprécier ? Au vu de normes professionnelles ? D'une réglementation ? Des résultats au regard des objectifs initiaux ? Des impacts au regard de la raison d'être de l'intervention publique ?

Dans d'autres exemples d'évaluation, il est apparu qu'un même fait était considéré comme un dysfonctionnement du point de vue administratif et comme une mise en œuvre efficiente du point de vue des acteurs de terrain. Il est frappant de voir jusqu'à quel point peuvent aller les différences de points de vue, de normes, donc de jugements.

Par exemple, l'existence de multiples procédures de concertation entre acteurs est généralement jugée *a priori* positivement par des personnes "du milieu", alors que d'un point de vue "d'économiste" c'est *a priori* une dépense de temps sans production de résultats.

L'application du principe de "distanciation" de l'évaluateur par rapport à ses propres normes de jugement sur ce qu'on observe, passe donc, notamment, par le respect de la référence constituée par le cahier des charges.

Pour revenir au cas, citer des exemples concrets dans les rapports est très souhaitable, mais délicat. La lecture d'un rapport d'évaluation peut être biaisée lorsqu'on donne un exemple, soit choquant, soit brillant : l'arbre peut cacher la forêt. Il faut être attentif à ce que les exemples reflètent les conclusions d'ensemble de l'évaluation.

➤ *Commentaires de L. TOUGARD (FASILD)*

La question du mélange des genres, entre évaluation et contrôle, est d'une grande actualité. Quand l'IGAS (corps d'inspection du ministère des affaires sociales) s'empare de l'évaluation, quand la Cour des comptes remet un rapport sur la politique d'intégration, on voit se profiler l'attribution de l'évaluation des politiques publiques à des corps de contrôle

et d'inspection. Comme l'ont montré les journées de la SFE de Bordeaux, la mise en place de la LOLF ne fait que renforcer cette situation, en donnant la priorité au contrôle de gestion.

Comment la charte constitue-t-elle un rempart contre ce genre de dérive ?

- principe de distanciation : « les personnes participant à l'évaluation informent les autres partenaires de tout conflit d'intérêt éventuel » ; les inspecteurs chargés de l'évaluation ont-ils, avant même de commencer l'évaluation, évoqué le risque de confusion des genres et l'éventualité d'un conflit de positionnement ? « le processus d'évaluation est conduit de façon autonome par rapport aux processus de gestion » : les inspecteurs n'ont-ils pas, d'emblée, imaginé le conflit entre leurs deux rôles, notamment du fait de leur responsabilité permanente de contrôle de la bonne gestion interne ?
- principe de respect des personnes : « ils s'interdisent de révéler l'origine nominative des informations, sauf accord des personnes concernées » ; cette clause ouvre la possibilité, exige même, d'en parler avec les personnes défaillantes dans la gestion du dispositif expérimental ; on peut alors imaginer que les questions ne soient débattues qu'à ce seul niveau.

Ces deux principes viennent toutefois se heurter à celui de la transparence qui rappelle que « l'intégrité des résultats doit être respectée ».

L'application serait-elle différente si aucun évaluateur n'avait également été inspecteur ?

Il est fréquent, voire habituel, qu'une évaluation révèle des dysfonctionnements internes relevant de la gestion interne, notamment lors de l'analyse de l'efficacité ; à l'occasion de cette analyse, l'évaluateur est amené à identifier les causes des écarts entre objectifs et réalisation ; bien souvent, certaines de ces causes relèvent d'une mauvaise gestion interne. La question est donc la même ; c'est pourquoi, d'ailleurs, l'évaluation fait si peur. Ce qui diffère, c'est le positionnement. Un évaluateur va énoncer les causes sans nommer les personnes ou les services, mais plus en termes de procédures ; un inspecteur interne va se poser la question de la sanction.

De mon point de vue, il n'y a pas à formuler de proposition pour améliorer la charte, mais plutôt à rappeler que contrôle et inspection diffèrent fondamentalement de l'évaluation tant dans leurs objectifs, que leurs méthodes et leurs responsabilités. Peut-être la charte devrait-elle rappeler que l'évaluation n'est pas le contrôle et que les responsabilités ne sont pas interchangeables.

➤ *Commentaires de V. HUBAULT*

Pour éclairer la suite du débat, il convient de préciser que le dysfonctionnement observé entrait bien dans le champ de l'évaluation. Il était grave (violences et départ de nombreux adolescents). D'une part, on ne pouvait attendre d'établir le rapport final. Et, même rendu anonyme, il aurait été repéré. De plus, le Comité de pilotage était présidé par le chef de l'Inspection et y siégeait le directeur dont l'établissement en cause relevait !

➤ *Débat général*

Mettre en évidence que la pratique de certains services est meilleure, cela fait partie de l'évaluation. Si une mauvaise pratique relève de la faute professionnelle, il faut le prendre en compte en le rendant anonyme.

Il est possible de demander à des inspecteurs de faire des évaluations sur des terrains autres que ceux où ils sont en position habituelle d'inspection. Il faut aider à construire un rapport plus exigeant entre évaluation et contrôle, comme l'ont fait, par exemple, les Pays-Bas et la Suède.

L'évaluateur ne doit dénoncer en aucun cas. C'est essentiel pour l'image de l'évaluation et pour la sincérité de ceux qui y sont impliqués. Il faut donner l'assurance que, quoi qu'on découvre, on ne le dénoncera pas.

Une évaluation peut être confiée à un inspecteur : il doit être capable de différencier son propos, selon qu'il mène une activité ou l'autre. En revanche, il peut rencontrer un problème d'image, si les évalués continuent de le percevoir comme contrôleur. Exemple d'une opération dans les transports routiers, où l'on s'était cru bien avisé en recourant aux inspecteurs des lois sociales parce qu'ils connaissaient le domaine : ils ont été très mal perçus.

Que faire d'une demande d'évaluation que l'on suspecte d'être une demande d'audit déguisée ? Face à une telle situation, nous avons vérifié qu'il y avait malentendu, et nous avons écrit explicitement qu'on ne ferait pas d'audit. On a expliqué ce qu'est une évaluation, sa finalité, sa visée d'intérêt général.

L'efficacité est toujours dans le champ de l'évaluation, à la différence de l'inspection, qui recherche la conformité. Par l'efficacité, un acteur se justifie très bien d'un écart à la règle.

Il y a souvent glissement de l'évaluation de programme à l'évaluation de la compétence professionnelle. C'est toujours de l'évaluation ; mais c'est très différent !

Si le dysfonctionnement ne relève pas d'une insuffisance professionnelle mais d'une faute pénale, alors toute personne impliquée dans l'évaluation est légalement tenue à signalement. (code pénal : quiconque a connaissance de faits ... est responsable pénalement.)

➤ *Synthèse*

Dans la mesure où il y a des choses qui relèvent du pénal, l'évaluateur a les devoirs de tout citoyen. Cet extrême étant balisé (c'est à dire hors pénal, qui n'est pas traité ici), on attend la non-dénonciation. C'est notamment une condition d'accès au terrain.

On voit des exemples dans lesquels l'application la plus scrupuleuse de la règle ne garantit pas la meilleure performance. Si ce n'était pas un inspecteur qui avait découvert le dysfonctionnement, est-ce que ça aurait changé quelque chose ?

Faut-il interdire l'évaluation aux inspecteurs et aux auditeurs ? Ou, plus simplement, leur demander de changer de posture quand ils pratiquent l'évaluation ? L'important est d'éviter la confusion des rôles dont un indicateur évident est la façon de traiter les dysfonctionnements.

4 - Apports du débat et nouvelles questions

Qui décide qu'un travail donné aura le statut d'une expertise ou d'une étude, ou, au contraire, le statut d'une évaluation, avec les devoirs que cela implique ? Il faut d'abord reconnaître que le terme « évaluation » exerce une certaine fascination et que la LOLF véhicule l'idée que tout doit être évalué. A contrario, est-il possible qu'un commanditaire choisisse une autre dénomination plutôt que de devoir appliquer les principes de la charte ? On n'en est pas encore là ! Celui qui choisit d'utiliser le terme « évaluation » est-il conscient

que son choix va créer des attentes en termes de jugement sur la valeur de l'action publique et de mise en jeu de l'intérêt général ? Si c'est le choix de l'évaluation qui est fait, est-on prêt à en assumer les implications ? Ces implications, peut-être encore imparfaitement traduites dans la charte, comportent les principes de pluralité, de transparence et de respect des personnes, qui ont été particulièrement discutés lors de la réunion débat.

➤ *Pluralité*

Le respect du principe de pluralité se mesure-t-il à l'implication des parties prenantes dans le processus d'évaluation ? C'est l'option choisie par le Conseil Général des Ponts, où l'on ne parle d'évaluation que s'il y a un comité de pilotage comprenant des personnes extérieures à l'administration. Au contraire, le respect du pluralisme ne devrait-il pas s'apprécier a posteriori, en vérifiant que les parties prenantes ont le sentiment que leur point de vue a été pris en compte, quelle que soit la manière dont on est parvenu à ce résultat ?

En ce qui concerne la mise en œuvre du principe de pluralisme, les participants au débat se posent de multiples questions pratiques. A quel moment le pluralisme doit-il s'appliquer : celui du questionnement, de la structuration du référentiel, de la formulation du jugement, de la diffusion des conclusions ? Quels points de vue doivent être pris en compte : ceux des destinataires de la politique, des lobbies, des décideurs ? Est-il toujours sage d'impliquer beaucoup d'acteurs dans un comité pluraliste ?

➤ *Transparence*

Une autre exigence de l'évaluation est la transparence. Jusqu'où doit-elle aller en ce qui concerne la diffusion des résultats ? " *La diffusion publique des résultats d'une évaluation est souhaitable. Les règles de diffusion des résultats sont établies dès le départ* ". Une application prudente de cette double disposition de la charte peut conduire à la décision, dès le départ, de ne pas diffuser ; tout en se laissant, à vrai dire, ... la possibilité de le faire ultérieurement si les résultats ne sont pas dérangeants ! Solution peu satisfaisante au regard de la charte elle-même.

Ce dilemme traduit le fait qu'il n'y a pas une demande évidente de transparence à l'heure actuelle. Les organisations de citoyens ne se jettent pas sur les rapports d'évaluation qui restent trop indigestes. On attend des médias qu'ils empoignent les résultats des évaluations pour en faire un enjeu de politique citoyenne, mais c'est oublier qu'ils répondent aux attentes de leur public (ou des lobbies qui les financent ?). Finalement, les responsables sont plus enclins à jouer la communication politique que la transparence.

Comment mettre en œuvre le principe de la charte dans ce contexte ?

- en accompagnant les rapports de synthèses accessibles à tous et très communicantes,
- en mettant en valeur les faits pour décourager la distorsion des conclusions,
- en utilisant Internet,
- en surveillant l'utilisation de l'évaluation et en réagissant aux abus,
- en s'inspirant des autres professions qui arrivent à faire prévaloir leurs principes (statisticiens, environnement...).

➤ *Respect des personnes*

Le principe de transparence entre-t-il en conflit avec celui du respect des personnes ? La question se pose dans le cas fréquent, voire habituel, d'une évaluation qui révèle des dysfonctionnements dans la mise en œuvre de l'action publique et qui est susceptible de mettre en cause des personnes. C'est une des raisons pour lesquelles l'évaluation fait peur. Cette situation pose la question du mélange des genres entre évaluation et contrôle/inspection, une question aussi vieille que l'évaluation elle-même, que la mise en place de la LOLF peut renforcer en mettant l'accent sur le contrôle de gestion.

Dans cette situation, les positionnements divergent. Un inspecteur doit envisager une sanction individuelle alors qu'un évaluateur va plutôt questionner l'efficacité des procédures sans nommer les personnes ni les services. L'application du principe de respect des personnes ne va cependant pas sans cas de conscience :

- Un point de vue est que l'évaluateur ne doit dénoncer en aucun cas. C'est essentiel pour l'image de l'évaluation et pour la sincérité des informations recueillies. Il faut donner l'assurance que, quoi qu'on découvre, on ne le dénoncera pas.
- Un autre point de vue est que si le dysfonctionnement va au-delà de l'insuffisance professionnelle et touche à la faute pénale, alors toute personne impliquée dans l'évaluation est légalement tenue à signalement (code pénal : « *quiconque a connaissance de faits ... est responsable pénalement* »).

➤ *Perspectives*

Comme le montre ce compte rendu, le débat a apporté au moins autant de nouvelles questions que de réponses. Il y a donc matière à le poursuivre sous des formes qui restent à trouver.

Le débat conduit-il à réexaminer les principes de la charte ou à mieux les promouvoir, ou aux deux à la fois ? Il donne certainement matière à argumenter toutes les options et le lecteur est invité à se faire sa propre opinion sur le sujet.

Finalement, le débat n'a approfondi que quelques uns des principes de la charte. Deux autres « cas de conscience » sont, d'ores et déjà, prêts à être soumis à un nouveau débat et il ne fait aucun doute que d'autres cas viendraient facilement s'y ajouter. Les perspectives sont donc bonnes pour renouveler cette expérience.

Charte de l'évaluation des politiques publiques et des programmes publics

*adoptée par l'assemblée générale de la SFE
le 28 octobre 2003*

• *Préambule*

La présente charte exprime les valeurs partagées par les membres de la Société Française de l'Évaluation (SFE). Elle marque une étape dans la réflexion de la SFE et devra être confrontée aux pratiques de ses membres en vue d'être complétée ultérieurement. Les utilisateurs de la charte pourront décliner ce texte de référence dans des documents adaptés à leurs besoins spécifiques, notamment dans le domaine des normes de qualité.

• *Attendus*

L'évaluation cherche à fonder des jugements sur la valeur des actions publiques.

Elle vise à donner plus de cohérence à l'action publique, à améliorer son efficacité et à mobiliser tous ses acteurs.

L'évaluation contribue à la prise de décision publique ; elle concourt à moderniser la gestion de l'administration et des services publics et à rendre plus efficace la dépense publique ; elle contribue simultanément au développement de la responsabilité et des pratiques de compte rendu interne ou externe, ainsi qu'à l'apprentissage organisationnel.

En cela, l'évaluation participe au débat démocratique à tous les niveaux de gouvernement.

Les enjeux de l'évaluation des politiques et programmes publics dépassent ceux de ses protagonistes directs et concernent l'ensemble des citoyens ; ces évaluations engagent l'intérêt général ; c'est pourquoi elle sont régies par quelques principes spécifiques :

Principes de référence

● Principe de pluralité

L'évaluation prend en compte de façon équilibrée les différents points de vue légitimes qui ont été exprimés sur l'action évaluée.

Cette prise en compte de la pluralité des points de vue se traduit -chaque fois que possible- par l'association des différentes parties prenantes au processus d'évaluation, ou par tout autre moyen approprié

● Principe de distanciation

L'évaluation est conduite de façon impartiale. Les personnes participant à l'évaluation à titre professionnel informent les autres partenaires de tout conflit d'intérêt éventuel.

Le processus d'évaluation est conduit de façon autonome par rapport aux processus de gestion et de décision. Cette autonomie préserve la liberté de choix des décideurs publics.

● Principe de compétence

Les personnes participant au processus d'évaluation à titre professionnel mettent en œuvre des compétences spécifiques en matière de conception et de conduite de l'évaluation, de qualité de la commande, de méthodes de collecte de données et d'interprétation des résultats. Elles ont le souci d'améliorer et de mettre à jour leurs compétences.

Elles s'appuient sur des méthodes reconnues, notamment celles en usage dans la communauté internationale de l'évaluation. Elles utilisent ces méthodes en explicitant leurs limites.

● Principe de respect des personnes

Les participants au processus d'évaluation respectent les droits, l'intégrité et la sécurité de toutes les parties concernées.

Ils s'interdisent de révéler l'origine nominative des informations ou opinions recueillies, sauf accord des personnes concernées.

● Principe de transparence

La présentation des résultats d'une évaluation s'accompagne d'un exposé clair de son objet, de ses finalités, de ses destinataires, des questions posées, des méthodes employées et de leurs limites, ainsi que des arguments et critères qui conduisent à ces résultats.

La diffusion publique des résultats d'une évaluation est souhaitable. Les règles de diffusion des résultats sont établies dès le départ. L'intégrité des résultats doit être respectée, quels que soient les modalités ou les supports de diffusion retenus.

● Principe de responsabilité

La répartition des rôles entre les différents acteurs de l'évaluation est établie dès le départ de façon à ce que toutes les fonctions de celle-ci soient bien prises en charge (définition du mandat, pilotage du processus, enquêtes et analyses, formulation du jugement et des recommandations, diffusion des résultats).

Les personnes et institutions impliquées dans le processus d'évaluation mobilisent les moyens appropriés et fournissent les informations nécessaires à la conduite de l'évaluation.

Elles sont conjointement responsables de la bonne application des principes énoncés dans cette charte.

Bibliographie sur les standards de l'évaluation

A l'origine cette bibliographie accompagnait l'article de Thomas Widmer, (cf. article p. 25) ; elle a été complétée et commentée par Manuela Loué alors en stage au Cabinet EVALUA.

Les ouvrages non disponibles en ligne ne sont pour la plupart pas commentés.

Barbier, J.C. "Devising and Using Evaluation Standards: The French Paradox" in Schwartz, R., and Mayne, J (eds), *Quality Matters: Seeking Confidence in Evaluation, Auditing and Performance Reporting*. New Brunswick : Transaction Publishers (2004), p. 21-40 : *Retour sur l'expérience du CSE / CNE et interrogation sur la capacité du système à garantir la qualité des évaluations : la France comme cas particulier à cause de son système administrativo-politique.* (version française adaptée dans Documents de travail CEE, n°23, Avril 2003).
<http://www.cee-recherche.fr/fr/publicationspdf/03doc23.pdf>

Barbier J.-C., et Perret B., (2000), "Ethical Guidelines, Process and Product Quality Standards, What For An SFE (French Evaluation Society) Perspective", paper presented at the European Evaluation Society Conference, Lausanne, October, 12th -14th.
http://www.europeanevaluation.org/docs/63_barbierperret.pdf

Beywl, W. "Standards for Evaluation: On the Way to Guiding Principles in German Evaluation." In C. Russon (ed.), *The Program Evaluation Standards in International Settings*. Kalamazoo : Evaluation Center, Western Michigan University, 2000 : *dans le cadre de la réflexion du groupe de travail de la DeGEval sur les standards de l'évaluation, W. Beywl démontre la nécessité d'adopter des lignes directrices. La deuxième partie définit les utilisations possibles de ces standards dont la profession a besoin, pour permettre aux évaluateurs comme aux non initiés de juger de la valeur d'une évaluation. Afin de préciser davantage ce que devront être les standards mis en place par la DeGEval, l'auteur compare, dans la troisième partie, ceux du Joint Committee, de l'American Evaluation Association et de l'Australasian Evaluation Society.*
<http://www.wmich.edu/evalctr/pubs/ops/ops17.pdf>

Beywl, W., and Taut, S. "Standards: Aktuelle Strategie zur Qualitätsentwicklung in der Evaluation." *Vierteljahrshefte zur Wirtschaftsforschung*, 2000, 69, 358-370 : *selon les auteurs, l'évaluation, située entre le "nouveau management public" et la démarche qualité, atteint ses objectifs dans la mesure où les citoyens y participent activement. Selon les auteurs, il est possible de partir des standards du Joint Committee pour améliorer la qualité des évaluations de politiques - sous réserve d'adaptations culturelles. Les possibilités d'intégrer ces standards dans des corps de règles, nationaux et internationaux, sont étudiées : les auteurs appellent à la conception de standards allemands et européens.*
http://www.diw.de/deutsch/produkte/publikationen/vierteljahrshefte/docs/papers/v_00_3_3.pdf

Bussmann, W. "Evaluation and Grassroots Politics: The Case of Switzerland." *Knowledge and Policy*, 1995, 8(3), 85-98.

Bussmann, W. "Democracy and Evaluation's Contribution to Negotiation, Empowerment and Information." *Evaluation*, 1996, 2(3), 307-319.

Conseil scientifique de l'évaluation. Petit guide de l'évaluation des politiques publiques. Paris: La Documentation française, 1996 : *ce guide rassemble approches, définition des objectifs, principes méthodologiques et présentation d'outils scientifiques concernant l'évaluation des politiques publiques. Le Conseil Scientifique de l'Evaluation (CES) s'interroge sur les critères de qualité d'une évaluation – notamment dans une sous-partie intitulée : « Les critères d'une bonne évaluation ». Le guide fournit donc un certain nombre de critères permettant d'en apprécier la méthode, le déroulement, ou le résultat.*

Deleau M. , Nioche J.P., Penz P., Poincard R. Evaluer les politiques publiques.Méthodes, déontologie, organisation. Rapport du groupe de travail du Commissariat général du Plan. Paris,mai 1986, La Documentation française, 181 pages.

Réflexion sur les instruments disponibles ou à créer pour développer l'évaluation des politiques publiques dans le contexte financier du 9^e plan,ce texte comprend l'examen des expériences étrangères et françaises (depuis le début des années 80),recense les instruments disponibles , analyse les conditions techniques de leur mise en œuvre et formule un certain nombre de recommandations pour faire une bonne évaluation de politiques publiques ; parmi celles-ci le rapport avance les éléments d'un code professionnel s'inspirant de règles suivies en matière de contrôle administratif comme de production scientifique.

European Commission Directorate General for Budget (ed.). Evaluating EU Activities: A Practical Guide for the Commission Services. Brussels: European Commission, 2003 : *ce guide pratique, en anglais, analyse la portée de l'évaluation au sein de la Commission européenne. Il donne des conseils pour concevoir, organiser et conduire une évaluation, pour rédiger le rapport final ou encore pour en exploiter les résultats. L'annexe A du document, qui reprend les standards censés encadrer les différentes étapes d'une évaluation, peut être utilisée comme un aide-mémoire, notamment par les décideurs (services de la Commission).*

http://europa.eu.int/comm/budget/evaluation/Key_documents/evalguides_fr.htm (en français).

European Commission Structural Funds (ed.). Evaluating Socio-Economics Programmes. Luxembourg: Office for Official Publications of the European Communities, 1999.

Evaluation Research Society Standards Committee. "Evaluation Research Society Standards for Program Evaluation." In P.H. Rossi (ed.), Standards of Evaluation Practice. New Directions for Program Evaluation, no. 15. San Francisco: Jossey-Bass, 1982.

Fontaine, C., and Monnier, E. "Evaluation in France." In J.-E. Furubo, R. Rist, and R. Sandahl (eds.), International Atlas of Evaluation. New Brunswick, N.J.: Transaction, 2002. *Les auteurs présentent le cas français en comparaison internationale ; ils pensent que petit à petit la pratique de l'évaluation en France, notamment sous l'influence européenne, convergera avec ce qui se passe chez nos voisins.*

Horber-Papazian, K., and Thévoz, L. "Switzerland Moving Towards Evaluation." In R.C. Rist (ed.), Program Evaluation and the Management of Government: Patterns and Prospects Across Eight Nations. New Brunswick, N.J.: Transaction, 1990.

Joint Committee on Standards for Educational Evaluation. Standards for Evaluation of Educational Programs, Projects, and Materials. New York: McGraw-Hill, 1981.

Joint Committee on Standards for Educational Evaluation. The Program Evaluation Standards. (2nd ed.) Thousand Oaks, Calif.: Sage, 1994 : *liste résumée des standards du Joint Committee, sans autre texte ni analyse.* <http://www.wmich.edu/evalctr/jc> (résumé)

Keith, W., The Canadian Evaluation Society (CES) Experience in Developing Standards for Evaluation & Ethical Issues, Paper presented at the Fifth European Conference on Evaluation of the Structural Funds, "Challenges for Evaluation in an Enlarged Europe," Budapest, June 2003 : *cet article présente le rôle du Joint Committee et de la Société Canadienne d'Évaluation (Canadian Evaluation Society, CES), notamment en ce qui concerne l'élaboration de standards. L'analyse des écrits de Stufflebeam, Fournier et Patton souligne l'utilité de ces standards et rappelle les diverses utilisations que peuvent en faire évaluateurs, commanditaires, enseignants, chercheurs, et autres praticiens. Les trois annexes listent les standards du Joint Committee, les principes et les lignes directrices du CES.*

http://europa.eu.int/comm/regional_policy/sources/docconf/budapeval/work/keith.doc

Kraan, A. and van Adrichem, H. "The Netherlands Court of Audit and Meta-research: Principles and Practice" in Schwartz, R., and Mayne, J (ed). Quality Matters: Seeking Confidence in Evaluation, Auditing and Performance Reporting. New Brunswick: Transaction Publishers (2004) : *Retour sur une expérience de plus de vingt ans dans le domaine du contrôle de la qualité des évaluations produites par les ministères.*

Läubli Loud, M.M. "Setting Standards and Providing Guidelines: The Means Toward What End?" Evaluation, 2004, 10, 237-245 : *après une réflexion sur la professionnalisation de l'évaluation en Suisse, l'article analyse l'utilité et les limites des standards élaborés par la SEVAL. En partant de l'exemple du SFOPH (Swiss Federal Office of Public Health) qui les transforme en checklist, il en définit les utilisations pratiques, effectives ou potentielles :*

- 1) aide-mémoire pour élaborer le calendrier d'une évaluation ;
- 2) support pour l'enseignement ;
- 3) outil de promotion de la qualité des évaluations ;
- 4) base pour le développement de la culture de l'évaluation.

<http://www.bag.admin.ch/cce/publikationen/workingpapers/e/budapest03.pdf>

Leca, J. 1993. "Sur le rôle de la connaissance dans la modernisation de l'Etat et le statut de l'évaluation." Revue Française d'Administration Publique 66, avril-juin:185-196.

Monnier E., 1992, Evaluations de l'action des pouvoirs publics, Economica, 2è ed, Paris.

Neale, J., Owen, J.M., and Small, D. "Encouraging the Use of Codes of Behavior in Evaluation Practice." Evaluation and Program Planning, 2003, no. 26, 29-36 : *selon les auteurs, malgré l'énergie consacrée par les associations professionnelles de l'évaluation au développement de codes de conduite, il y a bien des raisons de penser que beaucoup d'évaluations sont conduites sans y faire référence.*

Les auteurs présentent la stratégie de formation qu'ils ont utilisée pour encourager les évaluateurs à donner sens à des codes de conduite et à les intégrer dans leur travail. Le regroupement des standards en catégories permet aux évaluateurs et commanditaires de se concentrer sur les règles les plus pertinentes pour chaque étape d'une évaluation. <http://www.sciencedirect.com/> (30 \$)

Patton, M. Q. 1997. Utilization-Focused Evaluation, The New Century Text. Thousand Oaks: Sage.

Perret B. L'évaluation des politiques publiques. Paris, 2001, La Découverte, collection Repères (n°329), 124 pages.

Situant les méthodes, les dispositifs et les procédures existant en France, l'auteur revient sur l'utilisation de standards d'évaluation dans les organisations internationales, en lien avec l'émergence de nouvelles formes de gouvernance. Ces standards comportent une double dimension de principes déontologiques et de qualité méthodologique pour les évaluations : les plus connus et les plus influents sont cités, et le rôle de la SFE rappelé.

Toulemonde J., « Evaluation culture(s) in Europe : Differences and Convergence between National Policies », Vierteljahrshefte zur Wissenschaftsforschung, 3/2000, p. 350-357. *L'auteur souligne que, bien loin d'être universelles, les approches nationales de l'évaluation persistent ; il peut s'agir d'étapes dans l'évolution des pratiques, en fonction des institutions existantes.*

Toulemonde, J., Summa-Pollitt, H., Usher, N. "Triple Check for Top Quality or Triple Burden?: Assessing EU Evaluations" in Schwartz, R., and Mayne, J (ed). Quality Matters: Seeking Confidence in Evaluation, Auditing and Performance Reporting. New Brunswick : Transaction Publishers (2004) : *point sur le développement et l'usage des standards de qualité pour l'évaluation des politiques communautaires. Élaboration du « Guide des bonnes pratiques » de la Commission européenne. Analyse de la pratique d'assurance qualité de la DG Agriculture (fonctionnant sur la base de la grille MEANS). Perspectives de la Cour des Comptes Européenne en ce qui concerne le contrôle de la qualité des évaluations de la Commission.*

Schwab, O. "The Standards of the German Evaluation Society – and Why They Are of (No?) Use for Structural Funds Evaluations in Germany." Paper presented at the Fifth European Conference on Evaluation of the Structural Funds, "Challenges for Evaluation in an Enlarged Europe," Budapest, June 2003 : *à partir des travaux d'un groupe de réflexion sur l'évaluation des fonds structurels européens, l'auteur définit les objectifs des standards DeGEval et tente de comprendre les limites à leur connaissance et leur usage.*

La majeure partie de l'article explique comment utiliser ces principes, à la fois

- 1) *cadre de référence facilitant le dialogue entre évaluateur et commanditaire ;*
- 2) *garants de la qualité des évaluations ;*
- 3) *bases pour développer une culture d'évaluation.*

http://europa.eu.int/comm/regional_policy/sources/docconf/budapeval/work/schwab.doc

Schwartz, R., and Mayne, J. (ed). Quality Matters: Seeking Confidence in Evaluation, Auditing and Performance Reporting. New Brunswick : Transaction Publishers (2004) : *fruit d'un travail collectif international, cet ouvrage traite de la qualité de « l'information évaluative », produite dans un contexte d'évaluation, d'audit de performance et d'indicateurs de performance de type LOLF (ou GPRA aux Etats Unis). Treize contributions décrivent les systèmes d'assurance qualité mis en place dans différents contextes, les critères utilisés et les problèmes rencontrés.*

Stufflebeam, D.L. "Professional Standards and Principles for Evaluations." In T. Kellaghan and D.L. Stufflebeam (eds.), International Handbook of Educational Evaluation. Norwell, Mass.: Kluwer, 2003.

Stufflebeam, D.L. "A Note on the Purposes, Development, and Applicability of the Joint Committee Evaluation Standards." American Journal of Evaluation, 2004, 25,99-102.

Summa, H., and Toulemonde, J. "Evaluation in the European Union: Addressing Complexity and Ambiguity." In J.-E. Furubo, R. Rist, and R. Sandahl (eds.), *International Atlas of Evaluation*. New Brunswick, N.J.: Transaction, 2002.

United Kingdom Evaluation Society. "Guidelines for Good Practice in Evaluation." N.d. : *la Société d'Évaluation britannique (UKES) définit des lignes directrices utilisables pour des évaluations touchant à n'importe quel domaine. Afin d'en faciliter l'accès et l'appropriation, elles les classe en quatre groupes correspondant aux perspectives de l'évaluateur, du commanditaire, de tout autre participant au processus, et de ceux qui réalisent des auto-évaluations. Les standards sont présentés comme un "mémo" des étapes essentielles, comme la définition explicite et partagée des objectifs de l'évaluation ou les dispositions à prendre concernant l'utilisation des informations.* http://www.evaluation.org.uk/Pub_library/Good_Practice.htm

Weiss C.H., 1993 (1973), « Where Politics and Evaluation Research Meet », *Evaluation Practice*, 14(1), p. 93-106. *Cet article essentiel ne pose pas directement la question des standards, mais il montre de façon claire, que l'évaluation doit se poser constamment la question de ses rapports avec la politique. C'est l'une des conditions de sa qualité.*

Weiss Carol H., 1998 [1972], *Evaluation*, 2nd edition, Prentice Hall, Upper Saddle River. *Un manuel classique, qui traite notamment des standards de qualité. Lecture essentielle.*

Widmer, T. *Meta-Evaluation: Kriterien zur Bewertung von Evaluationen*. Bern: Haupt, 1996.

Widmer, T. "Evaluating Evaluations: Does the Swiss Practice Live Up to the 'Program Evaluation Standards'?" In C. Russon (ed.), *The Program Evaluation Standards in International Settings*. Kalamazoo: Evaluation Center, Western Michigan University, 2000 : *T. Widmer reprend dans cette article une méta-évaluation qui confronte 15 évaluations suisses aux standards définis par le Joint Committee. Après avoir décrit la méthode, il analyse les résultats qui révèlent notamment la nécessité de mettre en place des standards susceptibles d'améliorer l'image professionnelle des évaluateurs. La dernière partie est une analyse critique des critères du Joint Committee au regard de leur application à la situation suisse.*

<http://www.wmich.edu/evalctr/pubs/ops/ops17.pdf>

Widmer, T. *EES Policy on Standards, Guidelines, Codes and Ethics in Evaluation*. Zürich: European Evaluation Society, 2004b : *ce document bref (1 page) liste et explique brièvement les 7 options prises par la Société Européenne d'Évaluation (European Evaluation Society, EES) dans le cadre de sa politique relative aux « standards, lignes de conduite, codes et principes de l'évaluation », et notamment :*

- 1) *Le soutien aux Sociétés Nationales d'Évaluation pour le développement continu des standards ;*
- 2) *Le rôle de plateforme et de régulateur des échanges sur le sujet, entre autres par le biais d'un site ;*
- 3) *La mise en place d'un groupe de travail ;*
- 4) *Le développement des échanges internationaux.*

http://www.europeanevaluation.org/docs/standardspolicy_v2_18022004.pdf

Widmer, T. "Instruments and Procedures for Assuring Evaluation Quality: A Swiss Perspective." In R. Schwartz, J. Mayne, and J. Toulemonde (eds.), *Assuring the Quality of Evaluative Information: Prospects and Pitfalls*. New Brunswick, N.J.: Transaction, forthcoming.

Widmer, T., Landert, C., and Bachmann, N. Evaluation Standards of SEVAL, the Swiss Evaluation Society. Geneva/Berne: SEVAL, 2000 : *la SEVAL définit les standards qui doivent non seulement augmenter « la crédibilité, la compétence, et la fiabilité » des évaluations, mais aussi servir de base à des méta-évaluations. Elle les classe en quatre groupes : utilité, faisabilité, déontologie et précision. La liste de la première annexe – ou « aperçu fonctionnel » – en propose un autre classement qui permet un repérage rapide des standards prioritaires selon les étapes du processus, depuis la définition et la planification de l'évaluation jusqu'à la présentation du rapport final.*

http://www.seval.ch/fr/documents/SEVAL_Standards_2001_fr.pdf

Williams, K., de Laat, B., and Stern, E. The Use of Evaluation in the Commission Services: Final Report. Paris: Technopolis France, 2002 : *Cette étude analyse les mécanismes qui jouent sur l'utilisation des résultats des évaluations dans la mise en œuvre des activités de la Commission européenne, parmi lesquels elle relève l'influence des standards de l'évaluation. S'appuyant tant sur des études de cas que sur une comparaison des standards américains, canadiens et australiens, les auteurs soulignent le rôle potentiel des standards pour la qualité et le management des évaluations. La dernière partie formule des recommandations, dont certaines ont trait à la mise en valeur des standards.* http://europa.eu.int/comm/budget/evaluation/pdf/use_of%20evaluation_final_report.pdf

New Directions for Evaluation Volume 2004, Issue 104 Special Issue : International Perspectives on Evaluation Standards. Issue Edited by Craig Russon, Gabrielle Russon. Publié en ligne le 13 Jan 2005 : *cet ouvrage trace un panorama historique et conceptuel de l'appropriation ou de la création de standards de l'évaluation par les organisations nationales. Il s'attache en détail au développement de ces normes dans trois contextes socioculturels différents : l'Europe de l'Ouest, l'Afrique et l'Australie. Il analyse également la co-existence de ces pratiques nationales avec celles d'organisations internationales - comme la Commission Européenne pour ne citer qu'un exemple - qui consistent davantage à formaliser des standards qui tiennent compte des divergences culturelles.* <http://www3.interscience.wiley.com/cgi-bin/jissue/109865496>

LES CAHIERS MEANS (1999)

Volume 1 : Conception et conduite d'une évaluation

Volume 2 : Choix et utilisation des indicateurs pour le suivi et l'évaluation

Volume 3 : Principales techniques et outils d'évaluation

Volume 4 : Solutions techniques pour évaluer dans un cadre de partenariat

Volume 5 : Evaluation transversale des impacts sur l'environnement, l'emploi et les autres priorités d'intervention

Volume 6 : Glossaire de 300 concepts et termes techniques.

La série des cahiers Means, dont un numéro traite directement de la qualité des évaluations, analysée à travers une grille utilisée pour les rapports rendus à la Commission européenne, constitue un ouvrage de référence essentiel qui illustre l'état de l'art international.

Une association pour développer l'évaluation et promouvoir son utilisation

La Société Française de l'Évaluation (SFE) a été fondée en juin 1999 à Marseille. Elle a pour vocation générale de contribuer au développement de l'évaluation et de promouvoir son utilisation dans les organisations publiques et privées, et, aussi au sein des collectivités en charge de l'élaboration et de la mise en œuvre des politiques publiques. Elle s'efforce, notamment, de faire progresser les techniques et les méthodes et de favoriser le respect de règles éthiques et procédurales propres à garantir la qualité des évaluations ainsi qu'un usage approprié de leurs résultats.

C'est un lieu de rassemblement, d'échanges et de formation, ouvert à tous les acteurs de l'évaluation -les commanditaires, les praticiens et les utilisateurs- des secteurs tant publics que privés.

Progresser ensemble, ouvrir des lieux de débats, de réflexion et d'échanges

- affirmer l'évaluation comme une démarche volontaire au sein des administrations, d'état ou territoriales, et des différents partenaires institutionnels,
- valoriser l'apport de l'évaluation à la préparation, au pilotage stratégique et à la réalisation des politiques publiques,
- travailler sur la connaissance et œuvrer pour faciliter les échanges de "bonnes pratiques",
- rendre intelligible l'évaluation auprès de tous -décideurs, agents d'exécution et usagers- et œuvrer pour la reconnaissance de la spécificité de l'évaluation et sa distinction des métiers du contrôle, de l'audit ou de l'inspection.

Concrètement, l'action de la SFE tend à faire reconnaître l'évaluation, non seulement comme une pratique de bon sens mais, plus encore, comme un atout pour la conduite et l'optimisation de l'action publique et un stimulant des relations partenariales entre les différents acteurs.

En 2004, la SFE compte environ 250 adhérents, dont une cinquantaine d'institutionnels. Elle est animée par un Bureau composé de 12 membres d'origines diverses (fonctionnaires d'État, fonctionnaires territoriaux, chargés d'études et de recherche, consultants, associatifs).

Avec ses membres et ses partenaires, la SFE développe des actions nationales et en région

- L'organisation de colloques et de réunions de travail, seule ou en coopération avec d'autres instances publiques ou privées.
Depuis 1999, les "**Journées Françaises de l'Évaluation**" ont permis d'approfondir :
 - en 2000, « *Évaluation et gouvernance* » à RENNES
 - en 2001, « *L'Évaluation des politiques publiques à l'épreuve des territoires* » à ISSY-les-MOULINEAUX
 - en 2002, « *L'Évaluation au service de la stratégie publique ?* » à LILLE
 - en 2003, « *Évaluation et Développement durable* » à LIMOGES
 - en 2004, « *Évaluation et politiques sociales* » à BORDEAUX.
- La création de **groupes de travail thématiques** : « *Évaluation, territoires et collectivités territoriales* », « *Europe et évaluation des politiques publiques* », « *Qualité de la commande* », « *Standards et déontologie* », « *Évaluation qualitative en réseau* », « *Évaluation des politiques de l'emploi et de l'insertion professionnelle* », « *Métiers de l'évaluation* », « *Performance et évaluation* », « *Aide au développement* ».
- Le développement de **clubs régionaux de l'évaluation** regroupant localement les adhérents et les sympathisants de l'évaluation : *Club Rhône-Alpes*, *Café de l'évaluation en Nord-Pas-de-Calais*, *Club PACA*, *Club Midi-Pyrénées*, *Club Aquitaine*, *Club transfrontalier Lorraine-Belgique-Luxembourg*, *Club Bretagne/Pays-de-la-Loire*.
- L'élaboration d'une **Charte de l'évaluation** relative aux standards de référence et à la déontologie en matière de pratiques évaluatives et d'évaluations de politiques et de programmes publics.
- La conception et diffusion d'une information spécifique sur la vie de la SFE, de ses groupes de travail et clubs, sur l'évaluation en général et les travaux des Journées annuelles (Actes) sur le site web de l'association www.sfe.asso.fr, au moyen d'un bulletin de liaison et maintenant de Cahiers de l'évaluation.

En charge d'évaluation ou tout simplement soucieux d'une gestion rigoureuse et transparente, ses nombreux adhérents (institutionnels, universitaires, chercheurs, organismes d'Etat, collectivités territoriales, commanditaires, cabinets, associations, particuliers, étudiants, ...) partagent les objectifs de la SFE. Ils lui apportent, en outre, leur expérience, leur savoir, leur réflexion et leur action. Ils participent ainsi à l'enracinement d'une véritable et durable culture de l'évaluation dans la conception, la mise en œuvre, l'analyse des résultats des politiques publiques, et contribuent à l'inclusion de l'évaluation au sein de la stratégie et de la décision publique.

Première livraison d'une nouvelle série éditée par la Société Française de l'Évaluation (SFE), ce cahier produit par le groupe « standards de qualité et déontologie » entend faire le point sur les débats en cours depuis plusieurs années déjà, retracer les questions rencontrées sur ce sujet, en France comme à l'étranger, préciser les acquis et les positions des différents acteurs : tout ce qui peut avoir trait aux évolutions de la charte de l'évaluation et aux pratiques qui en découlent prend, dès lors, un relief accru.