

## PRÉSENTATION

Ce "Petit guide" est le fruit de l'expérience accumulée par le Conseil scientifique de l'évaluation dans son activité prescrite par le décret du 22 janvier 1990, au terme duquel il est chargé de "favoriser le développement des méthodes d'évaluation et de définir une déontologie en la matière". Le Conseil a évidemment utilisé l'expérience des avis rendus dans le cadre de la procédure d'évaluation interministérielle, mais les informations et préconisations contenues dans ce guide ne sauraient être de ce fait considérées comme pertinentes pour les seules évaluations interministérielles : il existe des questions fondamentales d'approche de l'évaluation et des attitudes qu'il est recommandé d'adopter dans tous les cas de figure.

Bien entendu, les solutions concrètement élaborées sur le terrain peuvent varier en fonction des contextes ; comme le Conseil scientifique l'avait noté dans son deuxième rapport (décembre 1992) : "les buts légitimes des évaluations peuvent être très divers et il est suprêmement important de les identifier au préalable pour savoir à quel type d'opération l'on a affaire avant de proposer un contrôle scientifique de sa faisabilité et de sa qualité. (...) Toute connaissance, et la connaissance évaluative en particulier, ne sont pas des modèles intellectuels abstraits qu'il suffirait d'appliquer comme une recette à tous les milieux et toutes les organisations sans considération de leurs histoires et pratiques". Cependant, les modèles d'analyse et d'action qui sont proposés dans ce "Petit guide" signalent des problèmes que l'on rencontre sous des formes diverses dans la plupart des cas d'évaluation, quels qu'en soient les niveaux et les objectifs.

Le Conseil scientifique espère avoir ainsi contribué à une meilleure collaboration entre le travail de connaissance scientifique jaugé par le test de vérité et le travail de réforme pratique jaugé par le test d'utilité. Ce premier texte devrait être complété par un second volume consacré à un examen plus approfondi de questions à caractère technique fréquemment rencontrées dans les démarches d'évaluation.

Jean LECA

# 1. Introduction

## 1.1. But poursuivi par le présent guide

Le présent guide est consacré à l'évaluation des "actions" et "activités" publiques, vocables délibérément imprécis pouvant désigner une large gamme d'objets : politiques, programmes, actions, mesures, réglementations, projets, activité d'une administration ou d'un corps de fonctionnaires, service(s) produit(s) par un organisme ou un établissement public, etc. L'évaluation individuelle des personnels est en revanche ici clairement exclue. Par commodité, on emploiera le plus souvent dans ce qui suit le terme *politique*, référence habituelle dans le contexte français de l'évaluation publique, dans un sens générique qui, sauf précision contraire, est susceptible d'inclure les différents objets mentionnés ci-dessus. On se limitera, pour l'essentiel, à l'évaluation ex-post ou concomitante, l'évaluation a priori posant, dans l'ensemble, des problèmes méthodologiques de nature différente. Cela étant, certains des principes et instruments que l'on va présenter peuvent s'appliquer à l'évaluation prévisionnelle.

Dans le cadre ainsi précisé, ce guide tente de définir des règles méthodologiques de portée générale. Il porte, certes, la marque du contexte dans lequel il a été élaboré : le Conseil scientifique de l'évaluation, auteur collectif de l'ouvrage, a eu à examiner des projets et rapports produits dans le cadre de la procédure interministérielle d'évaluation instaurée par le décret du 22 janvier 1990 (évaluations de grande ampleur, décidées par le gouvernement, concernant des politiques complexes à objectifs et acteurs multiples). Instruit par sa pratique, le Conseil scientifique a porté une grande attention à des questions qui, sans être spécifiques, revêtent une acuité particulière dans ce genre d'opérations. Le lecteur devra en tenir compte et, dans quelques cas, adapter les préconisations présentées à ses besoins propres. L'expérience prouve cependant que les problèmes rencontrés dans les cas les plus complexes ne sont pas absents d'évaluations apparemment plus simples, ou se voulant plus pragmatiques dans leur démarche.

Ce "Petit guide", qui ne traite que sommairement des techniques de la recherche évaluative, s'adresse d'abord aux agents publics impliqués dans des opérations d'évaluation au titre d'utilisateurs, de commanditaires ou de membres d'"instances d'évaluation". Il peut également être utile aux experts et chercheurs susceptibles d'être impliqués dans une opération d'évaluation, en leur donnant des informations sur le contexte dans lequel ils auront à travailler. Son but principal est de fixer les principes généraux de la démarche sans entrer dans le détail des différentes techniques de collecte et de traitement de l'information. Il s'agit d'abord, en d'autres termes, de répondre à deux grandes questions : à quoi peut servir l'évaluation et comment s'y prendre pour l'organiser et la mener à bien ?

## 1.2. Un éventail de pratiques administratives et cognitives

Qu'est-ce que l'évaluation ? il n'est pas si facile de répondre précisément à cette question. Le terme est en effet employé pour désigner des opérations qui n'ont à première vue pas grand rapport entre elles :

- a) Un service extérieur de l'Etat met en place un "tableau de bord" de son activité, regroupant dans un cadre formalisé des données sur les moyens humains et financiers mobilisés par les différentes actions dont il a la charge, ainsi que sur quelques paramètres physiques de cette activité (nombre de dossiers traités, populations bénéficiaires, type de prestations ou de services produits...etc), voire des *indicateurs* de performance ou de qualité du service.
- b) Un organisme de contrôle ou un corps d'inspection ministériel réalise une enquête sur l'activité d'un organisme sous tutelle, en s'efforçant d'analyser à la fois son fonctionnement et les résultats obtenus.
- c) Une équipe de sociologues mène une recherche sur l'impact d'une nouvelle mesure sociale (en application ou en projet) sur la situation de ses bénéficiaires.
- d) Un économiste estime l'impact économique d'une nouvelle mesure fiscale en utilisant un modèle économétrique.
- e) Un organisme public demande à un consultant de mener un audit sur sa structure, son fonctionnement et ses missions.
- f) Le Premier ministre demande à une personnalité ou à une commission de lui remettre un rapport sur un problème concernant l'action de l'Etat ou sur un problème général de société.

Cette énumération, non exhaustive, montre que l'on se trouve face à un continuum de pratiques impliquant la *collecte et le traitement d'informations* sur l'action publique (que s'est-il passé ?), des préoccupations *normatives* (a t-on bien fait ?), et/ou des préoccupations *instrumentales* (comment faire mieux ?) liées au bon fonctionnement et à l'efficacité des administrations et services publics. Il est clair que, dans chacune de ces opérations, une partie ce que l'on appelle ici évaluation est concerné. Mais on pourrait aussi bien, dans beaucoup de cas, parler de contrôle (ou d'analyse) de gestion, de contrôle de conformité, de conseil, d'expertise, de recherche appliquée ou d'audit.

## 1.3. Les spécificités de l'évaluation

Pour spécifier le type d'opérations visé par le présent guide, le plus simple est de recenser les principales exigences et préoccupations qui doivent impérativement être prises en compte dans une démarche d'évaluation. Celles-ci peuvent être résumées en deux points :

- l'ambition de mesurer, ou du moins d'apprécier le plus objectivement possible (en se conformant aux exigences de la méthode scientifique et en prenant de l'indépendance vis à vis de celui qui conduit la politique), les effets de la politique évaluée sur la société(1), et/ou de comprendre les logiques de son fonctionnement ;

- le souci d'aider le commanditaire politique (et, au delà, l'ensemble des lecteurs du rapport d'évaluation) à porter un jugement de valeur sur la politique évaluée (sur l'atteinte de ses objectifs et, éventuellement, sur leur validité ; sur son efficacité, son efficience, sa cohérence, etc...)

Une troisième préoccupation, de nature quelque peu différente, est souvent mise en avant. Elle n'est pas constitutive de l'évaluation au même titre que les deux précédentes, mais se révèle néanmoins fort importante dans bon nombre de cas :

- le souci de contribuer pratiquement à l'amélioration de la politique évaluée en éclairant ses différents protagonistes sur le sens, les conditions et les conséquences de leurs actions et décisions.

#### **1.4. Commentaire sur quelques définitions**

De nombreuses définitions de l'évaluation ont été proposées.

a) Certaines définitions considèrent l'évaluation comme une activité principalement cognitive.

La seule définition "officielle" est celle qui figure dans le décret du 22 janvier 1990 : *"Évaluer une politique, c'est rechercher si les moyens juridiques, administratifs ou financiers mis en oeuvre permettent de produire les effets attendus de cette politique et d'atteindre les objectifs qui lui sont fixés"*. Cette définition insiste à juste titre sur le fait que les effets de la politique doivent être appréciés en référence à ses objectifs. Il faut pour cela qu'ils aient été formulés de manière suffisamment précise et opératoire pour être confrontés à l'observation empirique de la réalité sociale, ou du moins qu'ils puissent être reconstitués et/ou interprétés dans cette perspective. C'est ainsi que le Conseil scientifique de l'évaluation a compris la définition du décret : le *référentiel*(cf 3.1.5.) de l'évaluation n'est jamais complètement déductible des textes officiels et sa construction fait partie intégrante du travail d'évaluation.

Une autre définition de l'évaluation avait été proposée dans un rapport du Plan de 1985 (rapport Deleau) : évaluer une politique, c'est *"reconnaître et mesurer ses effets propres"*. Cette définition ne présuppose pas l'existence d'objectifs, mais elle néglige le fait que les effets d'une politique sont en général extrêmement nombreux, complexes et hétérogènes : ils ne sont pas identifiables sans une grille de lecture préalable et sont, qui plus est, rarement mesurables. Elle a cependant le mérite de mettre l'accent sur la notion d'"effets propres" : ce que l'on attend de l'évaluation, ce n'est pas seulement de constater si des objectifs ont été atteints, mais de chercher à

---

(1) Les effets de la politique sur la société, ou l'"état de la société", sont à prendre au sens large, incluant l'état de la nature, la situation économique, etc.

identifier l'ensemble des effets d'une politique et de faire la part entre les effets imputables à l'action publique considérée et l'influence de facteurs exogènes.

b) La définition du rapport Viveret (1989), en revanche, met l'accent sur la dimension normative de l'évaluation : "*évaluer une politique, c'est former un jugement sur sa valeur*". Le terme d'évaluation contient le mot "valeur", et il est aisé de constater que des valeurs de référence sont à l'arrière plan de la plupart des évaluations. La décision d'évaluer est toujours liée à la volonté d'argumenter sur la réussite ou le bien-fondé d'une action. Reste que la décision finale sur la poursuite ou la réforme d'une politique publique est toujours un acte politique : on verra que l'évaluation ne peut qu'y contribuer, dans des conditions et des limites qui doivent être soigneusement précisées.

c) D'autres définitions soulignent la dimension instrumentale, utilitaire de l'évaluation, qui n'apparaît que de manière implicite dans les définitions précédemment citées. Pour Freeman et Rossi, par exemple, l'évaluation doit se préoccuper de "l'utilité, de la mise en oeuvre, de l'efficacité et de l'efficience des mesures qui ont pour but d'améliorer le sort des membres de la société".

On retiendra de cette brève discussion que l'évaluation est une activité à la fois, et en proportions diverses, cognitive, normative et instrumentale.

### **1.5. Evaluation, recherche, contrôle, audit**

A la lumière des développements précédents, il est possible de préciser le positionnement de l'évaluation par rapport à des "genres voisins" : recherche, contrôle, audit.

- L'évaluation se différencie de la *recherche pure* par sa prise en compte *explicite et a priori* de finalités normatives et instrumentales. Certes, comme l'a fortement souligné la sociologie de la science, la recherche, même fondamentale, n'est jamais exempte de préoccupations pratiques liées à son contexte social. Il n'en demeure pas moins que la crédibilité du chercheur, particulièrement dans le domaine des sciences sociales, repose sur une revendication d'autonomie par rapport au politique, autonomie qui implique en particulier le libre choix des questions et des grilles d'interprétation. Le seul but affiché est le progrès de la connaissance pour elle-même. Dans le cadre de l'évaluation, au contraire, le chercheur, sans renoncer aux exigences méthodologiques de sa discipline, accepte d'être intégré dans un dispositif et de considérer comme pertinentes (quitte à les discuter ou les reformuler) les questions et valeurs de référence d'un commanditaire politique.

- L'évaluation se différencie des différentes formes de *contrôle* (contrôle de conformité, contrôle et analyse de gestion) et de l'*audit organisationnel* par le type de point de vue adopté pour apprécier l'action publique. Le contrôle et l'audit se réfèrent à des *normes internes au système analysé* (règles comptables, juridiques, ou normes fonctionnelles), tandis que l'évaluation essaye d'appréhender d'un point de vue principalement *externe* les effets et/ou la valeur de l'action considérée.

L'évaluation peut néanmoins intégrer l'audit du fonctionnement des services chargés de la mettre en oeuvre, ou une étude sociologique ayant le même objet, dans la mesure où ce fonctionnement influe sur les effets que l'on cherche à mesurer et à expliquer. Il faut rappeler également que, contrairement à certaines formes d'audit, l'évaluation n'a pas pour objet de porter un jugement sur la manière dont les agents individuels, y compris les responsables hiérarchiques, remplissent leur mission.

### **1.6. Les fonctions et finalités sociales de l'évaluation**

Il est courant dans la littérature consacrée à l'évaluation de classer les pratiques évaluatives selon les buts (souvent supposés incompatibles entre eux) que poursuivent leurs initiateurs, et les résultats à en attendre. On oppose ainsi parfois l'évaluation "récapitulative" qui a pour but d'aboutir à une appréciation globale et distanciée, et l'évaluation "formative" qui vise prioritairement à éclairer les acteurs impliqués dans la mise en oeuvre d'une politique sur le contexte et les conséquences de leurs pratiques. Selon un autre point de vue, on oppose parfois l'évaluation "managériale" qui cherche à rendre plus efficiente la gestion publique et l'évaluation "démocratique" qui a pour ambition d'accroître la qualité et la transparence du débat public. Il n'est pas certain que ces catégories soient pertinentes en tant que principe de classement des opérations d'évaluation : l'expérience tendrait plutôt à montrer que les évaluations réussies remplissent une pluralité de fonctions. Il n'en demeure pas moins utile de distinguer différentes finalités qui peuvent avoir plus ou moins d'importance selon les cas considérés :

a) une finalité "déontologique" : rendre des comptes aux responsables politiques et aux citoyens sur la manière dont une politique a été mise en oeuvre et sur les résultats qu'elle a obtenus. Cette dimension recouvre donc à la fois l'amélioration de l'"imputabilité" des systèmes d'action publique (notion correspondant au terme anglais *accountability*, au sens de capacité à rendre des comptes sur son action à une autorité de niveau supérieur ou à un public démocratique), la finalité informative et la finalité "démocratique" de l'évaluation.

b) une finalité gestionnaire : répartir plus rationnellement les ressources humaines et financières entre différentes actions, améliorer la gestion des services chargés de les mettre en oeuvre.

c) une finalité décisionnelle : préparer des décisions concernant la poursuite, l'arrêt ou la refonte d'une politique.

d) une finalité d'apprentissage et de mobilisation : contribuer à la formation et à la motivation des agents publics et de leurs partenaires en les aidant à comprendre les processus auxquels ils participent et à s'en approprier les objectifs.

En outre, sans que l'on puisse parler de finalité, l'un des rôles manifestes de l'évaluation est de favoriser les convergences de vue et la coopération entre des acteurs politiques ou administratifs autonomes engagés conjointement dans la conduite d'une action.

Les questions posées, le choix des méthodes et des informations à collecter, et, finalement, le type d'arguments à développer dans le rapport d'évaluation, dépendront pour partie des buts organisationnels et politiques poursuivis par le commanditaire et de la nature des bénéfices escomptés de l'évaluation (cf 3.1.3.).

### **1.7. Classement des "qualités" d'une politique évaluée**

Fondamentalement, l'évaluation consiste à tenter de répondre à un ensemble de questions relatives à une politique, sa mise en oeuvre et ses effets. Ce questionnement peut être orienté de différentes manières en fonction du contexte et des buts poursuivis par le commanditaire. Pour aider celui-ci à former son jugement, l'évaluation cherche à apprécier dans quelle mesure la politique évaluée possède les principales "qualités" qui caractérisent idéalement une "bonne" politique :

- **Cohérence** (dans la conception et la mise en oeuvre) : les différents objectifs sont-ils cohérents entre eux ? les moyens juridiques, humains et financiers mis en place sont-ils adaptés à ces objectifs ?

- **Atteinte des objectifs** : Dans quelle mesure les évolutions constatées de la réalité sociale sont-elles conformes aux objectifs de la politique ?

- **Efficacité** : Dans quelle mesure les effets propres de la politiques sont-ils conformes à ses objectifs ? (on revient en 3.2.2 et au chapitre 8 sur la notion d'"effets propres").

- **Efficienc**e : Les ressources financières mobilisées par la politique ont-elles été bien utilisées ? les résultats de la politique sont-ils à la mesure des sommes dépensées ?

- **Impact** (effectivité) : Quelles sont les conséquences globales de la politique pour la société ? ces conséquences sont-elles bénéfiques ?

- **Pertinence** : Une politique se justifie très généralement par l'identification d'un "problème de société" auquel les pouvoirs publics se sentent tenus de faire face. Une politique sera dite pertinente si ses objectifs explicites sont adaptés à la nature du (des) problème(s) qu'elle est censée résoudre ou prendre en charge. La question de la pertinence est la plus délicate et la plus "politique" que l'évaluation ait à examiner. Il est parfaitement légitime de l'exclure du champ de l'évaluation, si tel est le choix du commanditaire.

## 1.8. Actualité et avenir de l'évaluation

L'évaluation répond à un besoin de rationalité, mais aussi de transparence : quoi de plus normal, en apparence, pour un responsable politique ou administratif que de chercher à connaître les conséquences de ses décisions ou de son action afin d'en assumer pleinement la responsabilité ? Dans une conception démocratique du fonctionnement de l'État et de la gestion publique, une telle exigence paraît s'imposer. Pourtant, l'expérience montre que l'évaluation n'est pas une démarche naturelle dans les services publics. Les obstacles auxquels elle se heurte peuvent être mis au compte de difficultés pratiques : une évaluation approfondie coûte cher et prend du temps. Les décideurs justifient souvent leur scepticisme vis à vis de l'évaluation par le fait qu'ils sont confrontés à des échéances de court terme et qu'ils doivent faire des choix qui ne peuvent attendre les résultats de longues études. Les calendriers électoraux sont cependant plutôt de moyen terme, et la gestion dans l'urgence et l'improvisation n'est pas une fatalité. Une meilleure maîtrise du temps de la décision donnerait à l'évaluation tout son sens et toute son utilité.

Des résistances plus profondes existent, plus ou moins consciemment liées à la volonté de préserver des zones d'opacité dans l'action publique. Cette opacité étant elle-même perçue comme un facteur d'autonomie par les fonctionnaires et les élus. Il est certain que le commanditaire de l'évaluation court le risque de ne pas maîtriser totalement les processus de changement qu'elle est susceptible d'enclencher.

En dépit de ces obstacles, l'évaluation se développe depuis une dizaine d'années, dans notre pays comme dans l'ensemble des pays développés. Les raisons qui paraissent expliquer ce développement sont les suivantes :

- les difficultés budgétaires et la crise de légitimité de l'action publique rendent nécessaire une justification plus solidement argumentée des interventions et des dépenses publiques. Le lien entre évaluation et décision budgétaire est rarement direct et contraignant. Il s'agit davantage d'éclairer partiellement des choix complexes que de les réduire à un calcul coût-avantages. Mais l'évaluation n'en est pas moins porteuse d'efficacité et d'efficacité publique, par définition.

- La décentralisation, la construction européenne et, plus généralement, l'intensification de la coopération économique mondiale ont accru les interférences entre niveaux de pouvoir juridiquement autonomes. La complexité qui en résulte accroît les besoins d'information commune, de coordination et de régulation. Le processus d'évaluation peut être un moment intense d'échange d'information et de coopération entre acteurs juridiquement autonomes. Ce besoin de coordination ne concerne d'ailleurs pas que les collectivités publiques : d'autres acteurs, entreprises, associations, organisations socio-professionnelles, sont de plus en plus fréquemment impliqués en tant que partenaires dans la mise en oeuvre des politiques publiques. Cette évolution se traduit notamment par la multiplication de procédures innovantes qui tentent d'instaurer, par un recours étendu à la technique

du contrat, de nouvelles formes d'action collective et de nouveaux mécanismes d'incitation et de régulation. Il est significatif que les actions de ce type suscitent fréquemment une demande d'évaluation.

- Les agents publics doivent suivre le mouvement général de modernisation, d'ouverture et de responsabilisation : ils ont besoin pour cela de formation continue et d'un surcroît de motivation. Visant à confronter les administrations à l'impact social de leur action, l'évaluation est pour les fonctionnaires l'occasion de vérifier leurs pratiques, de référer leur engagement professionnel à des objectifs sociaux, de se préoccuper des finalités et des résultats concrets de leur action. L'administration ne sera "une affaire qui marche" (*going concern*) que si les agents administratifs sont adaptés et efficaces. Cette condition est nécessaire, mais elle n'est pas suffisante : il faut aussi que le "sens de la marche" soit le bon. Le développement de l'évaluation semble de nature à répondre à cette double nécessité.

# Première partie :

## Conception et pilotage de l'évaluation

### 2. Vue d'ensemble sur une opération d'évaluation

#### 2.1. L'évaluation comme procédure formalisée

##### 2.1.1. Produire des connaissances appropriables

L'évaluation au sens du présent guide est une activité de nature *institutionnelle*, qui a vocation à s'intégrer à la gestion publique et au fonctionnement du système politique. Il est commode d'opérer une distinction entre l'évaluation ainsi comprise et la "recherche évaluative" (*evaluation research*), recherche en sciences sociales appliquée à l'étude des politiques et de leurs effets sur la société. L'évaluation s'appuie sur la recherche évaluative mais ne s'y limite pas.

L'évaluation a pour finalité de contribuer à l'élaboration d'un jugement de valeur, de préparer une décision, d'améliorer pratiquement la mise en oeuvre d'une politique ou le fonctionnement d'un service. Dans tous les cas, il faut que le commanditaire et les autres destinataires de l'évaluation (y compris, dans certains cas, le grand public) puissent *s'approprier* les résultats et connaissances produites, les intégrer à leur propre vision de la réalité. A cette fin, trois conditions doivent être réunies :

- l'évaluation doit *répondre aux questions* que se pose le commanditaire (ce qui implique un *projet d'évaluation* ciblé et un *cahier des charges* précis) ;
- les informations et raisonnements développés doivent être *crédibles* et, dans toute la mesure du possible, *compréhensibles* par l'ensemble des destinataires ;
- les jugements de valeur portés par l'évaluation doivent être perçus par eux comme *fondés sur des arguments légitimes*(1).

C'est en fonction de ces trois impératifs que doit être conçue toute démarche d'évaluation. Le projet d'évaluation doit s'efforcer de créer dès le départ les conditions d'une bonne réception de ses résultats, gage de leur impact sur les décisions et pratiques administratives.

L'idée de base est la suivante : il est plus facile de s'entendre sur des questions et sur une méthode que sur des réponses. Distinguer rigoureusement les différentes étapes de l'évaluation (les temps du questionnement, du choix des méthodes, de la collecte et du traitement des données, et enfin celui de la formulation des réponses et conclusions), permet de construire pas à pas les éléments d'un *monde commun*, ensemble de connaissances et de références partagées par les évaluateurs, les

---

(1) Cette exigence de légitimité de l'argumentation n'imposant pas à l'évaluateur de ne formuler que des conclusions qu'il estime acceptables pour le commanditaire.

commanditaires et les autres acteurs concernés.

*Exemple : Afin d'évaluer les "politiques d'insertion des adolescents en difficulté" (1993), il est apparu nécessaire de recenser les différents sens que les textes officiels, les chercheurs en sciences sociales et les acteurs de terrain donnent du terme "insertion". L'une des études a notamment mis en évidence les différences essentielles dans la manière de concevoir l'insertion selon qu'on est opérateur dans un canton rural, dans un bassin d'emploi économiquement sinistré ou en banlieue parisienne (1).*

C'est pour cette raison, fondamentale du point de vue de l'utilité sociale de l'évaluation, que l'on met délibérément l'accent dans ce guide sur le caractère *procédural* de la démarche.

### **2.1.2. Deux niveaux de méthode**

Compte tenu de ce qui vient d'être dit, toute démarche d'évaluation fait appel à deux registres d'élaboration méthodologique :

- les méthodes de la recherche évaluative, c'est à dire les méthodes de recherche en sciences sociales applicables à l'étude des politiques publiques. Ces méthodes sont peu spécifiques : on y retrouve l'ensemble de la "boite à outils" du statisticien, de l'économiste, du sociologue, voire du psychologue, de l'anthropologue ou de l'historien.

- les principes méthodologiques et déontologiques applicables au "management" politique et organisationnel de l'évaluation (y compris le fonctionnement du *dispositif* mentionné en 3.1.8.), et à l'élaboration du *projet d'évaluation*, cadre conceptuel dans lequel devra s'inscrire le travail de rassemblement des données et la recherche évaluative.

Les chapitres 2 à 4 concernent principalement ce dernier aspect. Quant aux méthodes de la recherche évaluative, on en aborde les principaux aspects aux chapitres 5 à 8, étant rappelé que le présent guide n'a pas vocation à se substituer aux ouvrages plus techniques de méthodologie des sciences sociales.

## **2.2. Caractéristiques générales d'une opération d'évaluation**

La première étape de l'évaluation est l'identification de son objet, de ses protagonistes, et la compréhension de ses enjeux.

### **2.2.1. Qui demande l'évaluation, et pourquoi ?**

Hormis le cas où l'évaluation résulte d'une procédure à caractère systématique (tel que l'examen des établissements d'enseignement supérieur par le Comité National d'Evaluation), elle fait toujours l'objet d'une décision spécifique de la part d'un commanditaire. Ce commanditaire est le plus souvent une autorité politique ou administrative hiérarchiquement responsable de la politique, mais ce peut être une instance *ad hoc* (Comité interministériel de l'évaluation, Comités régionaux d'évaluation...), ou un organisme de contrôle (Cour des comptes, corps d'inspection...). Dans le cas des politiques partenariales, surtout lorsqu'elles sont contractualisées (contrats de Plan État-Région, contrats de ville), la décision d'évaluer est généralement prise en commun par les partenaires. Il faut toutefois noter que le commanditaire officiel de l'évaluation n'est pas nécessairement son *initiateur* réel. Celui-ci (direction technique, corps d'inspection, instance consultative...), sera peut-être l'interlocuteur pertinent dans la phase d'élucidation du projet d'évaluation.

*Exemple : l'évaluation de la politique d'insertion des adolescents en difficulté a été décidée par le Comité interministériel de l'évaluation en 1990 à la demande du Ministre des finances, à l'initiative de l'Inspection générale des finances et de l'Inspection générale des affaires sociales (1).*

L'analyse des positionnements de l'initiateur et du commanditaire par rapport à l'action évaluée fait apparaître une grande variété de situations, depuis l'évaluation totalement intégrée au management d'une activité jusqu'à l'évaluation demandée, commanditée ou réalisée par un organisme indépendant du décideur.

En plus des initiateurs et commanditaires de l'évaluation, il faut tenir compte dans son déroulement de ses autres destinataires potentiels et des différents "porteurs d'enjeux" (*stakeholders*), acteurs sociaux dont les intérêts peuvent être affectés par les suites données à l'évaluation (associations, organisations professionnelles, groupes de pression...).

L'identification des acteurs, la reconstitution des étapes successives ayant conduit à la décision d'évaluer, les motifs invoqués, les formes prises par le processus décisionnel, etc...peuvent être autant d'éléments importants pour saisir les finalités réelles d'une opération d'évaluation, anticiper ses difficultés et fixer ses priorités.

### **2.2.2. Des objets de nature diverse**

L'évaluation peut concerner différents types d'objets : action, activité, service ou organisme. Les deux types les plus fréquemment rencontrés sont les politiques et les programmes, mais on peut aussi évaluer des projets, des pratiques, etc...

#### ***Politiques et programmes***

Les textes français relatifs à l'évaluation font presque toujours référence à l'*évaluation des politiques publiques*. Dans le monde anglophone, en revanche, comme d'ailleurs dans la plupart des autres pays, c'est la notion d'*évaluation de programme* qui domine. La différence n'est pas seulement d'ordre sémantique : elle est révélatrice de conceptions différentes de l'action publique. Le terme de programme désigne, en principe, une séquence d'action limitée dans le temps et, surtout, définie précisément dans ses moyens et ses objectifs opératoires (on parlera, par exemple, d'un programme de 20 000 stages pour chômeurs de longue durée). La notion de politique, quant à elle, renvoie plutôt à un ensemble complexe de programmes, de procédures et de régulations concourant à un même objectif général (la "politique de l'emploi", par exemple, est constituée d'un nombre important de tels éléments). En règle générale, l'évaluation paraît plus simple à concevoir dans le cas d'une action individualisée, de nature programmatique ou régulatrice, dans la mesure où les résultats attendus, les moyens et les mécanismes d'action sont immédiatement lisibles dans le dispositif. D'un autre côté, en choisissant des objets trop limités, on court parfois le risque de négliger les interférences avec d'autres actions publiques visant les mêmes populations ou s'attaquant aux mêmes problèmes. La démarche d'évaluation est, il est vrai, plus complexe lorsqu'elle prend pour objet une politique publique composite, et les aspects procéduraux soulignés précédemment prennent alors davantage d'importance. Mais, précisément, l'imbrication croissante des problèmes et des interventions est l'une des difficultés majeures que doit affronter l'action collective, et l'évaluation doit l'y aider.

### ***Grands projets***

Les *grands projets d'investissement* (une ligne de TGV, par exemple) peuvent être considérés comme des types particuliers de programme, mais ils n'en présentent pas moins des caractéristiques spécifiques du point de vue de l'évaluation :

- les décisions les concernant ont un caractère irréversible, ce qui implique que l'accent soit mis davantage sur l'évaluation *a priori* que sur l'évaluation *ex-post* ;
- ils permettent la production de biens et services, à caractère marchand ou non, qui donnent souvent lieu à des actes de consommation individualisés (prendre le train, se rendre dans une bibliothèque ou un musée), d'où l'utilisation plus systématique des méthodes d'évaluation économique (coût-avantages), sans négliger pour autant l'étude des externalités sociales, urbanistiques et environnementales.

### ***Pratiques professionnelles***

Certaines *pratiques professionnelles spécialisées* concourant à la production de biens et services publics peuvent donner lieu à évaluation. On évalue par exemple l'activité des organismes de formation ou les pratiques médicales. Dans ce dernier cas, l'évaluation porte principalement sur les stratégies diagnostiques et thérapeutiques (par exemple la pratique des frottis vaginaux pour prévenir le cancer

de l'utérus). Il s'agit, sur la base d'arguments scientifiques, de recherches évaluatives et de discussions contradictoires ("conférences de consensus"), d'aboutir à une "référence médicale", c'est à dire à une norme ayant vocation à régir la prescription du traitement ou de la technique évaluée.

### ***Activité des services et organismes***

Par extension, l'activité d'un service ou d'un organisme peut également être évaluée, la frontière avec l'audit étant alors, dans certains cas, plus difficile à tracer. On peut citer l'exemple des évaluations réalisées par le Comité National de la Recherche, ou encore l'évaluation du réseau des établissements culturels en Europe par le Ministère des affaires étrangères.

### **2.2.3. Temps de l'évaluation et "cycle de vie" de la politique évaluée**

L'évaluation peut être réalisée à différents moments de l'histoire d'une politique. On distingue principalement :

- l'évaluation *ex ante*, étude prospective de la faisabilité et de l'impact d'une mesure projetée ou en préparation ;
- l'évaluation *ex post*, qui vise à tirer des enseignements rétrospectifs sur une politique parvenue à maturité et "routinisée" ;
- l'évaluation *concomitante* (ou "chemin faisant").

Il n'est jamais facile d'harmoniser le temps de l'évaluation et celui qui rythme la vie d'une politique (calendrier électoral et budgétaire, événements sociaux susceptibles de modifier l'*agenda* politique, etc...). Dans un schéma idéal, les décisions politiques publiques les plus importantes devraient faire l'objet d'une évaluation *ex ante*, d'une évaluation concomitante intégrée au dispositif de pilotage de l'action, et d'une évaluation-bilan (*ex post*), débouchant sur une réactualisation des objectifs et des moyens. En pratique, il faut admettre que les retombées de l'évaluation sont souvent indirectes et longues à se manifester. L'expérience nord-américaine tend à prouver que le stock d'études et d'évaluations accumulé sur une question donnée constitue une *ressource pérenne* que l'on peut mobiliser à tout moment, pendant des années, pour préparer des décisions dans des délais relativement brefs.

### **2.3. Mandat, projet, cahier des charges**

La procédure qui part de la décision du commanditaire pour aboutir à la réalisation pratique des travaux d'évaluation peut prendre des formes variées. Elle doit être marquée par des étapes formelles matérialisées par des documents écrits. On distingue généralement :

- le *projet* d'évaluation, à savoir l'ensemble des réflexions qui précèdent la phase de

collecte et de traitement des informations. Le projet n'est pas nécessairement consigné dès le début dans un document précis et exhaustif. Son élaboration peut, selon les cas, précéder partiellement la rédaction du *mandat* d'évaluation, auquel cas celui-ci doit y faire référence. Lorsqu'une "instance d'évaluation" est constituée (cf 3.1.8.), sa tâche prioritaire est de valider et de s'approprier les premiers éléments du projet d'évaluation, de les compléter et de les préciser dans la ligne du mandat. Dans certains cas, les opérateurs ou "chargés d'évaluation" (cf 3.1.11) peuvent être, de fait, associés à la phase finale de son élaboration. Une fois définitivement précisés, les principaux éléments du projet devraient faire l'objet d'un document de synthèse, texte de référence qui servira de fil conducteur pour la rédaction du rapport final et de charte à l'instance d'évaluation.

- le *mandat* d'évaluation est le document par lequel le commanditaire politique confie officiellement la responsabilité de l'évaluation à un *évaluateur* (individu ou organisme), ou à une *instance d'évaluation* (en la personne de son président). Le mandat est généralement bref : il insiste sur les questions qui intéressent particulièrement le commanditaire, décrit le dispositif institutionnel et matériel de l'évaluation sans entrer dans le détail des méthodes, en se référant éventuellement à des documents complémentaires. Il peut en revanche fixer des délais pour la remise du rapport d'évaluation.

- l'expression de *cahier des charges* est parfois employée pour désigner le projet ou le mandat. Il vaudrait mieux lui conserver son sens précis de document servant de base à une relation contractuelle entre le commanditaire d'une étude spécifique et un opérateur. Dans le contexte de l'évaluation, le cahier des charges peut concerner soit un sous ensemble des travaux d'étude programmés dans le projet d'évaluation, soit l'opération d'évaluation dans son ensemble (lorsque sa conduite est confiée à un opérateur).

#### **2.4. L'importance du travail préparatoire**

Le travail de rassemblement et de traitement des informations qui constitue la substance d'une opération d'évaluation doit obéir à un schéma précis. Réduit à l'essentiel, le projet d'évaluation décrit un cheminement logique qui part des problèmes et interrogations qui ont motivé l'évaluation pour formuler un ensemble de questions et indiquer par quels moyens on va tenter d'y répondre. On ne saurait trop insister sur la nécessité de respecter ce cheminement. Trop d'évaluations sont orientées par le choix *a priori* d'une méthode ou par la nature de l'information immédiatement disponible, au détriment de la pertinence du questionnement.

De plus, avant de prendre la décision, coûteuse, de procéder à une évaluation, il est indispensable de se demander si elle est réalisable. L'emploi usuel du terme *faisabilité* ne doit cependant pas induire en erreur : il est toujours possible d'approfondir l'analyse d'un problème ou de compléter le stock des connaissances disponibles sur une politique et ses effets ; mais il peut se faire que les efforts

nécessaires soient disproportionnés aux bénéfices prévisibles ou aux exigences de délais imposés par la commande. Cette exigence de "parcimonie", qui est parfois citée comme l'un des principes déontologiques de l'évaluation, peut justifier la réalisation d'une étude de faisabilité, ou d'une pré-évaluation. Outre le fait qu'elle apporte souvent des éléments de réponse aux questions du commanditaire, l'étude de faisabilité permet de choisir entre différentes approches et de se prononcer en toute connaissance de cause sur l'opportunité d'entreprendre de longs et coûteux travaux de recherche.

### **3. Le projet d'évaluation**

#### **3.1. Les étapes du projet**

L'élaboration d'un projet d'évaluation comporte différentes étapes. La présentation que l'on en fait ici est volontairement très analytique. En pratique, ces étapes peuvent se recouvrir, ou être abordées dans un ordre différent.

##### **3.1.1. Définition du champ de l'évaluation**

L'élaboration du projet commence par un effort de *définition de la politique à évaluer*, c'est à dire du champ (institutionnel, temporel, géographique...) des activités et mesures à prendre en compte.

*Exemple : Il n'y a pas de définition univoque de ce qu'il faut entendre par "Politique de la montagne". Les mesures susceptibles d'être comprises dans cet ensemble sont très nombreuses et des choix sont nécessaires (jusqu'où faut-il prendre en compte les politiques sectorielles non spécifiques -agriculture, transport, tourisme- ayant une forte incidence sur le développement des zones considérées ?). Autre problème : cette politique a une longue histoire (plus d'un siècle) : quelle période faut-il prendre en considération ? jusqu'où faut-il remonter pour analyser les objectifs poursuivis par les pouvoirs publics ? (2)*

##### **3.1.2. Repérage des objectifs, officiels ou implicites.**

Les objectifs officiels sont parfois explicités dans les textes qui instituent une politique (par exemple, dans l'exposé des motifs d'une loi ou d'un décret). Cette source principale doit être confrontée à tout document susceptible d'éclairer sur les intentions des décideurs : déclarations de responsables politiques ou administratifs, compte rendus des débats parlementaires qui ont précédé l'adoption d'une loi, etc. Il arrive fréquemment que les objectifs initiaux d'une politique soient infléchis par des décisions ultérieures plus discrètes, notamment par des arbitrages budgétaires.

##### **3.1.3. Identification des finalités et enjeux de l'évaluation**

Au vu du contexte dans lequel a émergé la demande d'évaluation (cf 2.2.1.), des attentes exprimées par le commanditaire et de la phase du "cycle de vie" de la politique où intervient l'évaluation, il s'agit d'identifier le type de résultats et de retombées pratiques que l'on peut en attendre, ainsi que ses conséquences potentielles pour les différents protagonistes. Celles-ci peuvent être très différentes selon que l'évaluation a un enjeu limité (réaménagement technique d'une procédure, reformulation de certains objectifs...), qu'elle peut aboutir à modifier substantiellement l'allocation de ressources budgétaires, déboucher sur l'arrêt d'un programme ou, au contraire, sur sa généralisation, etc.

### **3.1.4. Hypothèses de départ**

Il est nécessaire de procéder à un *diagnostic provisoire* sur "ce que l'on pense", en bien ou en mal, de la politique à évaluer. Ce diagnostic peut s'appuyer sur la perception immédiate par les acteurs des *réussites*, *échecs* ou *dysfonctionnements* de la politique, ainsi que sur l'analyse des *évolutions du contexte* technique, social ou juridique qui paraissent rendre nécessaire une refonte de la politique, ou une reformulation de ses objectifs.

Il peut sembler paradoxal d'anticiper ainsi sur les conclusions : l'évaluation ne doit-elle pas aborder son objet sans idées préconçues, et n'a-t-elle pas précisément pour objet d'identifier ce qui marche et ce qui fait problème ? En fait, les idées a priori du commanditaire (ou d'autres protagonistes de la politique) doivent être considérées comme des *hypothèses à confirmer, à infirmer ou à reformuler*. Ces idées constituent à la fois un point de vue et une connaissance à prendre en compte dès le début de l'évaluation, ainsi qu'une indication permettant d'identifier les domaines dans lesquels l'évaluation est le plus susceptible de produire des résultats utiles et d'avoir des retombées pratiques.

### **3.1.5. L'élaboration d'un référentiel**

L'une des questions majeures rencontrées au cours de l'élaboration d'un projet est celle du *référentiel* : à quels critères devra-t-on confronter les résultats observés ? au regard de quelles normes pourra-t-on se prononcer sur le bon fonctionnement et la réussite de la politique ?

A première vue, il semblerait que les objectifs de la politique constituent le seul référentiel légitime de l'évaluation : tel est d'ailleurs le point de vue adopté dans la définition de l'évaluation contenue dans le décret du 22 janvier 1990 (cf introduction). En pratique, cependant, les choses ne sont pas si simples, et plusieurs problèmes peuvent surgir :

- comme on l'a vu précédemment, les objectifs officiels sont à compléter et interpréter par la prise en compte d'objectifs implicites ;
- tous les objectifs ne sont pas de même niveau, ils ne sont généralement pas

indépendants et certains peuvent être reliés à d'autres par des rapports de cohérence ou de causalité plus ou moins bien fondés;

- ils peuvent être, au contraire, partiellement contradictoires ;
- leur pondération et leur formulation peuvent avoir évolué au cours du temps ;
- tous les objectifs ne sont pas également susceptibles d'être traduits en effets directement observables (qu'est ce, par exemple, que "l'équilibre social d'une région" ?) et en indicateurs.

Enfin et surtout, on a insisté ci-dessus sur l'importance des *attentes du commanditaire*. de l'évaluation. La prise en compte de celles-ci conduit presque toujours à porter un regard nouveau sur les objectifs de la politique. L'évaluation est souvent l'occasion pour des responsables de modifier ou de réinterpréter les objectifs de l'action publique, voire même de faire émerger de nouveaux objectifs.

*Exemple : l'évaluation de la "politique de protection des zones humides" a abouti à l'affirmation par l'Etat d'un objectif écologique qui n'avait pas été initialement pris en compte dans bon nombre des mesures examinées (aides publiques et réglementations concernant notamment les domaines suivant : aménagement rural, infrastructures, urbanisme et tourisme) (3).*

En dehors de ce cas extrême, il est fréquent que l'évaluation conduise à modifier la hiérarchie des objectifs ou promouvoir des objectifs latents.

*Exemple : l'évaluation de l'informatique administrative était centrée sur la question de l'efficacité des dépenses consenties pour équiper les administrations en ordinateurs. Le commanditaire s'interrogeait au départ sur le "retour sur investissement" mesuré par l'efficacité en terme de gains de productivité comparée aux ressources allouées. La réflexion préalable au lancement des études a permis d'aboutir à une conception plus complexe de l'efficacité, prenant en compte des données telles que la qualité du service rendu au public, les conditions de travail des agents, la sécurité des procédures, etc (4).*

On retiendra de cette discussion que la notion de référentiel renvoie à la nécessité d'interpréter les objectifs de la politique en fonction du contexte et des priorités actuelles du commanditaire politique, et de les reformuler pour les transformer en critères pertinents d'évaluation. Les phases d'identification des finalités de l'évaluation et de diagnostic provisoire décrites précédemment ont précisément pour objet de fournir les bases de cette contextualisation.

### **3.1.6. Thématiques et registres de l'évaluation**

Parallèlement à la construction d'un référentiel, il s'agit de choisir les domaines et les thématiques qui feront l'objet des questions les plus précises et des investigations les plus lourdes.

*Exemple : la politique de la montagne comprend différents volets : agricole (aides*

*spécifiques à l'agriculture de montagne), industrie, commerce et artisanat (aides à la diversification des activités économiques), tourisme, environnement (parcs naturels...), transports (politique de désenclavement), etc. Tous ne pourront pas être traités de manière également approfondie par l'évaluation (2).*

Cet effort de "ciblage" concerne également les *registres* de l'évaluation, c'est à dire le choix et l'interprétation concrète des "qualités" de la politique -cohérence, efficacité, efficience, impact ou pertinence (cf introduction)- que l'on cherchera à évaluer en priorité.

*Exemple : dans l'évaluation des Fonds Régionalisés d'Initiatives Locales pour L'emploi (FRILE), il avait été décidé dans un premier temps (1991) d'évaluer le procédures d'instruction des dossiers, le partenariat inter-administratif et les conditions de délégation des crédits. Le but était donc d'évaluer la cohérence de l'action effectivement mise en oeuvre par rapport à ses objectifs. La mesure de l'efficacité en terme d'emploi a été abordée dans une évaluation ultérieure (1993) (5).*

### **3.1.7. Le questionnement**

On entend par questionnement une *liste de questions*, classées logiquement et hiérarchisées, aussi complètes et précises que possible, dont la réponse commande la réponse aux questions générales évoquées précédemment. On revient ci-après (3.2.) plus en détail sur l'organisation du questionnement.

*Exemple : l'une des questions posées par la Commission nationale d'évaluation du Revenu minimum d'insertion concernait la pertinence du "contrat d'insertion" comme moyen de renforcement des processus d'insertion sociale et professionnelle. Cette question générale renvoie à une multitude de questions plus précises sur la mise en oeuvre et le contenu des contrats, l'engagement des institutions et des acteurs locaux, le comportement et le devenir des allocataires bénéficiant de ces contrats, etc (6).*

**Remarque importante** : la transformation de questions générales orientées vers l'action, telles que se les pose le commanditaire dans le cadre de ses responsabilités administratives ou politiques, en hypothèses vérifiables exprimées sous forme de questions ayant un sens précis pour des chercheurs en sciences sociales, constitue l'étape clef de l'élaboration du projet d'évaluation.

Cette liste de questions est évidemment susceptible d'être modifiée et enrichie au cours de l'évaluation. Si tel est le cas, il est souhaitable que soit gardée trace dans le rapport final des raisons qui ont motivé ces changements.

### **3.1.8. Le dispositif d'évaluation**

On entend par dispositif d'évaluation le système des relations à établir dans le cadre de l'évaluation entre ses différents acteurs et participants : commanditaires, protagonistes de la politique à évaluer (cf 2.2.1.), groupe de suivi ou de pilotage, *instance d'évaluation* (cf 4.1.) , chercheurs et chargés d'étude, sans oublier, le cas échéant, l'*instance de régulation scientifique* mentionnée en 4.1.4.

Selon la dimension et la variété des questions d'évaluation, le dispositif peut exiger la constitution d'une instance d'évaluation chargée, entre autres tâches, de synthétiser les différentes études (cf 4.1.), ou requérir seulement une équipe réduite en lien direct avec le commanditaire. Toutes les formules intermédiaires existent ou sont envisageables (groupes de suivi ou de pilotage dont le rôle peut être diversement conçu). *Une évaluation que l'on souhaite rapide doit comprendre un dispositif resserré*, mais cela suppose que les questions du commanditaire aient été formulées de manière suffisamment précise, par lui-même ou par les auteurs du projet, et que la nature limitée du problème à traiter permette d'emblée de cibler les investigations à mener. Quel que soit le dispositif choisi, celui-ci doit permettre d'imputer clairement l'évaluation à un auteur (collectif ou individuel), qui en assume la responsabilité. Cette clarification des rôles est nécessaire, car l'évaluation doit satisfaire, en plus des critères de qualité du produit, des critères formels concernant la conduite de la procédure et l'établissement des protocoles de recherche évaluative.

### **3.1.9. Le bilan informationnel**

Au vu du questionnement, il convient de réaliser un recensement sommaire des informations immédiatement disponibles, données administratives, études et recherches relatives au domaine considéré, ainsi que des informations qui pourraient être obtenues sans nouvelles dépenses et dans un délai bref. Ce bilan aboutit à fournir des premières réponses à une partie des questions posées, à prévoir que des réponses pourront être facilement apportées à certaines autres, et à dresser un état des principales lacunes.

### **3.1.10. Le programme d'étude**

En fonction du point précédent, et des moyens dont elle dispose, l'instance d'évaluation élabore une *programmation des nouvelles investigations qui apparaissent nécessaires*. Celles-ci peuvent être :

- 1) la compilation d'études préexistantes ;
- 2) l'exploitation de données administratives ;
- 3) l'analyse de documents existants (textes à caractère juridique, coupures de presse, documents cartographiques, audio-visuels, etc) ;
- 3) des enquêtes ou investigations *ad-hoc* permettant de recueillir de nouvelles données :
  - enquêtes statistiques (par questionnaire, par entretien, par téléphone) auprès de bénéficiaires ou d'acteurs clefs de la politique (individus, entreprises ou

collectivités) ;

- le suivi d'un panel (de bénéficiaires, le plus souvent), interrogé à intervalle régulier pendant une durée suffisamment longue (pouvant atteindre plusieurs années) ;

- des entretiens approfondis avec des acteurs ou autres "personnes ressources" ;

- des monographies ou études de cas ;

4) des auditions par l'instance.

Les données ainsi collectées n'ont pas toutes le même statut (3.3.2), et elles visent à répondre à des questions de nature différente. Ainsi, lorsque les données de base

n'existent pas au départ, l'élaboration d'indicateurs de résultats chiffrés nécessite l'interrogation d'un échantillon statistiquement représentatif de bénéficiaires. En revanche, quelques études de cas bien choisies peuvent permettre de comprendre les mécanismes d'action de la politique et d'en décrire qualitativement les effets.

On traite à partir du chapitre 5 de quelques-unes des questions techniques posées par le choix et la mise en oeuvre de ces outils.

### **3.1.11. Le choix des opérateurs**

C'est à l'instance d'évaluation, lorsqu'elle existe, qu'il revient de choisir les opérateurs ou "chargés d'évaluation", sur la base de cahiers des charges spécifiques, le cas échéant par voie d'appel d'offre. Dans le cas contraire, le commanditaire aura néanmoins intérêt à solliciter divers avis, voire constituer un comité d'appel d'offre. Les compétences auxquelles peut faire appel l'évaluation sont très diverses, elles concernent à la fois le domaine particulier dont traite l'évaluation, la connaissance de certains champs théoriques ou techniques d'analyse et la familiarité avec les démarches d'évaluation. Il est souvent intéressant de combiner le regard distancié et indépendant d'experts n'appartenant pas à l'administration et non impliqués dans la conduite de la politique évaluée, et la connaissance plus approfondie qu'ont de celle-ci ses acteurs directs, ainsi que les services d'étude de l'administration ou les corps d'inspection.

**Remarque importante :** Dans le contexte de l'évaluation, la position des chercheurs n'est pas identique à celle qui est la leur dans les conditions habituelles de la recherche. Il est important qu'ils puissent trouver un juste équilibre entre les grilles d'interprétation propres à leur discipline et les interprétations les plus à même d'être discutées par l'instance d'évaluation dans le cadre de sa propre problématique.

## **3.2. Retour sur l'orientation et la structure du questionnaire**

Le coeur du projet d'évaluation est constitué par un ensemble de questions et par le choix des méthodes qui permettront d'y répondre. Comme on l'a rappelé ci-dessus, ces questions essayent de traduire les objectifs généraux de l'évaluation en termes de *faits vérifiables*, de *processus observables* et de *grandeurs mesurables*. Le projet s'efforcera de classer ces questions en fonction des registres d'évaluation et des aspects de la politique qu'elles permettent d'appréhender.

Une autre manière de raisonner est de considérer une politique comme une construction théorique, au sens où elle implique une *représentation a priori* des mesures mises en oeuvre, du comportement des acteurs, de l'enchaînement des mécanismes d'action et des effets produits sur la société. Le questionnaire évaluatif peut alors être vu comme le moyen de tester les différents aspects de cette théorie.

### 3.2.1. Mise en vigueur de la politique

Toute politique comporte des *intentions et orientations générales* et un *dispositif opératoire*. Celui-ci repose sur un schéma a priori des opérations que doivent réaliser les acteurs de terrain. En pratique, ceux-ci sont confrontés à une réalité complexe dont le décideur central n'a qu'une vue très partielle. Ils sont donc amenés à prendre des décisions non anticipées, à interpréter les intentions des concepteurs de la politique. Il peut également se produire qu'ils aient des objectifs ou une stratégie propre, inspirés par leur éthique professionnelle ou leurs intérêts corporatistes, ne coïncidant pas forcément avec ceux du décideur central. Le niveau le plus élémentaire de l'évaluation consiste à analyser ces écarts entre *l'application concrète des décisions* et l'ensemble des normes qui définissent la politique. Cette partie de l'évaluation comporte elle-même deux aspects :

- l'examen de la conformité aux normes juridiques formelles, aux délais prescrits, aux règles budgétaires, etc. On est ici proche du domaine traditionnel du contrôle de régularité pratiqué par les corps d'inspection et la Cour des comptes ;
- l'identification des écarts entre les objectifs affichés et la réalité des opérations effectuées sur le terrain au titre de ces objectifs.

Une fois des écarts constatés, il reste à les expliquer, à apprécier leurs conséquences et, le cas échéant, à en tirer des conclusions normatives. Celles-ci peuvent aussi bien consister à *reformuler des normes qui se seraient révélées inapplicables* qu'à renforcer le contrôle exercé sur les acteurs de terrain, ou améliorer leur formation et leur information. Dans tous les cas, l'écart manifeste un manque de cohérence formelle entre les différents éléments -objectifs, dispositifs réglementaires, moyens humains, matériels et financiers qui constituent la politique.

*L'exemple type d'un tel décalage concerne la sélection par les fonctionnaires locaux des bénéficiaires effectifs d'une politique : pour de multiples raisons (à commencer par le fait que la connaissance a priori des besoins ne peut être qu'imparfaite), le profil des bénéficiaires réels d'une mesure n'est jamais exactement conforme à l'intention politique initiale. On le constate aussi bien dans le cas des aides aux entreprises que pour certaines mesures sociales.*

L'étude de la mise en vigueur (*implementation*) de la politique est particulièrement importante pour une évaluation concomitante intégrée au pilotage d'une nouvelle action.

### 3.2.2. Quels résultats observer et mesurer ?

En fonction du référentiel choisi (cf 3.1.5.), le projet doit s'efforcer d'*identifier les effets de la politique que l'on souhaite observer et, si possible, mesurer*. Il est très important, à cet égard, de ne pas confondre *l'atteinte des objectifs* et les *effets propres* de la politique.

### ***Atteinte des objectifs et indicateurs de résultats***

La mesure de *l'atteinte des objectifs* répond à l'une des attentes principales des commanditaires : elle renseigne, en un certain sens, sur la réussite ou l'échec de la politique. C'est en général ce que permettent d'apprécier les *indicateurs de résultats*.

Rappelons qu'un *indicateur* est une mesure dont les valeurs fournissent une indication incomplète mais utile sur un phénomène dont on cherche à apprécier l'évolution. Il est en général nécessaire de disposer de plusieurs indicateurs pour cerner un phénomène complexe.

*Exemple : l'évolution du taux de chômage est l'un des indicateurs de la situation du marché du travail, et donc un indicateur de résultat de la politique de l'emploi. Si l'on veut se faire une idée plus complète de l'évolution de l'emploi et du chômage, il faut rassembler et mettre en regard un nombre suffisant d'indicateurs éclairant divers aspects du phénomène (chômage des jeunes, chômage de longue durée, nature des emplois créés, etc...).*

Il faut être conscient du fait qu'un ensemble d'indicateurs de résultats, si complet soit-il, ne permet pas, en règle générale, d'évaluer l'efficacité d'une politique. Les variables représentatives des objectifs de la politique subissent toujours l'influence d'une multitude de facteurs exogènes (par exemple, dans le cas du chômage, de la conjoncture économique).

### ***Efficacité, effets propres, impact***

L'évaluation de l'efficacité d'une politique nécessite la mesure de ses *effets propres*, à savoir les modifications de l'état de la société dont elle est la cause. En d'autres termes, l'évaluation de l'efficacité suppose de pouvoir répondre à la question : que se serait-il passé si la politique n'avait pas existé ? Il est toujours beaucoup plus difficile de répondre à cette question qu'à celle de l'atteinte des objectifs. On trouvera ci-dessous (cf 8.) une présentation succincte des principales techniques utilisées pour tenter d'identifier et de quantifier des *liens de causalité* entre une politique et ses effets sociaux. Aussi difficile soit-elle, la question de la causalité ne peut-être évacuée de l'évaluation.

**Remarque :** les effets d'une politique sur la société ne se limitent pas aux effets attendus ou souhaités, conformes à ses objectifs (les effets qui correspondent à la notion d'efficacité). L'évaluation doit également se préoccuper des effets collatéraux (*side effects*) et des *effets pervers* de la politique. On emploie parfois les termes d'*impact* ou d'*effectivité*, pour désigner l'ensemble le plus large des effets de la politique.

*Exemple : L'évaluation de la politique d'informatisation de l'administration a mis en évidence le fait que les gains de productivité permis par l'informatique se traduisent rarement par des réductions d'effectif (contrairement à l'un des principaux objectifs visés). En revanche, divers autres effets ont été mis en évidence : absorption de*

*tâches les plus lourdes et plus complexes, diminution de la charge de travail des agents d'exécution, amélioration de la qualité et de la sécurité du service (4).*

### ***Efficiencia au sens de "coût-avantages"***

La question de l'*efficiencia* est plus rarement abordée, en pratique, que celle de l'efficacité. Prise dans son sens le plus strict, elle suppose la réalisation d'un *bilan coût-avantages*, en termes monétaires. Outre le fait que l'information de base nécessaire pour calculer le coût réel d'une politique fait souvent défaut, la traduction en termes monétaires des modifications apportées à l'état de la société est une opération d'une grande difficulté pratique et soulève parfois des problèmes théoriques quasiment insolubles (qu'en est-il, par exemple, du prix de la vie humaine ?). *L'économie publique* est la discipline qui a pour objet de fonder rationnellement différentes méthodes de calcul de la valeur économique des *biens collectifs* et autres avantages produits par l'action publique.

### ***Efficiencia au sens de "coût-efficacité"***

Dans un sens plus restreint, l'évaluation de l'efficiencia peut se limiter à une approche *coût-efficacité*, qui consiste à comparer les coûts nécessaires à la réalisation d'un certain niveau de résultat pour différentes politiques (on peut aussi dire que l'on raisonne en *coût d'opportunité* : on aurait pu faire autre chose avec les mêmes fonds publics).

*Exemples :* - il est usuel de comparer les différents programmes de lutte contre le chômage en terme de "coût de l'emploi créé" ou de "coût du chômeur évité". Dans l'évaluation des aides à la localisation des activités dans les zones de conversion (1994), on a procédé à une comparaison du coût de l'emploi créé grâce aux aides publiques entre les différentes zones de conversion étudiées (7).

- les mesures prises pour économiser l'énergie peuvent être évaluées en comparant le "coût de la Tep (tonne d'équivalent pétrole) économisée" par différents moyens.

**Remarque importante :** en toute rigueur, seul un jugement porté sur l'efficacité (rapportée aux principaux objectifs) d'une politique, permet de fonder rationnellement une décision relative à la poursuite ou à l'abandon de celle-ci. De même, seule une comparaison globale et précise de leur efficacité pourrait permettre de se prononcer rationnellement sur la répartition de ressources budgétaires entre différentes politiques. Dans la pratique, il faut faire preuve de prudence et de pragmatisme dans l'utilisation de l'évaluation à des fins d'aide à la décision budgétaire.

### **3.2.3. L'étude des mécanismes d'action**

Dans le paragraphe précédent, les politiques publiques sont implicitement considérées comme des "boîtes noires", dont on observe les entrées et les sorties. En pratique, l'évaluation ne peut se désintéresser de l'intérieur de la boîte, de ses mécanismes internes. Pour prendre une comparaison dans le domaine médical, on ne se contente pas de constater l'efficacité d'un médicament : on cherche toujours à comprendre *comment* il agit (en identifiant ses récepteurs dans l'organisme).

Dans le cas d'une politique publique, il s'agit principalement d'identifier les processus micro-économiques ou micro-sociaux induits par les mesures prises : effets de liaison et d'apprentissage, effets prix, modifications qualitatives du comportement des acteurs, mobilisation de ressources humaines ou matérielles jusque là inemployées, réorganisation du travail, gains de productivité, etc. Ou, en négatif, encouragement de décisions non optimales du point de vue économique (désincitation au travail, investissements non rentables...), coûts de transactions, etc... Compte tenu de la difficulté de mesurer rigoureusement les effets propres, l'analyse des mécanismes d'action peut permettre dans bien des cas d'accéder à une connaissance qualitative de l'efficacité d'une mesure.

*Exemple : dans le cas d'une aide à l'emploi, on interrogera directement les employeurs (et, si possible, d'autres acteurs ou observateurs de l'entreprise), pour comprendre comment l'aide est prise en compte dans les pratiques de gestion du personnel, afin d'apprécier l'importance de l'"effet d'aubaine" (utilisation d'une aide pour réduire le coût d'une embauche qui aurait été réalisée même en l'absence d'aide).*

#### **3.2.4. Contexte de la mise en oeuvre et conditions de généralisation**

L'efficacité d'une politique est le résultat d'une *interaction* entre une intervention ou l'apport d'une ressource *exogène* et une dynamique sociale *endogène*. De ce fait, il est fréquent qu'une politique produise des effets différents selon les lieux où elle est mise en oeuvre : ces différences peuvent relever de multiples causes (caractéristiques socio-démographiques des bénéficiaires, dynamiques socio-économiques locales, motivation et compétence des acteurs politiques et administratifs locaux). En outre, il est fréquent que l'efficacité d'une politique dépende de circonstances variables au cours du temps : une mesure peut être en synergie avec la conjoncture économique, ou au contraire être rendue inopérante par celle-ci.

L'évaluation doit s'efforcer d'identifier les facteurs contextuels qui interagissent avec la politique. D'une part ceux qui *conditionnent une mise en oeuvre correcte des mesures* (par exemple, les acteurs de terrain comprennent-ils les intentions formulées dans les directives émanant d'une administration centrale ?), et, d'autre part, ceux qui en *modifient les effets*.

*Exemple : l'évaluation de la politique d'aménagement des rythmes de vie des enfants*

*a révélé d'importantes disparités locales dans l'ampleur, la nature et les résultats des actions menées, tant en terme de modification effective des horaires scolaires que d'offre d'activités non scolaires (culturelles et sportives, notamment). Selon les évaluateurs, cette diversité tient à plusieurs causes : volonté politique et moyens financiers des collectivités locales, implication et compétence des différents partenaires (acteurs administratifs locaux, associations, professionnels...) (8).*

### **3.3. Dès le projet, penser aux lecteurs du rapport**

#### **3.3.1. Expliciter les choix et les limites de l'évaluation**

Aucune évaluation ne peut traiter de manière approfondie l'ensemble des questions que se pose le commanditaire. Le projet d'évaluation doit donc préciser quelle priorité est accordée à chacune d'entre elles, en fonction des retombées attendues (décisionnelles ou autres) de l'évaluation, de l'information disponible et des moyens qu'il est envisagé de consacrer à l'évaluation.

En outre, bien que cela puisse être plus délicat, l'identification de certains pré-supposés, axiomes ou valeurs sociales que l'évaluation devra prendre en compte sans les remettre en cause, peut-être utile à l'explicitation de ce que l'on en attend.

*Exemple : dans l'évaluation de la politique de protection des zones humides, le caractère nécessaire de la préservation des zones humides pour le maintien de la qualité des ressources naturelles (eau, paysages, biodiversité, ressources halieutiques) était considéré comme une donnée non discutable (3).*

*A contrario*, il est important d'identifier les principales incertitudes des commanditaires, ainsi que les éléments de jugement préalable que l'on peut considérer comme les moins assurés, et à propos desquels on peut attendre de l'évaluation qu'elle apporte un savoir neuf et modifie les perceptions ou opinions préétablies.

#### **3.3.2. Sélection des informations et stratégie argumentaire**

L'évaluation d'une politique complexe nécessite le recours à des informations hétérogènes, de nature et d'origine variée. Toutes ces informations ne présentent pas les mêmes garanties de fiabilité et d'objectivité : l'opinion des acteurs, par exemple, ne peut être mise sur le même plan que les faits sociaux observés par des méthodes plus objectives. Dans certains cas, une hiérarchisation des informations s'impose en fonction de leur statut et des méthodes par lesquelles elles ont été obtenues. Les raisonnements que l'on pourra développer sur la base de ces informations n'auront pas tous la même robustesse. Le poids relatif des différentes données doit faire l'objet d'une réflexion préalable : avant de décider d'affecter des ressources à telle ou telle investigation (monographie, enquête statistique, panel d'acteurs...), il faut se

demander si les informations recherchées sont susceptibles de fournir des *arguments nouveaux*, de nature à compléter et préciser les éléments de réponse au questionnement évaluatif que l'on possède déjà.

Pour prendre un exemple fréquemment rencontré, il peut s'avérer peu judicieux de réaliser une étude coûteuse pourtant théoriquement nécessaire pour répondre aux questions posées si l'on pense que ses résultats ne modifieront pas substantiellement le savoir déjà disponible. Une étude de cas approfondie apportera parfois des informations inédites sur les effets intermédiaires et les mécanismes d'action d'une politique, alors qu'une enquête statistique représentative auprès des bénéficiaires n'aurait fait progresser que marginalement la connaissance de ses effets sociaux.

*Exemple : le projet d'évaluation des structures d'insertion par l'économie (associations intermédiaires, entreprises d'insertion, régies de quartier) envisageait la réalisation d'une enquête lourde et représentative au plan national sur le devenir des personnes passées dans ces structures (en interrogeant à plusieurs reprises un panel de 1000 individus). Après réflexion, il est apparu à l'instance que cette enquête apporterait peu d'informations nouvelles par rapport aux enquêtes plus légères réalisées par le Ministère du travail, et qu'il était préférable de mettre l'accent sur une approche territorialisée (4 sites) permettant d'analyser les trajectoires individuelles à la lumière du contexte local (9).*

La réflexion préalable sur l'importance relative des différents types d'informations et d'arguments doit en outre faire droit à une exigence de pédagogie. Pour qu'un rapport d'évaluation soit pris en considération par ses destinataires, il faut que ceux-ci puissent adhérer à la démarche logique qui sous-tend la recherche. Concrètement, cela signifie que plusieurs mécanismes de conviction doivent être mis en oeuvre (cf 5.2), et qu'il est imprudent de s'en remettre à un seul type de raisonnement, fut-il le plus pertinent d'un point de vue strictement scientifique.

## **4. Conduite de l'évaluation et valorisation des résultats**

### **4.1. Présentation d'un dispositif type : l'instance d'évaluation**

L'évaluation n'est ni pure connaissance, ni pur mécanisme politique. La difficulté de trouver un équilibre et une cohérence entre ces deux dimensions implique d'examiner avec soin les modalités de "pilotage" de l'évaluation. Celles-ci doivent permettre un ajustement réciproque des contraintes de l'action publique et des exigences d'une démarche cognitive rigoureuse, sans sacrifier l'une à l'autre. Comme on l'a indiqué plus haut (cf 3.1.8.), le dispositif d'évaluation doit être adapté à l'ampleur et à la nature du problème à traiter, ainsi qu'aux impératifs de temps et de moyen. Sur la base des opérations dont il a eu à connaître, le Conseil scientifique préconise un dispositif type, qui lui paraît le mieux à même d'organiser une *zone d'autonomie* de l'évaluation entre la science et l'action. La pièce centrale de ce dispositif est l'"instance d'évaluation", groupe chargé par le commanditaire d'assumer la fonctions de maître d'oeuvre de l'évaluation.

#### 4.1.1. Rôle de l'instance d'évaluation

L'instance d'évaluation est plus qu'un comité de pilotage : elle doit bénéficier d'une large responsabilité dans le cadre d'un mandat écrit du commanditaire. Plus précisément, ses missions sont de deux ordres :

a) **Piloter** les travaux d'évaluation, c'est à dire, d'abord, mener à son terme l'élaboration du projet d'évaluation, traduire le questionnement évaluatif dans le cahier des charges des études et recherche, passer commande et suivre les différents travaux d'étude, auditionner des "personnes ressources", responsables administratifs, experts ou autres "témoins", voire se déplacer collectivement sur le terrain...

b) **Intégrer** les matériaux de l'évaluation, c'est à dire rassembler des documents, réceptionner les recherches études commanditées, valider leurs résultats, interpréter ces résultats au regard des autres informations collectées (auditions ou informations préalablement disponibles...), répondre aux questions posées par le projet d'évaluation, formuler des conclusions générales et, le cas échéant, des propositions, rédiger le rapport final.

L'instance est en général le lieu où, par la délibération, des conclusions raisonnables sont déduites d'études analysées et interprétées. Celle-ci doit être conçue comme un *arbitre* entre les différents points de vue et non comme un *médiateur* de différents intérêts qu'il conviendrait d'accommoder.

#### 4.1.2. Composition et fonctionnement

Il n'y a pas de composition type : l'instance doit permettre la confrontation du point de vue des responsables directs de la mise en oeuvre de la politique et de points de vue extérieurs (experts du domaine, acteurs sociaux concernés par la politique en tant que partenaires ou "bénéficiaires"). Il est néanmoins impossible de concevoir l'instance comme un lieu où tous les points de vue seraient en permanence présents. Ceux-ci peuvent d'ailleurs être pris en compte d'une autre manière au cours de l'évaluation (cf, par exemple, 6.7.).

Il est important que l'instance se considère comme une véritable structure de travail et non comme un lieu de négociation. L'intervention d'une *instance indépendante de régulation scientifique*, de type conseil scientifique (comme il en existe au niveau national et dans plusieurs régions), chargée d'émettre un avis sur les projets et les rapports d'évaluation, peut aider une instance d'évaluation indépendante à se fixer des règles de fonctionnement compatibles avec les exigences de l'évaluation. Celles-ci impliquent de ne pas sacrifier la rigueur scientifique à la nécessaire prise en compte du pluralisme des points de vue. Dans son premier rapport annuel, le CSE soulignait à ce propos : "*Un équilibre est à établir entre une connaissance précise des préoccupations et des analyses des administrations qui facilitera la réappropriation des conclusions par ces dernières, et une pratique de réflexion qui ne se cantonne pas dans les seules hypothèses qui fondent ou sous-tendent les*

*pratiques décisionnelles".*

#### **4.1.3. Les phases du travail de l'instance**

Le travail de l'instance s'organise selon le schéma suivant :

- appropriation du mandat
- finalisation du projet (formulation définitive du questionnement)
- élaboration d'un cahier des charges pour les études évaluatives
- choix des opérateurs
  
- suivi des études
- réception et discussion des rapports d'étude
- synthèse de leurs résultats
- rédaction du rapport
- gestion des suites de l'évaluation (communication, dialogue avec les différents acteurs concernés par les résultats de l'évaluation)

Parallèlement au déroulement des travaux d'étude, l'instance d'évaluation peut procéder à des auditions (d'experts, de responsables politiques ou administratifs), ou à des visites sur site, afin de compléter son information.

#### **4.1.4. Rôles possibles d'une instance de régulation scientifique**

L'intervention d'une instance de régulation scientifique (groupe d'experts mandatés pour veiller à la qualité des évaluations, en position de totale indépendance par rapport aux commanditaires) peut aider les instances à fonctionner efficacement et à mettre en oeuvre une méthodologie rigoureuse. On peut faire jouer à un tel organisme trois rôles différents :

- 1) l'assistance méthodologique aux rédacteurs des projets d'évaluation ;
- 2) la participation d'un des ses membres à l'instance d'évaluation ;
- 3) l'émission d'avis formels sur les projets et rapports d'évaluation (cf 4.4.2.).

**Remarque :** - les missions 1 et 2 sont compatibles entre elles, ainsi que les missions 1 et 3, les missions 2 et 3, en revanche, le sont difficilement.

- selon les dispositifs, les avis de l'instance scientifique sont rendus publics ou réservés au seul bénéfice de l'instance d'évaluation.

#### **4.2. Valorisation des résultats**

Les résultats de l'évaluation peuvent être valorisés de plusieurs manières : réunions de travail avec les décideurs et acteurs de la politique évaluée, conférences, actions de formation, articles de presse, etc. Rien n'interdit, en outre, de réutiliser une évaluation terminée dans un travail ultérieur : aux Etats-Unis, se développe la

pratique d'études récapitulatives tirant profit des multiples évaluations réalisées au cours des années précédentes dans un domaine donné de l'action publique.

Il n'en demeure pas moins qu'une partie non négligeable de la qualité et de l'utilité d'une évaluation repose sur la qualité du rapport d'évaluation.

### **4.3. Le rapport d'évaluation**

A la lumière des considérations qui précèdent sur la stratégie d'argumentation, on peut formuler un certain nombre de recommandations sur la structure et le mode d'écriture des rapports d'évaluation.

La démarche évaluative se caractérise par la recherche d'une synergie entre une exigence de qualité scientifique et des finalités de régulation et d'apprentissage social par la connaissance. En conséquence, un rapport d'évaluation devrait se caractériser par un effort important pour articuler rigoureusement les jugements et les faits : il s'agit, idéalement, de référer toutes les assertions normatives à des raisonnements explicites, eux même appuyés sur des observations dûment documentées. L'écriture du rapport doit en outre tenir compte de la diversité des modes de lecture de ses destinataires (en fonction, notamment, de leur disponibilité et de leur degré de connaissance du domaine). Il est souhaitable à cette fin de ménager plusieurs entrées et de permettre plusieurs niveaux d'appropriation du rapport (résumé, annexes, bibliographie, etc).

#### ***Structure du rapport***

Le rapport d'évaluation comprend des parties méthodologiques, des descriptions, des rappels de conclusions de travaux précédents, des analyses appuyées sur des informations nouvelles, et enfin des interprétations élaborées par l'instance d'évaluation. Ces différents types de texte doivent être, autant que faire se peut, distingués. En particulier, les imputations de causalité (telle action publique produit tel effet propre) et les jugements normatifs doivent être clairement identifiés comme tels et argumentés en référence aux études réalisées dans le cadre de l'évaluation. Le défaut le plus couramment constaté à cet égard est l'insertion de jugements de valeur non argumentés dans le cours d'un développement descriptif.

Sans qu'il soit utile de proposer un plan type, tout rapport d'évaluation devrait comprendre les éléments suivants :

- des **éléments d'information suffisants sur la politique évaluée**, l'historique des mesures examinées, l'analyse des textes établissant ces mesures, l'exposé des objectifs du gouvernement, des éléments de cadrage statistique, etc.

- un **exposé méthodologique** rappelant les différentes étapes de l'évaluation depuis l'élaboration du projet initial jusqu'à la rédaction du rapport final. L'évaluation est une démarche continue et cohérente qui part d'une commande politique pour aboutir à des réponses argumentées à un certain nombre de questions. Le cheminement suivi par les porteurs du projet, les principaux choix méthodologiques effectués par l'instance et les évolutions successives du projet

doivent être lisibles dans le rapport.

*Remarque : il arrive fréquemment qu'une thématique jugée importante au stade de projet soit finalement abandonnée. Il importe de savoir si cet abandon est du au manque d'informations disponibles, au fait qu'une question s'est révélée non pertinente au cours de l'évaluation, ou au désintérêt des membres de l'instance.*

- un **bref compte rendu des études réalisées** : le rapport ne doit pas chercher à restituer systématiquement le contenu des différents travaux réalisés. Il doit en revanche s'efforcer de mettre en évidence l'apport propre de chacune de ces études dans les réponses apportées par l'instance au questionnaire évaluatif.

- **des développements informatifs** : le rapport d'évaluation comporte des parties descriptives, éventuellement extraites des études soumises à l'instance, qui, sans répondre directement au questionnaire évaluatif, constituent par elles-mêmes des informations utiles aux décideurs.

- **des raisonnements bien identifiés** : le cœur du rapport de l'instance d'évaluation est constitué par l'exposé des arguments et raisonnements conduisant aux conclusions. La rédaction doit en être claire, fut-ce au prix d'un certain didactisme ; le style allusif doit être évité ; les différentes questions abordées séquentiellement, dans un ordre logique. Lorsque l'argumentation s'appuie sur des résultats d'études précis, les renvois aux documents de base doivent être systématiques.

- **des recommandations dont le statut doit être clarifié** :

Les rapports d'évaluation comprennent généralement des recommandations ou des propositions de réforme de la politique évaluée. Le statut de ces prises de positions doit être précisé : s'inscrivent-elles directement dans le prolongement des réponses apportées au questionnaire évaluatif, ou ne font-elles que refléter l'opinion des membres de l'instance ? Faut-il les considérer comme des recommandations fermes techniquement nécessaires, comme des propositions de nature politique ou comme de simples orientations devant donner lieu à de nouvelles études ? Il n'est pas exclu que le rapport établisse une distinction entre les préconisations qui font l'unanimité de l'instance d'évaluation et celles pour lesquelles de fortes divergences se sont manifestées. Dans tous les cas, la nature du lien entre ces diverses formes d'affirmations normatives et le contenu de l'évaluation doit apparaître clairement au lecteur, ce qui implique au minimum des références explicites aux raisonnements développés dans le corps du rapport.

#### **4.4. Les critères d'une bonne évaluation**

##### **4.4.1. Définition et usages des critères de qualité**

Les critères de qualité de l'évaluation, parfois appelés "standards" ou "valeurs" de

l'évaluation, sont des caractéristiques typiques d'une "bonne évaluation", susceptibles de fournir des références pour juger la méthode, le déroulement ou le résultat d'une évaluation. Ces critères concernent à la fois la qualité scientifique des connaissances produites, le respect de règles déontologiques dans les rapports entre les différents protagonistes de l'évaluation, et enfin les éléments qui conditionnent l'utilisation pratique des résultats de l'évaluation. Elles peuvent être utilisées ex-ante, comme principes guidant l'élaboration d'un projet d'évaluation, ou ex-post, comme critères d'appréciation d'une évaluation terminée. Dans les deux cas, il s'agit de répondre à la question : qu'est ce qu'une bonne évaluation ?

En général, une norme exprime un consensus de base à partir duquel on peut porter un jugement sur la qualité d'une pratique. Il faut toutefois noter que le terme "norme" peut prendre deux significations distinctes. Une norme peut être un critère d'excellence à viser quand on doit juger de la qualité d'une évaluation; elle peut aussi correspondre à une exigence minimale qui devrait s'imposer à toute évaluation. La notion de "standardisation" est issue de ce deuxième sens. En pratique, les deux sens sont présents, y compris dans les avis du Conseil scientifique : il y a en effet à la fois l'idée que certaines exigences minimales doivent être satisfaites pour que les résultats d'une évaluation soient utilisables, mais il y a également l'idée de comparer les évaluations à un idéal-type de la bonne évaluation, dans le but de faire progresser la qualité des pratiques évaluatives.

L'une des possibilités de juger de la qualité est de réaliser des *méta-évaluations*, revues comparatives portant sur un ensemble d'évaluations (relatives ou non à un même type de politique ou de programme). Les critères de qualité peuvent servir à évaluer de manière comparative un ensemble de documents d'évaluation. La réalisation de telles études peut être un moyen pour faire progresser à la fois les pratiques et la théorie de l'évaluation, mais aussi pour diffuser un certain nombre de connaissances générales sur les politiques publiques et leurs effets. Les méta-évaluations sont une pratique courante aux Pays-Bas et aux Etats-Unis.

---

#### **Les critères de qualité aux Etats-Unis et au Canada**

La question des normes de qualité occupe une place importante dans les débats théoriques consacrés à l'évaluation, notamment outre atlantique. Cet effort de normalisation est inséparable de la constitution d'un milieu professionnel spécialisé qui cherche à se donner des critères d'identification et de régulation interne. Les objectifs poursuivis par les associations professionnelles d'évaluateurs sont de deux ordres :

- permettre une labellisation des prestations de services vendues sous le nom d'évaluation. Il s'agit de favoriser l'extension du marché de l'évaluation en donnant aux clients potentiels des informations sur les spécifications du "produit" qu'ils sont en droit d'exiger sous ce label. Cette préoccupation de protection du consommateur conduit à mettre l'accent sur l'aspect déontologique des normes de qualité.
- constituer l'évaluation en tant que discipline autonome, susceptible d'être enseignée, et de donner lieu à l'accumulation de connaissances spécialisées. Sur ce double plan

professionnel et scientifique, cette autonomisation est d'ailleurs loin d'être acquise, du fait que l'évaluation recoupe à la fois des disciplines scientifiques plus anciennes (l'économie, la statistique, les sciences humaines, la gestion), des professions et des fonctions (auditeur, inspecteur, consultant, chercheur...).

### **Exemple 1 : les objectifs des normes, selon la société canadienne d'évaluation (SCE)**

" Précisons d'abord les objectifs que peut poursuivre l'établissement de normes, objectifs qui ne sont pas sans se recouper:

1. *Protéger et promouvoir la fonction de l'évaluation par la détermination des champs et des types d'activités de l'évaluation;*
2. *Protéger les intérêts du public par la démonstration de l'utilité sociale de l'évaluation;*
3. *Améliorer la qualité de la pratique évaluative.*

*Comme la SCE souhaite que les normes constituent le noyau du développement professionnel, nous avons décidé, pour amorcer notre démarche, de considérer **l'amélioration de la qualité de la pratique évaluative comme l'objectif principal du développement et de l'adoption de normes.** La promotion de la fonction et la protection de l'intérêt public deviennent ainsi des objectifs secondaires".*

### **Exemple 2 : les critères proposés par le "joint committee on standards for educational evaluation", aux Etats-Unis**

Parmi les systèmes de normes ayant exercé une influence sur la pratique de l'évaluation, on peut citer celui du "Joint committee on standards for educational evaluation", élaboré en 1981. Il comprend 30 standards, prenant en compte les aspects sociaux de l'évaluation, qui peuvent être regroupés en quatre catégories :

**Les critères d'utilité** cherchent à garantir qu'une évaluation répondra aux besoins pratiques d'information des destinataires de l'évaluation. Ces critères incluent l'identification des destinataires, la crédibilité de l'évaluateur, le diversité et la sélectivité de l'information collectée, la capacité d'interprétation normative, la clarté du rapport, la diffusion du rapport, le fait que le rapport soit remis à temps et l'impact de l'évaluation.

**Les critères de faisabilité** cherchent à garantir qu'une évaluation sera réaliste, prudente, diplomatique, et frugale. Ceci inclut des procédures pratiques, la viabilité politique, et un bon rapport coût/efficacité.

**Les critères déontologiques** cherchent à garantir qu'une évaluation sera conduite dans la légalité, dans le respect de l'éthique, et en prenant en considération le bien-être de ceux qui sont impliqués dans l'évaluation, aussi bien que de ceux qui sont affectés par ses résultats. Ceci inclut des obligations formelles, la gestion de conflits d'intérêt, une communication complète et franche des informations, le droit du public de savoir, le respect des droits de l'homme, les interactions humaines, un rapport rédigé de manière équilibrée, et le sens des responsabilités pour ce qui concerne l'affectation des deniers publics.

**Les critères d'exactitude** cherchent à garantir qu'une évaluation utilisera une information techniquement adéquate sur les caractéristiques de l'objet étudié, particulièrement lorsque ces caractéristiques déterminent sa valeur ou son mérite. Ceci inclut l'identification de l'objet, l'analyse du contexte, l'explicitation des objectifs et des procédures du programme, l'utilisation de sources d'information justifiables, des mesures valides, des mesures fiables, un contrôle

systematique des données, une analyse de l'information quantitative, une analyse de l'information qualitative, des conclusions justifiées et un rapport objectif.

En résumé, une bonne évaluation est utile à ses destinataires, simple à mettre en oeuvre, conduite dans le respect des règles éthiques, et techniquement exacte

-----

#### **4.4.2. Les avis rendus par le Conseil scientifique de l'évaluation sur les évaluations interministérielles**

En France, il n'existe pas à ce jour de milieu professionnel certifié et agréé de l'évaluation. Il ne peut donc y avoir de définition collective des critères d'une "bonne pratique" d'évaluation. Le Conseil scientifique de l'évaluation a considéré que la définition de tels critères entrerait dans le cadre de ses attributions. Au titre de l'article 8 du décret du 22 janvier 1990, le Conseil doit en effet rendre un avis sur la qualité des évaluations financées par le Fonds national de développement de l'évaluation. Il est en outre chargé au terme de ce même décret de "favoriser le développement des méthodes d'évaluation et de définir une déontologie en la matière".

Les avis rendus par le Conseil scientifique à l'issue des évaluations ont une double finalité :

- une finalité de **validation des résultats de l'évaluation**. Il ne s'agit pas pour le Conseil scientifique de se prononcer sur la pertinence des conclusions, et encore moins sur l'opportunité des recommandations faites par l'instance d'évaluation, mais d'apprécier, d'une part, le degré de cohérence entre le rapport d'évaluation et les différentes études réalisées à cette fin, et, d'autre part, le lien logique entre les recommandations et l'ensemble des constats et analyses du rapport. Il s'agit avant tout de fournir aux décideurs une garantie extérieure sur la valeur scientifique des arguments mis en avant par l'instance.

- une finalité de **capitalisation de l'expérience méthodologique**. L'évaluation n'étant pas à proprement parler une discipline scientifique, son apprentissage ne peut se faire que par accumulation d'expériences. Il est donc important que des enseignements soient tirés des avancées réalisées et des difficultés rencontrées par les évaluateurs.

La procédure d'émission de ces avis fait intervenir un expert extérieur au Conseil, auquel est confiée la tâche de vérifier la cohérence entre le rapport d'évaluation, les rapports d'études intermédiaires et le projet d'évaluation. Une rencontre a ensuite lieu entre le Conseil scientifique et l'instance d'évaluation. Avant de rendre public le texte de son avis, celui-ci est communiqué au président de l'instance pour observations éventuelles. L'avis définitif est ensuite publié avec le rapport d'évaluation.

#### 4.4.3. Les cinq critères retenus par le CSE

La qualité d'une évaluation ne se résume pas à sa validité scientifique. Il faut également prendre en compte les finalités sociales de l'évaluation. C'est par rapport aux trois dimensions évoquées dans l'introduction (dimensions cognitive, normative et instrumentale) que doit être appréciée la qualité d'une évaluation.

**Utilité-pertinence :** c'est le fait que l'évaluation produise de l'information compréhensible et utile, non seulement pour les décideurs mais aussi pour les autres protagonistes de la politique publique. Pour cela, le rapport d'évaluation doit répondre de manière directe et compréhensible aux questions posées dans le projet. Le critère peut également inclure le respect des délais. Enfin, est également pris en compte l'intérêt de l'apport cognitif pour la société. Les réflexions produites par l'évaluation peuvent en effet contribuer de manière diffuse à former le jugement de l'opinion publique sur la politique évaluée et à son information sur les problèmes de société directement liés à la politique. Elles peuvent également faire progresser la recherche et la réflexion politique sur les problèmes liés à la politique évaluée.

**Fiabilité :** l'évaluation doit être "digne de confiance" (*trustworthiness*). Cela renvoie à l'exactitude des informations collectées et à la valeur scientifique des imputations causales sur lesquelles se fonde le jugement évaluatif. Il convient en particulier d'être attentif aux biais que les techniques de collecte et de traitement des informations peuvent introduire dans la production des conclusions.

**Objectivité :** l'objectivité s'entend comme le fait que les conclusions de l'évaluation n'ont pas été influencées par les préférences personnelles ou les positions institutionnelles des responsables de l'évaluation (chargés d'étude ou membres de l'instance), ou du moins que ces préférences ont été suffisamment explicitées ou contrôlées pour que l'on puisse supposer qu'une autre évaluation répondant aux mêmes questions et utilisant les mêmes méthodes conduirait aux mêmes conclusions. Il s'agit notamment d'apprécier la rigueur et l'honnêteté du travail de *qualification* (telle évolution, par exemple, est-elle "rapide", "normale" ou "lente" ?) et d'interprétation des données qui permet de passer de l'observation au jugement.

**Possibilité de généralisation :** il s'agit de savoir dans quelle mesure les conclusions portées sur les effets d'une action publique à partir d'observations faites sur un certain type de terrain pendant une période donnée peuvent être étendues à d'autres situations, à d'autres contextes, voire à des politiques différentes. Le plus souvent, les résultats d'une évaluation sont en effet utilisés pour éclairer des décisions relatives à des actions et à des contextes qui différeront en bien des points de ceux sur lesquels l'évaluation a porté.

**Transparence :** outre l'exigence d'un exposé complet et rigoureux des méthodes mises en oeuvre, ce critère inclut l'idée que l'évaluation doit expliciter son propre "mode d'emploi" et ses limites : positionnement par rapport à d'autres évaluations

possibles sur le même sujet, rappel des questions auxquelles il n'est pas ou incomplètement répondu, anticipation des objections possibles, etc . Cet effort de lucidité et de réflexivité est d'autant plus nécessaire que les évaluations sont généralement imparfaites, qu'elles laissent de nombreuses questions dans l'ombre, et que les résultats auxquels elles parviennent n'ont pas tous la même portée ni la même robustesse.

## Deuxième partie : Les outils de l'évaluation

*Les évaluateurs peuvent avoir accès à l'ensemble des instruments habituels des sciences sociales. Ceux-ci sont exposés dans des guides, manuels et traités spécialisés propres à chaque discipline (économie, statistique, sociologie, notamment). On se contente dans ce qui suit d'évoquer quelques questions de base directement en rapport avec l'évaluation.*

### **5. Informations mobilisées et registre de l'argumentation**

Un rapport d'évaluation se présente comme le développement d'une argumentation dont le but est de répondre à un ensemble de questions et d'hypothèses relatives à la politique évaluée. Concernant la nature de cette argumentation, le pragmatisme est de règle : contrairement à une recherche scientifique menée dans un cadre disciplinaire déterminé, qui privilégie un nombre limité de raisonnements, l'évaluation cherche à utiliser toutes les "heuristiques", et à s'adapter aux conditions réelles de la délibération et de la prise de décision. Comme toute discussion ou réflexion menée à des fins pratiques, l'évaluation n'exclut a priori aucun élément d'information concernant son objet, quelle que soit sa nature (quantitative ou qualitative) et son origine, dès lors qu'il semble pertinent. Cet éclectisme n'est pas synonyme de confusion : contrairement à une discussion ordinaire, menée sans méthode formelle, l'évaluation s'efforce de ne pas mélanger les différents types d'arguments, mais de les hiérarchiser, de les pondérer, et de relier chacun à des conclusions spécifiques.

#### **5.1. Le débat quantitatif/qualitatif**

Il est fréquent d'opposer les méthodes d'évaluation selon qu'elles s'appuient principalement sur des informations quantitatives ou qualitatives.

- les informations quantitatives répondent à des questions sur la valeur prise par une variable, sur l'intensité d'un phénomène social (combien ? jusqu'à quel point ? quelle proportion ? quel taux de croissance ? etc). Elles prennent souvent la forme d'*indicateurs* (cf 3.2.2.).

- les informations qualitatives s'expriment par des mots : on a besoin de toute la richesse et de la précision du langage pour décrire les événements, les situations, les processus sociaux qui constituent la réalité d'une politique. Ces données sont tirées d'une grande variété de sources : entretiens, étude de cas, analyse de documents...

La distinction quantitatif-qualitatif n'a cependant rien d'absolu. Les données qualitatives peuvent être transformées en données quantitatives (par exemple grâce à l'analyse lexicale) et, réciproquement, les chiffres reposent presque toujours sur une conceptualisation préalable de la réalité (définition des variables, nomenclatures...).

Les évaluateurs, et les lecteurs d'un rapport d'évaluation, accordent spontanément beaucoup d'importance aux données quantitatives : elles semblent posséder un fort pouvoir d'objectivation de la réalité et se prêter mieux à la comparaison (dans le temps et dans l'espace) et à l'agrégation (globalisation de constats partiels). Quel que soit le phénomène social rencontré au cours d'une évaluation, la question de sa *mesure* peut être posée. Il convient cependant de souligner les *limites de la quantification* dans le domaine des sciences sociales : compte tenu des conventions qu'il faut adopter pour mettre en forme la réalité sociale, l'impression d'objectivité et de précision qui émane des chiffres est souvent illusoire. En outre, une tendance existe actuellement parmi les spécialistes de l'évaluation à revaloriser l'importance des résultats qui peuvent être obtenus sur la base d'informations qualitatives (y compris pour apprécier l'efficacité d'une politique).

## **5.2. Le statut des assertions : description, démonstration et opinion**

Qu'elles soient quantitatives ou qualitatives, les données s'intègrent dans des assertions qui visent à *qualifier la réalité* et à former le jugement des lecteurs du rapport. On peut distinguer trois registres principaux d'argumentation, qui contribuent de manière complémentaire à assurer la crédibilité des conclusions de l'évaluation.

### **5.2.1. Le registre du constat**

Toute évaluation développe d'abord des constats qui semblent simplement induits par l'observation directe de la réalité (même si tout constat comprend nécessairement un mixte d'inductions et de déductions) :

- constats quantitatifs : il peut s'agir de simples comptages physiques ou financiers : coût d'une politique ou d'un programme, nombre de bénéficiaires, de dossiers traités, etc. Cette quantification peut être obtenue par exploitation de données exhaustives (par exemple comme sous produit d'une gestion administrative), ou sur la base d'une enquête portant sur un échantillon statistiquement représentatif du phénomène analysé.

- constats qualitatifs : la description du processus socio-organisationnel initié par la

politique, en terme de dispositifs institutionnels, d'événements, d'activités, de situations et de comportements, voire d'anecdotes ou d'histoires personnelles, constitue toujours une composante importante de l'évaluation. Cette description est nécessaire à la bonne interprétation des données globales. De surcroît, elle constitue une source d'illustrations concrètes d'effets et de mécanismes analysés par d'autres méthodes.

### 5.2.2. Le registre démonstratif et interprétatif

Parmi les questions posées dans le projet d'évaluation, un certain nombre appellent des réponses plus élaborées que de simples constats : on veut tester des explications (pourquoi ?) ou un mode de compréhension (comment ?). Répondre à ce type de questions, consiste à faire une *démonstration* ou proposer une *interprétation théorique* des informations dont on dispose. Le registre démonstratif concerne le test d'hypothèses, la recherche et la discussion des preuves de la réalité d'un phénomène ou d'un mécanisme non directement observable, notamment lorsque l'on a l'ambition de mesurer les *effets propres de la politique*, qui, comme on l'a vu précédemment, ne se confondent pas avec les modifications constatées de l'état de la société. En ce domaine, la démonstration doit s'efforcer d'apporter la preuve la mieux testée d'une relation causale, et, si possible, sa quantification.

En pratique, la démonstration présente souvent un caractère moins formalisé. Pour répondre de manière raisonnée à certaines questions, il n'est d'autre solution que d'*interpréter* les données collectées (presque toujours lacunaires ou ambiguës) en faisant appel de manière éclectique à des raisonnements de diverses natures : rationalisation intuitive du comportement des acteurs ou schémas théoriques plus complexes inspirés de corpus scientifiques établis (théories économiques ou sociologiques).

Toute évaluation est, pour une bonne part, un travail sur les représentations, et il arrive fréquemment que l'un des principaux résultats obtenus soit une réinterprétation des rapports entre les objectifs et les stratégies de mise en oeuvre de la politique évaluée, ce qui conduit à une reformulation de la *théorie d'action*(1) des décideurs qui a pu inspirer la politique en question.

*Exemple : l'évaluation d'une politique de transfert de technologie au profit des entreprises (Région Rhône-Alpes) a conduit l'évaluateur à proposer une nouvelle manière de se représenter l'efficacité de la politique : "Le concept de "transfert de technologie a inspiré à son début la politique régionale. La ressource était dans l'offre publique et il fallait la passer à l'entreprise pour qu'elle la transforme en richesse économique. Cette vision linéaire a poussé la région à aider l'offre puis à solliciter la demande pour qu'elle s'adresse à l'offre. En fait, la réalité de l'innovation est plus complexe : les ressources technologiques n'existent pas en tant que telles, elles se créent dans un processus d'apprentissage entre une entreprise au contact du marché et des compétences externes qu'elle sollicite" (10).*

---

(1) L'expression "Théorie d'action" désigne les idées (souvent implicites) qui inspirent les concepteurs et/ou les acteurs d'une politique quant à ses mécanismes d'action et relations de cause à effet entre les mesures prises et leur impact social attendu.

### **5.2.3. Le registre de l'opinion**

Le point de vue des groupes d'intérêts concernés par une mesure constitue parfois la principale information disponible pour évaluer sa réussite :

*Exemple : une Direction départementale de l'équipement a évalué ses prestations aux communes en se fondant principalement sur l'opinion des élus locaux (11).*

Par ailleurs, l'opinion des acteurs chargés de mettre en oeuvre une politique est souvent indispensable pour apprécier sa cohérence et pour identifier les dysfonctionnements qui peuvent nuire à son efficacité. Quant à l'opinion des experts, elle joue un rôle déterminant dans tous les domaines où les normes de l'action publique reposent sur des connaissances techniques spécialisées.

Les opinions émises sur une politique sont rarement indépendantes des positions et intérêts défendus par leurs auteurs. Si importants que soient, dans la pratique, les arguments d'opinion, l'instance d'évaluation devra donc toujours s'efforcer de les corroborer par d'autres sources, et, surtout, éviter de reprendre à son compte sans précaution le point de vue d'une catégorie particulière d'acteurs ou d'experts (en particulier, pour ces derniers, lorsqu'ils ont été impliqués, directement ou non, dans l'orientation de la politique évaluée).

## **6. Les techniques de collecte de l'information**

### **6.1. Mobiliser l'information préexistante**

L'information susceptible d'être mobilisée dans le cadre d'une évaluation n'a rien de spécifique. Comme on l'a rappelé ci-dessus, il est nécessaire de procéder au bilan des informations existantes avant d'entreprendre de nouvelles investigations. Celles-ci peuvent être :

- des études et recherches portant sur des sujets connexes ou apparentés ;
- des statistiques publiées ou immédiatement disponibles ;
- des informations de toutes natures détenues par les gestionnaires de la politique.

L'évaluation est facilitée par l'existence de dispositifs de suivi, du type contrôle ou analyse de gestion. La nécessité de développer de tels outils fait d'ailleurs fréquemment partie des recommandations formulées à l'issue d'une évaluation.

### **6.2. Des enquêtes et investigations spécifiques**

Les données préexistantes répondent rarement à la totalité des questions que l'on est amené à se poser dans le cadre d'une évaluation. Ainsi, dans le cas d'une politique d'aide (subvention, prestation sociale), on peut s'attendre à ce que les administrations gestionnaires détiennent un fichier des bénéficiaires, comptabilisent les moyens

humains et financiers engagés et les actes administratifs effectués, etc, mais qu'elles ne disposent d'aucune information sur les effets produits par l'aide. On est donc presque toujours conduit à envisager des investigations spécifiques, dont les principaux types sont les suivants :

- études documentaires et historiques
- enquêtes statistiques
- entretiens
- études de cas ou monographies
- groupes d'acteurs ou d'experts
- auditions

### **6.3. Etudes documentaires et historiques**

Toute évaluation d'une politique présentant une certaine ancienneté doit commencer par une étude historique. Le but de cette étude est de connaître le contexte dans lequel les décisions concernant la politique ont été prises, afin de mieux comprendre la nature des objectifs poursuivis. Quelle que soit la nature de l'action évaluée, il est nécessaire de procéder à un recensement des textes officiels et des traces écrites des processus décisionnels.

La recherche documentaire, ou l'appel à la mémoire des acteurs, peut parfois s'intégrer dans le déroulement de l'évaluation, pour reconstituer une séquence d'événements dont la compréhension est nécessaire à l'analyse des mécanismes de la politique.

*Exemple : Dans le cadre de l'évaluation de la politique de réhabilitation du logement social, les participants aux "groupes locaux d'évaluation" se sont efforcés de reconstituer l'histoire des opérations de réhabilitation dont ils avaient été les protagonistes (12).*

### **6.4. Les enquêtes**

En fonction des moyens dont on dispose, du type d'information recherché et de la représentativité souhaitée, on procédera à une enquête par téléphone, par questionnaire postal ou par entretien. Ces différentes procédures peuvent être couplées. Il est envisageable, par exemple, d'interroger par téléphone les non-répondants à une enquête postale lorsqu'on veut contrôler le biais induit par les non-réponses.

#### **6.4.1. Structure et présentation du questionnaire.**

Quelle que soit la technique d'enquête choisie, celle-ci suppose l'élaboration d'un questionnaire. La réussite d'une enquête repose en partie sur la qualité du questionnaire : une grande attention doit être apportée à la formulation et à l'ordre des questions.

Un bon questionnaire comprend une majorité de questions fermées (qui n'admettent qu'un petit nombre de réponses, généralement moins de cinq), et quelques questions ouvertes, qui permettent à la personne interrogée de s'exprimer plus longuement et plus librement. Les questions fermées ont pour objet de constituer une information quantitative, les questions ouvertes une information plus qualitative mais pouvant cependant être exploitée. Chaque question fermée ne doit traiter que d'un seul problème, et doit être posée de manière discriminante (de manière à maximiser les chances de faire apparaître des contrastes au sein de la population).

On distingue quatre types de questions :

- les questions de souhait ou d'intention (ce que veut la personne interrogée) ;
- les questions de connaissance (ses informations et opinions) ;
- les questions de comportement (ce qu'elle fait) ;
- les questions d'attributs ou de statut (ce qu'elle est)(1).

La formulation des questions d'opinion ne doit pas induire de biais (ne pas suggérer implicitement une réponse) et respecter la sensibilité des personnes interrogées. Les réponses proposées doivent être équilibrées. Si l'on veut connaître, par exemple, la satisfaction d'une personne vis à vis d'une mesure, il faut que la personne interrogée ait le choix entre un nombre égal de réponses positives et négatives.

Pour être acceptable, un questionnaire ne doit pas être trop long, et il ne doit pas être perçu comme inquisitorial. Pour cette raison, il est préférable de ne pas placer les questions d'attributs en tête. De même, les questions les plus délicates doivent être placées vers la fin du questionnaire, juste avant les questions d'attribut.

#### **6.4.2. Le suivi d'un "panel"**

Le suivi d'un panel (typiquement, d'un groupe d'individus bénéficiant d'une mesure sociale) vise à appréhender les effets de l'action publique dans leur dimension temporelle. La technique consiste à répéter la même enquête auprès d'un même échantillon, à intervalle régulier (6 mois ou plus), pendant une durée totale de plusieurs années. On peut ainsi saisir l'impact dynamique d'une intervention publique sur la trajectoire sociale des individus. Outil de base pour l'évaluation des politiques sociales, le suivi d'un panel se heurte à plusieurs difficultés pratiques :

- les contraintes de temps et de coût ;
- la difficulté de constituer un échantillon représentatif et de conserver cette représentativité au cours du temps. Il faut en effet obtenir des personnes un accord de principe pour se soumettre à des interrogations répétitives. Il faut également prévoir des procédures de remplacement des personnes qui disparaissent du panel

---

(1) D'après Hervé DÉRY, *Industrie Canada, présentation au Congrès annuel de la Société Canadienne d'évaluation, mai 1994.*

suite à un changement de leur situation (déménagement, abandon du programme...). La solution peut consister à prendre un échantillon suffisamment important au départ. Dans le cas de l'évaluation du RMI, le panel des bénéficiaires est passé de 2000 à 1000 en deux ans, suite à ces diverses causes de défection.

Une alternative intéressante au panel peut consister à reconstituer *ex-post* la trajectoire des individus qui ont participé à un programme, ou qui participent à celui-ci depuis une durée suffisamment longue. La constitution d'un fichier peut poser des problèmes pratiques et déontologiques : il faut en effet nécessairement utiliser des fichiers de gestion nominatifs sur la base de critères souvent considérés comme confidentiels (par exemple, dans le cas de l'évaluation des politiques d'insertion des adolescents en difficulté, le fait d'avoir été pris en charge par la Protection Judiciaire de la Jeunesse plusieurs années auparavant). En outre, toute interrogation portant sur le passé pose des problèmes spécifiques de fiabilité (sélectivité de la mémoire, reconstruction subjective des événements...).

## 6.5 L'entretien

Le recours à l'entretien répond à différentes nécessités :

- préparer une enquête par voie téléphonique ou postale en testant et précisant les questions auprès d'un petit nombre d'individus ;
- améliorer la qualité des réponses à un questionnaire, ce qui est d'autant plus nécessaire que ce questionnaire porte sur une réalité complexe ;
- recueillir un matériau verbal non formaté.

Dans le cadre d'une évaluation, l'entretien est particulièrement utile pour interpréter le comportement des acteurs, appréhender leur représentation de la réalité, comprendre leurs objectifs personnels et la manière dont ils perçoivent ceux de l'action publique. Il s'agit notamment de mettre en évidence les repères normatifs implicites à partir desquels ils agissent et prennent leurs décisions. Cela vaut aussi bien pour les usagers et partenaires de la politique que pour les agents chargés de la mettre en oeuvre à différents niveaux.

Au début de l'entretien, l'interviewer doit en indiquer clairement le cadre et les finalités. Il doit ensuite s'efforcer d'en fixer le cadre général (la liste et l'ordre des sujets à aborder), tout en laissant une grande initiative à l'interviewé dans la manière d'aborder ces différents points. Il a souvent été observé que les lieux et circonstances de l'entretien peuvent influencer sur son déroulement : un entretien portant sur le travail ne donnera pas les mêmes résultats selon qu'il est mené au domicile de l'interviewé ou dans son bureau.

On distingue les entretiens directifs ou semi-directifs, selon le caractère plus ou moins structuré du questionnement et le temps laissé à la personne pour s'exprimer librement. La durée de l'entretien et le nombre de personnes interrogées sont très variables, selon, par exemple, que l'on a affaire à des acteurs clefs de la politique ou

à un échantillon de bénéficiaires.

## **6.6. Les études de cas**

L'étude de cas (ou monographie) peut concerner un ou plusieurs sites de mise en oeuvre d'une politique (le plus souvent), une opération, un événement (par exemple un incident critique), voire l'itinéraire d'une personne. L'étude est elle-même constituée de matériaux divers : données quantitatives, entretiens, documents de nature historique, etc. La spécificité de l'étude de cas réside cependant dans la possibilité qu'elle offre d'appréhender la dimension humaine et événementielle de la réalité sociale. Les études de cas sont particulièrement utiles pour analyser une situation complexe, par exemple lorsque la mise en oeuvre d'une politique publique résulte des initiatives combinées d'un grand nombre d'acteurs différents. L'étude de cas est alors la seule méthode qui permet de décrire finement le processus réel, d'identifier de manière concrète et détaillée ses mécanismes d'action et les conditions de sa réussite, de mettre en évidence les différents types d'effets produits.

Les risques que comporte la généralisation des observations effectuées ne doivent pas être sous-estimés. Une grande attention doit être portée au choix du ou des sites. Dans un protocole d'évaluation, la réalisation de monographies est souvent conçue comme le complément d'investigations plus systématiques (enquêtes statistiques).

Dans cette logique de complémentarité avec d'autres sources, l'approche monographique peut être utilisée pour étudier des "cas polaires", c'est à dire des situations présentant le maximum de contrastes, afin de fournir un éclairage étendu sur la diversité des phénomènes que l'on souhaite analyser. Elle peut aussi constituer la pièce essentielle d'une pré-étude d'évaluation, en aidant l'instance d'évaluation à formuler des hypothèses pertinentes et à repérer les ressources informationnelles disponibles.

## **6.7. Les groupes d'acteurs et d'experts**

### **6.7.1. Finalités**

Les techniques de travail en groupe peuvent être utilisées dans deux perspectives : la première est de dégager un consensus portant sur la réponse à des questions complexes auxquelles les travaux de recherche ne fournissent pas de réponse univoque (les participants ont alors le statut d'experts) ; la seconde est de caractériser de manière différentielle les attitudes et opinions de catégories particulières d'acteurs. Au cours d'une évaluation, les groupes d'experts ou d'acteurs peuvent être utilisés à différents stades, en amont ou en aval des investigations quantitatives. Leur principal intérêt par rapport à l'entretien est de recueillir en un temps relativement bref le point de vue de plusieurs personnes.

L'objet des techniques de travail en groupe est de neutraliser autant que faire se peut les effets psycho-sociaux liés au fonctionnement du groupe, tout en maximisant la

stimulation réciproque des participants. Il s'agit de mettre en situation chaque membre du groupe de restituer l'information dont il est porteur dans un langage qui fasse sens pour les autres. L'avantage pour l'évaluateur est alors de disposer d'un matériau conceptuellement homogène et pertinent.

### **6.7.2. Les différentes techniques**

Parmi les différentes techniques de travail en groupe, on peut citer :

- la méthode *Delphi* : consultation d'un groupe d'experts par voie postale avec interrogations successives visant à réduire progressivement les désaccords entre les personnes consultées.

- la méthode du *groupe nominal* : les différents experts sont réunis dans une même pièce et donnent tour à tour leur opinion sur les questions posées, mais ils ne communiquent pas entre eux.

- les *groupes de discussion* : on se contente d'enregistrer une discussion libre, sur un sujet donné, sans intervention de l'animateur, et on procède ensuite à une analyse de contenu.

- L'*Abaque de Régnier* : Le groupe, physiquement réuni dans un même lieu, est confronté à une liste de questions fermées auxquelles il répond de manière non verbale en utilisant un code de sept couleurs (deux nuances de vert pour l'accord, deux nuances de rouge pour le désaccord, l'orange pour l'hésitation, le blanc signifie que le participant manque d'information pour répondre à la question, le noir signifie qu'il récuse les termes de la question et refuse de participer à la discussion).

- la méthode des *groupes silencieux* consiste à réunir des groupes représentatifs de différentes catégories d'acteurs auxquels on demande de hiérarchiser par un système de notation un ensemble prédéfini d'affirmations relatives à un problème donné (par exemple un liste de causes du chômage). La comparaison des hiérarchies proposées par les différents groupes peut s'avérer particulièrement instructive.

### **6.8. La perspective ethno-sociologique**

La perspective ethno-sociologique repose sur l'idée que la verbalisation opérée rétrospectivement par les acteurs ne reflète que très imparfaitement ce qu'ils vivent et la manière dont ils agissent dans le contexte social de leur activité habituelle. Nombre d'attitudes et de gestes "routinisés" sont en effet spontanément considérés comme non pertinents par leurs auteurs, voire même échappent à leur conscience. La seule manière pour le chercheur d'appréhender la réalité d'une situation complexe impliquant des interactions sociales régies par des règles informelles est d'observer directement les processus relationnels qui en constituent la trame.

*Exemple : Dans le cadre de l'évaluation de l'accueil des populations défavorisées dans certains services publics, on a procédé à l'observation en situation des comportements et des échanges verbaux entre les usagers et les agents de guichet, afin d'apprécier la manière dont les services publics s'adaptent aux caractéristiques des publics accueillis (13).*

L'observation directe peut dans certains cas modifier sensiblement l'interprétation de certaines données. Ainsi, l'observation des attitudes réelles face à un téléviseur allumé (nombre de personnes qui suivent effectivement le programme, occurrence éventuelle d'autres activités telles que la lecture ou des tâches ménagères...) fournit une information très différente de celle que donne l'audimat. Il n'est pas toujours nécessaire pour le chercheur d'assister réellement aux scènes qu'il souhaite étudier. Le recours à des enregistrements (son ou image) peut s'avérer une solution plus praticable. La principale difficulté est alors d'obtenir l'accord des intéressés, tout en évitant que la conscience d'être observé modifie leur comportement.

## **7. Le traitement des informations par la statistique**

### **7.1. Définitions et principes méthodologiques**

Rappelons qu'on peut définir la statistique comme *"l'ensemble de techniques d'interprétation mathématique appliquées à des phénomènes pour lesquels une étude exhaustive de tous les facteurs est impossible à cause de leur grand nombre et de leur complexité"*. (Grand Robert).

Ces techniques font appel à la notion mathématique de *probabilité*, que nous ne chercherons pas ici à définir avec précision. La théorie des probabilités permet de conduire des raisonnements rigoureux et d'effectuer des calculs à propos d'événements incertains (par exemple, nous ne connaissons pas la date de notre mort, mais il est possible de calculer notre espérance de vie).

L'utilisation des statistiques dans le contexte de l'évaluation se justifie d'un double point de vue :

- l'incapacité de l'esprit humain à prendre en compte simultanément un trop grand nombre de données ;
- le fait que les processus sociaux n'apparaissent pas à nos yeux comme des processus déterministes (par exemple, suivre un stage de formation professionnelle ne garantit pas avec certitude que l'on va retrouver un emploi).

Le lien entre les deux aspects (nombre de données à prendre en compte, maîtrise de l'incertitude), est la *loi des grands nombres* qui fonde la possibilité de connaître plus précisément les caractéristiques d'un phénomène aléatoire en multipliant les observations (ce qui permet de prévoir, par exemple, qu'en jouant à pile ou face un grand nombre de fois, le rapport entre les piles et les faces se rapproche progressivement de 0,5). Cette loi est à la base de toutes les techniques d'*estimation*

et d'*inférence statistique*.

Dans la pratique, trois types de mise en oeuvre des techniques statistiques peuvent être rencontrés : la description résumée des données, l'induction et le test d'hypothèses théoriques.

### **7.1.1. Description résumée des données**

Le niveau élémentaire de la pratique statistique est la *statistique descriptive*, c'est à dire les techniques qui permettent de résumer et de présenter de manière synthétique l'information contenue dans un ensemble de données. Ces techniques recouvrent notamment le calcul et l'interprétation des paramètres caractéristiques de la distribution des valeurs prises par une variable observée : paramètres de centralité (moyenne, médiane), et paramètres de dispersion (écart type, quantiles, etc), ainsi que les techniques graphiques (tableaux, histogrammes, graphiques...).

### **7.1.2. Analyse des données et typologies**

Face à un ensemble de données, différentes techniques (analyse factorielle, classification hiérarchique) permettent de rechercher une structure sous-jacente : par structure, on entend soit une typologie, soit des rapports de corrélation entre variables.

*Exemple : l'évaluation d'une aide de l'Agence Nationale pour la Valorisation de la Recherche a montré que les entreprises aidées sont de plus grande taille et plus dynamiques que la moyenne (14).*

Une typologie obtenue par l'analyse des données n'est pas seulement utile pour décrire commodément une réalité complexe, elle peut également, parfois, servir de base à la formulation d'une nouvelle hypothèse.

*Exemple : une étude portant sur les jeunes fréquentant des missions locales a dressé une typologie de "parcours d'insertion" en appliquant à un corpus de récits suscités et mis en forme par les conseillers des missions, des méthodes d'analyse lexicale et de classification hiérarchique. On en tire un nouvel éclairage sur les causes de l'exclusion professionnelle et sur la pertinence des réponses institutionnelles qui lui sont apportées. La typologie des unités textuelles distingue par exemple deux catégories de situations initiales -"faible milieu scolaire, milieu défavorisé" et "conflits familiaux et sociaux" qui traduisent le repérage par les conseillers de besoins différents chez les jeunes (15).*

### **7.1.3. Techniques d'inférence et vérification probabiliste des hypothèses**

Les techniques statistiques peuvent être utilisées pour tester empiriquement la validité d'une relation suggérée par des raisonnements théoriques ou l'intuition, c'est à dire vérifier qu'elle n'est pas contredite par la réalité observée. En pratique, on ne

dispose pour cela que d'un nombre limité d'observations, et on ne peut donc parvenir à aucune certitude. Il est toujours possible, en théorie, que l'échantillon (d'entreprises, d'individus, etc) ne soit pas *représentatif*, c'est à dire qu'il ne reflète pas fidèlement la réalité dont il a été extrait. Le respect des règles de l'art (tirage aléatoire, représentativité géographique, méthode des quotas, etc) permet de réduire ce risque, mais, même dans le meilleur des cas, un doute subsiste. A défaut de certitude, la statistique permet de quantifier le *degré de confiance* qui peut être accordé à telle ou telle *inférence* (estimation d'une grandeur ou d'une relation, affirmation d'un lien de causalité) déduite d'un nombre limité d'observation. C'est l'objet des tests statistiques (Test du Chi<sup>2</sup>, Test de Student...), couramment employés par les praticiens, mais dont la présentation rigoureuse et la justification théorique sont relativement complexes. Il importe cependant d'en bien saisir la signification : quand on dit par exemple que "la différence (pour un indicateur donné) entre deux groupes est statistiquement significative au niveau 0.05", cela veut dire que, compte tenu des observations effectuées, il y a moins de 5% de chance pour que le phénomène observé ait la même distribution dans les deux groupes.

La théorie statistique permet donc de donner un contenu précis et quantitatif à la notion de représentativité. Dans le contexte de l'évaluation, on est cependant souvent amené à se satisfaire d'une conception plus "qualitative" de la représentativité. A l'extrême, il n'est pas rare que l'on cherche à tirer des enseignements de portée générale sur la base d'un petit nombre d'études de cas. La validité d'une généralisation dépend autant de la variabilité du phénomène observé que du nombre d'observations : un même phénomène observé dans cinq départements de caractéristiques contrastées a des chances de se reproduire dans la grande majorité des départements. Un échantillon de petite taille qui reflète les principaux contrastes dans les situations que l'on souhaite observer suffit souvent pour tirer d'utiles conclusions.

*Exemple : les dix groupes locaux constitués dans le cadre de l'évaluation de la politique de réhabilitation du logement social ont abouti à des conclusions suffisamment convergentes pour que l'on puisse les considérer comme représentatives de la situation de l'ensemble du pays (12).*

## **7.2. Les cadres de rangement de l'information**

### **7.2.1. Nomenclatures et concepts.**

Toute mesure repose sur une structuration préalable de la réalité, définition d'unités statistiques, nomenclatures, etc. La construction de ce cadre conceptuel résulte d'un compromis entre les différentes représentations sociales de la réalité que l'on veut mesurer et les exigences de clarté et de permanence liées à la pratique de l'observation statistique ou administrative (que l'on pense par exemple à la définition d'un "chômeur", d'une "agglomération", ou d'un "quartier", ou encore de la définition d'un indicateur de "taux de délinquance"). Dans le contexte de l'évaluation, il faut s'efforcer de construire des données qui reflètent la nature du problème à traiter par

la politique, même si ce problème est formulé de manière volontairement floue et ambiguë dans l'exposé des objectifs de la politique. Bien souvent, notamment dans le domaine social, une politique est identifiée par sa "population cible" (les bénéficiaires présumés : par exemple, les jeunes en difficulté), sans que ses concepteurs aient vérifié que la catégorie de population ainsi désignée a une quelconque pertinence et une signification claire pour les acteurs sociaux et qu'elle est repérable par les systèmes d'information existant. Les statistiques fournies par les administrations reflètent le fonctionnement des institutions plus que l'évolution des problèmes qu'elles ont à traiter.

*Exemple : lors de l'évaluation de l'accueil des populations défavorisées dans certains services publics, il est vite apparu que la catégorie "populations défavorisées" visée par les promoteurs de l'évaluation ne correspondait à aucun regroupement de catégories administratives et statistiques connues (étrangers, chômeurs de longue durée, personnes âgées...) (13).*

Dans le même ordre d'idée, l'analyse du comportement des acteurs (cf 8.3.2.) conduit fréquemment à l'établissement de nouvelles distinctions au sein de catégories constituées *a priori*. Il peut se faire, par exemple, que l'attitude des élus locaux vis à vis d'une politique devienne un critère pertinent pour définir des catégories de communes répondant aux besoins de l'évaluation.

### **7.2.2. Cohérence des nomenclatures, "triangulation" des données.**

Cette structuration du système d'information doit viser la mise en cohérence, ou la confrontation, des différentes sources d'information. Le rapprochement d'informations d'origines différentes, quantitatives ou qualitatives, portant sur les mêmes unités statistiques et sur les mêmes problèmes, peut se révéler particulièrement fructueux. A cette fin, les concepts et nomenclatures, le cadre de restitution de l'information, les modalités et les dates d'observation, doivent être harmonisés et adaptés au questionnement évaluatif

*Exemple : L'évaluation du "Revenu minimum d'insertion". Peu de temps après le vote de la loi du 1<sup>o</sup> décembre 1988 instituant le dispositif et prévoyant son évaluation, la Commission nationale d'évaluation du RMI est mise en place par le décret du 6 septembre 1989. Elle est chargée "d'apprécier les effets du RMI au regard des objectifs fixés par la loi". Pour remplir sa mission, la Commission a pu s'appuyer sur un groupe de coordination scientifique animé par le Commissariat général du plan. Les discussions au sein de ce groupe ont fait apparaître très vite à l'ensemble des acteurs concernés (chercheurs, chargés d'étude, statisticiens...) la nécessité de disposer de statistiques de référence conceptuellement pertinentes et surtout comparables d'un organisme à l'autre. La commande passée par la Commission nationale d'évaluation a accéléré la mise en place de ces systèmes d'information, qui ont été structurés par deux instruments : un "module de coordination statistique" (définition des concepts et de nomenclatures), et un*

*"dictionnaire des sources statistiques sur le RMI" (présentation synthétique des sources disponibles par thème) (16).*

Le couplage entre une enquête statistique par questionnaire et un nombre limité d'entretiens approfondis ou d'études de cas (cf 6.6.) portant sur un sous-échantillon de la même population constitue un exemple typique de mise en cohérence de plusieurs sources d'information complémentaires. L'enquête statistique fournit un cadrage quantitatif, les études de cas une vision plus qualitative des phénomènes étudiés qui permet d'interpréter les résultats de l'enquête. C'est ce qu'on appelle parfois la "triangulation" des données.

***Exemple:** L'évaluation, au Canada, d'un programme de prévention des abus et négligences envers les enfants. Dans un premier temps, on a procédé à une enquête par questionnaire auprès d'un échantillon de 175 familles et d'un échantillon témoin équivalent, en utilisant une grille quantitative censée mesurer la qualité de l'interaction parents enfants. Dans un second temps, on a réalisé 55 entretiens approfondis au domicile des familles (17).*

## **8. Le problème de la causalité et l'analyse des systèmes complexes**

L'identification et la mesure des effets de l'action publique constitue souvent, pour le commanditaire, l'enjeu principal de l'évaluation. C'est aussi, en général, la question la plus difficile à traiter. Toute action visant à modifier l'état de la société interfère avec les évolutions endogènes de celle-ci, et avec une multitude d'autres facteurs exogènes, y compris l'impact d'autres actions publiques. Dans la pratique de l'évaluation, on a souvent affaire à des systèmes complexes au sein desquels il est difficile de quantifier des relations de causalité. Le but que doit se fixer l'évaluation est alors de comprendre le fonctionnement de ces systèmes, afin de décrire l'enchaînement des mécanismes qui aboutissent aux effets constatés.

### **8.1. Les difficultés de l'expérimentation**

#### **8.1.1. Position du problème**

Pour prouver et quantifier une relation de cause à effet entre une politique et certaines évolutions constatées de l'état de la société, il faudrait théoriquement pouvoir répondre à la question : *que se serait-il passé si, toutes choses égales d'ailleurs, la politique n'avait pas été mise en oeuvre ?*

Poser le problème en ces termes oriente immédiatement vers l'idée d'*expérimentation* : en toute rigueur, seule une expérience *reproductible* à

*l'identique* mettant en parallèle deux situations qui ne diffèrent que sous l'angle du phénomène étudié permettrait de se prononcer sur l'"effet propre" d'une action sur la réalité. Tel est, on le sait, le modèle de la vérification d'une hypothèse dans les sciences de la nature. Notons que ce mode d'établissement de la preuve reste valide si la relation entre la cause et l'effet observé n'est pas strictement déterministe mais seulement "probabiliste", comme c'est souvent le cas en biologie (compte tenu de la variabilité des caractéristiques individuelles des êtres vivants).

### 8.1.2. Les limites de la méthode expérimentale dans les sciences de l'homme et de la société

Rien n'interdit, en principe, de transposer cette démarche à l'évaluation d'un programme public. En pratique, on doit alors constituer deux échantillons aléatoires de "cibles" du programme (personnes, entreprises...), l'un d'entre eux étant soumis au "traitement" défini par celle-ci (la participation à une activité de formation, le bénéfice d'une aide), et l'autre non. Les paramètres représentatifs de l'effet attendu (par exemple, les compétences acquises dans tel ou tel domaine) sont mesurés pour les deux échantillons *avant et après* le traitement. La comparaison de l'évolution constatée de ces paramètres (sur la base de tests statistiques adéquats) permet alors de se prononcer sur l'effet propre du programme.

Bien qu'elle soit assez largement mise en oeuvre dans les pays anglo-saxons pour évaluer les programmes sociaux et éducatifs, l'expérimentation avec groupe de contrôle soulève, dans la plupart des situations réellement rencontrées, des difficultés qui en limitent la portée. Il est rare que l'on puisse raisonner "toutes choses égales d'ailleurs", c'est à dire comparer deux situations sur lesquelles toutes les variables exogènes agissent exactement de la même manière. Les limites et obstacles rencontrés sont de diverses natures :

- l'expérimentation peut poser un problème éthique : a-t-on le droit de faire jouer le hasard lorsque l'intérêt des individus est en jeu ? le problème est classique dans le cas de l'expérimentation médicale (et il a fait l'objet d'un traitement juridique spécifique), mais il existe également pour des programmes sociaux. Dans le contexte français, la force du principe d'égalité devant le service public et l'attachement aux droits acquis interdisent pratiquement d'attribuer des droits sociaux à titre expérimental.

- En dehors de toute considération éthico-juridique, l'impossibilité de constituer de manière aléatoire deux groupes comparables tient le plus souvent à la nature même de l'action évaluée. Il est très fréquent, par exemple, que le bénéfice d'une aide publique soit accordé à des individus (ou à des entreprises) qui le demandent expressément, et que les candidats soient sélectionnés par l'administration selon des critères qui interdisent pratiquement la constitution d'un groupe témoin comparable. Ce mode de recrutement induit un *biais de sélection* impossible à neutraliser.

*Exemple : les entreprises qui demandent à bénéficier d'une aide à la modernisation (du type Fonds de Développement des Petites et Moyennes Industries), sont sélectionnées en fonction de la qualité économique et technique de leurs projets. Les entreprises bénéficiaires présentent donc dès le départ des caractéristiques spécifiques qui les différencient des entreprises non bénéficiaires (cf annexe).*

- Un autre problème est celui de la *reproductibilité*. La généralisation des résultats d'une "expérience sociale" à la mise en oeuvre "routinisée" d'un programme

suppose que l'on contrôle l'effet du contexte particulier de cette expérience. Or, s'agissant d'un processus social, c'est très difficile. Le comportement des acteurs impliqués dans la mise en oeuvre du programme est une variable à la fois particulièrement importante et particulièrement susceptible d'être influencée par les circonstances. Le fait de participer à une expérience est un facteur de mobilisation, voire de "surmotivation" qui améliore sensiblement les résultats du programme. C'est ce qu'on appelle l'"effet Hawthorne", que l'on caractérise parfois de la formule "toute expérience commence par réussir".

- L'expérimentation pose enfin des problèmes de *délais*, et donc (notamment) de coût.

Pour toutes ces raisons, le schéma expérimental est rarement appliqué avec toute la rigueur requise, et il ne constitue pratiquement jamais, du moins en France, la pièce essentielle du protocole d'évaluation d'une politique publique. Une exception doit toutefois être faite pour le domaine éducatif, qui présente, du point de vue de la mise en oeuvre d'un protocole expérimental, plusieurs caractéristiques favorables :

- la taille des populations est importante ;
- les élèves n'ayant pas la possibilité de choisir leur établissement en fonction du programme qui y est expérimenté, il n'y a pas de biais de sélection ;
- le ministère de l'Education nationale réalise depuis plusieurs années une évaluation systématique du niveau acquis dans les matières de base (lecture et mathématiques) à l'entrée en CE2 et à l'entrée en sixième, ce qui fournit un matériau de base pour effectuer des comparaisons rigoureuses entre différentes sous-populations.

*Exemple : dans le cadre de l'évaluation de la politique d'aménagement des rythmes de vie des enfants, il a été procédé à une comparaison de l'évolution des acquis des élèves dans les matières de base entre deux échantillons d'établissements comparables et tirés au hasard, l'un composé d'établissements bénéficiant de la politique et l'autre d'établissements n'en bénéficiant pas. Les résultats obtenus ont établi de manière robuste l'absence d'effet significatif de la politique sur le niveau scolaire moyen des élèves, tout en suggérant un effet faiblement positif pour les enfants les plus en difficulté (8).*

La méthode expérimentale pourrait cependant être utilisée dans d'autres domaines, pour évaluer des actions d'ampleur limitée qui n'affectent le comportement et la situation de leurs "cibles" que sur un point très précis. On peut penser, entre autres, aux actions de communication en matière de prévention, à l'efficacité de mesures de contrôle, etc.

*Exemple : le département de l'environnement du New Jersey a évalué grâce à un protocole expérimental l'efficacité relative de plusieurs stratégies de communication d'un message de prévention sanitaire (concernant en l'occurrence la pollution*

domestique par le radon) (18).

### 8.1.3. Protocoles pseudo-experimentaux

La plupart du temps, pour des raisons pratiques, les comparaisons effectuées dans les évaluations portent sur des *groupes non équivalents*, dont le mode de constitution n'est pas neutre vis à vis des effets attendus du programme. C'est notamment le cas lorsqu'on compare la situation d'entreprises ayant bénéficié d'une aide publique avec des entreprises comparables qui n'en ont pas bénéficié. Il est généralement admis que le fait de participer au programme (bénéficiaire d'une aide), résulte d'une trajectoire ou d'un comportement particulier corrélé avec les évolutions que l'on cherche à favoriser (par exemple, ce sont, en règle générale, des entreprises plutôt dynamiques qui bénéficient des aides à l'innovation).

Il est clair que l'on ne peut tirer aucune conclusion définitive sur l'efficacité propre d'un programme en comparant l'évolution de deux groupes non équivalents. La comparaison de bénéficiaires et de non bénéficiaires est néanmoins utile pour d'autres aspects de l'évaluation : connaissance du mode de sélection des dossiers, motivations des bénéficiaires, mécanismes d'action, et, donc, identification des leviers d'amélioration de l'efficacité. Dans certaines situations favorables, il est possible d'analyser rigoureusement l'effet du programme dans le cadre d'un protocole pseudo-expérimental, en modélisant l'impact des principales variables exogènes qui n'ont pu être neutralisées par un tirage aléatoire. On est alors ramené aux problèmes de la modélisation évoqués ci-dessous.

## 8.2. Causalité et modélisation

Lorsque l'expérimentation est impossible, la principale méthode alternative pour prouver et quantifier une relation de causalité est l'estimation statistique d'un *modèle mathématique* schématisant l'action de la politique sur la société.

### 8.2.1. L'estimation d'un modèle

Par modèle, on entend ici une relation mathématique -dont la forme est définie *a priori* sur la base d'arguments théoriques- entre les effets attendus de la politique et différentes variables explicatives dont l'une représente l'action du programme. Pour donner une idée simple de ce que l'on entend par *estimation* d'un modèle, on peut dire qu'il s'agit du *calcul des valeurs les plus probables* des paramètres (ou coefficients) du modèle au vu des valeurs observées des différentes variables. Les techniques d'estimation des modèles font appel à des notions mathématiques relativement complexes.

Les seuls modèles dont les techniques d'estimations sont parfaitement maîtrisées sont les modèles linéaires (de la forme  $Y=a_1X_1+a_2X_2+a_3X_3\dots$ , Y étant la variable expliquée,  $X_1, X_2\dots$  les variables explicatives,  $a_1, a_2\dots$  les coefficients que l'on cherche

à estimer). Ces modèles sont très largement utilisés en économie, d'où le vocable *économétrie* parfois utilisé pour désigner cette branche particulière de la

statistique mathématique. On parle également de *régression multiple*, ou d'*analyse multivariée*. Par différents artifices (changement de variables), il est possible de traiter une large gamme de modèles par les méthodes d'estimation des modèles linéaires.

Pour modéliser un phénomène, il faut bien entendu disposer d'un ensemble de données qui s'y prête. Dans le contexte de l'évaluation, les deux principaux types de données rencontrées sont les *séries temporelles de données agrégées* (par exemple les séries économiques tirées des comptes nationaux) et les *données individuelles* (les mêmes variables sont observées pour un grand nombre d'individus). Dans le cas des séries temporelles, la variable représentant l'effet du programme est généralement, comme les autres variables, une variable continue.

*Exemple : pour évaluer l'effet d'une mesure d'abaissement du coût du travail sur l'emploi, on peut utiliser une relation entre l'emploi et le coût du travail estimée sur des séries temporelles agrégées.*

Dans le cas des données individuelles, la variable représentant l'effet de la politique est une *variable binaire* (prenant arbitrairement la valeur 1 pour les individus participant au programme, 0 dans le cas contraire). Le coefficient de la variable binaire, s'il est statistiquement significatif, fournit alors directement une mesure de l'effet du programme. Dans le cas le plus fréquent, les données de base sont issues d'un *panel* regroupant des individus participant ou non au programme.

*Exemple : L'évaluation des zones d'éducation prioritaires (ZEP) par la Direction de l'évaluation et de la prospective (DEP) au Ministère de l'Education Nationale, utilise ce type de méthode. Les données sont tirées de deux échantillons nationaux de 2000 élèves entrant en CE2 et en 6ième, dont on a mesuré le niveau en français et en mathématique, et pour lesquels sont renseignées les principales variables socio-démographiques susceptibles d'influencer de manière évidente les résultats scolaires, ainsi que l'appartenance ou non de l'établissement à une ZEP. L'application du modèle linéaire général permet d'obtenir un ensemble de relations entre le niveau en mathématique et en français et les différentes variables explicatives, chaque coefficient estimé représentant l'influence d'une variable en supposant les autres variables fixées à leur valeur de référence<sup>1</sup>. Si l'appartenance à une ZEP semble peser significativement et de manière négative sur les résultats en français et en mathématiques au niveau CE2, cet effet semble s'estomper "toutes choses égales d'ailleurs" au niveau de l'entrée en 6ième, soit deux ans plus tard s'il s'agissait des mêmes individus. On peut donc raisonnablement en déduire que l'effort pédagogique particulier réalisé dans les zones d'éducation prioritaires aboutit bel et bien à une réduction du handicap scolaire des enfants issus d'un milieu défavorisé qui en bénéficient (19).*

---

<sup>1</sup> "L'évaluation des zones d'éducation prioritaires", dossiers Education et formation n°14, Ministère de l'Education nationale, Septembre 1992, Paris.

## 8.2.2. Les pièges de la modélisation

Les tests statistiques (cf 7.1.3.) permettent de vérifier que le modèle estimé fournit une représentation de la réalité compatible avec les observations effectuées. Il importe cependant de remarquer qu'une bonne adéquation du modèle à l'observation (même lorsqu'on peut vérifier son pouvoir prédictif) ne constitue en aucun cas une preuve suffisante de la validité des relations causales qu'il contient.

*Exemple : Un lecteur (ingénieur en retraite) écrivait récemment à un quotidien la lettre suivante :*

*"Les statistiques montrent qu'il y a une relation entre les taux d'épargne, de croissance et de chômage. On peut représenter cette relation par la formule empirique :*

*augmentation annuelle du chômage = 0.075 (taux d'épargne - taux de croissance/0.15)*

*Suivent différentes courbes montrant que cette formule fournit une taux de chômage calculé en bon accord avec le chômage observé. Et le lecteur de conclure : "D'après cette formule, pour réduire le chômage, on pourrait agir sur le taux d'épargne ou sur le taux de croissance. Le premier moyen est plus facile". Suivaient des considérations sur les moyens de réduire les taux d'intérêt pour agir sur le taux d'épargne.*

Cet exemple présente l'intérêt de mettre en évidence les principales chausse-trappes de la modélisation. Ce que l'auteur cité n'a pas compris c'est qu'une *corrélation ne doit pas être confondue avec un lien de causalité*. Plus précisément, dans ce cas, bien d'autres hypothèses peuvent être envisagées que celle mentionnée par l'auteur (ce qui n'interdit pas de lui accorder une part de vérité), par exemple :

- l'augmentation du chômage peut être la cause de la hausse du taux d'épargne, et non sa conséquence ;
- les deux grandeurs peuvent être influencées par des variables non prises en compte dans le modèle, ce qui peut expliquer leur corrélation.

En bref, les méthodes d'estimation statistique ne donnent des résultats fiables que lorsqu'on les applique à un modèle *convenablement spécifié*, cohérent et reposant sur de solides arguments théoriques (ce qui signifie, notamment, que l'on est certain d'avoir pris en compte les principales variables explicatives, et que celles-ci ne sont pas influencées par la variable expliquée).

## 8.3. Les limites de l'approche causale et l'analyse des systèmes sociaux

### 8.3.1. Le problème de la circularité

En dehors des difficultés pratiques de l'expérimentation et de la modélisation, l'approche causale en matière de phénomènes sociaux se heurte à un obstacle plus fondamental : la fréquence des phénomènes de *circULARITÉ*. Par circularité, on entend deux phénomènes en relation pouvant être simultanément cause et effet l'un de l'autre.

*Exemple : l'évaluation des effets structurants de l'autoroute Paris-Bordeaux. "Les évaluateurs ont réalisé plusieurs monographies concernant des villes situées aux abords de l'autoroute. L'hypothèse de causalité testée était la suivante : plus la ville est proche de l'autoroute, plus le développement économique est important. Pour commencer, l'évaluateur a découvert que les villes les plus proches avaient connu un développement plus vigoureux et un meilleur dynamisme politique. Mais il est rapidement apparu que ces villes disposaient d'un bon potentiel de développement local avant la construction de l'autoroute. Elles étaient donc mieux préparées à profiter des avantages offerts par l'autoroute en combinant des investissements publics et privés adaptés. De plus, on a remarqué que le tracé de l'autoroute avait été lui-même influencé par le dynamisme de ces villes. Dans ce cas, l'autoroute apparaît comme un effet et non plus comme une cause. Lorsque les causes et les effets sont aussi étroitement imbriqués, ils forment un système circulaire qui interdit toute analyse d'impact (20)."*

Les systèmes circulaires simples du type de l'exemple précédent ne constituent eux-mêmes qu'un cas particulier. La plupart du temps, on a affaire à des systèmes relationnels complexes, dans lesquels les "boucles de rétroaction" peuvent être multiformes et difficiles à identifier. C'est très généralement le cas du système social constitué par les acteurs d'une politique publique, dont les intérêts, attitudes et décisions dépendent mutuellement les uns des autres selon des configurations qui diffèrent selon la politique considérée.

### **8.3.2. L'analyse des systèmes d'action**

Toute politique publique mobilise un nombre plus ou moins élevé d'individus, de groupes, d'institutions. Les acteurs ne se comportent pas nécessairement de façon passive et neutre. Au contraire, leurs savoir-faire, leurs routines, leurs intérêts, leurs modes de relation d'échange et d'interdépendance exercent un impact tout à fait significatif sur les objectifs, sur les réalisations, sur les effets générés par l'intervention publique, notamment au moment de sa mise en oeuvre sur le terrain. Il s'agira donc pour l'évaluation d'en comprendre de façon concrète la nature et la portée.

*Exemple : le transport routier de fret. Voulant combattre les accidents de la route, la puissance publique contrôle réglementairement les conditions techniques et sociales de la conduite des poids lourds. Or, pour les véhiculeurs, et au quotidien, la fraude s'avère fournir une solution -en termes de coûts, donc de prix- pour gérer un problème crucial -trouver du fret dans un contexte de dépendance forte à l'égard des auxiliaires et des affréteurs qui, pour leur part, échappent largement à l'emprise*

*de la réglementation. D'autre part, les forces de police et les corps administratifs chargés d'appliquer les règles adoptent des comportements plus ou moins stricts, préférant dans certains cas assouplir les sanctions plutôt que d'en imposer la lettre. Le véhiculaire, qui forme la population-cible de l'action étatique, est en fait un acteur dont la rationalité dépend d'un système d'action économique sur lequel l'intervention publique n'a pas de prise (21).*

Identifier les acteurs qui comptent pour la réussite ou pour l'échec d'une politique publique, expliquer comment leur rationalité et leurs logiques d'action spécifiques affectent le cours des choses, constituent des tâches et des étapes significatives de l'évaluation. Des modes d'analyse existent, grâce notamment à l'apport de la sociologie des organisations. Celle-ci traite une politique publique comme une intervention sur un tissu social dont elle repère les divers acteurs, analysant leurs relations mutuelles et les règles concrètes qui structurent les comportements des acteurs, par delà les classifications institutionnelles ou statistiques. L'étude des ces *systèmes d'action*, notamment par le recours à la technique du *sociogramme* et à l'aide de la *grille d'analyse des stratégies* comportementales, formalise la manière dont les parties prenantes gèrent leurs interdépendances et leur intervention dans le champ de la politique publique. Elle permet aussi de recenser si et comment de tiers acteurs, ne relevant pas formellement du dispositif public d'intervention, jouent en fait un rôle décisif pour les effets réellement générés par ce dispositif. Elle s'appuie principalement sur l'observation des pratiques sur le terrain, collectant des données qualitatives au moyen de techniques d'intervention en profondeur, d'analyse documentaire et d'observation participante auprès de groupes d'acteurs interdépendants.

## Index des principaux termes techniques

Abaque de Régnier, 6.7.2.  
Analyse des données, 7.1.2.  
Analyse multivariée, 8.2.1.  
Argumentation (registres d'), 3.3.2., 5.5.2.  
Atteinte des objectifs, 1.7.  
Audit, 1.5.  
Biais de sélection, 8.1.2.  
Bilan informationnel, 3.1.9.  
Cahier des charges, 2.3.  
Causalité, 8.  
Chargés d'évaluation, 3.1.11.  
Circularité, 8.3.1.  
Cohérence, 1.7.  
Commanditaire, 2.2.1.  
Conditions de généralisation, 3.2.4.  
Constat, 5.2.1.  
Contexte de la mise en oeuvre, 3.2.4.  
Contrôle, 1.5.  
Coût-avantages, 3.2.2.  
Coût-efficacité, 3.2.2.  
Coût d'opportunité, 3.2.2.  
Critères de qualité d'une évaluation, 4.4.  
Démonstration, 5.2.2.  
Delphi (méthode), 6.7.2.  
Dispositif d'évaluation, 3.1.8.  
Économie publique, 3.2.2.  
Économétrie, 8.2.1.  
Effets propres, 3.2.2.  
Efficacité, 1.7., 3.2.2.  
Efficience, 1.7., 3.2.2.  
Enquêtes, 6.4.  
Entretien, 6.5.  
Ethno-sociologie,  
Étude de cas, 6.6.  
Étude documentaire, 6.3.  
Évaluateur, 2.3.  
Évaluation (démocratique, formative, managériale, récapitulative), 1.6.  
Évaluation (ex ante, concomitante, ex post), 2.2.3.  
Expérimentation, 8.1.  
Faisabilité, 2.4., 4.4.1.  
Fiabilité, 4.4.3.  
Groupe nominal, 6.7.2.  
Groupes silencieux, 6.7.2.  
Impact, 1.7., 3.2.2.  
Inférence statistique, 7.1.  
Initiateur, 2.2.1.  
Instance d'évaluation, 4.1.

Instance de régulation scientifique, 4.1.4.  
Mandat, 2.3.  
Mécanismes d'action, 3.2.3.  
Méta-évaluation, 4.4.1.  
Mise en vigueur (*implementation*), 3.2.1.  
Modèles, 8.2.1.  
Monographie, 6.6.  
Nomenclatures, 7.2.1.  
Objectifs (de la politique), 3.1.2., 3.1.5.  
Objectivité, 4.4.3.  
Opinion, 5.2.3.  
Panel, 6.4.2.  
Pertinence, 1.7.  
Porteur d'enjeu (*stakeholder*), 2.2.1.  
Politique, 2.2.2.  
Possibilité de généralisation, 4.4.3.  
Programme, 2.2.2.  
Projet d'évaluation, 2.3., 3.  
Pseudo-expérimental (protocole), 8.1.3.  
Qualitative(information), 5.1.  
Quantitative (information), 5.1.  
Questionnaire, 6.4.1.  
Questionnement, 3.1.7.  
Rapport d'évaluation, 4.3.  
Recherche évaluative, 1.5., 2.1.1.  
Référentiel, 1.4., 3.1.5.  
Registres de l'évaluation, 3.1.6.  
Régression multiple, 8.2.1.  
Reproductibilité, 8.1.2.  
Sociologie des organisations, 8.3.2.  
Sociogramme, 8.3.2.  
Tests statistiques, 7.1.3.  
Théorie d'action, 5.2.2.  
Transparence, 4.4.3.  
Triangulation, 7.2.2.  
Utilité-pertinence, 4.4.3.  
Variable binaire, 8.2.1.

## Références des exemples cités

- (1) Comité interministériel de l'évaluation, Commissariat général du plan, *L'insertion des adolescents en difficulté*, La Documentation française, 1993.
- (2) Évaluation de la politique de la montagne, décidée par le Comité interministériel de l'évaluation en 1993, non encore terminée.
- (3) Comité interministériel de l'évaluation, Commissariat général du plan, *Les zones humides*, La Documentation française, 1994.
- (4) Comité interministériel de l'évaluation, Commissariat général du plan, *L'informatique de l'État*, La Documentation française, 1992.
- (5) Conseil scientifique de l'évaluation, *L'évaluation en développement 1993*, La Documentation française, p. 151.
- (6) *RMI, le pari de l'insertion*, rapport de la Commission nationale d'évaluation, La Documentation française, 1992.
- (7) Évaluation des politiques d'aides à la localisation des activités dans les zones de conversion, Avis du Conseil économique et social, Juillet 1994.
- (8) Comité interministériel de l'évaluation, Commissariat général du plan, *L'aménagement des rythmes de vie des enfants*, La Documentation française, 1993.
- (9) Évaluation des l'insertion par l'activité économique, décidée par le Comité interministérielle de l'évaluation en 1993, non terminée.
- (10) Comité régional d'évaluation de la région Rhône-Alpes, Evaluation de la politique de transfert de technologie au profit des entreprises, rapport non publié.
- (11) Christina NYRUP, "La DDE et les maires, une évaluation des prestations aux communes", in *Les cahiers du Management n°14*, Ministère de l'Aménagement du Territoire, de l'Équipement et des Transports, Octobre 1995.
- (12) Comité interministériel de l'évaluation, Commissariat général du plan, *La réhabilitation du logement social*, La Documentation française, 1993.
- (13) Comité interministériel de l'évaluation, Commissariat général du plan, *Les services publics et les populations défavorisées*, La Documentation française, 1993.
- (14) P. CORBEL et Y. PECCIA-GALETTO, "Aider les innovateurs" in *L'innovation technologique, les chiffres clés*, Dunod, 1993.
- (15) Valérie BAUDOIN et Anne-Lise AUCOUTURIER, "Histoires d'insertion", *Travail et emploi n°65*, La Documentation française 1995.

- (16) "Dictionnaires des sources statistiques sur le RMI", Commissariat général du Plan-Commission nationale d'évaluation, note ronéotée, 1991, Paris.
- (17) F. OUELLET, D.DURAND et R.MASSE, "Une stratégie de triangulation pour l'évaluation des effets d'un programme de soutien aux nouveaux parents", in *Revue canadienne d'évaluation de programme*, Avril-Mai 1992, P.69.
- (18) Neil D. WEINSTEIN et alii, "Evaluating personalized risk messages" in *Evaluation review*, Juin 1992, Sage-London, p. 235.
- (19) "L'évaluation des zones d'éducation prioritaires", *Dossiers Education et formation n°14*, Ministère de l'Education nationale, Septembre 1992, Paris.
- (20) Extrait du Cahier n°2 du programme MEANS, *Identifier les effets structurants des interventions communautaires*, Union Européenne - Politique Régionale et Cohésion, p. 45.
- (21) "Une illustration : la politique du transport routier de fret", dans MENY Y. et THOENIG J.C., *Les politiques publiques*, Paris, PUF 1989, pp. 258-267.

## Ouvrages généraux sur l'évaluation des politiques publiques

### En français :

- Rapports annuels du Conseil scientifique de l'évaluation :
  - *L'évaluation de l'expertise à la responsabilité (1992)*
  - *L'évaluation en développement (1993)*
  - *L'évaluation en développement 1993 (1994)*
  - *L'évaluation en développement 1994 (1995)*
  - *L'évaluation en développement 1995* (à paraître en juin 96)

(Ouvrages publiés à La Documentation française, Paris)

- Commissariat général du plan, *Outils, pratiques et institutions pour évaluer les politiques publiques*, Actes du séminaire Plan-ENA, La Documentation française 1990.
- DELEAU (Michel) (dir.) : *Évaluer les politiques publiques : méthodes, déontologie, organisation*, rapport du groupe de travail "Méthodes d'évaluation des politiques publiques", Commissariat général du Plan, Paris, La Documentation française, 1986.
- MONNIER (Eric) *Évaluation de l'action des pouvoirs publics*, Paris, Economica, 2ème édition, 1992.
- NIOCHE (Jean-Pierre), POINSARD (Robert), *L'évaluation des politiques publiques*, Paris, Economica, 1984.
- "L'évaluation en question", *Revue française d'administration publique n°66*, , Institut International d'Administration Publique, Avril-Juin 1993.
- RAY (Jean-Claude) (dir.), *Les politiques sociales transversales : une méthodologie d'évaluation de leurs effets locaux*, rapport d'un groupe de travail sur l'évaluation des politiques sociales, Commissariat général du Plan, Paris, La Documentation française, 1985.
- VIVERET (Patrick) *L'évaluation des politiques et des actions publiques*, Rapport au Premier ministre, Paris, La Documentation française, 1989.

### Revue

- *La Revue canadienne d'évaluation de programme*, éditeur R.V. Segworth, département de sciences politiques, Laurentian University, Sudbury, Ontario (Canada)

### Quelques ouvrages en langue anglaise :

- Center for the Study of Evaluation, *Evaluator's Handbook*, Sage publications, Londres

1987.

- CHEN (Huey-Tsyh), *Theory-Driven Evaluation*, Newbury Park, Sage publications, 1990.
- GUBA (E), LINCOLN (S), *Fourth génération Evaluation*, Sage 1987
- MAYNE (J.), BEMELMANS-VIDEC (M.L.), HUDSON (J.) et CONNER (R.), *Advancing Public Policy Evaluation*, North-Holland, Amsterdam 1992.
- PALUMBO (Dennis) (dir.), *The politics of program evaluation*, Beverly Hills, Sage Publication, 1987.
- RIST (Ray C.) (dir.), *Program evaluation and the management of government : patterns and prospects across eight nations*, New Brunswick, Transaction publishers, 1990.
- ROSSI (Peter H.), FREEMAN (Howard E.), *Évaluation : a systematic approach*, Sage Publications, 1993.

### **Revues**

- *Evaluation, the international journal of theory, research and practice*, Sage publications, en association avec The Tavistock Institute, Londres.
- *Evaluation review, a journal of applied social research*, Sage periodical press, Thousand Oaks, USA.

## ANNEXE I

### **Trois cahiers des charges d'évaluation**

On trouvera ci-après, à titre d'illustration, trois cahiers des charges qui reflètent des démarches d'évaluation proches de celle qui est proposée dans le présent guide. Ils concernent respectivement :

- le service inter administratif du logement (SIAL), évaluation réalisée en 1994 et 1995, à la demande du Ministre de la fonction publique, sous la responsabilité du Comité pour la réorganisation et la déconcentration de l'Administration (CRDA).
- le Fonds de Développement des Petites et Moyennes Industries (FDPMI), évaluation réalisée en 1995 à l'initiative du Ministère de l'industrie.
- la politique du tourisme rural dans la région Bretagne, évaluation réalisée en 1993 du Comité Régional d'Evaluation de Bretagne.

## L'ÉVALUATION DU SERVICE INTER ADMINISTRATIF DU LOGEMENT

Comité pour la réorganisation et la déconcentration de l'Administration (CRDA).

### Origine du projet d'évaluation

Le projet fait suite à une commande de M. ROSSINOT précisée dans sa **lettre du 4 Mai 1994** où il est demandé aux administrations concernées (Intérieur, Affaires Sociales, Logement, Equipement, Budget) de procéder à l'évaluation de l'expérimentation d'un "pôle de compétences mis en place à LYON pour gérer les problèmes d'attribution de logements sociaux".

Il est également indiqué que cette évaluation se fera dans le cadre d'une instance présidée par S. VALLEMONT, Président du Comité pour la réorganisation et la Déconcentration de l'Administration (CRDA).

Cette décision de procéder à une évaluation fait **suite à une réunion de travail** qui s'est tenue le 14 avril 1994 au ministère de la Fonction Publique à laquelle la Préfecture du Rhône était venue exposer l'expérience menée dans le Rhône. L'intérêt suscité par la démarche a conduit le Ministère de la Fonction Publique à proposer cette évaluation.

### Champ de l'évaluation

Le champ de l'évaluation est circonscrit à l'action menée par le SIAL en matière d'attribution de logements sociaux : il ne s'agit pas d'évaluer l'ensemble de la politique du logement dans le Rhône.

L'évaluation porte essentiellement sur le caractère interministériel de l'expérience conduite. Le SIAL n'est étudié que comme "la préfiguration d'une organisation administrative susceptible d'être développée dans l'avenir". Le coeur de la réflexion à mener concerne donc le fonctionnement de l'interministériel.

### Description de l'action à évaluer

Le SIAL a été mis en place en Octobre 1992 dans le cadre de l'élaboration du plan triennal d'action pour le logement des personnes défavorisées.

Ses missions (cf. note de présentation du 3 août 1994) visent à la fois à :

1. constituer un **dispositif interministériel** de compétence en matière de logement ;
2. mettre en place un **pôle d'accueil** de la demande des populations les plus défavorisées ;
3. constituer au plan départemental un **lieu de connaissance** et d'analyse de cette demande ;
4. optimiser la **gestion de logements réservés au Préfet** dans le cadre du contingent préfectoral ;
5. amener le service à être un **partenaire reconnu** et incontournable dans la politique d'attribution des logements sociaux.

Le SIAL comprend **18 agents** d'origines diverses :

- 1 chef de service issu de l'ANPE
- 9 agents du cadre national de Préfecture
- 3 contractuels DDASS
- 1 agent de la DDASS
- 1 agent de la DDE
- 3 venant d'association ou d'un bureau d'étude (CREPAH)

Exposé des motifs de l'évaluation  
Prédiagnostic

Une base de diagnostic est fournie par la note rédigée le 3 Août 1994 par le SIAL où un premier bilan de l'action du SIAL est effectué.

Ces éléments seront complétés par une analyse de l'historique du SIAL et du contexte ayant présidé à sa création. Cette analyse de l'évolution du SIAL devra permettre de dégager les premières hypothèses de travail.

### **Les objectifs de l'évaluation**

Les objectifs assignés à l'évaluation sont explicités dans la lettre de M. ROSSINOT du 4 Mai dernier. Ils concernent à la fois :

1. l'efficacité de la démarche, en termes de service rendu aux usagers et aux partenaires ;
2. les conditions de sa pérennisation ;
3. la transposabilité éventuelle de la démarche.

### **Les attentes locales à l'égard de l'évaluation**

**Les attentes des acteurs locaux** (Préfecture, DDE, DDASS, SIAL) telles qu'elles se dégagent d'entretiens préalables menés auprès d'eux se fondent sur une conviction forte : celle que le SIAL représente une tentative pratique pour instaurer un fonctionnement interministériel. L'évaluation de cette expérience doit donc permettre de montrer selon eux :

1. comment la notion d'interministérialité se concrétise, au-delà des seuls discours ou déclaration de principe ;
2. les conséquences qu'en tirent les administrations centrales, en terme de pilotage d'actions interministérielles.

Leurs attentes concernent donc simultanément :

1. une reconnaissance de l'action entreprise ;
2. un bilan de l'action engagée, sur la base d'un diagnostic contrasté, permettant à la fois de dégager les points forts et les points faibles de la démarche ;
3. la levée par l'administration centrale de certains blocages, notamment en termes de moyens ;
4. des propositions visant à améliorer le dispositif mis en place et à le "faire vivre".

Il est noter que les attentes des responsables locaux se situent davantage dans une perspective d'amélioration que dans une perspective de réforme. La suppression du dispositif apparaîtrait d'ailleurs selon eux comme un "désengagement auprès des partenaires".

Toute proposition par l'instance consistant à supprimer le dispositif devrait être assortie, selon eux, de propositions alternatives aboutissant à un service et des résultats au moins autant efficaces que le SIAL.

### **Problématique et hypothèses**

Il est difficile en l'état et prématuré de dresser de véritables hypothèses. On peut néanmoins recenser les principales questions que pose une évaluation du SIAL. Un premier diagnostic permettra de les affiner et de dégager des hypothèses d'étude.

Celles-ci s'organisent en fait autour de trois axes principaux :

*1/ le fonctionnement interne de l'Etat,*

*2/ le positionnement de l'Etat à l'égard de ses différents partenaires, notamment la communauté urbaine de Lyon, les communes, le département, les associations et les organismes HLM,*

*3/ l'amélioration du service rendu aux usagers.*

Sur le fonctionnement interne de l'Etat

**Les questions posées concernent essentiellement la plus-value apportée par le SIAL et sa valeur ajoutée par rapport à une organisation traditionnelle.**

Cette interrogation renvoie à plusieurs questions :

1. quelles sont les améliorations apportées par le SIAL par rapport à la situation antérieure ?
2. l'organisation mise en place répond-elle aux objectifs fixés ? est-elle cohérente avec la nature des problèmes à traiter ?
3. quelles sont les synergies opérées entre le SIAL et les autres services intervenant dans le même domaine (DDE, DDASS) ?
4. quels sont les modes de coopération qui se sont noués ?  
au niveau de la conception même du SIAL  
au niveau des moyens mobilisés  
au niveau des modes de travail en commun
5. quel est l'apport essentiel par le SIAL :  
rôle d'animation et d'interface ? lieu de connaissance et de production d'informations ? mise en réseau des différents partenaires ?.....

**En termes de perspective, on peut également s'interroger sur les aspects suivants :**

1. en quoi l'organisation actuelle est-elle adaptée ? comment doit-elle évoluer ?
2. quels sont les modes d'expertise et les compétences à développer par le SIAL ?
3. comment préserver et entretenir la dynamique mise en place ? susciter les conditions d'un véritable travail interministériel qui dépasse la mise en commun des moyens ?
4. comment éviter une bureaucratisation et une routinisation de l'activité du SIAL ?
5. quelles sont les conditions de pérennisation de la démarche engagée par le SIAL ?
6. en quoi ce mode d'organisation préfigure-t-il un véritable fonctionnement interministériel de l'Etat au niveau local ?
7. l'organisation actuelle est-elle transposable :  
dans d'autres départements ?  
à d'autres domaines ?  
à quelles conditions ?

Sur le positionnement de l'Etat à l'égard de ses partenaires (communes, département, associations et organismes HLM)

les questions concernent à la fois :

**l'efficacité** de l'action entreprise : "le SIAL a-t-il permis une gestion plus efficace et moins coûteuse pour l'Etat qu'elle ne l'aurait été par les moyens traditionnels ?" (cf. p. 1 de la lettre de M. Rossinot) ;

**l'image** de l'Etat auprès de ses interlocuteurs.

Il en découle une série d'interrogations :

le fonctionnement interministériel a-t-il favorisé un meilleur traitement des problèmes ? dans des délais plus rapides ? a-t-il contribué à mettre en oeuvre des procédures originales ?

l'expérience du SIAL a-t-elle permis un meilleur positionnement de l'Etat en matière de politique de logement ? A-t-elle accru la capacité de négociation et l'influence de l'Etat ?

Sur l'efficacité du SIAL auprès des usagers

Concernant l'amélioration du service rendu aux usagers, l'analyse s'attachera à distinguer :

la qualité de l'accueil mis en place (accueil personnalisé, liens avec les travailleurs sociaux, professionnalisation de cet accueil, etc ....),

la réponse aux demandes des usagers,  
modalités de traitement de la demande,  
catégories de population concernées,  
délais moyens de satisfaction, etc ....

sur ce dernier point, l'analyse s'appuiera sur les données statistiques disponibles au SIAL.

## Méthode

Il sera procédé à une analyse préalable permettant de se constituer des **éléments de référence et des critères** indispensables à toute évaluation.

La constitution de cet état de référence s'appuiera notamment sur :  
une reconstitution de l'historique du SIAL et des motifs qui ont présidé à sa création. Cette reconstitution s'appuiera sur des documents d'archives, notamment les résultats d'un questionnaire adressé par le SIAL à ses interlocuteurs à l'issue d'un an de fonctionnement et des entretiens auprès de quelques acteurs clés.

des expériences menées dans d'autres départements qui permettront d'avoir des bases de comparaison. Ces informations seront collectées par les membres de l'instance sur la base de rapports disponibles dans leur propre administration ou de leur propre expérience.

L'approche utilisée sera essentiellement **qualitative**, fondée sur des **entretiens approfondis** auprès de trois catégories d'acteurs :

1. les agents du SIAL ;
2. les agents des différents services de l'Etat impliqués dans la politique du logement (à la DDASS, la DDE et à la Préfecture) ;
3. les partenaires extérieurs (organismes HLM, associations, conseil général, élus, communauté urbaine de Lyon ....). Ces entretiens pourront être complétés par l'**analyse plus ponctuelle et précise de quelques procédures** déployées par le SIAL.

## Dispositif d'évaluation

### *composition de l'instance :*

l'instance comprend des représentants des différents ministères (Equipement, Logement, Affaires Sociales, Intérieur, Budget) (cf. liste jointe) ;

L'instance est présidée par S. VALLEMONT (CRDA)

### *calendrier :*

le délai prévu de 18 à 24 mois paraît long et peu compatible avec le souci local d'exploiter les retombées de l'évaluation pour relancer la dynamique du SIAL. On peut envisager un calendrier plus resserré visant la remise d'un rapport à l'été 1995,

le travail pourrait s'organiser selon le calendrier suivant :

mi-octobre : 1ère réunion de l'instance

étude des matériaux collectés par les membres de l'instance (informations sur des expériences analogues, rapports d'inspection disponibles sur le même sujet...) ;  
présentation de l'historique du SIAL ;  
mise au point définitive du cahier des charges et élaboration de quelques hypothèses ;  
mise au point des grilles d'entretien.

novembre : collecte des informations sur le terrain

décembre : 2ème réunion de l'instance d'évaluation

restitution des informations collectées  
premiers constats

janvier : point d'étape avec les acteurs locaux

février - mars : investigations complémentaires (sur quelques procédures précises par exemple), mise au point de l'analyse.

avril - mai : mise au point des propositions

juillet : remise du rapport

juin : élaboration du rapport définitif

juillet : présentation du rapport d'évaluation.

***organisation du travail :***

la phase de collecte des informations pourrait s'organiser de la façon suivante : il pourrait être constitué des binômes responsables de la conduite des entretiens auprès des différents interlocuteurs.

On peut suggérer que les interviews soient effectués :

à la DDASS et à la DDE, par M. FLORENNE, Inspecteur de l'Administration et M. ROCHARD, Inspecteur Général des Finances

au SIAL et à la Préfecture, par M. LAPILLONNE, Ingénieur Général des Ponts et Chaussées et M. VINCENT, Inspecteur Général des Affaires Sociales

auprès des partenaires, par Mme CHANUT, Chargée de mission auprès du Président du CRDA et Mme VARAGNE, Administrateur Civil à la Direction de l'Habitat et de la Construction du Ministère du Logement.

**EVALUATION  
DES DISPOSITIFS PUBLICS  
D'INCITATION ET D'ACCOMPAGNEMENT  
DE L'INVESTISSEMENT MATERIEL DES PMI**

**Cahier des Charges**

**DIRECTION DE L'ACTION REGIONALE  
ET DE LA PETITE ET MOYENNE INDUSTRIE  
Sous-Direction du Développement  
Industriel et Technologique Régional**

L'évaluation des politiques publiques répond à une double nécessité :

- d'une part, **l'impératif démocratique** qui impose l'information des élus et de la Nation sur le bon emploi des deniers publics et les effets des politiques menées,
- d'autre part, **un impératif technique** qui rend nécessaire l'identification des sources de dysfonctionnement et des lieux d'amélioration des dispositifs actuels et futurs.

Le présent projet d'évaluation du dispositif d'aide à l'investissement matériel des PMI s'inscrit dans cette double nécessité.

**1 - objectifs de l'étude**

L'étude poursuit un double objectif :

- d'une part, de manière rétrospective, évaluer le dispositif d'aide à l'investissement matériel en faveur des PMI mis en place par le ministère chargé de l'industrie depuis 1984. Ce volet de l'étude doit caractériser l'impact des Plans productifs régionaux et du Fonds de redéveloppement industriel sur les PMI effectifs jusqu'en 1994.
- d'autre part, de manière prospective, sur la base des enseignements de la première partie de l'étude, et compte tenu des éléments de similitude entre le dispositif antérieur à 1994 et le dispositif actuel d'aide à l'investissement des PMI, simuler les effets du Fonds de développement des PMI et en proposer des pistes d'amélioration pragmatiques.

**2 - nature et historique du dispositif d'aide à l'investissement matériel**

Le dispositif français d'aide à l'investissement matériel des PMI comprend de nombreuses aides visant chacune, des objectifs et/ou des populations spécifiques. Ces aides relèvent de plusieurs ministères comme ceux de l'Agriculture, de l'Aménagement du territoire ou de l'Industrie. Certaines mesures sont définies et mises en place en partenariat avec les collectivités locales et l'Union européenne.

Jusqu'au 31 Décembre 1993, les principales mesures d'aide à l'investissement matériel gérées par le ministère chargé de l'industrie étaient les Plans productifs régionaux (PPR) et le Fonds de redéveloppement industriel (FRI). Ces mesures étaient de portée géographique limitées. Dans un souci de clarification et de rationalisation, le Comité interministériel d'aménagement du territoire du 12 Juillet 1993 a décidé de substituer une nouvelle mesure -le Fonds de développement des PMI (FDPMI)- aux anciens PPR et FRI. Le FDPMI se substitue aussi aux Opérations pilote productive (OPP) du ministère de l'industrie et à l'Aide à l'investissement industriel en zone rurale (AIIZR) du ministère chargé de l'aménagement du territoire. Mis en place en Janvier 1994, le FDPMI est très largement inspiré des PPR dont il constitue, dans une large mesure, une extension à l'ensemble du territoire.

## 2.1 Les plans productiques régionaux :

Les deux premiers Plans productiques régionaux (PPR) sont des actions pluriannuelles massives décidées fin 1984 dans le cadre du redéveloppement et de la reconversion du Nord Pas de Calais et de la Lorraine. Ils ont été introduits à cette date dans la première génération des Contrats de plan Etat-régions (CPER) concernés. Ces CPER prévoyaient dans leur forme originale des actions de développement de la productique plus restrictives et moins massives.

Les PPR ont pour objectif immédiat de moderniser massivement les PMI en leur apportant des ressources financières à bon marché, l'objectif final recherché étant la dynamisation de l'emploi industriel dans un but de reconversion des bassins concernés.

Les PPR sont des procédures déconcentrées et contractualisées. L'aide consiste en une subvention d'équipement limitée à 20% du montant d'un programme d'investissement en productique. L'assiette éligible est limitée à 20 MFHT. Elle peut comprendre une part d'investissement immatériel.

Par la suite, respectivement en 1989 et 1991, un PPR a été mis en place en Limousin puis en Auvergne.

Le bilan quantitatif global de PPR s'établit comme suit :

|                            | NORD P.C.    |              | LORRAINE     |              | LIMOUSIN  | AUVERGNE  |
|----------------------------|--------------|--------------|--------------|--------------|-----------|-----------|
| période                    | 1984<br>1988 | 1989<br>1993 | 1984<br>1988 | 1989<br>1993 | 1989-1993 | 1991-1993 |
| nombre de dossiers         | 655          | 821          | 467          | 435          | 286       | 231       |
| montant des programmes     | 3250M<br>F   | 8900M<br>F   | 2160M<br>F   | #1500M<br>F  | 550 MF    | 665 MF    |
| montant des aides allouées | 484.4M<br>F  | 576.2<br>MF  | 248.6M<br>F  | 275.5<br>MF  | 86.3 MF   | 92.6 MF   |
| dont Etat                  | 270.9        | 241.3        |              | 135.3        | 38.5      | 45.0      |
| dont région                | 105.2        | 113.2        |              | 33.9         | 35.4      | 33.4      |
| dont U.E.                  | 108.3        | 221.7        |              | 106.3        | 12.4      | 14.2      |

## 2.2 Le fonds de redéveloppement industriel :

Le Fonds de redéveloppement industriel (FRI) a été créé en 1989. La vocation initiale telle qu'elle ressort de la décision de création du Premier ministre est de "soutenir la modernisation technologique et le renforcement des fonds propres des PMI situées dans des bassins touchés par des restructurations industrielles".

La circulaire de Juillet 1989 qui porte modalités techniques d'utilisation du FRI précise l'objet du fonds : "Le Fonds de redéveloppement industriel est essentiellement destiné à permettre le développement des PMI et des sociétés du tertiaire industriel en situation saine qui sont situées dans des bassins en difficulté qui ne disposent pas d'un dispositif de conversion satisfaisant".

Les notes circulaires d'instruction accompagnant les délégations de crédits aux préfets précisent en outre :

- que la priorité doit être donnée aux programmes d'investissement ayant une incidence positive en terme d'emploi,
- que le programme doit s'inscrire dans une démarche stratégique globale de développement.

Le FRI est une procédure déconcentrée. Annuellement, une dotation au titre du fonds est déléguée aux préfets de région. Cette dotation est accompagnée de la liste limitative des zones d'emploi de la région sur lesquelles le FRI peut intervenir.

L'aide consiste en une subvention d'équipement limitée à 20% du montant d'un programme d'investissement de modernisation. L'assiette éligible, qui peut contenir une part de dépenses de formation, est limitée à 20MFHT.

Le bilan quantitatif global du FRI s'établit comme suit :

|                            |                                    |
|----------------------------|------------------------------------|
|                            | toutes régions, sauf Ile de France |
| période                    | X° plan, 1989 -1993                |
| nombre de dossiers         | 1830                               |
| montant des programmes     | # 5000 MF                          |
| montant des aides allouées | 757 MF                             |
| dont Etat                  | 429 MF                             |
| dont région + U.E.         | 328 MF                             |

### 2.3 Le fonds de développement des PMI :

En janvier 1994, sur la base de premiers éléments locaux d'évaluation, les PPR ont été généralisés à toutes les régions -à l'exception de l'Ile de France- sous l'appellation Fonds de Développement des PMI (FDPMI). A partir de cette date, le Fonds de Développement des PMI s'est substitué aux mesures antérieures d'aide à l'investissement matériel, notamment aux PPR et au FRI.

La circulaire du 13 janvier 1994, qui définit les modalités de mise en place du FDPMI, précise que ce fonds "sera orienté prioritairement vers les zones défavorisées (zones en conversion industrielle, zones rurales, quartiers sensibles)". La circulaire précise par ailleurs que le FDPMI est destiné à favoriser l'élévation technologique et la compétitivité des PMI. Seuls les programmes s'inscrivant dans un projet global de développement sont éligibles.

Le FDPMI est contractualisé, au titre du XIème plan, avec toutes les collectivités régionales, sauf l'Ile de France. Dans les zones éligibles aux fonds structurels européens, le Fonds Européen de Développement Economique Régional (FEDER) co-finance le FDPMI.

L'aide consiste en une subvention d'équipement dont le taux varie en fonction de la localisation de l'entreprise et, éventuellement, de sa taille. Le taux d'aide maximal est de 30% en zone de Prime à l'Aménagement du Territoire (PAT). Il est de 75% dans les D.O.M.

Le budget prévisionnel global du FDPMI s'établit comme suit :

|                            |                                    |
|----------------------------|------------------------------------|
|                            | toutes régions, sauf Ile de France |
| période                    | XI° plan, 1994 -1998               |
| nombre de dossiers         |                                    |
| montant des programmes     |                                    |
| montant des aides allouées | # 5 400 MF                         |
| dont Etat                  | # 1 900 MF                         |
| dont région + U.E.         | # 3 500 MF                         |

### 3 - résultats attendus de l'étude

Le cahier des charges embrasse un champ volontairement vaste afin d'appréhender les divers aspects du dispositif étudié. Néanmoins il est précisé au consultant que le coeur de l'étude doit être consacré aux effets des dispositifs d'aide à l'investissement matériel des PMI, notamment aux effets directs sur les entreprises, aux coûts globaux de ces dispositifs et aux effets à moyen terme du FDPMI. Les éléments globaux sur l'investissement et les aides (§4.2.1.) ainsi que la comparaison internationale (§4.2.8.) doivent être traités comme des éléments de contexte et de situation.

#### 3.1 Au titre de l'évaluation des procédures passées :

L'étude doit évaluer les effets micro-économiques et macro-économiques des mesures de soutien à l'investissement matériel des PMI mises en oeuvre depuis 1984 (PPR et FRI), sur fonds publics (Etat, Union européenne et, dans la mesure du possible, régions). Les résultats obtenus seront notamment comparés aux objectifs affichés, aux moyens mis en oeuvre et au coût du dispositif.

Le degré de réalisation de chacun des objectifs initiaux des mesures d'aide sera caractérisé. Les conséquences connexes sur les PMI et sur les économies locales seront mises en évidence.

Le coût global à moyen terme de ces mesures d'aide, tant pour les finances publiques que pour l'économie sera évalué à l'échelle nationale et à l'échelle locale.

Le dispositif de mise en oeuvre des aides sera étudié afin de dégager les points forts à réutiliser et des pistes éventuelles d'amélioration.

#### 3.2 Au titre des propositions pour le dispositif actuel :

L'étude présentera une simulation des effets prévisibles, micro et macro-économiques, à court et moyen terme, du FDPMI. Elle évaluera l'investissement de l'Etat à moyen terme nécessaire à sa mise en oeuvre.

A l'expérience des conclusions de l'évaluation des PPR et du FRI, mais en tenant compte des incidences dues à des contextes économiques et à des structures d'aide différents, l'étude mettra en évidence les points forts et les points faibles de la procédure FDPMI (tant dans ses effets que dans sa mise en oeuvre). Des pistes seront recherchées pour améliorer les différents ratios : résultats/objectifs, résultats/moyens.

Le coût global et les effets prévisibles du FDPMI seront comparés à ceux pouvant être attendus d'une mesure générale sur l'environnement des entreprises, notamment une baisse de l'impôt sur les sociétés, de la taxe professionnelle et une baisse des charges sociales.

## **4 - définition des paramètres de l'étude**

### **4.1. Champ de l'étude, éléments méthodologiques :**

Afin d'apprécier l'apport spécifique, pour les entreprises, du dispositif d'aide, l'étude comparera les méthodes de décision et les effets des investissements d'un échantillon d'entreprises aidées à ceux d'un échantillon d'entreprises non aidées ayant cependant investi. Pour certains aspects, des comparaisons complémentaires avec un échantillon d'entreprises n'ayant pas investi et un échantillon d'entreprises dont la demande d'aide a été refusée, seront nécessaires.

Pour la partie de l'étude recherchant les effets économiques locaux, le consultant comparera l'état et l'évolution des bassins et des régions éligibles avec des bassins et des régions étalons non éligibles. Une attention particulière sera apportée aux biais qui peuvent être introduits par la qualité intrinsèque de la zone (infrastructure, dispositifs de formation, densité du dispositif d'aide, attractivité du climat...).

Dans tous les cas, le consultant rapprochera les effets constatés sur les PMI et sur le tissu industriel des effets recherchés par l'action publique. Il recherchera les causes des écarts mis en évidence et proposera des pistes pour les réduire. Les effets "autonomes" ou "pervers", induits indépendamment de la volonté publique, y compris hors du champ industriel, pourront être identifiés et caractérisés.

Toutes les questions devront être traitées de manière à dégager une typologie des effets et des comportements. La restitution des conclusions sera validée par le comité de pilotage.

#### **4.1.1. Echantillonnage**

Les différents échantillons seront représentatifs du tissu des PMI nationales potentiellement éligibles au dispositif. Ils devront permettre de caractériser les différents effets des aides sur ces entreprises.

L'échantillon final des entreprises aidées sera constitué d'environ 500 entreprises ayant bénéficié entre 1984 et 1994, d'une ou plusieurs aides à l'investissement matériel (PPR ou FRI). Ces entreprises seront localisées dans plusieurs régions dont le Nord-Pas-de-Calais, la Lorraine, le Limousin, l'Auvergne, la Champagne-Ardenne, Midi-Pyrénées, Rhône-Alpes, les Pays de la Loire, la Franche-Comté.

Il appartiendra au consultant de proposer un échantillon de référence, constitué de PMI investissant de manière autonome, sans aide publique. Cet échantillon devra être cohérent avec l'échantillon des entreprises aidées afin de permettre notamment des comparaisons par taille d'entreprise, par secteur d'activité, par région et par principaux ratios micro-économiques.

Pour traiter la partie macro-économique de l'étude, le consultant proposera une ou plusieurs "régions ou bassins étalons" sur lesquelles le dispositif d'aide était faible ou inexistant qui devront être comparés de manière cohérente avec un échantillon de bassin sur lesquels l'implication du dispositif était maximal.

La taille et la définition des échantillons étudiés par chaque méthode relève de la responsabilité du consultant. Il devra s'engager sur leurs représentativités en considération des contraintes définies par le présent cahier des charges. Les échantillons seront validés par le comité de pilotage.

#### **4.1.2. Methodologie et faisabilité**

La méthode d'investigation fera appel à des dispositifs d'enquêtes complémentaires dont les résultats devront être croisés. Dans tous les cas, des entretiens approfondis de validation avec quelques membres des échantillons devront être prévus pour confirmer les tendances des enquêtes systématiques.

Sous réserve des propositions du consultant, les enquêtes auprès des chefs d'entreprises seront complétées par des enquêtes auprès des partenaires et des acteurs économiques entourant l'entreprise. Le cas échéant, dans un cadre de stricte confidentialité, des analyses de dossiers typiques d'entreprises pourront s'avérer nécessaires pour confirmer des axes de conclusion.

Les éventuelles divergences dans les conclusions des différentes méthodes d'investigation (enquêtes, analyses de dossiers...) devront être commentées et expliquées.

Préalablement au lancement de l'évaluation proprement dite, le consultant testera sa méthode et ses questionnaires sur un échantillon témoin d'entreprises. La validité de la méthode d'investigation auprès de l'environnement des entreprises (banques, élus, institutionnels...) sera testée dans les mêmes conditions. Les résultats des tests seront présentés au comité de pilotage pour validation des méthodes.

Toute investigation au titre de la présente étude est soumise à l'approbation préalable du comité de pilotage. Les méthodes d'analyse, de recoupement, de modélisation prospective... envisagées par le consultant seront présentées au comité de pilotage.

Les recherches prospectives ne se limiteront pas à une prolongation des tendances constatées par le passé. Elles devront prendre en compte les dynamiques industrielles émergentes et les évolutions prévisibles et/ou supposées de l'environnement économique des PMI. L'étude présentera des conclusions dans le cadre de scénarios dont la prévisibilité sera évaluée.

Compte tenu des multiples aspects abordés par l'étude ( rétrospectif/prospectif, micro/macro-économique, effets/mise en oeuvre...) le consultant devra veiller à maintenir en permanence une forte cohérence entre les différentes parties de l'étude qui doivent se conforter mutuellement et, surtout, ne pas constituer une juxtaposition d'éléments disparates..

#### **4.1.3. Typologies**

Le consultant devra rechercher des typologies mettant en relation des groupes d'entreprises homogènes (au titre d'un ou plusieurs caractères) avec les principaux effets du système d'aide. Sur certains aspects, des typologies seront recherchées en fonction des réactions et attitudes des entreprises. Le consultant comparera les classes d'entreprises ainsi révélées avec des regroupements d'entreprises communément admis, notamment des regroupements fonction de la localisation géographique, des principaux secteurs d'activité, des tranches d'effectif et des principaux ratios économiques (VA/eff., Investissement/eff., CA export, Résultat/CA...).

Le consultant étudiera la concordance des typologies révélées ou confirmées par l'étude, avec les typologies dégagées dans des études antérieures (par exemple MEI 92, FRAC, ARC, PMI 93...).

#### **4.2. Contenu minimum du rapport final :**

##### **4.2.1. Des éléments situant l'investissement et les aides aux PMI**

Cette partie de l'étude doit présenter des éléments de référence permettant de caractériser la matière sur laquelle intervient le dispositif d'aide évalué. Il s'agit essentiellement pour le consultant de recueillir et d'homogénéiser les données existantes.

##### **4.2.1.1. Caractérisation de l'investissement des PMI françaises**

La présentation des chiffres globaux sera complétée par une analyse effectuée par segment (notamment par taille d'entreprise, par secteur d'activité, par région et par principaux ratios micro-économiques).

L'étude pourra notamment répondre aux questions suivantes :

- Quels sont le niveau et l'évolution, sur la période, de l'investissement productif des PMI françaises?
- Dégager une typologie et une répartition de ces investissements par grands thèmes industriels, par degré de modernité et intensité technologique et par fourchette de coûts.
- Dégager une typologie des motivations de l'investissement (remplacement, stratégie à moyen-long terme, réponse court-terme à une commande ponctuelle, opportunité ...).

#### **4.2.1.2. Bilan quantitatif global du dispositif**

Cette partie doit présenter une synthèse générale quantitative et des éléments statistiques globaux sur le dispositif.

L'étude pourra notamment répondre aux questions suivantes :

- Quelle est la pénétration des aides à l'investissement matériel par région? Nombre de dossiers proposés, aidés, refusés? Montant des aides allouées, montant des assiettes éligibles et des assiettes retenues, montant des programmes?
- Quelle est la part respective d'entreprises en création et d'entreprises en développement dans les entreprises aidées? Quelle est la part d'entreprises sous-traitantes à plus de 70% dans les entreprises aidées ?
- Caractériser les entreprises aidées en fonction de leur taille, de leur secteur d'activité, de leur marché (déclin, expansion, export ...), de leur positionnement dans leur secteur (plutôt performantes ou non).

#### **4.2.2. Une évaluation des effets micro-économiques de l'aide sur les PMI.**

L'étude montrera en quoi l'octroi d'une aide à l'investissement matériel modifie la compétitivité globale d'une entreprise réalisant un programme d'investissement.

Elle évaluera l'effet de l'aide sur le déclenchement de la décision, sur le contenu des projet et, plus généralement, sur le fonctionnement de l'entreprise. Au-delà de l'aspect strictement financier, cette partie de l'étude caractérisera la "valeur ajoutée qualitative" d'une aide pour une entreprise bénéficiaire. L'étude ne se limitera pas aux seuls effets positifs du dispositif d'aide, elle recherchera si le dispositif introduit des effets pervers et, le cas échéant, elle les caractérisera.

##### **4.2.2.1. Effet sur la décision d'investir**

Cette partie de l'étude doit évaluer le caractère incitatif du dispositif. Elle doit apprécier dans quelle mesure le dispositif modifie les anticipations des chefs d'entreprise.

L'étude pourra notamment répondre aux questions suivantes :

- L'aide est-elle une condition nécessaire à l'investissement, est-elle un élément de déclenchement de la prise de décision d'investir? L'aide modifie-t-elle le calendrier d'investissement? Les entreprises n'ayant pas investi auraient-elles investi si elles avaient bénéficié d'une aide ? Quel type de mesure aurait emporté leur décision d'investir ? Quelle est la motivation du recours à une aide ?
- Sans aide, le projet était-il économiquement rentable? Sans aide, sa rentabilité financière était-elle suffisante pour emporter la décision d'investir?
- Qui sont les demandeurs et les décideurs des investissements dans l'entreprise (direction, cadres, maîtrise, ouvriers)? Sont-ils les mêmes dans le cas d'un projet aidé ou non aidé ?

##### **4.2.2.2. Effet sur le contenu du programme**

Cette partie de l'étude doit caractériser le contenu des programmes d'investissement aidés par rapport aux programmes non aidés. Elle doit mettre en évidence l'apport du dispositif sur la méthode de définition d'un programme d'investissement, ainsi que sur le contenu qualitatif et quantitatif de celui-ci.

L'étude pourra notamment répondre aux questions suivantes :

- La typologie des investissements aidés et non aidés est-elle différente ?
- L'aide modifie-t-elle le contenu du programme ? Le niveau technologique des projets aidés est-il différent

des autres ?

- La notion de projet stratégique d'entreprise est-elle présente à des degrés différents suivant que l'entreprise est aidée ou non?

- L'investissement entraîne-t-il un développement global de l'entreprise plus important lorsque celle-ci est aidée?

- L'aide incite-t-elle à un coût d'investissement plus important ? Le surplus est-il chiffrable?

- La prise en compte des contraintes de protection de l'environnement est-elle modifiée par l'octroi d'une aide ?

- Les programmes aidés et les programmes non aidés différent-ils dans leur volet formation et adaptation des salariés ? caractériser cette différence (contenu, montant,...).

- Le mode de définition d'un programme d'investissement est-il différent suivant que l'entreprise est aidée ou non (cahier de charge, recours à un conseil extérieur, mise en concurrence...) ? Le délai de définition est-il différent ?

- Le délai de mise au point du matériel est-il différent suivant que l'entreprise est aidée ou non ? Le mode de suivi du projet est-il différent ?

- L'aide modifie-t-elle le volume et la qualité des projets futurs ? Le taux d'investissement des entreprises aidées évolue-t-il différemment de celui des entreprises non aidées ? La démarche méthodologique imposée à l'occasion d'un dossier d'aide est-elle reprise lors d'investissements ultérieurs?

- L'origine géographique des matériels est-elle différente suivant que l'investissement est aidé ou non ?

- La dimension commerciale est-elle appréhendée différemment dans un projet aidé ou non aidé ?

#### **4.2.2.3. Effet sur le montage financier du programme**

Cette partie de l'étude doit mettre en évidence si un accompagnement public entraîne une modification significative du plan de financement du programme.

L'étude pourra notamment répondre aux questions suivantes :

- Indépendamment des subventions allouées, les plans de financement des projets aidés et des projets non aidés présentent-ils des caractéristiques différentes ? Quelles sont les parts d'autofinancement, de prêt bancaire, de capital risque, de crédit-bail ?

- Dans le cas où l'aide publique s'impute sur une opération financée en crédit-bail, distinguer les interventions sous forme de reversement immédiat des interventions en réduction de loyer.

- L'aide crée-t-elle un effet de levier significatif pour l'obtention des prêts bancaires et des prise de garanties ? L'aide modifie-t-elle les taux pratiqués par les banquiers ?

#### **4.2.2.4. Degré de réussite des programmes d'investissement**

Cette partie de l'étude doit mettre en évidence si un programme d'investissement aidé atteint avec plus de précision ses objectifs qu'un programme équivalent non aidé. Par ailleurs, elle doit montrer si les effets globaux sont différents suivant qu'il y a ou non aide.

Le consultant devra proposer des éléments pertinents pour caractériser l'efficacité des programmes d'investissement, indépendamment de leurs objectifs initiaux (VA/effectif, CA, ...). Les biais introduits par les

règles d'éligibilités des procédures (bonne santé financière,...) seront corrigés.

L'étude pourra notamment répondre aux questions suivantes :

- Suivant que le programme est aidé ou non, le degré de réalisation des objectifs initiaux est-il différent? Les effets globaux des programmes sont-ils différents d'une population à l'autre? Quels en sont les motifs ?
- Y-a-t-il une typologie des causes d'échec des programmes d'investissement? Y-a-t-il des populations particulièrement exposées aux échecs? Y-a-t-il un taux (post-investissement) de défaillance d'entreprise différent suivant que le projet est aidé ou non?
- L'apport (en terme de C.A. identifiable induit par l'investissement) est-il différent entre un programme aidé et un autre ? Mesurer cet apport.
- Le gain de performance dû à un investissement est-il plus important lorsque l'entreprise est aidée? Les handicaps de compétitivité sont-ils mieux résolus lors d'un accompagnement public?
- L'évolution des ratios économiques de l'entreprise, notamment du rapport "valeur ajoutée/effectif", est-elle différente suite à un investissement est aidé ou non ?

#### **4.2.2.5. Effet sur les marchés de l'entreprise**

Cette partie de l'étude doit montrer si l'aide de l'Etat favorise une évolution qualitative des marchés de l'entreprise.

L'étude pourra notamment répondre aux questions suivantes :

- Les entreprises faisant appel à l'aide publique sont-elles positionnées sur des marchés captifs ou concurrentiels ? sur des marchés en déclin, stables, en développement, en redéveloppement (recherche d'une part importante de marchés nouveaux) ? L'évolution de la position de l'entreprise sur ses marchés est-elle différente dans le cas d'un investissement aidé ou non aidé ?
- Quel est le positionnement export des entreprises aidées ? Ce positionnement évolue-t-il différemment suite à un investissement aidé ou non aidé ?
- La structure du carnet de commande d'une entreprise aidée (nombre de clients, poids respectifs, localisation régionale, nationale ou internationale des clients...) évolue-t-elle différemment de celle d'une entreprise non aidée?

#### **4.2.2.6. Effet sur les ressources humaines**

Cette partie de l'étude doit montrer si l'aide de l'Etat favorise une modification quantitative et/ou qualitative du facteur travail dans l'entreprise.

L'étude pourra notamment répondre aux questions suivantes :

- L'investissement matériel entraîne-t-il des incidences directes et indirectes sur l'emploi dans l'entreprise (structure d'effectif, niveau d'effectif, politique de formation,...)? Ces incidences sont-elles différentes suivant que l'investissement est aidé ou non ?
- L'incidence des programmes d'investissement sur le mode d'organisation de l'entreprise est-elle différente d'un groupe à l'autre ?

#### **4.2.2.7. Effet sur la structure financière de l'entreprise**

Cette partie de l'étude doit montrer si l'octroi d'une aide lors d'un programme d'investissement modifie de manière significative la structure financière immédiate et future de l'entreprise.

L'étude pourra notamment répondre aux questions suivantes :

- L'aide modifie-t-elle significativement les ratios financiers de l'entreprise? Modifie-t-elle sa cotation Banque de France ou SFAC? L'incidence d'un investissement sur ces ratios est-il significativement différent suivant que celui-ci est aidé ou non?
- Les aides ont-elles touché des entreprises plutôt performantes dans leur catégorie (taille - secteur - ratios) ou non ? A l'issue du programme les entreprises sont-elles mieux classées dans leur segment?
- L'aide a-t-elle un effet significatif sur la trésorerie de l'entreprise ?

#### **4.2.2.8. Effet sur l'environnement de l'entreprise**

Cette partie de l'étude doit apprécier les effets induits d'une aide sur les relations de l'entreprise avec son environnement. Sur ce chapitre, l'environnement de l'entreprise devra obligatoirement être interrogé.

L'étude pourra notamment répondre aux questions suivantes :

- L'aide est-elle un élément modifiant sensiblement les décisions des partenaires financiers (banquiers, actionnaires,...) ? Quels éléments de l'aide facilitent plus particulièrement la mobilisation de ces partenaires?
- L'investissement, suivant qu'il est aidé ou non, a-t-il une incidence différente sur les clients, sur les fournisseurs,... ?
- La communication externe et interne de l'entreprise sur son programme d'investissement est-elle différente suivant que le programme est aidé ou non ?
- L'octroi d'une aide est-il connu des salariés, des clients, des concurrents, des institutionnels? La perception d'une entreprise par ses salariés, les institutionnels locaux, ... est-elle différente suivant qu'elle fait appel ou non à l'aide publique ?
- L'investissement aidé incite-t-il la concurrence à investir? Est-il facteur de transfert de technologie? La réaction de la concurrence est-elle identique dans le cas d'un projet non aidé?
- L'investissement entraîne-t-il des effets significatifs sur le recours à la sous-traitance par l'entreprise, notamment la sous-traitance de service? Ces effets sont-ils différents si l'investissement est aidé ou non?

#### **4.2.3. Une évaluation des effets macro-économiques de l'aide**

Cette partie de l'étude doit permettre de qualifier l'efficacité de l'aide à l'investissement matériel comme dispositif de redéveloppement local.

Le consultant proposera des indicateurs permettant d'évaluer le degré et la qualité du redéveloppement local.

##### **4.2.3.1. Effets locaux : emploi et développement**

Cette partie de l'étude doit permettre d'apprécier si le dispositif induit localement une dynamique aboutissant à diminuer la fragilité du tissu industriel, à le renforcer de manière significative ou encore à limiter ou inverser une tendance à la désindustrialisation. Les effets des aides à l'investissement matériel devront être isolés de ceux découlant directement d'autres types d'aides, notamment des procédures d'aide à l'emploi. Le consultant mettra en évidence les synergies entre le dispositif d'aide à l'investissement et les autres aides. Il recherchera les effets amplificateurs que crée ce dispositif sur les autres aides. Le cas échéant, le consultant identifiera et caractérisera les effets "pervers" induits par le dispositif d'aide.

L'étude pourra notamment répondre aux questions suivantes :

- Les courbes d'évolution de l'emploi industriel local et de l'emploi total sont-elles modifiées par l'éligibilité de la zone à une aide à l'investissement ? La structure d'emploi (quantité, qualité, durée) et son évolution sont-elles modifiées ? Combien d'emplois créés, combien d'emplois consolidés par les investissements aidés ?
- L'aide favorise-t-elle plutôt la création d'entreprises ou le développement d'unités existantes?
- Peut-on évaluer les emplois induits par les emplois industriels créés et/ou soutenus ?
- Y-a-t-il globalement une substitution capital/travail ?
- Quel est le délai nécessaire à l'apparition d'inflexions sensibles (emploi local, richesse locale...) ? Y-a-t-il un seuil critique de volume d'accompagnement (montant d'aide et nombre d'agents chargés du développement) au dessous duquel les effets ne sont plus significatifs?- Quel est l'incidence d'un

investissement sur la sous-traitance de proximité (S/T capacité, S/T savoir-faire,...) ? Est-il différent dans le cas d'un investissement aidé ou non ?

- Quelle coordination existe-t-il entre les aides à l'investissement matériel et d'une part les autres aides locales à la conversion, d'autre part les aides à l'emploi ? Existe-t-il une synergie particulière entre ces dispositifs? Le dispositif d'aide à l'investissement amplifie-t-il les effets des autres aides ?

- Apprécier l'incidence du dispositif d'aide sur l'économie locale notamment en examinant les effets du dispositif sur les indicateurs économiques locaux comme le PIB, l'emploi, la FBCF, les rentrées fiscales (taxes professionnelles, mutations immobilières...), le volume des transactions immobilières, les statistiques du commerce local, etc. ?

- L'efficacité relative d'une aide en terme de développement local est-elle significativement différente suivant la "richesse" initiale du territoire" (en terme d'aménagement du territoire)?

- Le dispositif pérennise-t-il des déséquilibres structurels ou favorise-t-il, même indirectement, un développement territorial équilibré ?

#### **4.2.3.2. Retombées nationales**

Cette partie de l'étude doit permettre d'apprécier les conséquences globales du dispositif sur quelques thèmes importants comme la compétitivité des PMI, le marché intérieur, l'emploi, la richesse nationale...

L'étude pourra notamment répondre aux questions suivantes :

- Les investissements aidés sont-ils créateurs de richesse nationale ou influencent-ils seulement la répartition de celle-ci (CA gagné sur un concurrent national) ? Les investissements aidés sont-ils facteurs d'ouverture du marché ?

- Quel est le bilan national des emplois créés et des évolutions de CA par secteur et par tranche d'effectif rapportés aux montants des investissements aidés d'une part et non aidés d'autre part (exclure éventuellement l'Ile de France) ?

- Un même investissement est-il plus créateur d'emploi global lorsqu'il est aidé ?

- Un même investissement influe-t-il plus sur le PIB national lorsqu'il est aidé ?

- L'effet global de l'aide à l'investissement matériel sur le tissu des PMI (rentabilité, C.A. à l'exportation, parts de marchés, poids des charges dans la V.A. ...) est-il significatif ?

- L'aide stimule-t-elle l'offre nationale de bien d'équipement industriel ? Dans quels secteurs ?

- L'aide entraîne-t-elle une spécialisation de l'investissement différente de ce qui serait observé en l'absence de soutien public?

#### **4.2.4. Une évaluation du coût global à moyen terme pour l'état**

à Cette partie de l'étude doit permettre d'apprécier le coût final pour les finances publiques (comprises au sens large). L'analyse à moyen terme portera sur le niveau national et sur l'échelon local.

Le calcul du coût devra tenir compte des emplois financiers directs (montants des aides, frais de promotion, frais de gestion des organismes relais,...), indirects (coût administratifs Etat, collectivités et partenaires, coûts de dysfonctionnement,...) et des recettes induites (impôts et taxes, économies/dépenses ASSEDIC, URSSAF,...).

Les coûts de gestion des procédures feront l'objet d'une évaluation précise qui distinguera les coûts

fixes des coûts variables.

#### **4.2.5. Une évaluation du dispositif de mise en oeuvre des mesures d'aide**

Cette partie de l'étude doit permettre d'apprécier l'efficacité du dispositif d'instruction. Les points forts et les points faibles des procédures doivent être identifiés. Des pistes d'amélioration seront dégagées.

Chacun des points soulevés dans ce chapitre fera l'objet de proposition d'amélioration.

L'étude pourra notamment répondre aux questions suivantes :

- Quel est l'apport qualitatif du dispositif d'instruction et de décision? Les chefs d'entreprises sont-ils prêts à de nouveaux contacts avec les chargés de mission ?
- Quelles sont les attentes des chefs d'entreprise en matière d'information à l'occasion d'un projet d'investissement ? Les chefs d'entreprises considèrent-ils les avis des chargés de mission suffisants ou ressentent-ils le besoin d'une expertise indépendante pour apprécier le bien fondé du choix technologique et/ou valider les performances des équipements choisis ? Les chefs d'entreprise ressentent-ils le besoin d'un accompagnement particulier lors du déroulement d'un programme d'investissement? Ce service ressort-il de l'Etat ou d'un conseil extérieur?
- Les chefs d'entreprise considèrent-ils les dossiers d'instruction adaptés? Dans quelle mesure et à qui ont-ils eu éventuellement recours pour constituer le dossier ? Y-a-t-il une corrélation avec les cas de recours à un organisme extérieur pour la définition du projet ? Combien coûte à l'entreprise un dossier d'instruction (coût interne + coût externe) ? Dans les cas des aides aux taux les plus faibles (par exemple inférieures à 10% de l'assiette), serait-il pertinent d'alléger le contenu des dossiers ?
- L'existence du dispositif d'aide est-elle connue? ses objectifs sont-ils connus, compris et admis (par les chefs d'entreprise, les banques, ...)? Le fonctionnement du dispositif est-il connu (point d'entrée, mode d'instruction,...)? les chefs d'entreprises peuvent-ils citer les financeurs publics, les organismes chargés de l'instruction ?
- Les "non-clients" présentent-ils des caractéristiques particulières ? Sont-ils isolés de l'administration, des organismes consulaires...? Existe-t-il une typologie des "clients" du dispositif d'aide ? Des classes "d'attentes" des entreprises peuvent-elles être identifiées ?
- Par quels vecteurs d'information les chefs d'entreprise ont-ils connu le dispositif et son fonctionnement? Comment informer les chefs d'entreprise isolés?
- Quels sont les moyens mis en place pour promouvoir les procédures (y compris par les collectivités locales partenaires) et valoriser les opérations exemplaires?
- Le dispositif d'aide à l'investissement matériel est-il un élément positif (et substantiel) de l'image des DRIRE, du MIPTCE, des collectivités locales, auprès des chefs d'entreprise PMI ? (des banques ?).
- Comment le dispositif est-il perçu par ses acteurs (chargés de mission, administration locale, y compris TPG, collectivités locales, centrale, budget)?

#### **4.2.6. Une simulation des effets du FDPMI**

Cette partie de l'étude doit évaluer les effets à court et moyen termes du Fonds de développement des PMI. Par ailleurs, à l'éclairage des conclusions dégagées par les chapitres précédents, elle proposera des pistes d'amélioration de ce dispositif. Le consultant proposera en outre une méthodologie de suivi de l'efficacité du FDPMI basée sur quelques indicateurs simples et réactifs à court terme.

### **L'étude pourra notamment répondre aux questions suivantes :**

- Les effets micro et macro-économiques du dispositif PPR + FRI sont-ils transposables au FDPMI? Esquisser des prévisions d'effet à moyen et long termes du FDPMI.
- Existe-t-il un seuil de taux d'aide au dessous duquel les effets diminuent fortement? Est-ce qu'il existe une corrélation entre le niveau de l'aide et ses effets?
- La modulation des taux d'aide du FDPMI peut-elle avoir un effet significatif sur la localisation des créations d'entreprises ?
- La durée d'existence d'un système d'aide a-t-elle une influence sur ses effets? Y a-t-il une durée de vie du dispositif au-dessous duquel celui-ci devient inopérant?
- Quelles sont les conditions essentielles pour qu'une mesure d'aide à la modernisation des entreprises soit efficace?
- Quelles dispositions seraient de nature à accroître l'efficacité du FDPMI?

#### **4.2.7. Une appréciation du FDPMI**

Cette partie de l'étude doit permettre d'une part d'apprécier dans quelle mesure le FDPMI correspond aux attentes des chefs d'entreprises, d'autre part, de comparer le FDPMI à d'autres outils publics, à coût global égal pour l'Etat.

##### **4.2.7.1. Adhésion au FDPMI**

Le consultant mettra en évidence l'intérêt des chefs d'entreprises et des acteurs locaux du développement industriel pour le dispositif d'aide à l'investissement matériel, en comparaison d'autres aides publiques. Il dégagera les attentes prioritaires des chefs d'entreprises en matière générale d'aide publique.

##### **4.2.7.2. Présentation comparative du FDPMI**

Cette partie de l'étude sera évaluative. Dans les conditions d'un coût global égal pour l'Etat, elle fera une comparaison entre le FDPMI et des mesures générales de soutien à l'environnement des entreprises comme une baisse généralisée de l'impôt sur les sociétés, de la taxe professionnelle ou des charges sociales des entreprises et l'instauration d'un crédit-impôt investissement. Les thèmes de performance étudiés seront le volume d'investissement, l'emploi industriel et la balance industrielle.

Une comparaison avec les dispositifs attendus, dégagés des réponses au paragraphe 4.2.7.1., pourra être demandée au consultant.

#### **4.2.8. Une étude comparative des dispositifs équivalents dans les principaux pays industrialisés**

Cette partie de l'étude doit permettre de positionner de manière quantitative le dispositif français de soutien à la modernisation des entreprises par rapport à ceux en place dans les principaux pays industriels.

Les pays étudiés seront notamment la Grande-Bretagne, l'Allemagne, l'Italie, les USA et le Japon.

L'étude pourra notamment répondre aux questions suivantes :

- Quel est le niveau (et l'évolution) de l'investissement des PMI, par pays?

- Identifier les aides, préciser les moyens financiers alloués.

- Présenter ces dispositifs de manière homogène et comparable.

#### **5 - documents finals :**

Le consultant fournira :

- . un rapport final détaillé,
- . un document de synthèse ayant vocation à être diffusé auprès des décideurs publics et privés,
- . un document de large communication présentant en un maximum de quatre pages le bilan des mesures d'aide étudiées et les conclusions caractéristiques de l'évaluation.

Le consultant remettra l'ensemble des documents-sources permettant la reprographie des documents ci-dessus.

#### **6 - suivi de l'étude**

L'évaluation est placée sous la responsabilité directe de la DARPMI-SDDITR qui fait appel en tant que de besoin à Comité de pilotage pluridisciplinaire d'une dizaine de membres. Le consultant tiendra régulièrement informé la DARPMI-SDDITR et éventuellement le comité de pilotage de l'avancement de l'étude.

Les méthodes, les échantillons et les questionnaires écrits sont soumis à l'approbation de la DARPMI-SDDITR

#### **7 - durée de l'étude**

Le candidat proposera des délais pour la réalisation des différents items de l'étude. Un calendrier de présentation des lots sera proposé. La durée globale de l'étude ne dépassera pas huit mois.

#### **8 - cout de l'étude**

Le prix total de l'étude sera de l'ordre de 1.5 million de francs.

Les propositions devront chiffrer individuellement chacun des items 4.2.1 à 4.2.8

#### **9 - conditions de paiement**

Le candidat proposera des clefs de paiement compatibles avec le Code des Marchés Publics et notamment le Cahier des Clauses Administratives Générales (prestations intellectuelles).

#### **10 - propriété de l'étude**

Tous les droits sur l'étude, y compris la valorisation et l'exploitation des fichiers de données brutes et des résultats, sont la propriété exclusive du ministère de l'industrie, des postes et télécommunications et du commerce extérieur.

#### **11 - contenu de l'offre**

Le candidat devra faire ressortir dans sa proposition le maximum d'idées et de propositions d'actions concrètes qu'il conviendrait d'exploiter dans le cas où il serait sélectionné. Il doit présenter un projet homogène dans lequel les différents aspects étudiés doivent se conforter mutuellement. Le projet respectera les priorités annoncées au paragraphe 3 du présent cahier des charges.

Le candidat précisera quels seront les intervenants pour chacune des phases, fonctions ou actions qu'il sera amené à traiter ou proposer. Il précisera l'organisation qu'il entend donner à son équipe en termes de coordination technique et temporelle.

Les offres présentées doivent comprendre au minimum :

- un plan de travail détaillé indiquant séparément le travail requis pour les paragraphes 4.2.1 à 4.2.8 ainsi que les experts intervenants sur chacun de ces items,
- une présentation des méthodes d'enquête, d'analyse et des techniques de discrimination que le consultant envisage utiliser,
- une évaluation des documents finals qui seront présentés en fin de l'étude,
- une liste complète des experts qui devraient participer à l'étude, assortie d'une explication sur le rôle qu'ils doivent y jouer. Pour chaque expert, il sera joint : un curriculum vitae complet et une présentation de ses motivations par rapport au sujet de l'étude.
- une évaluation détaillée exposant toutes les dépenses , ventilées pour chacun des paragraphes 4.2.1 à 4.2.8 en frais de personnel (nombre de jours par catégorie de consultants et de secrétariats associés,...), frais de déplacement et de séjours, frais techniques (téléphone, reprographie...), frais de traitement informatiques, frais d'accès à des bases de données,...Les salaires seront répartis en salaires nets et charges sociales. Le budget sera présenté hors taxes et toutes taxes comprises.

## **12 - confidentialité**

Tout consultant ou expert participant à l'étude devra signer un engagement de stricte confidentialité, notamment sur l'ensemble des informations relatives aux entreprises qu'il sera amené à connaître dans le cadre de l'étude.

# EVALUATION DE LA POLITIQUE "TOURISME RURAL" ET DE SON INCIDENCE SUR LES ZONES ELIGIBLES AUX PROGRAMMES EUROPEENS

## Cahier des charges

### 1 - Préambule

La consultation a pour but de sélectionner un cabinet, un laboratoire de recherche universitaire ou toute structure d'étude, pour conduire l'évaluation de la politique en faveur du Tourisme rural menée dans le cadre du contrat de Plan Etat-Région et de son incidence sur les zones éligibles aux programmes européens.

Le présent document doit aider à la formulation des modalités de l'étude.

### 2 - Champ d'évaluation

Il s'agit d'évaluer une politique développée depuis 1984 dans le cadre des contrats de 9ème et 10ème plan Etat-région et qui s'appuie largement sur la structure des "Pays d'Accueil", entité géographique homogène de la taille de quelques cantons, qui s'est organisée de manière à mettre en valeur son potentiel touristique.

Il existe 19 pays d'Accueil en Bretagne qui sont regroupés au sein de la Fédération Régionale des Pays d'Accueil dont le siège est à Bourbriac (22).

Dans les zones éligibles aux financements européens (dites "zones Morgane" ou "zone OID" avant 1991), certaines actions bénéficient de modalités d'intervention particulières.

4 Pays d'Accueil sont situés en zone "Morgane"

8 Pays d'Accueil sont hors zone "Morgane"

7 comprennent des communes situées en zone "Morgane" et hors zone "Morgane".

### 3 - Description de la politique à évaluer

#### *a) Objectifs de la politique "Tourisme Rural"*

- Contribuer à la diversification des activités économiques
- Renforcer le professionnalisme des acteurs locaux
  
- Moderniser l'offre touristique
  - Politique de qualité
  - Amélioration de l'accueil
  - Réalisation d'équipements de loisirs structurants
  - Modernisation de l'hébergement

- Développer la mise en place d'outils de gestion et d'observation
- Améliorer la commercialisation des produits touristiques.

***b) Bénéficiaires***

- Hôteliers
- Communes
- Prestataires de loisirs
- Groupements de communes.

***c) Procédure d'instruction et modalités d'intervention***

Le pays d'accueil établit un programme pluriannuel d'actions qui hiérarchise ses propositions et précise les actions prioritaires pour lesquelles il souhaite le concours de l'Etat et de la Région.

Celles-ci concernent les domaines suivants :

- Investissements immatériels
  - Assistance technique : Animation et coordination d'opérations touristiques par un professionnel
  - Création promotion, commercialisation de produits touristiques. C'est à ce titre qu'un poste de conseiller technique a été financé depuis 1988 dans le cadre de l'opération Intégrée de Développement (OID) et prolongé par Morgane
  - Formation des professionnels du tourisme et implication dans l'organisation du Pays d'Accueil
  - Connaissance technique et économique des produits, marchés, résultats
  - Elaboration et utilisation d'outils de gestion du Pays d'Accueil.
- Investissements matériels
  - Modernisation ou création d'équipements d'hébergement
  - Equipements de loisirs structurants.

3 mesures bénéficient de modalités d'intervention spécifiques dans le cadre du programme Morgane, qui prolonge jusqu'en 1993 le programme OID (1988-91).

- Assistance à la création et la promotion de produits touristiques
- Aide au conseil et à l'innovation touristique (FACIT)
- Aide à la création et la mode
- Modernisation des équipements d'hébergements et de loisirs.

**4 - Motifs du projet d'évaluation**

L'évaluation par une meilleure connaissance des résultats et de l'impact de cette politique, mais aussi des conditions de sa mise en oeuvre, doit permettre à l'Etat et à la Région d'en améliorer le contenu et d'en faire évoluer les priorités dans l'optique de la préparation du futur Contrat de Plan-Etat-Région.

L'évaluation doit également permettre une mesure des résultats et des effets propres à la zone Morgane

afin d'en détecter les spécificités éventuelles et de répondre aux futures exigences d'évaluation de la Commission Européenne.

Cette étude s'inscrit, en outre, dans le cadre de la démarche du Comité Régional d'Evaluation des politiques publiques, qui vise à développer et coordonner les initiatives d'évaluation conduites localement. A ce titre le CRE émettra un avis sur le projet d'évaluation et ses résultats.

## **5 - Méthodes d'évaluation**

### *Contexte général de l'étude*

L'étude comportera trois phases qui pourront faire l'objet de réponses spécifiques et dont les conclusions devront être regroupées dans une synthèse finale.

Celle-ci devra permettre de répondre à la question suivante :

A-t-on réussi à organiser une véritable offre touristique de qualité répondant à la demande des clientèles d'aujourd'hui et de demain ?

**1ère phase** : étude statistique des aides accordées depuis 1984

Il s'agit de rendre compte de la réalité des investissements matériels et immatériels financés par l'Etat et la Région et du degré d'exécution des objectifs de la politique "Tourisme Rural".

Les informations concernant la politique "Pays d'Accueil" sont contenues dans une base de données constituée par la Région à partir de 1986, qui sera complétée pour les années antérieures et qui comprend les renseignements suivants :

- Année de financement
- Pays d'Accueil
- Localisation (Bretagne Centrale ou pas)
- Canton
- Commune
- Maître d'ouvrage
- Libellé de l'opération
- Type d'opération
- Nombre de lits prévus (équipements d'hébergements)
- Coût total de l'opération
- Financement :
  - . Fidar
  - . Etat
  - . Feder
  - . Région
- Date de notification de la subvention au bénéficiaire

- Crédits Région consommés
- Date du dernier paiement Région
  
- Observations Région

Les informations financières concernant la politique FACIT (Aide au conseil) sont localisées notamment dans le fichier informatique de suivi financier de la Région.

A partir de ces bases il conviendra d'examiner le contenu des aides accordées depuis 1984 et en particulier d'analyser les ratios et indicateurs suivants en indiquant pour chacun leur valeur dans les zones éligibles aux financements européens :

- Evolutions annuelle des aides (Nombre, coûts réel, coût prévisionnel)
- Répartition géographique des aides
- Typologie et répartition des aides par maître d'ouvrage (une typologie plus précise devra être définie)
- Nature et répartition des aides par type d'opération
- Nombre de lits prévus et évolution annuelle
- % participation des différents financeurs y compris départements et autofinancement
- % consommation des crédits
- Nombre d'études suivies de réalisation
- Effet de levier des études.

Cette étude statistique doit également permettre de constituer un échantillon représentatif des aides accordées, dans leur ensemble mais aussi sur les zones Morgane, qui permettra de délimiter 2 ou 3 pays d'accueil servant de références pour la suite de l'évaluation.

L'échantillon choisi sera complété par un certain nombre d'informations quantitatives disponibles dans les dossiers de demande ou chez les bénéficiaires, qui permettent de mesurer l'impact de cette politique.

- Variation de chiffre d'affaire (prévu-réalisé) due à l'investissement
- Variation de la fréquentation due à l'investissement
- Variation du résultat d'exploitation (prévu-réalisé) due à l'investissement
- Créations d'emplois liées à l'investissement.

## **2ème phase** : étude des effets de la politique

Il s'agit dans cette phase, d'évaluer par une enquête qualitative de terrain, les effets directs ou indirects (impact sur l'environnement économique) de l'ensemble des politiques en faveur du tourisme sur 2 ou 3 pays d'accueil tests, ainsi que l'adéquation des investissements réalisés et des dispositifs mis en place avec la demande de tourisme "vert".

Cette enquête s'adressera aux bénéficiaires identifiés dans l'échantillon, mais aussi aux responsables techniques des Pays d'Accueil, à des professionnels du tourisme ou à des élus locaux des zones concernées.

Elle doit permettre de répondre notamment aux questions suivantes :

- Les investissements réalisés ont-ils eu un impact sur la fréquentation et le chiffre d'affaire des bénéficiaires, sur leur rentabilité, sur l'emploi ou d'autres effets ?
- Comment les bénéficiaires, les élus, les professionnels, perçoivent-ils la politique publique en faveur du tourisme ?
- Perçoivent-ils des évolutions du secteur touristique qu'il serait nécessaire d'accompagner ou d'anticiper par une action des pouvoirs publics ?
- Dans quelle mesure ces aides ont-elles généré ou accéléré les investissements réalisés ?
- Comment les bénéficiaires, les élus perçoivent-ils l'action des pays d'accueil ?
  - . sur le plan de l'assistance technique
  - . sur le plan de la commercialisation des produits touristiques
  - . sur le plan de l'animation du secteur touristique
  - .(etc...).
- Existe t-il une véritable collaboration entre acteurs du tourisme ?
- Ont-ils le sentiment d'oeuvrer pour un objectif commun ?
- Les actions des pays d'accueil s'adaptent-elles à la demande touristique ?
- La politique touristique mise en oeuvre a-t-elle permis de renforcer le professionnalisme des acteurs, la qualité des produits, l'accueil ?

**3ème phase** : analyse du fonctionnement des structures

Les pays d'Accueil est le vecteur principal de la politique "Tourisme Rural", l'objectif est ici de mieux comprendre la manière dont fonctionnent ses structures, quels sont les moyens qu'elles mettent en oeuvre et en faire ressortir les caractéristiques principales.

Cette étude pourra s'appuyer sur les données financières et comptables concernant chacun des Pays d'Accueil, qui sont transmises régulièrement à la Délégation Régionale du Tourisme puis complétée par entretien portant sur l'activité de ces structures.

***a) Données financières et comptables***

- Nature de la structure juridique
- Budget total
- Part assistance technique/Budget Total
- Part masse salariale/Budget Total
- Part autres charges/Budget Total
- Structures des autres charges
- Part subventions Etat-Région, Département/Budget Total
- Part cotisations communales/Budget

Total

- Part cotisations professionnelles/Budget Total
- Par autofinancement/Budget Total
- Structure de l'autofinancement
- Relations avec les structures départementales et régionales
- Nombre de personnes salariées

***b) Activités des structures***

- communication interne
- Communication vers les marchés
- Commercialisation de produits
- Créations d'hébergements
- Equipements
- Accueil - animation
- Autres activités
- Relations entre les Pays d'accueil.

**6 - Dispositif d'évaluation**

***a) Suivi de l'évaluation***

Un groupe de suivi de l'évaluation sera formé, composé de :

- SGAR
- Conseil Régional
- Délégation régionale du Tourisme
- Comité Régional du Tourisme
- Un comité Départemental du Tourisme (Ex : Finistère)
- Fédération Régionale des Pays d'Accueil
- Chambre Régionale des Comptes

Celui-ci sera chargé :

- D'élaborer le cahier des charges définitif de l'évaluation
- De lancer une consultation auprès des structures d'études susceptibles d'y répondre
- De proposer le choix d'une équipe d'évaluation à l'issue de la consultation

- De suivre et d'orienter éventuellement les travaux d'orientation
- De faciliter les contacts entre les évaluateurs et le terrain
- De valider le rapport final d'évaluation
- De préconiser des réorientations pour la mise en oeuvre future de la politique.

***b) Rôle de l'Instance Régionale d'Evaluation***

- Comité Régional d'Evaluation

- Décide de l'engagement définitif de l'évaluation, après accord des personnes publiques dont les politiques sont évaluées et après avis de la commission scientifique
- Décide, en accord avec collectivités concernées, de la publicité éventuelle à donner aux rapports d'évaluation et des modalités de cette publicité.

- Commission scientifique

- Donne un avis sur le projet d'évaluation (Composition de l'équipe d'évaluation, méthode proposée)
- Donne un avis sur le rapport final d'évaluation.

***c) Chargés d'évaluations***

La structure d'étude répondant à la consultation pourra proposer selon ses compétences, d traiter tout ou partie du cahier des charges, le groupe se réservant la possibilité de choisir une équipe d'évaluation croisant plusieurs propositions afin d'aboutir à la meilleure réponse possible.

Dans cette hypothèse les structures d'études retenues devront se rapprocher afin d'élaborer en commun une synthèse de leurs travaux comme précisé plus haut (Contexte général de l'étude).

Elles devront témoigner dans une note annexe à leur réponse d'expériences acquises dans les domaines traités.

L'équipe d'évaluation avec à sa tête un rapporteur devra réunir les compétences suivantes :

- Connaissance du milieu du tourisme
- Capacité à faire de l'analyse statistique
- Capacité à élaborer un échantillon représentatif
- Capacité à l'analyse économique et financière
- Aptitude à mener une enquête de terrain
- Capacité à faire de l'analyse statistique

***d) Coordination des travaux d'évaluation***

La mise en oeuvre de l'évaluation nécessite des contacts étroits et une coordination entre le chargé d'évaluation et le comité de suivi, dans ce but, un des membres du comité de suivi sera l'interlocuteur direct du chargé d'évaluation.



***e) Calendrier***

Juillet 1992 Elaboration du cahier des charges

15 Septembre 1992 Date limite de réponse à la consultation

15 Octobre 1992 Début des travaux

15 Décembre 1992 Remise d'un rapport intermédiaire présentant les résultats des phases 1 et 2

31 Mars 1992 Fin de l'étude et remise du rapport définitif d'évaluation.

***f) Maîtrise d'ouvrage et financement***

Chacune des phases de l'évaluation pourra être conduite par un maître d'ouvrage différent.

Le financement de cette opération sera assuré conjointement par l'Etat (Ministère du Tourisme) et la Région Bretagne. Les maîtres d'ouvrages respectifs ainsi que la part de chacun des financeurs seront déterminés à l'issue de la consultation.

## **ANNEXE II**

### **Liste des membres et des rapporteurs du Conseil scientifique de l'évaluation**

Jean LECA (Président)

Maurice M. BASLÉ

Michel CALLON

Jacques FREYSSINET

Frédéric JENNY

Christian JOIN-LAMBERT

Danièle LAMARQUE

Jean-Paul MARCHETTI

Claude SEIBEL

André TANTI

Jean-Claude THOENIG

Rapporteur général : Bernard PERRET

Rapporteur général adjoint : Claire GUIGNARD-HAMON

# Sommaire

## 1. Introduction

- 1.1. But poursuivi par le présent guide
- 1.2. Un éventail de pratiques administratives et cognitives
- 1.3. Les spécificités de l'évaluation
- 1.4. Commentaire sur quelques définitions
- 1.5. Evaluation, recherche, contrôle, audit
- 1.6. Les fonctions et finalités sociales de l'évaluation
- 1.7. Classement des "qualités" d'une politique évaluée
- 1.8. Actualité et avenir de l'évaluation

## Première partie : Conception et pilotage de l'évaluation

### 2. Vue d'ensemble sur une opération d'évaluation

- 2.1. L'évaluation comme procédure formalisée
  - 2.1.1. Produire des connaissances appropriables
  - 2.1.2. Deux niveaux de méthode
- 2.2. Caractéristiques générales d'une opération d'évaluation
  - 2.2.1. Qui demande l'évaluation, et pourquoi ?
  - 2.2.2. Des objets de nature diverse
  - 2.2.3. Temps de l'évaluation et "cycle de vie" de la politique évaluée
- 2.3. Mandat, projet, cahier des charges
- 2.4. L'importance du travail préparatoire

### 3. Le projet d'évaluation

- 3.1. Les étapes du projet
  - 3.1.1. Définition du champ de l'évaluation
  - 3.1.2. Repérage des objectifs, officiels ou implicites.
  - 3.1.3. Identification des finalités et enjeux de l'évaluation
  - 3.1.4. Hypothèses de départ
  - 3.1.5. L'élaboration d'un référentiel
  - 3.1.6. Thématiques et des registres de l'évaluation

- 3.1.7. Le questionnaire
- 3.1.8. Le dispositif d'évaluation
- 3.1.9. Le bilan informationnel
  
- 3.1.10. Le programme d'étude
- 3.1.11. Le choix des opérateurs
  
- 3.2. Retour sur l'orientation et la structure du questionnaire
  - 3.2.1. Mise en vigueur (*implementation*) de la politique
  - 3.2.2. Quels résultats observer et mesurer ?
  - 3.2.3. L'étude des mécanismes d'action
  - 3.2.4. Contexte de la mise en oeuvre et conditions de généralisation
  
- 3.3. Dès le projet, penser aux destinataires de l'évaluation
  - 3.3.1. L'explicitation des choix, des présupposés, et donc des limites de l'évaluation
  - 3.3.2. Nature des informations rassemblées et statut des arguments

#### **4. Conduite de l'évaluation et valorisation des résultats**

- 4.1. L'instance d'évaluation
  - 4.1.1. Rôle de l'instance
  - 4.1.2. Composition et fonctionnement
  - 4.1.3. Les phases du travail de l'instance
  - 4.1.4. Rôles possibles d'une instance de régulation scientifique
  
- 4.2. Valorisation des résultats
  
- 4.3. Le rapport d'évaluation
  
- 4.4. Les critères d'une bonne évaluation
  - 4.4.1. Définition et usages des critères de qualité
  - 4.4.2. Les avis rendus par le Conseil scientifique de l'évaluation sur les évaluations interministérielles
  - 4.4.3. Les cinq critères retenus par le CSE

### **Deuxième partie : Les outils de l'évaluation**

#### **5. Nature des informations mobilisées et registre de l'argumentation**

- 5.1. Le débat quantitatif/qualitatif

## 5.2. Le statut des assertions : description, démonstration et opinion

### 5.2.1. Le registre du constat

### 5.2.2. Le registre démonstratif et interprétatif

### 5.2.3. Le registre de l'opinion

## **6. Les techniques de collecte de l'information**

### 6.1. Mobiliser l'information préexistante

### 6.2. Des enquêtes et investigations spécifiques

### 6.3. Etudes documentaires et historiques

### 6.4. Les enquêtes

#### 6.4.1. Structure et présentation du questionnaire.

#### 6.4.2. Le suivi d'un "panel"

### 6.5 L'entretien

### 6.6. Les études de cas

### 6.7. Les groupes d'acteurs et d'experts

#### 6.7.1. Finalités

#### 6.7.2. Les différentes techniques

### 6.8. La perspective ethno-sociologique

## **7. Le traitement des informations par la statistique**

### 7.1. Définitions et principes méthodologiques

#### 7.1.1. Description résumée des données

#### 7.1.2. Analyse des données et typologies

#### 7.1.3. Techniques d'inférence et vérification probabiliste des hypothèses

### 7.2. Les cadres de rangement de l'information

#### 7.2.1. Nomenclatures et concepts.

#### 7.2.2. Cohérence des nomenclatures, "triangulation" des données.

## **8. Le problème de la causalité et l'analyse des systèmes complexes**

### 8.1. Les difficultés de l'expérimentation

#### 8.1.1. Position du problème

#### 8.1.2. Les limites de la méthode expérimentale dans les sciences de l'homme et

de la société

8.1.3. Protocoles pseudo-experimentaux

8.2. Causalité et modélisation

8.2.1. L'estimation d'un modèle

8.2.2. Les pièges de la modélisation

8.3. Les limites de l'approche causale et l'analyse des systèmes sociaux

8.3.1. Le problème de la circularité

8.3.2. L'analyse des systèmes d'action

**Index des principaux termes technique**

**Références des exemples cités**

**Ouvrages généraux sur l'évaluation des politiques publiques**

**Annexe I : Trois cahiers des charges d'évaluation**

**Annexe II: Liste des membres et rapporteurs du Conseil scientifique de l'évaluation**