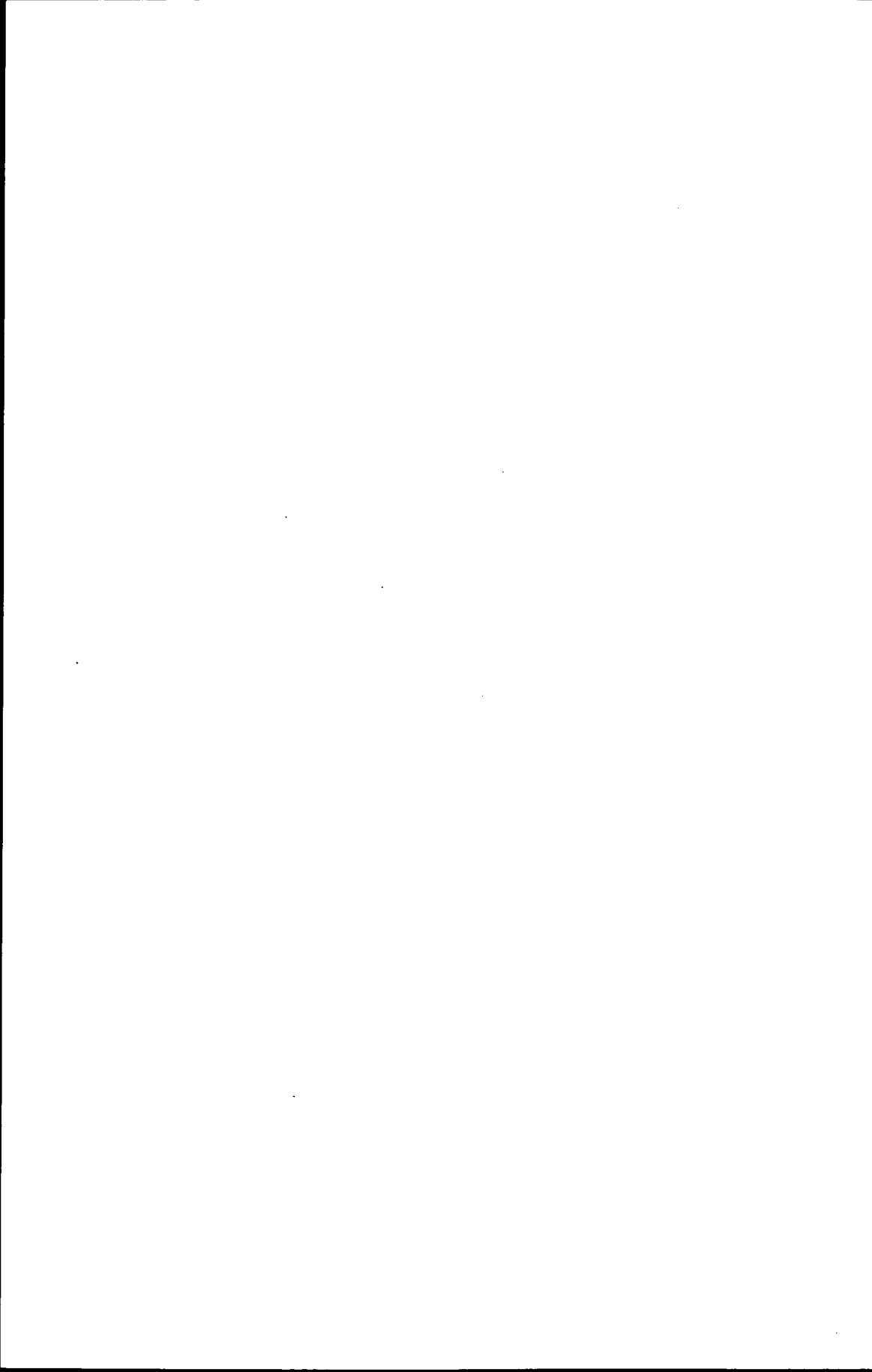


III

La mise en œuvre de l'évaluation



III

La mise en œuvre de l'évaluation

III - 1 Introduction

Les développements précédents ont concerné la **boîte à outils** de l'évaluation : plans de recherche alternatifs, techniques utilisables pour le traitement et l'interprétation des données collectées. Nous nous intéressons maintenant au **mode d'emploi** de cette boîte à outils. Autrement dit, il s'agit d'élaborer quelques propositions sur l'utilisation, **en situation**, de ces plans de recherche et de ces techniques : dans telle configuration d'évaluation, existe-t-il un meilleur plan de recherche ? Y-a-t-il au contraire certains arbitrages possibles etc. ? Telles sont les questions auxquelles un bon mode d'emploi devrait répondre.

Là encore il n'est pas envisageable, ni souhaitable, de fournir un guide *presse-bouton* qui aplanirait toutes les difficultés et résoudrait une fois pour toutes les problèmes que pose à l'évaluateur le choix de sa démarche. Chaque situation concrète appelle des ajustements spécifiques et même, souvent, l'utilisation combinée de plusieurs approches. On peut toutefois dégager quelques grands principes¹.

La mise en œuvre de l'évaluation est conditionnée par **deux facteurs principaux** :

— **la nature des questions posées** : s'interroge-t-on seulement sur la manière dont la politique a été appliquée (question descriptive aboutissant à une évaluation de mise en œuvre) ? Cherche-t-on en outre à apprécier ses effets spécifiques (question en forme de causalité demandant une évaluation de résultat) ? Veut-on, ou non, étendre les conclusions (problème de la validité externe) ?

— **les contraintes et limites** de diverses natures, qui sont des données pour l'évaluation, en matière notamment :

- de collecte de l'information ;
- de coût ;
- de délais ;
- d'insertion administrative.

Compte tenu des questions posées et des contraintes auxquelles il est confronté, l'évaluateur doit réaliser certains arbitrages « stratégiques », avant de procéder à l'évaluation proprement dite :

1. La démarche qui suit recoupe certains développements contenus dans l'ouvrage de méthode du GAO, *Designing evaluations* (op. cit.).

— arbitrages en matière d'objectifs de l'évaluation : à quelles questions est-il susceptible de répondre, et avec quel degré de validité ?

— arbitrages relatifs au sujet de l'évaluation, c'est-à-dire le champ examiné et le niveau auquel il se situe : faut-il chercher à évaluer des actions ponctuelles ou au contraire des politiques très globales ? Peut-on se placer à un niveau intermédiaire d'objectifs de l'action publique ?

Une fois ces arbitrages de nature stratégique effectués (paragraphe III-2), il reste à choisir le(s) plan(s) de recherche et la (les) technique(s) d'exploitation des données (paragraphe III-3). Au-delà de la phase d'évaluation proprement dite il convient encore de s'interroger sur « l'après-évaluation » (paragraphe III-4) : les conclusions obtenues permettent-elles de conforter, de réorienter, ou de remettre en cause certaines actions de la puissance publique ? Est-il souhaitable de poursuivre l'évaluation de la politique considérée en précisant le sujet et en se dotant de davantage de temps et de moyens ? Enfin, quelques règles simples de déontologie générale seront proposées (paragraphe III-5).

Les développements qui suivent fournissent ainsi des recommandations générales pour la conduite d'une évaluation. Comme il a été dit, ils ne peuvent viser à l'exhaustivité. En particulier, il serait artificiel d'émettre des conclusions figées sur la prise en compte de l'environnement « humain » de l'évaluation. On ne peut se borner, en la matière, qu'à une recommandation d'attention. Il est clair que l'évaluateur occupe souvent une position très particulière au sein d'un réseau d'acteurs : il est (ou il doit être — voir III-5) distinct de ceux qui décident la politique ou l'exécutent. Il peut alors être conduit à rechercher de l'information pertinente auprès de ces acteurs. Plus généralement, l'évaluation peut impliquer une multiplicité d'agents concurrents en termes de pouvoir ou de problématique, à l'intérieur ou à l'extérieur de l'administration. Deux écueils se présentent pour l'évaluateur. Le premier consiste évidemment à ignorer superbement cet environnement et à adopter l'attitude de celui qui sait. On risque, ce faisant, à la fois de rencontrer des difficultés importantes pour conduire l'évaluation et d'aboutir à un rejet de ses résultats. Le second est symétrique et consiste en l'abandon de toute position externe de l'évaluateur et, en particulier, de sa responsabilité finale, qu'il détient et ne peut céder : **émettre un jugement sur les résultats**. Cette démission conduit directement à une digestion de l'évaluateur par le système et l'acte d'évaluation dégénère en animation de groupe ou en vote de motion de synthèse.

Au total, et parce qu'il faudra bien dans de nombreux cas solliciter les acteurs pour analyser et comprendre la politique et répondre à ceux qui chercheront à exercer une pression sur l'évaluation, une voie moyenne peut être proposée. Cette voie suppose la fixation d'un **contrat** préalable et l'adoption d'un protocole adapté. Le contrat est d'autant plus facile à conclure que la procédure d'évaluation est définie à l'amont, dès la formulation de la politique à évaluer. Une telle démarche contractuelle peut être plus délicate à suivre lorsqu'il s'agit d'étudier des politiques bien « enracinées », génératrices de rentes pour certains acteurs. Dans ce cas, la recherche d'un consensus ne saurait servir d'excuse à l'immobilisme et les risques d'un conflit doivent être affrontés. De même, le protocole adopté, notamment à propos de la mise à contribution des acteurs pour produire et réunir des informations sur la politique, doit faire la part de la complexité du sujet traité, de

ses composantes subjectives. Si une attitude d'écoute est à recommander, il convient de ne pas perdre de vue l'objectif. Celui-ci n'est pas de faire comprendre aux acteurs que leurs préoccupations sont prises en compte, il est de produire un jugement sur la politique suivie.

III - 2 La stratégie de l'évaluation

En amont du choix proprement technique d'un plan de recherche particulier, la mise en œuvre de l'évaluation requiert l'examen de **certaines facteurs généraux d'environnement** et demande que soient fixées certaines **orientations stratégiques**. On traitera successivement des points suivants :

- les objectifs de l'évaluation et les contraintes d'environnement ;
- le niveau d'évaluation de la politique ;
- le choix des experts.

1) **Les questions auxquelles l'évaluation des politiques publiques peut être amenée à répondre sont diverses**. En allant du plus simple vers le plus complexe, on peut distinguer :

— des questions de nature **descriptive** :

- comment la politique a-t-elle été mise en œuvre ? et l'a-t-elle été ?
- sur quels instruments était-elle fondée ?
- quels étaient ses objectifs ?
- ces objectifs ont-ils été atteints ?

— des questions impliquant une analyse des **causalités** :

- si les objectifs de la politique ont été atteints, en quoi l'action publique en est-elle responsable ?
- en quoi les différents ingrédients de la politique publique ont-ils contribué à la réalisation de ces objectifs ?
- une autre politique aurait-elle été plus efficace ?
- peut-on généraliser les résultats de l'évaluation à une politique de même nature mais visant d'autres cibles (autres secteurs, autres groupes d'agents) ou se situant dans un autre contexte (une autre région par exemple) ?

— des questions aboutissant à un **jugement de valeur** :

- la politique a-t-elle eu des effets positifs ou négatifs ?
- que peut-on dire *ex post* de la pertinence de ses objectifs ?
- quels autres objectifs devraient éventuellement leur être substitués ?

Encadré n° 1

La mise en œuvre de l'évaluation par le General Accounting Office (*)

Le General Accounting Office (GAO) a été créé par le Budget and Accounting Act de 1921 qui en fait la contrepartie législative du Bureau du Budget (devenu Office of Management and Budget en 1970).

Le GAO est dirigé par un Comptroller General et un Deputy Comptroller General nommés pour quinze ans par le président des États-Unis avec l'accord du Sénat. Le Comptroller General n'est subordonné ni au Président ni au Congrès, et bénéficie du statut d'« Officer of the United States ».

Le principal domaine de compétence du GAO s'étend à toute question relative à l'utilisation des fonds fédéraux. Le GAO a un droit de regard sur tout document administratif et a toute latitude pour retenir le niveau qui lui paraît le plus adapté (simple contrôle financier, ou analyse d'efficacité).

Les évaluations réalisées par le GAO s'appuient sur des relations négociées avec les administrations, pour l'obtention des informations nécessaires et la mise en œuvre des réformes préconisées. La présence physique des évaluateurs du GAO au sein des administrations facilitent ces relations.

Le GAO est en fait très indépendant. Il accorde certes une certaine priorité aux demandes d'évaluation présentées par le Congrès, mais il a également ses propres priorités. Dans l'ensemble le GAO est très bien perçu par le Congrès.

Les autres compétences dont dispose le GAO ne sont pas sans lien avec l'évaluation : amélioration des systèmes comptables des administrations, interprétation des textes en matière de finances publiques ; conseil auprès du Congrès sur les questions législatives, réglementaires et juridiques ; compétence juridictionnelle contentieuse pour connaître notamment les litiges financiers portés par ou contre le gouvernement des États-Unis.

La « comprehensive evaluation » et sa mise en œuvre :

La « comprehensive evaluation » regroupe les différents types d'évaluations mises en œuvre par le GAO :

- l'audit des opérations financières et des systèmes de comptabilité ;
- l'audit d'« efficacité et d'économie » qui permet de savoir comment sont gérés les programmes ;
- l'évaluation portant sur l'atteinte des objectifs des programmes ;
- l'évaluation dégageant des approches alternatives à partir d'analyses coûts-avantages ;
- des « études spéciales » : recherches méthodologiques, promotion de nouvelles techniques de gestion.

(*) Voir la communication au groupe de Luc Rouban, « L'évaluation des politiques publiques aux États-Unis ».

Encadré n° 1 (suite)

La « comprehensive evaluation » comprend à la fois des évaluations visant à apprécier si une politique a été correctement mise en œuvre (process evaluations) et des évaluations ayant pour objet de mesurer l'effet d'une politique et son efficacité (impact evaluations). Le GAO a cherché à mieux articuler ces deux types d'évaluations bien qu'il ait tendance en raison du contexte institutionnel dans lequel il fonctionne (évaluations faites à la demande du Congrès, supervision de la gestion des administrations) à mettre davantage l'accent sur les « process evaluations ».

L'ampleur d'une évaluation réalisée par le GAO dépend :

- de l'intérêt qu'il trouve à étudier la question posée ;
- de l'existence préalable d'informations (notamment de rapports précédents d'évaluation) ;
- de l'éventualité d'un risque de conflit avec le Congrès ou l'Administration pour des raisons de nature politique.

Selon l'objet de l'évaluation, le GAO peut avoir recours à toutes les techniques possibles. Toutefois, le fait qu'il intervienne souvent « au second degré », à partir d'évaluations réalisées par les administrations responsables et dont il peut vérifier le protocole, explique qu'il ne fasse que rarement appel aux méthodes de quasi-expérimentation ou d'expérimentation pure, ou encore à l'analyse sociologique.

Les rapports d'évaluations réalisées par le GAO ne sont pas tous rendus publics : certaines études restent confidentielles.

La mise en œuvre de la « comprehensive evaluation » nécessite des moyens importants : le GAO dispose de 5 000 agents répartis entre l'état-major de Washington, 80 bureaux d'audits installés dans certaines administrations, 15 bureaux régionaux et 3 bureaux internationaux.

La sélection des objectifs de l'évaluation est faite à partir de vingt-neuf aires de recherche (« issue areas »). Chaque thème est affecté à un département du GAO qui établit un plan de travail sur une base de deux à quatre ans.

Le personnel du GAO est opérationnel à 80 %. Il comporte des spécialistes de différents domaines (comptabilité, droit, économie, gestion, informatique, etc.).

La plupart des départements du GAO établissent en outre des listes d'experts extérieurs sur des sujets spécialisés. Deux groupes d'experts, l'un composé de membres du secteur privé et de l'administration, l'autre d'universitaires et d'experts en sciences sociales, ont en outre un rôle consultatif.

Le prescripteur de l'évaluation attend bien évidemment de celle-ci qu'elle lui fournisse le maximum de réponses aux questions qu'il se pose sur l'efficacité de la politique étudiée. Il attend en outre que ces réponses lui soient apportées à temps pour que l'action menée puisse être, si nécessaire, infléchie ou révisée. Il est enfin tenu de veiller à ce que le coût de l'évaluation reste dans les limites qu'il juge raisonnable.

Le problème de l'évaluateur est donc « d'optimiser » la finesse et la robustesse de ses conclusions sans pour autant dépasser les limites qui lui sont fixées, notamment en matière de temps et de budget. En règle générale, aller au-delà du constat simplement descriptif et approfondir les analyses de causalité exige de recourir à des techniques sophistiquées s'appuyant sur des dispositifs relativement lourds et coûteux : le souci de validité entre donc naturellement en conflit avec les contraintes qui s'imposent à l'évaluateur.

Il serait injustifié de tirer de cette constatation des conclusions exagérément pessimistes. Certes, dans bien des cas, le poids des contraintes obligera l'évaluateur à se fixer des objectifs plus modestes qu'il ne l'aurait souhaité, soit en restreignant son champ d'investigation, soit en recherchant seulement des présomptions et non pas des preuves réfutables quant aux effets de la politique. Pour des raisons déontologiques, une telle acceptation de résultats incomplètement fiables serait inadmissible dans certains domaines d'action. En matière de thérapeutique notamment, il est inconcevable de décider la diffusion d'un médicament ou d'un traitement sans que les effets et l'efficacité en aient été décrits complètement dans le détail. Dans d'autres domaines, en revanche, celui des politiques sociales en particulier, il peut être tout à fait judicieux de se contenter d'une rigueur moins absolue² : ainsi, pour le décideur, l'évaluation de politique qui établit de simples présomptions est déjà significative, surtout si elle permet d'infléchir l'action en temps utile.

En perdant en validité³ pour gagner en rapidité, l'évaluation permet dans certains cas d'améliorer l'efficacité de l'action publique. A la limite, une évaluation menée avec des techniques sophistiquées, et qui aboutit alors que la politique considérée s'est déroulée complètement, peut n'avoir aucun intérêt autre que historique. A l'opposé, une évaluation de mise en œuvre menée assez près de l'action elle-même et avec des techniques rustiques peut permettre d'arrêter ou d'infléchir une politique ayant des effets pervers inattendus. Toutefois, cette dernière façon de faire n'est pas compatible avec la prise en compte d'effets indirects sur longue période.

Les observations précédentes ne signifient évidemment pas que l'exigence de rigueur doive être volontairement négligée : il convient au contraire que l'évaluation soit aussi rigoureuse que possible, dans le respect des contraintes dans lesquelles elle s'insère. Elles suggèrent plutôt une démarche de réponse graduée :

2. La démonstration totalement rigoureuse étant par ailleurs souvent impossible dans un milieu où les mécanismes en cause sont de nature aléatoire et évolutive.

3. Il est probable que les réflexes de rejet vis-à-vis d'évaluations perçues comme « sales » sont assez forts dans les milieux scientifiques français, moins familiers des démarches pragmatiques que leur homologues américains. D'où les reproches d'« académisme » formulés parfois par les décideurs administratifs à l'encontre d'évaluations réalisées par des centres d'études et de recherche.

— dans un premier temps, il convient de tirer tous les renseignements possibles des techniques les plus faciles à mettre en œuvre. La **diversification des plans de recherche**, non affectés par le même biais, est susceptible de pallier le faible degré de validité de chacun d'entre eux, pris individuellement, et peut permettre, dans certains cas, d'aboutir à des présomptions fortes quant aux effets de la politique ; une telle démarche a été mise en œuvre lors de l'évaluation du plan textile ;

— si ces présomptions paraissent trop sujettes à caution, une évaluation complémentaire menée avec des moyens plus importants (dispositif d'observation, approche expérimentale, utilisation de techniques statistiques complexes, etc.) peut être alors déclenchée. Cette deuxième phase exige des délais plus longs et peut nécessiter une certaine implication des services opérationnels concernés, ainsi qu'un appel plus massif à des compétences extérieures à l'administration (des laboratoires universitaires à caractère « permanent » dont on cherche à déterminer les ingrédients les plus appropriés (aides à des secteurs en difficulté, aides à l'emploi, etc.) à partir d'une analyse approfondie du comportement des acteurs.

Enfin, l'importance de la nature des informations disponibles va conditionner étroitement la mise en œuvre de l'évaluation :

— **le contexte le plus favorable correspond à l'existence d'un dispositif de suivi de la politique.** Bien que ce cas ne soit pas le plus fréquent dans notre pays, il n'est pas exceptionnel (insertion des jeunes de 16 à 18 ans, sécurité routière, périnatalité, TGV Sud-Est, travaux d'utilité collective, préretraites) et il semble se développer (souci de mettre au point un dispositif permanent en matière d'aides à l'emploi). Aux États-Unis, la mise en place d'un dispositif de suivi intervient presque systématiquement lors du lancement d'un nouveau programme social.

L'existence d'un tel dispositif permet de mener des analyses approfondies (recours aux méthodes statistiques) et présente l'avantage de fournir des observations d'unités individuelles diversifiées.

Certains de ces dispositifs autorisent en outre des comparaisons de groupes non équivalents (quasi-expérimentation menée dans le cadre des stages de formation des jeunes de 16 à 18 ans par exemple).

— **en l'absence de dispositif de suivi, il convient de ne pas négliger les informations provenant des systèmes d'information existants, en particulier les fichiers de l'administration.** Leur utilisation pose toutefois une série de problèmes : secret statistique, destruction de l'information après clôture de l'exercice (difficulté rencontrée notamment dans le cadre de l'évaluation de l'API), manque de normalisation, etc. S'ils donnent une bonne image de la population atteinte, ces fichiers ne permettent pas, en revanche, de constituer des échantillons représentatifs de l'ensemble de la population visée.

Par ailleurs, l'existence d'organismes relais qui ont en charge tout ou partie de l'action menée (le Crédit Agricole pour les prêts bonifiés à l'agriculture par exemple) introduit un degré de complexité supplémentaire et peut rendre plus difficile la mobilisation de l'information nécessaire à l'évaluation.

Le système statistique (INSEE, services statistiques des ministères) fournit des informations stables et qui peuvent être utiles à l'évaluation de politiques publiques même si elles ne sont pas collectées à cette fin. Les grands fichiers (répertoires d'entreprises, recensement) permettent en particulier de tirer des échantillons

représentatifs nécessaires à la mise en place d'un dispositif d'évaluation (l'enquête auprès des entreprises sur les aides à l'emploi menée dans le cadre de l'étude prioritaire RCB sur ce thème s'est appuyée par exemple sur un échantillon d'entreprises demandé à l'INSEE).

2) Le manque de transparence des objectifs de certaines politiques publiques a été évoqué à plusieurs reprises au sein du groupe pour illustrer les obstacles à l'évaluation. Ce manque de transparence peut avoir plusieurs origines :

— l'imprécision du contenu de la politique

La politique de santé, par exemple, ne fait pas directement référence, dans ses objectifs, à la notion de santé. Elle est définie en termes d'organisation et de gestion de l'offre de soins et de prévention. Or, mesurer les effets d'une politique d'offre de soins ne permet pas toujours d'en déduire son impact sur la santé des individus. C'est négliger les mécanismes complexes afférents à l'ajustement de l'offre à la demande. Ainsi, en l'absence d'un équipement lourd donné, un hôpital dirige ses malades vers un centre spécialisé non saturé et mieux qualifié. Lorsqu'il aura installé ce matériel, il permettra de mieux localiser les traitements mais ceci n'aura qu'un effet réduit sur les malades eux-mêmes.

— l'existence d'objectifs non affichés

Ce cas de figure est assez fréquent :

- les politiques agricoles ne comportent pas d'objectifs explicites sur le nombre des agriculteurs ; et pourtant le maintien d'un milieu rural constitue une préoccupation réelle et permanente des décideurs ;
- de même, les aides à l'emploi qui accompagnent les restructurations industrielles ont notamment pour vocation de limiter les conflits sociaux pouvant survenir à cette occasion ; cette ambition reste largement implicite ;
- quand aux objectifs de certaines politiques d'emploi, ils paraissent pour le moins ambigus. Suivant quel critère apprécier les stages destinés aux jeunes : l'insertion dans la vie active, l'amélioration de la qualification, la lutte contre la marginalisation irréversible, ou encore l'impact sur les statistiques du chômage ? Les résultats sont souvent limités vis-à-vis des deux premiers objectifs : les taux de placement après le stage restent faibles, les stages de qualification sont minoritaires. Les deux autres objectifs n'ont pas été affichés explicitement. L'un se situant dans le moyen terme échappe à des évaluations courtes, tandis que l'autre a probablement une priorité égale aux yeux des décideurs politiques, ce qui explique le recours systématique à de telles politiques dans la plupart des pays.

— la coexistence d'objectifs antinomiques, au moins à court terme

L'existence d'objectifs contradictoires est également assez fréquente. Elle se manifeste par exemple dans le cadre de politiques d'aides sectorielles à l'industrie qui cherchent d'une part à freiner l'hémorragie d'emplois dans un secteur et, d'autre part, à améliorer sa compétitivité, ce qui peut supposer une réduction d'effectifs. Cette ambivalence existait dans le cas du plan textile. Les politiques visant à améliorer la productivité des exploitations agricoles françaises vis-à-vis de l'étranger tout en limitant la diminution du nombre d'agriculteurs fournissent une autre illustration de ce phénomène.

Analyser les objectifs d'une politique est donc un exercice délicat ; l'évaluateur peut être conduit à cette occasion à reconstituer *ex post* et au prix de beaucoup d'efforts un regroupement et une cohérence qui n'existaient pas nécessairement auparavant. Bien que chaque expert ait souvent tendance à considérer ces difficultés comme propres au secteur dont il est chargé — pour le spécialiste de l'agriculture, l'action publique apparaît plus facile à cerner en matière d'emploi ou de santé... et réciproquement —, les travaux du groupe ont montré qu'il s'agit d'un problème général qui appelle la mise en œuvre d'instruments largement communs. L'analyse sous forme de structure d'objectifs — programmes (S.O. — P.)⁴, peut ainsi mettre en évidence les différents niveaux d'objectifs de la politique étudiée. Elle constitue sans doute un point de départ pour « débroussailler » un sujet complexe dont la cohérence interne n'apparaît pas au premier coup d'œil. Ce type de démarche a par exemple été retenu dans le cadre de l'évaluation des aides à l'élevage bovin ; il a permis d'établir l'existence de trois niveaux différents, deux d'objectifs et un d'action (cf. encadré 1 — partie I) :

- **les objectifs généraux** (finalités de la politique) ne sont pas spécifiques à l'élevage bovin : participation à la croissance, redressement du commerce extérieur, sécurité d'approvisionnement, augmentation du revenu des agriculteurs, équilibre du milieu rural ;
- **les objectifs opérationnels** portent plus spécifiquement sur la filière bovine : orientation des productions, organisation des producteurs, amélioration de la productivité du cheptel et de la qualité des produits, etc. ;
- à ces objectifs opérationnels correspondent **des actions** et des moyens qui ont été regroupés en programmes : amélioration génétique, prophylaxie, interventions sur les marchés, etc.

Cette décomposition — moyens → programmes → finalités — conduit à privilégier le niveau des objectifs opérationnels, ce qui permet d'éviter à la fois le caractère flou des évaluations de politiques globales et l'absence de vision cohérente de l'action publique qui risque de caractériser une évaluation menée à un niveau plus fin (celui des programmes élémentaires constituant la politique). Si le pouvoir explicatif de l'évaluation s'en trouve amélioré, il se peut toutefois que ce soit au prix d'un affaiblissement de sa problématique : l'évaluateur s'efforce de voir si les objectifs opérationnels ont été atteints et si la politique suivie en est responsable, il renonce par contre le plus souvent à juger *ex post* de la pertinence de ces objectifs. Prenons l'exemple du plan textile : il est certes plus facile d'appréhender ses effets en termes d'emplois dans le secteur textile (objectif opérationnel) que de mesurer son impact sur l'emploi global (finalité), notamment parce qu'un emploi créé a un endroit ou dans un secteur est peut-être un emploi délaissé ailleurs. Toutefois en se limitant aux effets internes au secteur textile, l'évaluation permet déjà de juger de l'adéquation du plan aux objectifs opérationnels visés (emplois et compétitivité du secteur) ; elle ne permet en aucun cas de dire si l'État a eu raison, du point de vue de ses objectifs globaux en matière d'emploi et de compétitivité de l'économie, de soutenir le textile plutôt qu'un autre secteur.

4. Pour reprendre une dénomination employée dans les premiers temps de l'opération RCB.

On peut également observer une certaine confusion entre les objectifs internes ou externes en particulier dans le domaine des politiques sociales. Ainsi par exemple, l'évaluation de tel programme éducatif sera très différente si l'on cherche à évaluer la façon dont les élèves assimilent le programme (point de vue interne) ou si on évalue l'utilité des capacités acquises pour la vie professionnelle (point de vue externe). De la même manière, il est différent d'évaluer les TUC par rapport à la situation d'emploi ou par rapport aux conditions de vie qu'ils permettent aux intéressés. Dans les deux cas les politiques visent simultanément les deux catégories d'objectifs mais l'évaluation peut légitimement privilégier l'un ou l'autre objectif. Le problème est de procéder à l'analyse préalable des objectifs et d'afficher les choix sur lesquels repose l'évaluation.

Ce type de difficultés risque d'apparaître chaque fois que la contribution des objectifs opérationnels aux finalités de l'action de l'État ne semble pas aller de soi.

3) **La réalisation de l'évaluation nécessite de recourir à des experts.** Le choix de ceux-ci doit respecter un certain nombre de principes :

— ces experts doivent être externes par rapport à la politique à évaluer. Ce principe de séparation entre l'évaluation et l'exécution (cf. partie III-5) est valable, que l'évaluation soit elle-même administrative ou non. Ainsi le service d'évaluation du ministère de l'Éducation nationale⁵ est-il directement rattaché au ministre, et donc indépendant des directions opérationnelles responsables de l'élaboration, des programmes d'enseignements. Toutefois, la réalisation de l'évaluation se fait avec la collaboration des enseignants et sur la base de critères issus des directives de l'Inspection générale.

— le groupe d'experts réunit des spécialistes de la politique à évaluer et des spécialistes des méthodes d'évaluation. Il convient d'une part d'éviter le danger d'imprécision de l'évaluation liée à une familiarité insuffisante avec les plans de recherche et les techniques utilisables et, d'autre part, de veiller à ce que le groupe puisse disposer de toutes les informations nécessaires sur la politique à évaluer. La présence, au sein du groupe, d'experts ayant une compétence technique pointue tant dans les différentes disciplines concernées (économie, droit, psychologie, sociologie,...) que sur le sujet traité, peut-être indispensable. Il en est ainsi, notamment, lorsque l'évaluation fait appel à la modélisation ou à des méthodes statistiques complexes. Il est en outre très souhaitable d'associer ces experts à la mise en place du dispositif d'évaluation, de manière à assurer une cohérence entre ce dispositif et les conditions de mise en œuvre de l'évaluation (type d'informations, nature des données).

— la désignation du groupe d'experts appartient à l'autorité qui a le pouvoir de déclencher l'évaluation ; le mandat confié au responsable du groupe doit définir implicitement le contenu minimal de l'évaluation sans pour autant limiter de façon excessive la réflexion, celle-ci doit pouvoir se développer sur tous les sujets et les problèmes et utiliser tous les instruments susceptibles d'affermir l'évaluation entreprise.

5. Le SPRESE : voir encadré n° 3, partie II.

Encadré n° 2

Mise en œuvre de l'évaluation et choix des experts : l'exemple du comité national d'évaluation des universités

Le Comité national d'évaluation des universités est un organe spécifique chargé d'évaluer les établissements d'enseignement supérieur de l'Éducation nationale. Cette évaluation, qui constitue la contrepartie de l'autonomie accordée aux établissements d'enseignement supérieur, est réalisée établissement par établissement et par thèmes transversaux.

Le Comité se compose de 15 personnalités nommées pour quatre ans et d'un secrétariat général.

Les établissements faisant l'objet d'une évaluation seront découpés en unités. Pour chaque unité, le Comité enverra un groupe de deux ou trois experts, choisis par lui-même parmi des enseignants, des chercheurs, des cadres d'entreprises et des personnalités étrangères. Ces experts devront posséder une certaine légitimité : ils doivent être jugés par les responsables des unités étudiées comme ayant toutes les capacités requises par l'évaluation.

Le Comité cherchera à réunir toute l'information possédée par l'administration centrale afin de fournir aux experts une première synthèse qui pourra éclairer sa mission.

III - 3 Le choix d'un plan de recherche

Une fois l'environnement « stratégique » de l'évaluation exploré et défini reste une question de logistique importante : **le choix d'un (ou de) plan(s) de recherche et des techniques afférentes**. Ce choix va dépendre de divers facteurs, dont les suivants :

- la nature descriptive ou causale de la question posée. S'agit-il d'une évaluation de mise en œuvre ou d'une évaluation de résultat ?
- le souci de généraliser les conclusions de l'évaluation ;
- les contraintes de délais et de budget de l'évaluation ;
- la possibilité de multiplier les observations en coupe transversale sur des sujets assez différents les uns des autres ;
- l'existence d'observations de même nature avant et après mise en œuvre de la politique ;

— la possibilité de trouver ou de constituer des groupes de comparaison et éventuellement de recourir à l'expérimentation avec tirage au sort ;

— l'existence d'une réflexion théorique permettant l'élaboration de modèles.

Cette liste de facteurs fait apparaître une multiplicité des cas de figure possibles et montre qu'il serait illusoire en l'état actuel des choses de chercher à associer de manière systématique un plan de recherche à une évaluation d'un type donné. Un certain nombre de propositions générales peuvent toutefois être suggérées.

○ En dehors des contraintes de coût, de délais et de mise en œuvre, le choix d'un plan de recherche dépend essentiellement des exigences de l'évaluation :

— lorsque la question posée est simplement descriptive, une bonne validité de la réponse pourra se satisfaire de dispositifs simples ; les plans de recherche sans comparaison avec des groupes témoins sont alors assez adaptés ;

— si l'on cherche en plus à savoir dans quelle mesure la politique est à l'origine des effets observés, le souci de validité conduit à comparer des groupes affectés et des groupes non affectés par la politique. Si ceci s'avère impossible, des analyses avant-après peuvent toutefois apporter certains enseignements d'autant plus précieux qu'ils résultent de l'observation de panels ;

— enfin, le souci de généraliser les conclusions de l'évaluation à d'autres cibles (autres secteurs, autres groupes d'agents) ou à d'autres contextes (autres pays, autres régions, etc.) demande que soit déterminé l'ensemble des facteurs qui peuvent interférer avec la politique et que les effets en soient mesurés. Ceci passe de préférence par des comparaisons entre groupes et peut donner lieu à une modélisation plus ou moins poussée. Pour certaines politiques caractérisées par l'impossibilité ou la grande difficulté des comparaisons entre groupes, des approches plus simples pourront tout de même mettre en évidence des facteurs particulièrement favorables (ou défavorables) à la réussite de la politique (l'exemple de la prophylaxie de la brucellose, évoqué précédemment, est significatif de ce point de vue).

● La deuxième conclusion est que des contraintes très fortes de coût et de délais conduisent à privilégier des approches rapides et peu approfondies et rendent difficile la mise en place de dispositifs lourds d'observation :

— dans ce cas, si la question posée est de nature descriptive, il se peut que ces contraintes imposent le recours à des études de cas, de préférence à des analyses basées sur des enquêtes statistiques qui nécessitent davantage de temps et de moyens ;

— si l'on s'interroge sur les liens de causalité entre la politique et les effets observés, de telles contraintes n'autorisent guère la mise en œuvre d'approches expérimentales ou quasi-expérimentales qui seraient par ailleurs les plus adaptées ; dans ce cas l'évaluation n'a pas lieu d'être entreprise ;

— le groupe a pu constater que les comparaisons de groupes témoins non équivalents issus de l'observation sont les plus fréquemment retenues en situation de pénurie relative de moyens. Ceci n'est pas le fruit du hasard : en dépit des critiques méthodologiques qui peuvent lui être adressées, ce plan de recherche réalise probablement le meilleur compromis entre le respect des contraintes et le souci de validité de l'évaluation. Si sa validité interne n'est en soi guère

satisfaisante, il permet toutefois de dégager des présomptions de causalité assez solides, surtout lorsqu'on peut multiplier les groupes témoins indépendants (ce fut notamment le cas pour la prophylaxie de la brucellose et à un moindre degré dans le cadre de l'évaluation du plan textile). A cet égard, pour l'évaluation de nombreuses politiques publiques, le recours à des comparaisons internationales ou régionales constitue sans doute un instrument précieux, mais à utiliser et surtout à interpréter avec prudence ;

— lorsque la comparaison entre groupes non équivalents n'est pas envisageable, l'analyse de séries temporelles (avant-après) peut constituer une solution de repli. Toutefois les corrélations qu'elle est susceptible de montrer entre des inflexions ou des ruptures de tendance d'une part et la mise en œuvre de la politique d'autre part ne prouvent pas vraiment l'existence de liens de causalité ;

— si l'analyse avant-après n'est pas non plus possible, il est nécessaire de se contenter d'analyses *ex post* reposant sur des enquêtes statistiques, sur des études de cas ou encore sur des avis d'experts. Cependant, ces techniques se limitent dans les cas les plus favorables à confirmer (ou plutôt à ne pas infirmer) l'existence d'un lien de causalité que l'on avait de fortes raisons de soupçonner au départ. Le recours à des avis d'experts peut également se révéler très utile à titre d'instrument de complément, de même que l'évaluation de synthèse consistant à tirer parti des études d'évaluation déjà réalisées.

• Les contraintes de délais risquent d'apparaître particulièrement **incompatibles avec le souci de généraliser les conclusions de l'évaluation**, à moins que des comparaisons de groupes non équivalents permettent de déterminer sous quelles conditions la politique évaluée peut contribuer efficacement aux objectifs poursuivis. Le recours simultané à plusieurs plans de recherche de mise en œuvre rapide, à l'image de ce qui a été réalisé dans l'évaluation du plan textile, peut également constituer un moyen de préserver la validité de l'évaluation en dépit de ces contraintes de délais.

III - 4 L'après-évaluation

Le rapport d'évaluation doit théoriquement fournir aux décideurs des éléments leur permettant de statuer en pleine connaissance de cause sur l'avenir de la politique étudiée : poursuivre cette politique, la renforcer, ou au contraire l'infléchir, voire y mettre fin. Ceci implique en premier lieu un effort de « lisibilité » de la part des auteurs du rapport. Ceci suppose en outre que l'évaluation possède une validité interne et externe suffisante et que les limites de cette validité soient précisées. Sinon, les résultats de l'évaluation risquent d'être utilisés de façon inappropriée, ou bien encore les divergences d'appréciation quant aux effets de la politique suivie ne seront pas vraiment atténués : ainsi, le bilan du

L'évaluation de la politique d'aide à l'élevage bovin (*)

L'élevage bovin se caractérise en France à la fois par son importance (près des deux tiers des exploitations agricoles sont concernées) et par les retards qui l'affectent (en particulier une productivité nettement inférieure à celle enregistrée en Allemagne et aux Pays-Bas), en dépit d'une longue politique de soutien des pouvoirs publics (de 1972 à 1983 l'effort financier de l'État est passé de 5 % à 14 % de la valeur de la production de la filière bovine).

La réalisation de l'étude prioritaire interministérielle d'évaluation confiée à Michel Colonna, ingénieur général du Génie rural et des Eaux et Forêts, s'appuie sur un groupe de travail composé pour l'essentiel d'ingénieurs généraux du Conseil général du Génie rural et des Eaux et Forêts (CGREF), de représentants du ministère de l'Agriculture et du ministère de l'Économie, des Finances et du Budget.

La démarche de l'étude d'évaluation :

1. Les niveaux d'objectifs considérés :

L'étude a distingué deux niveaux d'objectifs de la filière bovine : des **objectifs généraux** qui ne sont pas spécifiques à l'élevage bovin (contribution à la croissance économique et au commerce extérieur, maintien du cadre de vie rural, etc.), et des **objectifs opérationnels** (organisation de la filière bovine, amélioration de la productivité). A ces objectifs contribuent des actions regroupées en programmes (amélioration génétique, prophylaxie, interventions sur les marchés, etc. (cf. encadré n° 1, p. 30). Le travail du groupe a essentiellement porté sur les objectifs opérationnels.

Cette démarche n'a pas toujours pu être suivie, notamment lorsque coexistaient au sein d'un programme des actions d'orientation (associées à des objectifs explicites) et des actions d'accompagnement (destinées à répondre au coup par coup à certaines difficultés). C'est le cas notamment de la politique de modernisation des abattoirs publics.

2. Les niveaux d'efficacité distingués :

Trois niveaux d'évaluation de la politique d'élevage bovin ont été distingués :

- l'évaluation de la réalisation des actions ;
- l'évaluation du degré d'atteinte des objectifs ;
- l'évaluation de l'impact des mesures sur la satisfaction des bénéficiaires (à partir d'entretiens avec des acteurs locaux de la politique).

3. Les plans de recherche utilisés :

L'appréciation des effets de la politique a été réalisée essentiellement à partir de comparaisons de groupes témoins non équivalents :

(*) Voir la communication au groupe de Claude le Pen : présentation méthodologique de l'étude d'évaluation de la politique publique de l'élevage bovin.

Encadré n° 3 (suite)

- comparaisons internationales des performances de la filière bovine en France et chez nos principaux partenaires européens (Allemagne, Pays-Bas, Royaume-Uni) ;
- comparaisons régionales dans le cas de la prophylaxie de la brucellose (cf. encadré) ;
- comparaisons entre exploitants pour analyser les conséquences des plans de développement.

Lorsque ce plan de recherche n'a pu être mis en œuvre, le groupe a eu recours à des jugements d'experts (exemple : aide aux bâtiments d'élevage, monopole de zone dans le cas de l'insémination artificielle).

plan textile, mené par un groupe de travail interministériel a-t-il donné lieu à des interprétations différentes, en ce qui concerne l'efficacité du plan. La poursuite des travaux d'évaluation et le recours à des techniques et des plans de recherche plus performants peuvent, dans ce type de situation, être proposés comme un préalable aux décisions à prendre en ce qui concerne la politique analysée, surtout s'il s'agit de l'étendre à d'autres groupes. Toutefois, cette conclusion de l'évaluation en forme de poursuite du processus d'évaluation n'a pas lieu de servir de prétexte pour temporiser ou pour voiler le fait que la question posée n'appelle pas vraiment de réponse.

A l'image du contrôle administratif, **l'évaluation doit pouvoir donner lieu à un échange contradictoire** avec les services responsables de l'élaboration et de la mise en œuvre de la politique. Cela suppose que ces services disposent des éléments qui leur permettent de répondre point par point aux arguments avancés par le groupe d'évaluation (accès aux sources d'informations utilisées), et qu'ils soient à même de critiquer s'il y a lieu la méthodologie adoptée. Il se peut toutefois que cette dernière condition soit difficile à remplir lorsque l'évaluation repose sur des approches à caractère technique (la modélisation par exemple) à moins que ces services ne se dotent, avec l'appui de centres de recherches, d'instruments concurrents de même nature, mais reposant sur leur propre interprétation des mécanismes de la politique.

La **publicité** du rapport d'évaluation est un aspect sur lequel il n'est pas possible de fournir des recommandations systématiques, sauf à donner dans la naïveté. Ainsi aux États-Unis, le degré de diffusion des évaluations réalisées par le GAO est-il très varié. S'il est a priori souhaitable que les résultats soient, dans toute la mesure du possible, rendus publics, il apparaît légitime que la décision relève du **prescripteur** de l'évaluation. S'agissant de politiques très controversées, il se peut que l'éventualité d'une trop large diffusion incite les évaluateurs à une prudence exagérée. Enfin, certaines évaluations sont destinées par nature à une confidentialité plus marquée (celles notamment qui concernent des politiques touchant à la défense nationale). Une norme minimum peut consister à annoncer, avant l'évaluation, si celle-ci sera, ou non, publique.

Un schéma gradué peut également constituer une bonne solution dans certains cas. Il consisterait à séparer deux niveaux de publicité :

— un ensemble de données de bases déjà analysées, mises en perspective, mais sans qu'aucun jugement de valeur ne soit porté. Ce document serait partagé entre tous les partenaires (y compris le public) ;

— un rapport qui porte explicitement un jugement et qui peut dans certains cas rester confidentiel. Dans les deux cas le prescripteur dispose d'un droit d'antériorité qui lui permet de préparer sa « défense ».

III - 5 Pour une déontologie de l'évaluation des politiques publiques

A ce point de l'analyse il importe d'avancer les éléments d'un « code professionnel » propre à permettre l'exercice d'une évaluation conforme à ses finalités, en évitant les dangers précédemment mentionnés.

Ce code doit évidemment rester simple, eu égard aux variétés de situations concernées. Sa définition nous paraît pouvoir s'inspirer à la fois de certaines règles suivies en matière de contrôle administratif comme de production scientifique.

Deux grandes règles des contrôles administratifs semblent susceptibles d'avoir leur contrepartie dans le domaine de l'évaluation : le caractère contradictoire du contrôle, la séparation entre contrôleur et contrôlé. Plusieurs considérations justifient cette référence. Tout d'abord, mis à part les aspects plus spécifiquement techniques comme les finalités poursuivies, la conduite de l'évaluation soulève des questions dont certaines sont similaires à celles rencontrées par le contrôle administratif. On peut donc bénéficier de l'expérience acquise dans le cadre de cette activité. Les résultats en sont d'ailleurs probants. La mise en place des contrôles administratifs avait ainsi pour objet de garantir une administration « régulière »⁶. En dépit d'écarts ponctuels toujours fâcheux qui font les délices de la presse, on peut estimer, notamment par comparaison à la situation d'autres pays, que cet objectif a été globalement atteint en France.

L'absence de lien avec l'objet étudié, la confrontation avec des points de vue différents, sont des soucis qui se retrouvent également dans l'activité scientifique. Ils autorisent la production, le partage et l'accumulation des connaissances.

Nous proposons donc les deux règles fondamentales suivantes pour la conduite de l'évaluation :

Reproductibilité : l'évaluation doit être conduite suivant des protocoles qui la rendent reproductibles par d'autres évaluateurs.

6. Voir, sur ces points, l'introduction de l'ouvrage de Pierre Milloz, *Les inspections générales ministérielles dans l'administration française*, 1983, Éditions Economica.

Séparation : l'évaluateur, responsable des conclusions de l'évaluation, doit être une personne indépendante des agents (ou organisations) responsables de la formulation ou de l'exécution de la politique évaluée.

Chacune de ces règles, principalement la deuxième, moins généralement acceptée a priori que la première, appelle quelques commentaires :

1) La règle de **reproductibilité** correspond à l'impératif simple suivant : étant donné une évaluation réalisée par un certain évaluateur, tout autre évaluateur disposant des mêmes informations et suivant les mêmes procédures de recherche doit pouvoir aboutir au même résultat.

Cette règle passe en particulier par **l'affichage d'un protocole clair d'évaluation** précisant l'objet de l'évaluation, les sources d'information utilisées, les plans de recherche et les techniques employées. Cette exigence de transparence implique également le refus d'« arguments d'autorité » et constitue une garantie contre les risques de charlatanisme.

La fidélité à cette règle permet d'éclairer les utilisateurs de l'évaluation sur ses limites de validité. Elle autorise par ailleurs la comparaison d'évaluations différentes faites sur le même sujet.

2) La règle de **séparation** apparaît moins immédiate à certains. Dans sa discussion nous mettrons à part les pratiques consensuelles auxquelles le terme d'évaluation est parfois associé mais dans un sens différent de celui retenu ici et qui concernent plus des phases de formulation de politiques que de véritables appréciations *ex post*⁷.

Ce problème de séparation se pose évidemment dans le cas où une organisation a été impliquée de manière active dans la formulation ou l'exécution de la politique évaluée. Cette organisation peut alors avoir un avantage, ou un monopole, d'information. En outre, en particulier lorsque l'organisation concernée « produit » vraiment la politique évaluée (éducation, santé...), l'évaluation de cette politique peut être ressentie comme une évaluation de personnes. Dans ces conditions, convient-il, et si oui comment, d'impliquer l'organisation dans l'évaluation ?

Plusieurs justifications, générales ou particulières, d'une association entre évaluateurs et « évalués » ont été présentées. Ainsi D.T. Campbell (« Assesing... ») souligne le coût élevé qui peut être associé à une évaluation faite totalement à l'extérieur de l'organisation concernée. De même, P. Mondon souligne la nécessité d'une association étroite des enseignants au système d'observation mis en place par le ministère de l'Éducation nationale⁸. Mentionnons, dans un domaine voisin et au plan des pratiques, que dans certains ministères les fonctions d'inspection sont exercées par d'anciens « opérationnels »⁹.

7. Voir les exemples précédents des conférences-consensus proposées pour la formulation de recommandations de politique médicale ou des DRIR.

8. Ce système, à la fois transversal et longitudinal, est décrit dans l'encadré n° 3 de la partie II.

9. Il en est ainsi pour l'Agriculture et l'Équipement où les missions d'inspection sont traditionnellement confiées à des ingénieurs généraux (Génie rural et Eaux et Forêts, Ponts et Chaussées).

Il convient sans doute, en la matière, de ne pas avoir de position trop dogmatique. L'association des responsables de l'exécution de la politique de l'évaluation peut être nécessaire, à la fois pour préciser le passage des objectifs généraux aux objectifs opérationnels et pour éviter des réactions de rejet, soit du dispositif d'observation, soit des analyses de résultats. Au-delà de ces seuls objectifs, elle paraît présenter plus d'inconvénients que d'avantages, à la fois pour l'impartialité de l'évaluation et pour la diffusion de ses résultats. Reprenons ainsi le cas de l'éducation et supposons, scénario imaginé ici, qu'une étude fasse apparaître de meilleurs résultats pour les élèves du secondaire dont les professeurs proviennent de tel recrutement plutôt que de tel autre. Il est fort probable que, pour ne pas « désespérer » tel ou tel groupe d'enseignants, l'étude garde un caractère strictement confidentiel et ne reçoive pas certaines suites souhaitables (en matière de formation, de recrutement... ou de débat public !). L'auto-évaluation, si elle ne débouche pas nécessairement sur l'autosatisfaction, a par contre toute chance de conduire à l'autocensure, sauf à s'imposer des contraintes de publicité.

Au total, le principe de la séparation entre évaluation et exécution paraît devoir être affirmé. Sur ce point, nous rejoignons totalement la position prise dans le rapport sur l'évaluation des programmes de recherche¹⁰. Ces programmes doivent faire l'objet d'un suivi par des instances spécifiques, conseils de programme ou groupes d'évaluation et de prospective (GEP). Le rapport recommande que l'évaluation *ex post* soit confiée à des instances distinctes. Reprenons-en les considérations :

« Ce travail fait en continu (nota : de pilotage et de suivi par les GEP), outre son intérêt pour la conduite du programme, pour son bon déroulement, pour son adaptation continue, constituera la matière première indispensable de l'évaluation *ex post*. C'est pourquoi, la question a pu légitimement être posée de savoir s'il ne serait pas plus simple et plus efficace de déléguer totalement le processus d'évaluation à ces organes.

La réponse donnée est cependant négative : les GEP ou conseils de programme participent de façon trop directe à la conduite du programme pour pouvoir être investis d'une mission d'évaluation *ex post*.

Les hommes mêmes doivent être différents : il ne semblerait pas souhaitable de demander à un membre d'un GEP de faire partie, *intuitu personae*, d'un groupe d'évaluation. L'exigence d'un regard neuf, voire provocateur, se fait sentir, ce qui conduit à proposer de confier l'évaluation des programmes à des experts qui ne soient pas issus des instances du programme, mais qui devront travailler en étroite liaison avec elles... »

Mentionnons pour terminer que les justifications de l'évaluation en association, par des considération de « démocratie », ou l'appel à l'apprentissage qu'elle réalise, ne sont pas toujours dépourvues d'arrière-pensées. Elles peuvent, parfois, traduire le refus d'un regard extérieur, le souhait du maintien d'une structure close¹¹. Sur ce point, on peut citer le commentaire fait par R. Fraisse sur

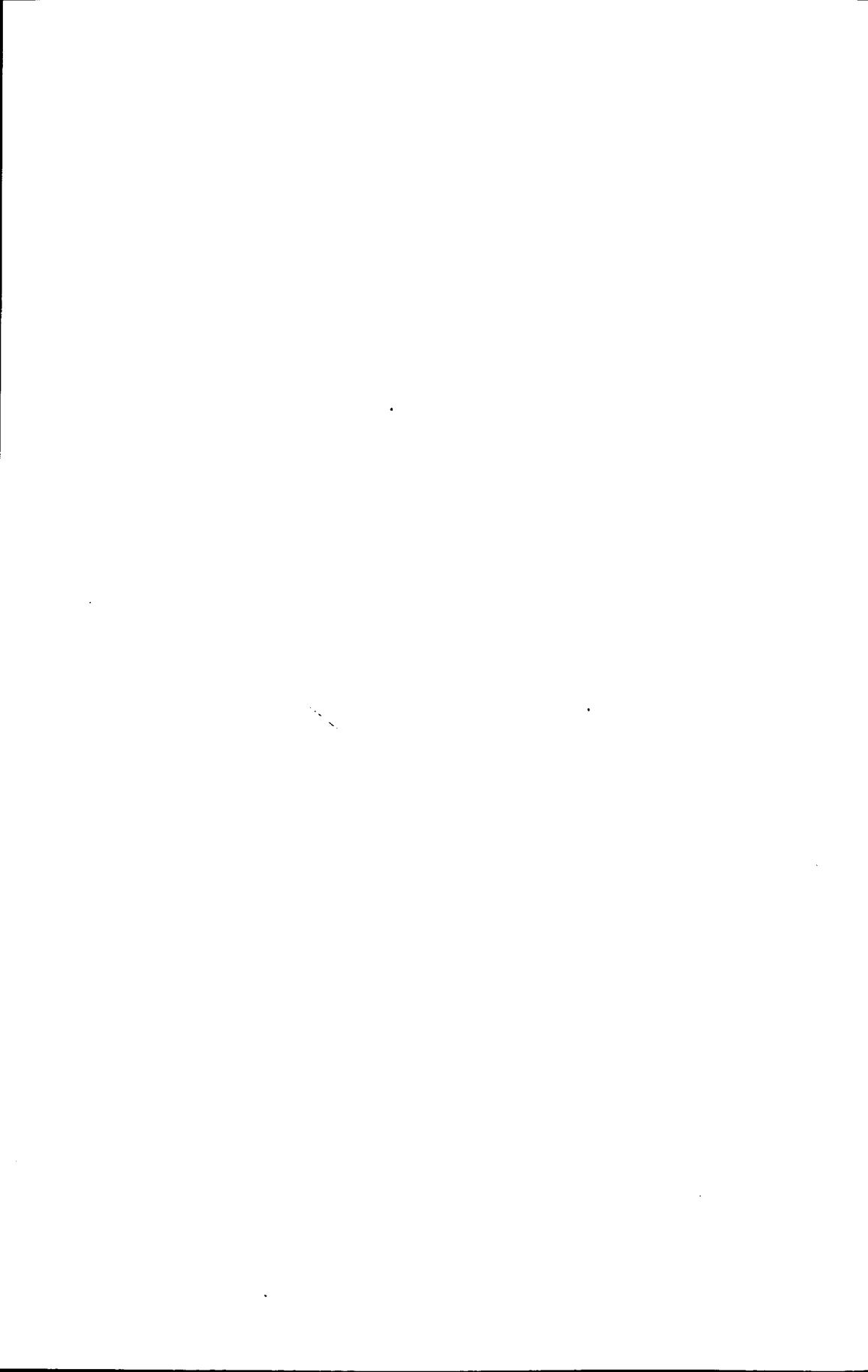
10. Voir en introduction p. 17.

11. Les partisans de l'auto-évaluation manifestent souvent une plus grande sollicitude pour l'organisation responsable de l'exécution de la politique publique que pour les citoyens qui en reçoivent les effets ou qui la financent.

l'évaluation-apprentissage¹², opposée à l'évaluation-sanction (*Compte-rendu du colloque : l'évaluation des politiques publiques*, Politique et Management Public, mars 1985).

12. Dans les termes du commentaire de R. Fraisse, l'évaluation-apprentissage concerne en fait l'auto-évaluation réalisée à des fins d'apprentissage.

« L'évaluation-apprentissage est trop facilement présentée comme la forme "progressive" de l'évaluation qui apprend à chacun à progresser en responsabilité et en autonomie. Cependant, il y a une utopie derrière la généralisation de l'évaluation-apprentissage : participant d'une conception naïve de l'autogestion, elle oublie volontiers l'existence et l'opinion des autres. Le dialogue avec des tiers est une nécessité : il doit pouvoir être un débat et le débat doit pouvoir accepter le conflit. Toute évaluation peut conduire à des ruptures dans l'action entreprise : la prise en compte de ces discontinuités appelle l'intervention des autres et la sanction de leurs jugements ainsi que la netteté de leurs avis, parfois brutaux parce qu'extérieurs. L'apprentissage, c'est bien aussi l'apprentissage du fait qu'on n'est pas seul... »



IV
Les domaines de l'évaluation



IV

Les domaines de l'évaluation

IV — 1 Un découpage des politiques publiques

Les développements précédents ont concerné la « boîte à outils » disponible pour réaliser une évaluation de politique publique (plans de recherche, techniques) et son « mode d'emploi » (mise en œuvre). Des exemples particuliers d'évaluation ont permis de les illustrer.

Au-delà de ces exemples, il importe de reconnaître si un découpage des politiques publiques par grands types ou catégories peut s'avérer utile vis-à-vis de la démarche évaluative. Plusieurs sont envisageables *a priori* : par secteur (agriculture, industrie, transports, etc.), par niveau (mesure élémentaire, programme...), en distinguant politiques ministérielles et interministérielles, etc.

Quelles qualités rechercher dans l'adoption d'un découpage particulier ? Tout d'abord, à des fins pédagogiques mais aussi opérationnelles, le découpage retenu doit pouvoir être mis en rapport avec les catégories conceptuelles, les nomenclatures, les sources existantes. Ces données peuvent présenter, pour tel cas particulier, des limites dont il convient de s'affranchir par des travaux spécifiques (enquête *ad hoc*, modélisation théorique, etc.). Il demeure que, pour obtenir un cadrage synthétique immédiatement utilisable, la référence à l'existant est nécessaire. Par ailleurs, compte tenu de l'objet du rapport, le découpage doit être discriminant vis-à-vis des aspects techniques et organisationnels de l'évaluation et permettre de dégager quelques conclusions globales simples, sur les plans de recherche, les techniques plus particulièrement adaptées (ou adaptables) aux diverses catégories de politiques définies.

Il serait vain de rechercher un découpage idéal excluant tout arbitraire ou redondance. Une assurance toutefois : un découpage sectoriel, de type ministériel, serait peu pertinent. En effet, la « spécificité » des politiques sectorielles ne doit pas être surestimée comme c'est fréquemment le cas¹. Ainsi, l'évaluation de l'efficacité (ou de l'inefficacité) d'une procédure d'aide aux entreprises ne soulève pas des

1. L'argument de spécificité est, presque systématiquement, opposé aux « outsiders » de tout poil, contrôleurs, analystes de politiques publiques, évaluateurs ou économètres, qui prétendent appliquer une méthodologie générale au domaine particulier considéré. S'il traduit, souvent, un simple réflexe de défense de pré carré, son usage peut également procéder d'une surestimation de l'importance des facteurs institutionnels... ou de la simple ignorance des autres domaines. L'examen, par exemple, de politiques sectorielles diverses (industrie, agriculture, transports...) tend plutôt à conforter la proposition suivant laquelle les mêmes causes produisent toujours les mêmes effets.

problèmes très différents, que ces entreprises appartiennent à l'agriculture ou à l'industrie manufacturière. Il en est de même pour les diverses réglementations conférant une rente à une profession particulière (taxis, pompes funèbres, transports aériens, etc.).

En définitive, il s'avère qu'un découpage traditionnel, inspiré de la théorie économique des finances publiques² constitue un choix adapté. On distinguera les domaines suivants : production de biens et services, interventions économiques financières (incitation, redistribution, taxation), réglementation, communication sociale, régulation macroéconomique.

Outre sa correspondance immédiate avec une conceptualisation existante de l'action publique, ce découpage présente plusieurs avantages.

Tout d'abord, pour les politiques publiques donnant lieu à des flux financiers enregistrés par des documents de comptabilité publique ou de comptabilité nationale, il permet un recouplement avec les nomenclatures comptables existantes³, soit en terme de catégories de dépenses, soit en termes de fonctions. Ainsi la ventilation du budget de l'État suivant la nomenclature fonctionnelle budgétaire fait immédiatement ressortir le pôle « production de biens et services », à fort contenu en main-d'œuvre (éducation, défense) et le pôle « incitation, redistribution », où, au contraire, les dépenses dites d'intervention publique sont dominantes (exemple : Sécurité sociale, santé, emploi). Le découpage retenu permet donc de délimiter quantitativement sans trop de difficultés les domaines d'action concernés. On reviendra sur ce point plus bas.

Un deuxième avantage du découpage retenu est d'être discriminant par rapport à l'acquis existant en matière d'évaluation ou d'approches connexes. Ainsi, le contrôle administratif a principalement concerné la production de biens et de services et la régularité de gestion des organisations publiques responsables. Il s'est beaucoup moins exercé sur les interventions économiques financières et, *a fortiori*, sur la réglementation⁴.

Enfin, cette approche satisfait à l'ambition initiale d'obtenir quelques conclusions globales sur les plans de recherche et les techniques utilisés ou utilisables pour chaque catégorie retenue. Le découpage proposé, inspiré *a priori* de la théorie des finances publiques, recoupe en effet des facteurs de différenciation décisifs pour le choix de plans de recherche : mise en œuvre d'une organisation publique complexe (production de biens et de services) ou pas (interventions économiques financières), niveau (régulation macroéconomique ou politiques micro), etc.

Au total, s'il demeure des difficultés liées au caractère non homogène des statistiques existantes⁵, à l'absence de comptes synthétiques (exemple : secteur hospitalier) ou, éventuellement, de données chiffrées (réglementation), le choix de catégories proposé permet de cerner, qualitativement et quantitativement, les

2. Voir par exemple les ouvrages de R. Musgrave.

3. Ce recouplement n'est pas fortuit : la nomenclature comptable s'inspire elle-même, en partie, de la théorie économique des finances publiques.

4. Ou, pour le moins, sa démarche interne laisse totalement dans l'ombre l'aspect essentiel dans ce cas, à savoir l'impact sur l'extérieur.

5. Ainsi, en matière de statistiques de comptabilité nationale, l'activité publique est retracée à la fois dans le compte des administrations, des grandes entreprises nationales, de l'industrie (pour le secteur public industriel), etc.

grands domaines d'intervention de l'État, et de manière pertinente pour l'évaluation.

En particulier, la référence au compte des administrations⁶ de la comptabilité nationale permet de préciser (ou de rappeler) plusieurs caractéristiques importantes sur le champ et l'organisation des interventions publiques que ces administrations (au sens de la comptabilité nationale) assument. On rappelle que ce compte ne retrace pas directement l'activité des entreprises publiques (y compris PTT) et du secteur hospitalier (sinon au travers d'opérations de répartition). Avec cette restriction en tête, les conclusions suivantes peuvent être obtenues (voir tableaux 1 et 2) :

— **l'essentiel des activités de production de biens et de services (non marchands) par les administrations est le fait des administrations centrales** (tableau 1). Appréciée au travers des salaires versés (ou éventuellement de la « consommation finale » au sens de la comptabilité nationale), la majeure partie correspond aux fonctions **d'éducation** (48 % des salaires versés par l'État) ainsi qu'aux **fonctions régaliennes** usuelles : défense (19,4 %), sécurité intérieure (11 %), administration financière (7 %).

Les autres fonctions de production de biens et de services sont assurées surtout à la **périphérie de l'État** (santé par exemple).

— pour chaque fonction, le **poids relatif des activités de production de biens et de services** et des activités **d'intervention économique** peut être apprécié en considérant le **rapport entre les dépenses** (au sens de la comptabilité nationale) et le **coût de production**. Un rapport élevé est en général caractéristique d'une importance marquée des « opérations de répartition ». Aux fonctions « de production » précédentes (enseignement, administration générale, justice et sécurité, défense nationale), pour lesquels le rapport dépense totale/coût de production est voisin de l'unité, s'opposent clairement les fonctions santé et interventions sociales (tableau 2).

— au plan global, la référence aux dépenses des administrations publiques au sens de la comptabilité nationale (tableau 2) montre que le **pôle « interventions économiques financières »** l'emporte largement sur le **pôle « production de biens et de services »**. Cette prédominance est essentiellement due au poids des administrations de **sécurité sociale** qui n'ont, en elles-mêmes, qu'un rôle de **tuyauterie redistributrice**⁷. La partie directement active (ou supposée telle) du pôle, correspondant par exemple aux diverses aides au secteur productif, est **plus modeste**. En l'identifiant aux « subventions d'exploitation versées » de la comptabilité nationale, elle représenterait environ 4,5 % des dépenses des administrations publiques. **L'essentiel de ces interventions « actives »** relève des **administrations publiques centrales**, les interventions redistributives étant surtout le fait, comme il a été dit, des administrations de sécurité sociale.

6. Plus précisément, à la présentation des dépenses et des coûts de production par fonction.

7. Avec, encore une fois, cette restriction que la partie production proprement dite (secteur hospitalier) est incluse, à ce niveau, dans une description en termes de transfert.

TABLEAU 1 : COÛTS DE PRODUCTION PAR FONCTION DES ADMINISTRATIONS PUBLIQUES
 (Année 1982 - Millions de francs courants - Comptes de la Nation)
 La répartition entre administrations est donnée entre parenthèses en pourcentage.
 Le pourcentage le plus important est en gras.

Administ. pub. centrales	Administ. pub. locales	Administ. pub. de Séc. soc.	Total	Fonctions
175 489 (87,4)	25 383 (12,6)		200 872	Enseignement, formation
5 119 (34,4)	9 781 (65,6)		14 900	Culture
2 241 (12,8)	2 883 (16,4)	12 436 70,8	17 560	Santé
4 450 (9,7)	14 742 (32,1)	26 684 58,2	45 876	Interventions sociales
1 013 (95,1)	52 (4,9)		1 065	Logement
2 158 (8,0)	24 660 (92,0)		26 818	Développement urbain
2 519 (29,9)	5 897 (70,1)		8 416	Développement rural
14 465 (68,3)	6 705 (31,7)		21 170	Transports, télé-communications
23 653 (71,5)	9 428 (28,5)		33 081	Action économique
39 587 (57,0)	29 840 (43,0)		69 427	Administration générale
9 086 (99,9)	7 (0,1)		9 093	Recherche
26 013 (83,2)	5 234 (16,8)		31 247	Justice, sécurité
4 718 (100,0)			4 718	Relations extérieures
404 (100,0)			404	DOM-TOM
148 275 (100,0)			148 275	Défense nationale
581 (60,1)	385 (39,9)		966	Non ventilé
459 771 (72,5)	134 997 (21,3)	39 120 (6,2)	633 888	TOTAL

Ce cadrage, pour incomplet qu'il soit, permet d'apprécier la surface des divers domaines de l'activité publique. Nous allons maintenant les passer en revue du point de vue de leur spécificité vis-à-vis des méthodes de l'évaluation.

**TABLEAU 2 — COÛTS DE PRODUCTION ET DÉPENSES TOTALES DES ADMINISTRATIONS
PUBLIQUES PAR FONCTION**

(source : rapport sur les Comptes de la Nation de l'année 1984 — INSEE Série C 124-125)
(Année 1982, Milliards de francs constants)

Opérations Administrations	Consommations inter- médiaires (1)	Rémunération salariés (2)	Impôts nets (3)	Coût de production (4) = (1)+(2)+(3)	Consommation finale (5)	Opérations courantes de répartition (6)	Opérations en capital (7)	Dépenses totales (8) = (5)+(6)+(7)
Enseignement, formation	34 437	163 583	2 852	200 872	189 388	9 994	13 567	212 949
Culture	5 972	8 805	123	14 900	21 366	6 988	11 663	40 017
Santé	3 639	13 155	766	17 560	17 532	215 413	2 751	235 696
Interventions sociales	13 552	30 858	1 204	45 614	42 697	676 010	6 303	725 010
Logement	720	339	6	1 065	1 247	37 533	8 236	47 016
Développement urbain	11 438	16 080	- 700	26 818	23 134	14 876	28 105	66 115
Développement rural	3 870	4 625	- 79	8 416	7 970	3 420	9 383	20 773
Transports, télécommunications	10 584	10 892	- 306	21 170	14 868	15 626	18 939	49 433
Action économique	15 603	16 481	997	33 081	26 461	40 784	15 172	82 417
Administration générale	19 490	49 509	690	69 689	60 746	1 724	11 323	73 793
Recherche	2 529	5 320	1 244	9 093	9 081	2 869	1 244	13 194
Justice, sécurité	3 934	26 790	523	31 247	32 470	621	2 197	35 288
Relations extérieures	816	3 833	69	4 718	4 764	7 375	430	12 569
DOM-TOM	154	245	5	404	419	9 493	1 134	11 046
Défense nationale	73 760	72 862	1 653	148 275	124 552	5 283	6 943	136 778
Non ventilé	563	400	3	966	1 794	68 340	2 550	72 684
TOTAL	201 061	423 777	9 050	633 888	478 489	1 116 349	139 940	1 834 778

IV - 2 La production de biens et de services

Les activités publiques correspondantes reposent sur la mise en œuvre de facteurs (travail, consommations intermédiaires, capital) au travers d'une organisation spécifique (entité administrative, établissement public, entreprise). Elles conduisent à fournir certains biens ou services à des « consommateurs » externes, publics ou privés.

Les activités couvertes par cette description présentent à la fois des caractéristiques communes fondamentales et des caractéristiques de différenciation.

Première caractéristique commune : le fait que les politiques publiques associées soient le fait d'organisations complexes faisant appel à une « main d'œuvre » abondante. Il en résulte que l'analyse « interne » de l'organisation peut être décisive pour l'évaluation. Le problème sera généralement différent pour les politiques incitatives ou réglementaires. Ce poids des aspects « organisation » est fortement discriminant pour les plans de recherche envisageables et, en particulier, dans bon nombre de cas, limite le recours à des expérimentations « vraies ».

Deuxième caractéristique commune, corollaire de la première : l'existence d'une forte tradition de contrôle administratif, essentiellement de régularité, fondée sur l'exploitation de documents comptables (voir I.3.).

On notera que cette tradition concerne surtout l'organisation de la production elle-même et que l'analyse de l'impact externe, qui se développe actuellement dans les contrôles administratifs, demeure encore limitée.

La troisième caractéristique générale tient également à l'existence d'une organisation lourde qui donne un poids particulier à la dimension temporelle. Il en résulte une continuité entre les opérations de suivi de résultats, stricto sensu, et des opérations d'évaluation de politiques publiques consécutives à l'introduction d'un changement finalisé (réforme de l'enseignement, nouvelle tarification, extension du secteur public, etc.). Cette continuité est très sensible au plan des méthodes. On notera en particulier l'importance générale d'un recours à des dispositifs de suivi statistique lourds, de champ cohérent entre périodes.

Mentionnons enfin deux autres caractéristiques, de différenciation.

L'une concerne la distinction marchand/non marchand. Elle renvoie à une difficulté spécifique aux activités non marchandes, à l'amont de l'évaluation proprement dite : celle de l'identification même du bien ou du service produit et de sa valorisation éventuelle.

L'autre caractéristique de différenciation concerne l'importance de l'interaction entre l'interne (organisation) et l'externe (milieu récipiendaire). Cette interaction peut jouer un rôle fondamental et complexe dans le cas de certains services non marchands, tel que l'éducation : les résultats obtenus dépendent à la fois des performances de l'organisation mais aussi des caractéristiques des bénéficiaires (exemple : paramètres socio-culturels). On conçoit que dans ce cas, le problème d'évaluation puisse être fortement multidimensionnel, d'où des difficultés pour réaliser de « vraies » expériences contrôlées.

1) La production « publique » de biens et services marchands couvre, en particulier, l'ensemble du secteur public productif (grandes entreprises nationales, secteur public industriel), les établissements publics à caractère industriel et commercial, certaines activités tarifées exercées par des administrations produisant essentiellement dans le non marchand.

Les questions d'évaluation (de résultats comme de politiques) les plus fréquentes concernent :

- l'évolution des résultats comptables ;
- les décisions sur les facteurs de production (emploi, capital) ou leur gestion ;
- les politiques de prix (tarifs en particulier).

Ces questions peuvent être étroites (évaluation de l'introduction d'une nouvelle forme de tarification) ou larges (évaluation du programme électronucléaire) ou même renvoyer aux statuts des organisations considérées (avantages liés à l'existence d'un monopole public).

On rappelle qu'il existe en la matière une tradition importante de contrôle administratif et de suivi. Les travaux correspondants sont parfois peu centrés sur l'efficacité proprement dite et comportent souvent des lacunes. Par exemple, il est rare que, dans l'évaluation des résultats de telle entreprise publique, le partage entre impact de la politique proprement dite de l'entreprise (efforts de productivité, choix d'investissement, choix de produits) et évolution du marché soit très clairement opéré.

L'acquis apparaît toutefois solide. Des directions de progrès prioritaires concernent l'amélioration des éléments de comparaison intertemporelle et externe et une meilleure séparation de l'effet des divers facteurs.

La première direction concerne à la fois les entreprises publiques « concurrentielles » et « non concurrentielles ».

Il conviendrait, tout d'abord, de disposer d'informations statistiques homogènes dans le temps, permettant de pratiquer de manière fiable un suivi ou des approches de type avant-après. L'expérience prouve aussi l'intérêt de comparaisons systématiques avec divers « groupes témoins », constitués par exemple d'entreprises intervenant dans des secteurs similaires en France et à l'étranger. Si ces groupes témoins ne peuvent pas être considérés comme totalement équivalents, de telles comparaisons sont très utiles pour répondre à des questions du type suivant :

— on observe une amélioration sensible des résultats de certains groupes chimiques publics : cette amélioration est-elle supérieure ou inférieure à celle des groupes chimiques privés ou étrangers ? Quelle est la signification des différences observées ?

— les coûts de production d'Air France sont-ils, ou non, supérieurs à ceux de ses concurrents étrangers ? si oui, pourquoi (productivité, rémunérations) ?

— dans quelle mesure l'effort électronucléaire de la France a-t-il bénéficié aux divers agents économiques nationaux (comparaison des tarifs industriels ou domestiques avec ceux pratiqués à l'étranger) ?

— etc.

L'existence de « groupes témoins » externes et le recours à des comparaisons intertemporelles fiables sont particulièrement intéressants pour les questions d'évaluation concernant la politique suivie en matière de facteurs de production (emploi, capital) ou de tarifs par les groupes publics en situation de monopole. En l'absence de concurrence par le marché, les risques de dérive sont notables.

La seule manière d'en apprécier la réalité est souvent de recourir à de telles comparaisons externes ou intertemporelles. Une problématique similaire s'applique à l'évaluation *ex post* des grands projets, à l'instar de l'analyse du TGV Sud-Est faite par la Cour des comptes.

Notre deuxième remarque concerne la prise en compte de l'environnement de l'entreprise. Elle est commune aux entreprises publiques « concurrentielles » et « non concurrentielles » pour ce qui est de l'évolution du marché : l'obtention par une entreprise publique de résultats corrects sur un marché conjoncturellement porteur n'est pas en soit un signe de réussite exceptionnelle. Elle se renforce dans le cas d'entreprises publiques fortement « régulées » par l'État, c'est-à-dire essentiellement les grandes entreprises nationales. L'intervention de l'État, soit au travers de contraintes de service public, soit par le biais de mesures de régulation macroéconomique ou industrielle, est un facteur à prendre en compte dans l'évaluation. Ainsi une évaluation de l'évolution des tarifs énergétiques ou téléphoniques ne saurait ignorer l'impact de la politique des prix ou les choix de financement de la filière électronique, le programme d'équipement nucléaire d'EDF répond à certains objectifs de lissage industriel, etc. Dans bien des cas, l'évaluation de l'entreprise publique se mue en évaluation du couple État tutelle-entreprise.

Au total, l'analyse rapide précédente suggère que pour le secteur public marchand, les progrès de l'évaluation des résultats et des politiques publiques concernent, de manière somme toute classique :

- la mise en place de systèmes d'information permanents permettant d'établir des comparaisons externes fiables, dans le temps ou dans l'espace ;
- une meilleure prise en compte des divers facteurs d'environnement (évolution des marchés, rôle de l'État tutelle) pesant sur les résultats.

Ces préalables satisfaits, les plans de recherche correspondants sont évidemment de type « non expérimental » (au sens de l'expérimentation vraie de la partie II).

A cet égard, les possibilités offertes par l'expérimentation vraie paraissent limitées du fait du poids des aspects organisationnels. Cette position a été très clairement exprimée par A. Rivlin⁸, qui souligne le caractère coûteux, ou impraticable, d'un recours à l'expérimentation, pour identifier des « fonctions de production » d'organisations. Lorsque trop de paramètres indépendants sont en cause, sans possibilité de moyenne statistique, la rigueur de l'expérience de laboratoire contrôlée, paraît difficile à atteindre. Il y a même un danger à s'appuyer sur des pseudo-expériences » dont les conclusions sont appliquées sans discernement sous couvert de vertus apparentes de rigueur. Cette objection à une pratique abusive de l'expérimentation se retrouvera avec une acuité accrue dans le cas d'organisations

8. Voir A. Rivlin (1974), *How can experiments be more useful ?* (op. cité)

« non marchandes » en forte interaction, pour leurs résultats, avec le milieu social, telle l'éducation.

Il demeure que pour des cas simples de politiques très élémentaires, modifiant un signal de l'entreprise à destination de l'extérieur (tarifs, qualité, nouveaux services...), l'expérimentation retrouve ses droits. Ainsi la mise en place sur l'ensemble du territoire de la tarification EJP (effacement jour de pointe) par EDF a-t-elle été précédée de quasi-expérimentations locales. De véritables expériences ont été pratiquées dans ce domaine à plusieurs reprises aux États-Unis⁹.

Rappelons enfin l'importance fréquente pour une évaluation portant sur des caractéristiques relativement globales (exemple : productivité du travail) d'avoir auparavant écrit une « histoire » suffisamment solide de l'entreprise concernée, les choix passés ayant souvent un poids déterminant.

2) Nous en venons maintenant aux politiques publiques de **production de biens et services non marchands**.

Leur évaluation soulève, pour partie, des problèmes identiques à ceux rencontrés dans le cas de la production publique marchande. L'existence d'une organisation complexe, support de la production et, par là, des politiques, implique que l'évaluation requiert des appareils de suivi statistique permettant des comparaisons intertemporelles ou entre institutions similaires.

Des difficultés supplémentaires et conjuguées affectent l'élaboration même de la procédure d'évaluation pour :

— **identifier le bien ou service produit ;**

— **identifier le « résultat » qui peut impliquer une interaction forte avec le milieu extérieur « consommateur » du bien ou du service.**

Nous ferons quelques observations succinctes sur chacun de ces points.

La première difficulté est bien connue. Un moindre mal consiste à définir des **indicateurs**, éventuellement multiples et apportant des éclairages complémentaires : soins délivrés par catégorie, pourcentage de diplômés, heures d'enseignement, publications, brevets, pourcentage d'affaires criminelles résolues, etc. On n'examinera pas davantage cette question du choix des indicateurs. Elle a fait l'objet d'une littérature abondante dans le passé. Notons que la seule utilisation d'indicateurs d'activité peut éclairer un diagnostic sur l'efficacité des organisations concernées mais non sur leur efficacité sociale. Rappelons aussi que la sélection *a priori* d'indicateurs peut, dans certains cas, induire des comportements qui nécessitent un redressement lors de l'évaluation (cf. politique anti-crime du gouvernement Nixon).

Le deuxième point important concerne la **dépendance des résultats par rapport à l'environnement extérieur**, et notamment aux « récepteurs ». Cette situation est caractéristique d'un grand nombre d'activités publiques non marchandes, par

9. On trouvera dans l'article de D.J. Aigner et E.E. Leamer, *Estimation of time of-use pricing response in the absence of experimental data : an application of the methodology of data transferability* (a. paraitre dans *The Energy Journal*), de nombreuses références sur de telles expérimentations tarifaires et une description de certaines d'entre elles (Southern California Edison Company, Los Angeles Department of Water and Power, Kansas City Power and Light).

exemple, l'éducation, la santé ou, à un degré moindre, la sécurité intérieure. On connaît la masse de travaux sur les liens entre échec scolaire et facteurs socio-culturels. De même, l'évolution, à un plan global, des taux de morbidité dépend de facteurs non directement sanitaires, etc. Ces diverses difficultés ont été illustrées dans les parties II et III.

Ces deux remarques ont une implication directe pour le choix des procédures. Le diagnostic renforce en un sens celui opéré dans le cas marchand.

Une première conclusion est négative : dans le cas où les politiques publiques associées mettent en jeu de manière complexe l'organisation « productrice » et le milieu récepteur, le recours à l'expérimentation vraie doit être envisagé avec prudence. L'éducation est un exemple particulièrement significatif : le nombre de « paramètres » concernés, la longueur du processus éducatif, son caractère cumulatif rendent illusoire une reproduction des conditions d'une expérience contrôlée en laboratoire. L'échec du programme américain d'évaluation mené dans les années 1970 fondé sur une approche systématique en terme de « planned variations » (on fait varier un à un les paramètres supposés importants : taille de la classe, etc.) est significatif¹⁰. Exemple actuel, la déconsidération dont souffre la recherche pédagogique tient sans doute aux déboires rencontrés dans l'introduction, parfois tonitruante, et chaque fois, bien sûr, assise sur le résultat d'expériences indubitables, de réformes pédagogiques successives et paraissant contradictoires : méthodes d'enseignement des langues, des mathématiques, de la lecture ; classes de transition puis soutien ; disciplines d'éveil, etc. En l'occurrence, la conjonction de références abusives à des expériences ponctuelles de portée limitée et d'applications massives n'a pas été sans coût. **La prudence s'impose donc : rien n'est pire en effet, dans des domaines fondamentaux tels que l'éducation ou la santé, que des pseudo-certitudes assises sur des pseudo-expériences non généralisables.** Par exemple, les résultats observés dans des lycées-pilotes où la population des enseignants et des élèves, les moyens disponibles, sont très spécifiques, ne peuvent être extrapolés sans danger à l'ensemble de la population scolaire. De même, l'introduction d'une modification globale du système de gestion des hôpitaux, sans identification préalable de leurs caractéristiques spécifiques, ne peut être que faiblement conclusive. Dans ces domaines, plutôt que de viser l'expérience contrôlée décisive (et, par là, la politique miracle), il semble préférable d'avoir une approche d'accumulation progressive de connaissances de type « muddling through ». **D'où l'intérêt de systèmes de suivi long du type de celui géré par le Service de la prévision, des statistiques et de l'évaluation du ministère de l'Éducation Nationale. D'où l'intérêt également d'expérimentations locales, ne prétendant pas au rang d'expériences décisives, mais dont les résultats puissent être analysés et confrontés de manière pragmatique et conduisent à cette accumulation de variance chère aux statisticiens.**

Dans le cas d'activités non marchandes, la démarche expérimentale retrouve ses droits dès lors que s'efface la double complexité d'un processus de production multidimensionnel et de fortes interactions avec l'extérieur. Il en est ainsi pour l'évaluation de politiques sanitaires relativement simples où l'extérieur constitue

10. Voir A. Rivlin, *How can experiments be more useful ?* (op. cité).

un milieu largement « passif » : fluorisation des eaux, campagne de vaccination, etc. Dès lors que ces politiques cessent d'être strictement « biochimiques », l'expérimentation demeure valide mais les difficultés de conception et d'interprétation sont accrues¹¹.

Une dernière remarque. On notera que les fonctions régaliennes les plus traditionnelles n'ont fait l'objet, en France, que d'un nombre très restreint d'évaluations. Il est ainsi pour l'administration générale¹², la justice et, bien évidemment, la défense. Comparée à d'autres pays, cette préservation d'un sanctuaire peut être, éventuellement, considéré comme archaïque. Aux États-Unis par exemple, les études d'évaluation sur la police ou la justice sont légion. Le GAO n'est restreint dans ses pouvoirs d'investigation que vis-à-vis de deux organismes : la Central Intelligence Agency et le Federal Reserve Board. On peut ainsi mentionner qu'une évaluation (de formulation) de politique de défense, portant sur le programme d'armes chimiques envisagé par le Pentagone et réalisée par le GAO à la demande du Congrès, a conduit au rejet de ce programme¹³.

IV - 3 Les interventions économiques financières : incitations, redistribution, taxation

Nous en venons maintenant au domaine principal de l'action publique en termes de flux monétaires. Le rapide cadrage précédent, établi à partir des données de comptabilité nationale, montre au sein de cette catégorie l'importance des administrations de sécurité sociale.

Par opposition avec la production de biens et de services et dans une perspective d'évaluation, une caractéristique notable des activités publiques afférentes est, en général, la (relative) simplicité de « l'interne »¹⁴ : les organisations concernées ont essentiellement un rôle d'application de textes ou de gestion administrative et comptable. Il n'en résulte pas qu'elles n'appellent ni audits ni évaluations (tout au contraire !). Les questions soulevées, de régularité (application des textes...) ou d'efficacité (délais de traitement des dossiers...), sont, toutefois, en général plus simples que pour des administrations d'éducation ou de santé.

11. Voir par exemple les difficultés mentionnées dans l'article de J.E. Harris, *Prenatal medical care and infant mortality* (1982, in V.R. Fuchs ed, *Economic Aspects of Health*, NBER, The University of Chicago Press, Chicago, Il).

12. On rappellera que les administrations centrales sont largement préservées des désagréments des contrôles.

13. Voir la communication d'E. Chelmsky au colloque sur le management public. (26-28 septembre 1984) sur le thème « Program evaluation in the United States : what influence does the congressional use of evaluation findings have on public management at the federal level ? »

14. Nous insistons sur les termes « en général » et « relatif ». Il va de soi que l'application de procédure d'aides soulevant des problèmes de reconnaissance d'éligibilité ou d'appréciation économique peut être délicate. Il en est ainsi notamment des politiques d'incitation à l'innovation. Dans ce cas, une étude du fonctionnement interne retrouve évidemment ses droits.

Les coûts de production (salaires, frais de fonctionnement etc.) engagés directement pour les interventions économiques apparaissent donc comme des coûts de mise en œuvre. Les politiques publiques affectent, soit le revenu des agents concernés, soit les prix (ou des variables équivalentes) auxquels ils sont confrontés : les « résultats » dépendront donc, au premier chef, des réactions des agents à ces signaux financiers. Cette description, valide en première approximation, explique le recours, largement développé dans certains pays mais non en France, à des modélisations microéconomiques des comportements pour l'évaluation. On reviendra sur ce point plus bas.

Rappelons que les politiques publiques concernées peuvent affecter les dépenses de l'État (il en est ainsi pour les diverses prestations sociales ou les aides versées aux entreprises¹⁵), ou ses recettes (la variation des taux d'imposition ou des assiettes est un instrument d'orientation éventuelle des décisions des agents). De manière générale, une politique de soutien à un secteur ou à un groupe d'agents particulier pourra comporter des actions pesant à la fois sur les dépenses (aides) et sur les recettes (fiscalités dérogatoires, parfois baptisées « dépenses fiscales »).

Enfin, nous avons retenu la distinction traditionnelle entre interventions incitatives et redistributives. Cette distinction permet dans certains cas de séparer des configurations extrêmes. Elle recouvre le plus souvent une distinction de finalités ou de résultats : la plupart des interventions économiques auront à la fois des impacts incitatifs et redistributifs (avec des chassés-croisés involontaires).

Des exemples d'évaluation de ces politiques ont été discutés précédemment (aides à l'emploi, aides à l'élevage bovin etc.). La tendance récente est à leur accroissement. Malgré ce progrès, l'acquis demeure limité. Certes des cadrages quantitatifs existent. On peut citer les travaux effectués au début des années soixante-dix sur les transferts État-industrie¹⁶ ou les rapports sur les aides annexés à la loi de finances. Pour nécessaires qu'ils soient, ces constats demeurent surtout comptables et n'autorisent aucune conclusion sur l'efficacité des dispositifs. De même les contrôles administratifs peuvent mettre en évidence des irrégularités graves d'application et permettent alors, de conduire à l'inutilité des interventions. En l'absence de telles irrégularités, quand « tout marche bien », ils n'autorisent pas en général un jugement définitif d'efficacité économique.

On peut donc estimer qu'il y a, dans ce domaine des interventions économiques financières, une carence importante en matière d'évaluation. Il existe pourtant de solides présomptions de l'inefficacité incitative de nombreuses interventions économiques présentées comme des leviers d'action importants pour l'État¹⁷. Il importe de rendre plus décisives ces présomptions, ou de les invalider. Dans le système actuel, on peut s'interroger sur l'efficacité incitative de nombreuses aides aux entreprises, compte tenu des « degrés de liberté » des systèmes auxquels elles s'appliquent (une aide à l'investissement peut fort bien se muer en distribution de dividendes). Symétriquement, de nombreuses aides sociales, ou la fiscalité, souvent appréciées dans une perspective redistributive, génèrent sans aucun doute

15. Sous de multiples formes on le sait : aides sectorielles, bonifications d'intérêt etc.

16. A. le Pors Les transferts État-Industrie Statistiques et Etudes financières, série orange, n° 2, 1971.

17. Divers rapports administratifs ont été ainsi consacrés à certaines catégories d'aides publiques.

des effets incitatifs non voulus¹⁸. Telles sont des questions auxquelles une évaluation plus méthodique permettrait d'apporter une réponse.

Ce constat suggère trois voies de progrès.

La première consiste en un développement du **recours aux méthodes d'expérimentation**. Ces méthodes paraissent convenir aux cas d'interventions économiques financières s'adressant à des groupes nombreux (aides aux ménages en particulier, aides générales aux entreprises) et où les phénomènes de « cannibalisme »¹⁹ sont limités. Le développement de telles pratiques suppose toutefois une remise en cause de certaines traditions d'égalitarisme formel.

On évoquera ensuite **l'association, à tout nouveau système d'intervention économique** (aides, plan sectoriel, etc.), **d'un dispositif d'évaluation prévoyant, en particulier, un bilan d'efficacité à horizon donné et une remise en cause éventuelle de la politique**. Ce mécanisme²⁰, permettrait d'éviter l'empilement de systèmes d'intervention reconduits sous la double pression des groupes récipiendaires et des administrations gestionnaires.

Enfin, l'évaluation des interventions économiques gagnerait souvent à s'appuyer sur une **modélisation microéconomique des comportements** avec estimation sur **données individuelles**. On a mentionné la maigreur des travaux réalisés en France dans ce domaine²¹. Le manque de jugements relativement assurés soit sur l'efficacité incitative de telle aide aux entreprises, soit sur l'absence d'effets pervers liés à un système particulier de prestations sociales ou de taxes découle, notamment pour les politiques existantes où l'expérimentation n'est plus possible, **du caractère le plus souvent comptable et non explicatif des constats**.

IV - 4 La réglementation

Du point de vue de l'évaluation, les **politiques de nature réglementaires ont des caractéristiques proches de celles reposant sur des interventions économiques financières** : relative simplicité des organisations administratives responsables²², dépendance quasi exclusive des résultats par rapport à la réponse, en termes de

18. De nombreux travaux anglo-saxons, procédant par expérimentation ou modélisation, ont été consacrés au thème de l'impact sur l'offre du travail d'aides sociales ou de taxes. Voir les références des actions publiques en matière d'emploi (1975-1985) », (communication aux Journées d'études des 3 et 4 octobre 1985, organisées par l'ANPE, le CGP et la délégation à l'emploi, ronéo).

Un exemple pour le cas de la France (avec une présentation générale et de nombreuses références) est fourni par l'étude de F. Bourguignon et B. Lemmenier *Fiscalité, para-fiscalité et offre de travail féminin* (1984, CREDOC, ronéo).

19. Dans le cas d'une procédure d'aide publique aux entreprises, ce terme désigne l'élimination d'entreprise non aidées par les entreprises aidées. A marché donné, de tels phénomènes peuvent rendre délicat l'extension à un plan global de conclusions éventuellement positives fondées sur la comparaison entreprises aidées — entreprises non aidées.

20. La « sunset legislation » aux États-Unis en fournit un exemple.

21. Voir les quelques études citées précédemment.

22. On fera les mêmes remarques que précédemment. Ce jugement sur la simplicité de l'interne est vrai à condition que la réglementation soit elle-même « relativement simple » à appliquer, ce qui n'est pas toujours le cas.

comportement, des agents concernés (et donc possibilité de détour et de neutralisation pour les systèmes récepteurs complexes). Les mêmes recommandations peuvent donc être formulées : développement de l'expérimentation, inclusion de procédures d'évaluation (avec possibilité d'extinction de la réglementation) dès la mise en place de nouveaux dispositifs réglementaires, recours à la modélisation microéconomique et à l'estimation sur données individuelles. Nous ne reviendrons pas sur ces divers thèmes.

Par rapport aux politiques précédentes, la réglementation se distingue toutefois par une caractéristique importante : l'absence fréquente de « trace » chiffrée générée directement par la mise en œuvre. Ainsi, le volume des aides à l'exportation est enregistré en comptabilité publique. Compte tenu de l'existence de dossiers d'aide, il peut être rapporté aux flux d'exportations concernés, mis en parallèle avec la situation des entreprises bénéficiaires, etc. Un premier cadrage quantitatif est donc réalisable à partir des données existantes. A l'opposé, l'évaluation des effets de la loi sur le prix unique du livre nécessitera des données spécifiques non générées automatiquement par l'application du dispositif réglementaire lui-même : évolution des chiffres d'affaires des libraires, des éditeurs, des ventes (en volume), du prix des livres, etc. Sauf mise en place d'un dispositif *ad hoc* de collecte, l'évaluation est impossible. Une telle situation est fréquente et l'acquis d'information en matière de politiques réglementaires est très en deçà de celui existant pour les interventions économiques.

Cette absence d'acquis concerne aussi les outils d'analyse conceptuelle comme les enseignements à tirer des contrôles administratifs. S'il existe des théories économiques ou politiques de la réglementation, les travaux réalisés dans ce domaine, notamment de nature appliquée, sont peu nombreux en France. Les enseignements de contrôles administratifs de régularité ne peuvent être entièrement conclusifs. Vérifier qu'une réglementation est appliquée est une chose. En apprécier l'impact en est une autre de nature très différente.

Au total, si l'analyse s'est développée dans ce domaine au cours des années récentes, la réglementation demeure un grand continent vierge de l'évaluation des politiques publiques.

L'accent devrait donc être fortement mis au cours des années à venir sur l'évaluation des politiques réglementaires, de manière d'autant plus volontariste qu'elles jouissent d'une immunité certaine. Ressenties comme moins directement coûteuses en termes budgétaires, elles ne suscitent pas les mêmes réserves *a priori* que des interventions économiques financières. Par ailleurs, les réglementations conduisant à des transferts entre agents n'ont de chance d'être remises en cause que si des groupes de pression suffisamment forts se mobilisent contre le groupe de pression bénéficiaire de la réglementation²³. Une telle situation s'observe parfois dans le cas de face à face entre producteurs également organisés. Les chances de remise en cause sont beaucoup plus faibles lorsque la réglementation transfère un surplus perçu sur des usagers dispersés et nombreux vers un secteur bien structuré. Il en est ainsi dans le cas des monopoles publics ou de certaines professions

23. Il y a là une des difficultés majeures en matière de « déréglementation ». Voir par exemple E.L. Thurow *The zero sum society*.

fermées²⁴. L'évaluation des politiques publiques afférentes est particulièrement justifiée pour de telles configurations. Une déréglementation partielle peut d'ailleurs constituer une approche expérimentale intéressante, dans certains cas (exemple : transport aériens).

Mentionnons enfin un dernier domaine où le développement d'une évaluation par expérimentations locales, pourrait présenter des avantages : celui de la **réglementation du travail**²⁵.

IV - 5 La communication sociale

Nous en venons maintenant à un domaine particulièrement immatériel des politiques publiques : celui de **l'action sur l'opinion**. Ce domaine concerne, bien sûr, l'ensemble des politiques publiques précédemment discutées. Il comprend aussi des actions spécifiques, les campagnes d'information de la puissance publique, pour lesquelles une évaluation systématique découle d'obligations réglementaires. Une double raison pour consacrer un court exposé à ce thème.

Sur le premier point, il va de soi que le contenu des politiques publiques en « communication sociale » peut être important. A un extrême, on trouve, dans les catégories précédentes, de multiples exemples de politiques symboles, coups médiatiques ou réponses à une émotion de l'opinion publique, dont les effets autres que de communication sont restreints. Une vision extrême est de considérer que toute politique publique **n'est que communication**. Nous n'adhérerons pas, dans ce rapport, à une telle vision, inexacte, sans nier que la référence à des objectifs de communication sociale puisse être fondamentale pour comprendre la **genèse** d'une telle politique particulière ou son **exécution**, et par là, constituer un aspect nécessaire de son évaluation.

Les campagnes publiques d'information proprement dites se sont développées de manière importante au cours des dix dernières années. On sait qu'une appréciation de leur impact est obligatoire. Il semblerait donc que la question de leur évaluation soit largement résolue. En fait, bien des difficultés demeurent²⁶.

Les pratiques actuelles restent en effet assez largement centrées sur la communication proprement dite. C'est **l'impact d'information** de la campagne qui est mesuré au travers d'indicateurs caractéristiques de la réception, de la

24. Des réglementations *a priori* moins conflictuelles, telles que les réglementations techniques, peuvent être également susceptibles de cette démarche d'évaluation. Notons qu'il peut également y avoir dans ce cas des transferts entre agents non nécessairement justifiés. Ainsi le choix des normes d'isolation thermique peut constituer un enjeu entre producteurs d'énergie, constructeurs, experts en économies d'énergie, etc.

25. Certaines difficultés rencontrées dans l'application des dispositions sur la réduction de la durée hebdomadaire de travail tiennent ainsi à une approche excessivement globale et ignorant le comportement des micro-agents.

26. Dans ce qui suit, on s'inspire largement de la communication faite par B. Dubois, Y. Evrard, D. Lindon, « L'évaluation des campagnes d'intérêt public » au séminaire organisé par l'ESSEC en 1980 sur le thème « L'État et la publicité » (note CESA, ronéo).

mémorisation ou de la compréhension. Par contre les effets sur les attitudes ou les comportements sont peu explorés. Les difficultés sont réelles. S'il est possible d'apprécier, par exemple, les modifications de comportement induites par des campagnes d'information routière (comptages statistiques, enquêtes *ad hoc*, etc.) la mesure des conséquences sur les attitudes et les comportements de campagnes anti-tabac, ou encore de campagnes en faveur de la tolérance envers les immigrés, peut se révéler très délicate. On retrouve, avec une acuité renforcée, les difficultés précédemment mentionnées :

- mauvaise connaissance du groupe cible avant la campagne (ce qui limite la portée d'analyse avant-après) ;
- possibilité d'effets secondaires non directement perceptibles (report du tabac vers une autre consommation nocive...) ;
- problème de la permanence des effets, etc.²⁷

Enfin, au-delà des politiques de communication sociale au sens strict, l'évaluation de l'impact de politiques publiques « classiques » sur les attitudes et les comportements des Français s'est également développée au cours de ces dernières années. Pour ce faire, l'utilisation de systèmes de suivi sociologique permet de repérer cet impact par rapport aux tendances longues. L'encadré n° 1 présente le système 3 SC²⁸ mis au point et géré par le Cofremca qui permet d'opérer de telles reconnaissances.

IV - 6 La régulation macroéconomique

Le terme de « régulation macroéconomique » renvoie aux diverses politiques publiques ayant pour objet d'influer sur l'évolution des grands agrégats macroéconomiques : produit national, emploi, solde budgétaire, indice des prix, solde commercial, etc. Les instruments utilisés peuvent concerner les politiques publiques précédemment évoquées. Inversement, ces politiques publiques, même lorsqu'elles ne sont pas conduites dans cette perspective, sont susceptibles de peser d'un poids notable sur l'évolution macroéconomique globale. On touche là au problème du caractère mixte des politiques ou, encore, de la multiplicité des objectifs et des effets.

L'évaluation des effets d'une politique de régulation macroéconomique (exemple : relance, dévaluation, blocage des prix) ou des effets macroéconomiques d'une politique spécifique (exemple : aides à l'emploi) a fait l'objet de plusieurs travaux. Ils recourent tous à la même approche : celle d'une simulation *ex post* de la politique, compte tenu des variables d'environnement macroéconomique observées

27. On se reportera, sur ces divers points, à l'article cité de B. Dubois et qui contient en outre, pour le cas des campagnes d'intérêt public, une comparaison des avantages et inconvénients des divers plans de recherche (groupe témoin, avant-après etc.).

28. 3 SC pour Système Cofremca de suivi des courants socio-culturels.

Encadré n° 1

Le système Cofremca de suivi des courants socio-culturels²⁹

* Il s'agit d'un outil sociologique d'observation et de mesure des changements de la société civile française.

* Il prend en compte l'ensemble de la population métropolitaine française de 15 ans et plus.

* Ce qui est observé :

- les sensibilités, les mentalités ;
- les motivations, les systèmes de valeurs ;
- les comportements et les mœurs, dans tout une série de domaines (alimentaire, habitat, loisir, travail, sexualité, etc.).
- les paradigmes, c'est-à-dire les grandes représentations sous-jacentes du monde, de la nature, de l'avenir, des institutions, des formes d'action, partagées à un moment donné par l'ensemble, ou par tel ou tel segment de population ;
- les structures informelles, relationnelles, la position de l'individu par rapport aux structures de son environnement et aux institutions (l'État, l'entreprise, les syndicats, la famille, etc.) ;
- les processus discrets et systèmes d'autorégulation non immédiatement visibles et qui assurent le fonctionnement d'une société.

* Comment :

- par combinaison de différentes approches de recherche sociologique et capitalisation dans le temps, permettant à la fois des analyses diachroniques et synchroniques ;
- analyse socio-historique par analyses secondaires de milliers d'enquêtes — périodes 1950-1970 — et travaux antérieurs de différents sociologues ;
- monographies ethno-sociologiques — de familles, d'entreprises, de groupes sociaux, d'innovations réussies ou échouées, d'événements sociaux, politiques, médiatiques, etc ;
- entretiens psychosociologiques en profondeur sur des segments de population particuliers, des précurseurs, des cohortes, etc ;
- enquête quantitative annuelle reproduite depuis 1974 sur un échantillon renouvelé chaque année de 2 500 Français, représentatif de la population de 15 ans et plus, et alimentant une banque de données socio-culturelles donnant lieu à de multiples entrées et procédures de traitement et structuration permettant de décrire quantitativement les phénomènes de changement et de diffusion des nouvelles valeurs dans la société française.

* Disponibilité et accès à l'information :

- par souscription annuelle donnant droit à l'ensemble des rapports et à l'insertion de questions spécifiques au souscripteur ;

29. Communication au groupe de M. Devez, directeur à la Cofremca.

Encadré n° 1 (suite)

- par séminaires de travail sur l'information socio-culturelle et ses implications dans un domaine donné ;
- élaboration et utilisation de scénarios d'évolution de la société française ;
- par extraction et traitement à la demande d'informations stockées en banque de données.

* Internationalisation

La Cofremca a suscité la création d'observatoires utilisant sa méthodologie dans 18 pays, dont tous les principaux pays européens, regroupés au sein d'une structure commune : RISC. Des programmes de recherche européens, permettant des comparaisons de nature et de rythme d'évolution, ont été lancés depuis 1981 (programme ACE — Anticiper le Changement en Europe).

* Applications principales :

- marketing des produits et services, marketing de la consommation ;
- recherche d'innovation produits/services ;
- marketing socio-politique — Marketing des institutions et services publics ;
- évaluation et pilotage des stratégies, des politiques et des actions ;
- sensibilisation — introduction à l'action systématique sur des ensembles complexes ;
- élaboration des stratégies, des desseins à long terme d'entreprises ou institutions intégrant les dynamiques sociétales ;
- pilotage catalytique, combinant action / évaluation / ajustement.

* Principaux programmes spécifiques dérivés :

- évolution des mœurs alimentaires des Français ;
- travail et changement social ;
- socio-médical ;
- habitat et cadre de vie ;
- international food and drink monitor ;
- médias.

(exemple : conjoncture internationale, niveau du dollar), avec un **modèle macroéconométrique**. Divers exemples de cette approche, portant sur le passé récent, figurent dans l'ouvrage A. Fonteneau et P.-A. Muet (1985), *La gauche face à la crise*, Presses de la Fondation nationale des sciences politiques³⁰. On trouvera dans l'encadré n° 2 un résumé de l'évaluation faite de la relance de 1981-1982.

Cette approche présente plusieurs avantages : facilité de mise en œuvre (dès lors qu'un modèle est disponible, bien sûr !), possibilité de décomposer l'interprétation causale (par exemple, les effets moindres que prévu d'une relance sont dus, pour tel montant, à un accroissement du taux d'épargne des ménages, pour tel autre, à un environnement international plus bas que projeté lors de la formulation de la politique etc.). L'expérience accumulée en la matière conduit à juger la **fiabilité numérique des résultats obtenus convenable, mais sur la seule période de validité des mécanismes des modèles macroéconométriques utilisés, c'est-à-dire 2 ou 3 ans**. L'évaluation avec des modèles macroéconométriques de politiques de « moyen terme » reposant sur des mécanismes « d'offre » plus inertes (exemple : politiques de restructuration industrielle, abaissement du coût relatif capital-travail...) apparaît, en l'état de l'art actuel, beaucoup plus sujet à caution.

30. Plusieurs études analogues ont été publiées dans la revue *Economie et Statistiques*.

LA RELANCE DE 1981-1982 : L'ÉVALUATION DE L'OFCE

L'évaluation de l'OFCE a consisté tout d'abord à répertorier les principales mesures de relance prises (revalorisation du SMIC, mesures sociales, actions en faveur de l'emploi et de l'investissement, mesures fiscales et parafiscales) et à les « traduire » en termes de variables exogènes du modèle OFCE-annuel. Divers types de simulation peuvent être alors effectués :

- avec un environnement international conforme aux prévisions retenues lors de la mise en œuvre de ces mesures ;
- avec l'environnement international observé.

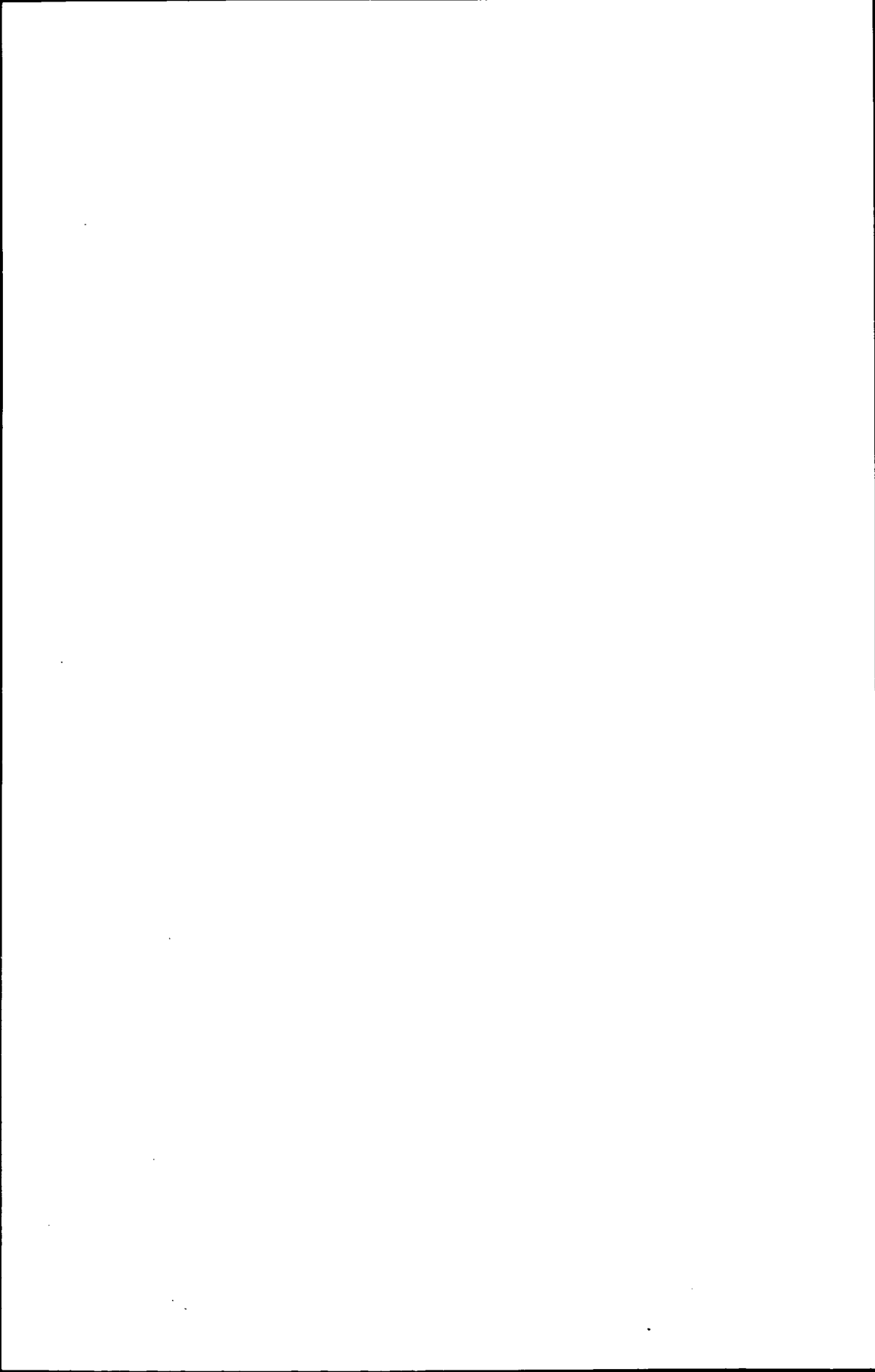
Par comparaison à l'évolution observée, on peut mettre en évidence la contribution des divers facteurs, soit de politique économique, soit d'environnement international, aux résultats atteints. A titre d'illustration on reproduit un tableau mettant en évidence l'impact de la détérioration de l'environnement de l'économie française en 1982 (par rapport à l'environnement prévu).

	Evolution observée* (1)	Impact :		Evolution simulée avec l'environnement prévu en 1982 (4) = (1)-(2)-(3)
		Baisse de la demande et de l'inflation mondiale (2)	Hausse du dollar par rapport à l'écu (3)	
<i>Variations par rapport à 1981 (%)</i>				
PIB (volume)	1,8	- 1,7	- 0,3	3,8
Consommation des ménages (volume)	3,4	- 0,3	- 0,5	4,2
Investissement total (volume)	0,7	- 1,9	- 0,9	3,5
Exportations (volume) ..	- 2,6	- 7,0	+ 0,3	4,1
Importations (volume) ..	6,8	- 4,5	- 0,8	12,1
Prix à la consommation	11,2	- 0,1	+ 2,0	9,3
<i>Niveau</i>				
Chômage (en milliers) ...	1 870	+ 90	+ 10	1 770
Solde des biens et services (MdF)	- 100	- 20	- 24	- 56

* Les volumes sont mesurés aux prix de l'année 1970. Les évolutions observées en 1982 pour l'investissement, les importations et, à un moindre degré, la consommation des ménages diffèrent de celles qui sont mesurées aux prix de l'année précédente. Ces divergences tiennent aux différences de pondération dans les deux modes d'évaluation.

V

**Les enseignements
de quelques expériences nationales**



V

Les enseignements de quelques expériences nationales

V - 1 Evaluation et esprit des peuples

Le mouvement évaluatif n'est pas seulement la mise en œuvre de certaines méthodes de recherche pour apprécier les effets des actions publiques dans tel ou tel domaine, c'est en même temps un ensemble de pratiques administratives et politiques qui définissent son organisation, l'usage qui en est fait et les conséquences qu'il a sur le fonctionnement des institutions. L'observation rapide d'un certain nombre d'expériences nationales, dont la nôtre, montre à quel point l'aspect apparemment universaliste des méthodes de recherche et des techniques statistiques se combine avec le particularisme de chaque système institutionnel et de chaque régime politique. Mis à part les régimes non démocratiques avec lesquels l'évaluation, dont la logique profonde est le doute scientifique, est par nature incompatible, les pratiques d'évaluation expriment l'« esprit des peuples » et leur façon d'utiliser plus ou moins intensément les acquis des sciences sociales comme instrument de démocratie.

Nous commencerons par l'exemple des États-Unis. L'ampleur et la diversité de l'expérience américaine lui confèrent en effet une place particulière. Son étude montre en outre que l'évaluation y est un mouvement fortement institutionnalisé et aux racines anciennes, c'est-à-dire tout le contraire d'une mode. Cela sera confirmé par une brève revue de quelques autres pays dans lesquels l'évaluation s'est implanté selon des voies originales.

On examinera ensuite le cas français en s'efforçant de répondre à deux questions. La première est de savoir pourquoi l'évaluation reste assez peu développée en France. La seconde consiste à s'interroger sur les possibilités d'un développement spontané de l'évaluation dans notre pays.

V - 2 L'exemple des États-Unis

1) La prise de conscience, à la fin des années 1960, de l'échec du Planning Programming Budgeting System (PPBS) est parfois citée comme étant à l'origine du mouvement évaluatif aux USA. En fait, si l'expérience du PPBS a joué un rôle important dans le développement de l'évaluation, celle-ci est issue principalement d'un mouvement bien antérieur et beaucoup plus profond constitué de la recherche en sciences sociales appliquée aux conséquences de l'action gouvernementale¹.

Ce courant, très ancien dans son inspiration, s'est exprimé pour la première fois en rapport direct avec l'action du pouvoir fédéral lors du New Deal. A partir de cette époque, de multiples recherches empiriques vont être menées dans les domaines de l'éducation et des programmes sociaux, bénéficiant des progrès méthodologiques, notamment en matière de mesure, offerts par les études psychologiques menées pour l'armée américaine pendant la guerre. Les années 1950 voient le développement d'évaluations importantes utilisant les méthodes d'enquête et d'analyse statistique mises au point progressivement dans les différentes disciplines des sciences sociales. En effet, ce courant d'évaluation est pluridisciplinaire, il s'enrichit des apports méthodologiques de la sociologie, la science politique, la psychologie, les statistiques, l'économie et le marketing. Il va se trouver considérablement renforcé par l'existence d'un débat permanent dans les années 1960 sur le bien-fondé des réformes lancées sous Kennedy et surtout sous Johnson (la « Great Society ») dans les domaines de l'éducation, la justice, l'aide sociale, l'urbanisme, les transports, le logement et qui vont entraîner une croissance très rapide des dépenses publiques.

Cette décennie est celle de la croissance vertigineuse et du déclin aussi rapide de l'effort de rationalisation connu sous le terme de **Planning Programming Budgeting System (PPBS)**. Ce deuxième courant qui alimente le mouvement évaluatif est inspiré essentiellement par les économistes, renforcés par les spécialistes de recherche opérationnelle. Le PPBS, système intégré d'administration, comporte une phase amont de fixation d'objectifs et de choix des projets reposant sur les techniques d'analyse de système et de calcul économique. Cette phase suppose un travail d'évaluation des effets des programmes publics. Il s'agit là du lien intellectuel le plus direct, même s'il n'est que très partiel, entre l'évaluation de politique et le PPBS. Une différence fondamentale existe néanmoins entre ces deux conceptions. L'évaluation opérée dans le cadre des études analytiques du PPBS est essentiellement une *évaluation a priori des effets attendus* de programmes alternatifs et non une observation et une mesure *a posteriori* des effets réels d'une

1. Pour cette analyse du mouvement évaluatif aux États-Unis nous nous inspirons d'une part de l'étude de E. Chelimsky, *L'évaluation de programmes aux États-Unis*, Politique et Management Public, n° 2, juin 1985, et d'autre part de l'analyse de L. Rouban présentée dans un rapport particulier effectué pour le groupe de travail sur *L'évaluation des politiques publiques aux États-Unis*, ronéo (Centre de recherches administratives et politiques, université de Rennes, décembre 1985). E. Chelimsky décrit notamment la constitution des deux courants à l'origine du mouvement évaluatif. L. Rouban étudie en particulier le lien entre l'évaluation et la vie politique américaine. Nous utilisons également F. Levy, *L'évaluation des politiques publiques aux États-Unis* in EPP. op. cit., ainsi que N. Beckman, ed. *Policy Analysis in Government, alternative to muddling through*, Public Administration Review, 37, n° 3 (1977), p. 221-262.

politique une fois qu'elle a été appliquée (même si le bouclage du système prévoyait *in fine* une comparaison prévisions-réalisations). En outre, l'évaluation que suppose le PPBS repose principalement sur le calcul économique alors que l'évaluation de politique puise plus largement dans différentes sciences sociales.

En 1965, sur la base du succès apparent de la mise en place du PPBS au département de la Défense, Lyndon B. Johnson décide sa généralisation à l'ensemble des départements ministériels et des agences fédérales. Peu de temps après apparaissent à la fois les premiers signes de déconvenue et les premières analyses² d'un échec qui sera consommé en 1971. Parmi les multiples causes de cet échec, certaines ont trait directement aux études analytiques. D'une part, les économistes ont découvert que leur conception normative de la rationalité n'était pas celle qui présidait aux décisions politiques et aux choix budgétaires, d'autre part, il est apparu que leurs ambitions analytiques étaient démesurées compte tenu de l'écart entre les moyens d'évaluation *a priori* dont ils disposaient notamment en matière d'instruments de mesure, de systèmes statistiques et de données disponibles sur les phénomènes en jeu, et ceux qui auraient permis d'effectuer des estimations réalistes des effets attendus des programmes. L'écart entre l'ambition des analystes et la faiblesse de l'information était d'autant plus accentué que les programmes en question correspondaient à des domaines et à des moyens d'actions nouveaux pour l'administration fédérale américaine.

Autrement dit, l'échec du PPBS a fait ressentir aux analystes comme aux autorités l'intérêt de développer l'analyse *a posteriori* et, pour cela de renforcer les systèmes statistiques et les bases de données, et d'une façon générale, la connaissance des processus en œuvre dans les principaux domaines d'intervention de l'État.

2) La lente et précautionneuse démarche de l'évaluation empirique des politiques par le courant de la recherche appliquée, qui s'appuyait sur la connaissance accumulée, l'expérience et la formation progressive des évaluateurs, s'avérait donc finalement plus efficace, y compris comme aide à la décision, que la démarche plus technocratique du PPBS. Mais cette expérience avait apporté, outre la conviction au sein de l'exécutif qu'il fallait poursuivre l'effort d'analyse des politiques, des méthodes telle que l'analyse coûts-avantages utilisables dans l'évaluation *a posteriori*, et aussi des modifications dans les procédures de préparation et d'exécution des budgets à travers des ensembles de programmes hiérarchisés qui constituent depuis autant d'unités d'analyse pour les évaluateurs.

Les années 1970 vont être celles du développement accéléré de l'évaluation et de son institutionnalisation. D'environ 20 millions de dollars à la fin des années 1960, les crédits fédéraux dévolus à l'évaluation sont décuplés 10 ans après. Encore ces chiffres ne représentent-ils qu'une partie de l'effort d'évaluation, puisque celle-ci est aussi le fait des États, des universités, de fondations, de cabinets privés, dont les sources de financement sont diverses.

Parmi les centaines d'évaluations menées annuellement apparaissent des évaluations de grande ampleur et d'une grande exigence méthodologique, notamment

2. Notamment le célèbre article de A. Wildavsky, *Restoring policy analysts from PPBS*, *Public Administration Review*, n° 29, 1969, p. 189-201.

des expérimentations dont l'une des plus célèbres est menée sur le principe de l'impôt négatif (voir partie II), ainsi que de vastes évaluations dans le domaine éducatif ou celui des programmes sociaux.

L'évaluation, et plus globalement l'analyse de politique, se consolident comme discipline à part entière dans le domaine académique. Des programmes doctoraux de formation d'analystes et de chercheurs se constituent dans les écoles et les universités, qui proposent également des programmes de formation continue aux fonctionnaires en activité. L'analyse des politiques devient, en 1975, une des matières obligatoires dans les programmes de maîtrise destinés aux futurs fonctionnaires³.

L'évaluation s'implante dans le système politico-administratif. La branche exécutive se dote de cellules d'évaluation dans les différents départements qui constituent d'importantes banques de données à usage d'évaluation et l'**Office of Management and Budget** devient un grand utilisateur d'évaluations. Il en est de même de la branche législative avec le **Congressional Budget Office** créé en 1974 pour évaluer les projets de lois de finance. Mais c'est surtout le **General Accounting Office (GAO)** qui constitue l'instrument principal d'évaluation auprès du Congrès (voir partie III). Le GAO développe, surtout à partir d'une loi de 1974 qui renforce le statut légal de l'évaluation, une importante activité évaluatrice à côté de sa fonction traditionnelle d'audit financier.

Le mouvement évaluatif, et plus généralement l'analyse des politiques, constitue aux États-Unis, dans les années 1980, un milieu professionnel de plusieurs milliers d'analystes répartis dans les universités, les centres de recherche, les fondations, les cabinets d'études, l'administration fédérale et locale et les services parlementaires. Ces spécialistes se retrouvent au sein de « sociétés savantes »⁴ et s'expriment et publient les résultats de leurs travaux dans plusieurs revues scientifiques spécialisées dont la plus importante est aujourd'hui le *Journal of Policy Analysis and Management*. Les résultats des études sont fréquemment repris dans la grande presse, et le débat politique lui-même est alimenté par les résultats des études. L'ampleur du mouvement et son institutionnalisation ne doivent pourtant pas masquer que celui-ci n'est à l'abri ni d'une certaine banalisation ni de certains aléas dus aux changements politiques et aux querelles d'écoles sur la nature et l'utilisation de l'évaluation.

3) Sous les présidences démocrates de Kennedy et de Johnson, s'engager à évaluer les programmes sociaux était un moyen de donner des garanties à certaines parties de l'opinion et du Congrès. D'autre part, l'administration qui ne vivait pas en situation de rationnement budgétaire voyait dans l'évaluation le moyen d'améliorer, au plan technique, ses modalités d'intervention. La préoccupation d'efficacité l'emportait donc sur celle d'efficience. Cette dernière va, au contraire, se renforcer sous l'administration Nixon. Celle-ci qui avait, d'une part, des éléments d'appréciation des résultats de certains programmes de la période

3. *Guidelines and standards for professional Master Degree program in public affairs-public administration*, National Association of Schools of Public Affairs/Public Administration, 1974.

4. Policy Studies Organization (science politique), Society for the Study of Social Problems (sociologie), Society for the Psychological Study of Social Issues (psychologie), Evaluation Research Society (pluridisciplinaire).

précédente, d'autre part, le souci de ne pas se laisser entraîner dans une vague supplémentaire de programmes publics, utilisa l'évaluation pour redéfinir son intervention davantage sous forme d'incitation que de prestations directes.

L'administration Nixon qui a lancé de grandes évaluations de nature expérimentale fut accusée de l'avoir fait également parce que c'est une forme d'évaluation qui permet d'attendre, éventuellement longtemps, ses résultats et peut donc servir d'alibi scientifique à l'attentisme.

La période de méfiance vis-à-vis de l'exécutif qui suit le départ de Nixon voit le **renforcement des pouvoirs évaluatifs du Congrès**. L'arrivée de Reagan, marquée par une volonté de rupture avec la croissance de l'intervention de l'État, se traduit à la fois par un coup d'arrêt donné à certains programmes sociaux et un ralentissement de l'effort d'évaluation mené par l'exécutif. Ce ralentissement va être en fait modéré et de courte durée. D'une part, parce que l'évaluation est pluraliste. Ainsi, un programme d'aide aux adolescents ayant quitté le domicile familial qui avait été lancé par Carter et qui était très critiqué par l'Administration Reagan fut-il évalué par le GAO à la demande du Congrès⁵. Ce programme ayant fait l'objet d'une évaluation très favorable, le Congrès, loin de le supprimer, augmenta son budget de 50 %. D'autre part, l'exécutif lui-même prit assez vite conscience du rôle que pouvait jouer l'évaluation pour réaliser ses projets de déréglementation, de réduction des programmes d'intervention et d'amélioration de la productivité dans les services publics. Ainsi, les crédits relatifs aux sciences sociales de la National Sciences Foundation, qui avaient été supprimés par l'administration Reagan dans un premier temps, ont été rétablis deux années plus tard. Et un effort considérable est fait pour accentuer la pratique de l'évaluation au niveau subfédéral (États et grandes municipalités).

Alors qu'en matière de système de gestion, on avait vu le PPBS de Kennedy/Johnson, puis le « Management By Objective » de Nixon et le « Zero Base Budgeting » de Carter, le développement sur moyenne période du mouvement évaluatif ne semble pas profondément affecté par l'alternance politique. Les changements d'orientation politique du pouvoir modifient les usages de l'évaluation et influence les types de démarche évaluative qui vont prévaloir, notamment du côté de l'exécutif. Toutefois, l'évaluation s'est déployée et installée au sein du système politico-administratif américain par l'action de conditions favorables qui tiennent à ce système lui-même mais aussi à des causes plus larges qui en font une composante de la démocratie américaine.

4) L'administration fédérale est constitutionnellement soumise à la double autorité de la Présidence et du Congrès. L'institutionnalisation de l'évaluation tant du côté du législatif que de l'exécutif peut largement s'interpréter comme un élément de la « course aux armements » entre les deux pouvoirs, cette compétition étant tempérée par l'existence d'instances intermédiaires comme le GAO dont les travaux d'évaluation contribuent au dialogue entre les deux branches du gouvernement.

5. U.S. General Accounting Office, Federally supported centers provide needed services for runaways and homeless youths, GAO/IPE — 83-7, 26 septembre, 1983, cité par E. Chelimsky, op. cit.

L'évaluation joue également un rôle d'intégration et de régulation dans un système administratif dont la cohésion n'est pas assurée, comme elle l'est en France, par de multiples dispositifs d'intégration tel qu'un droit administratif particulier, la planification nationale, le système des grands corps ou les inspections ministérielles.

Enfin, les valeurs culturelles propres à la société américaine font une grande place à l'empirisme et à l'approche inductive de la décision. De plus, l'utilisation de la science pour éclairer l'action publique est une valeur largement partagée et ce, depuis les origines de la démocratie américaine. Les sciences sociales jouissent d'un statut élevé notamment depuis la démonstration de leur apports pratiques lors de la Seconde Guerre mondiale. Elles vont donc pouvoir « entrer dans le siècle », dialoguer entre les différents pôles de pouvoir et servir de base au développement de l'évaluation.

V - 3 Autres expériences étrangères : R.F.A, Pays-Bas, Royaume-Uni, Suède, Suisse

Si les États-Unis ont incontestablement joué un rôle d'avant-garde dans le développement de l'évaluation, plusieurs expériences étrangères, parfois plus modestes, fournissent également des enseignements intéressants. On passera successivement en revue le cas de la République fédérale d'Allemagne, des Pays-Bas, du Royaume-Uni, de la Suède et de la Suisse.

1) **L'administration fédérale allemande**⁶ a connu, comme la plupart des pays occidentaux, un mouvement de rationalisation incluant la conception de mécanismes de planification et de contrôle dans la fin des années 1960. Ce mouvement a coïncidé avec la formation en 1965 d'une coalition entre les chrétiens démocrates et les socio-démocrates et l'acceptation par les premiers, sous l'impulsion des seconds, d'une plus forte intervention de l'État dans les activités économiques et sociales. Ce changement de point de vue est considéré comme l'origine, avec un décalage de temps, des activités d'évaluation qui se sont développées à partir de 1969.

6. Des compléments sur le cas de la R.F.A. peuvent être trouvés dans les documents suivants :

H.U. Derlien, *Program evaluation as an approach to the management of public policies in the Federal Republic of Germany*, Symposium on management of public policies, ENA, Paris, 27 octobre 1980.

H.U. Derlien, *Implementations probleme : Bürokratische ineffizienz oder politische Programm fehler ?* dans *Verwaltungsarchiv*, 1984.

H. König, *L'évaluation de la politique en République fédérale d'Allemagne*, conférence du Groupe européen d'administration publique, 3-6 septembre 1985, Louvain.

G. Färber, *Evaluation and financial control in the Federal Republic of Germany*, conférence du Groupe européen d'administration publique, 3-6 septembre 1985, Louvain.

Dès le début des années 1970, on constate dans plusieurs ministères la création d'unités de recherche socio-économique qui, dès leur origine, se consacrent à l'évaluation *a priori* et *a posteriori*. La plupart de ces évaluations sont accomplies sous forme de recherches sur contrat, mais ces unités spécialisées discutent les plans de recherche, fournissent des données, passent les contrats et suivent le travail des équipes de recherche. La création de ces unités a été encouragée par la réforme budgétaire de 1969 qui permettait au ministère des Finances d'exiger une étude économique des ministères présentant des programmes nouveaux. De telles unités furent créées au ministère de l'Agriculture, des Transports, de l'Aide au Développement, de la Technologie. Cette dernière s'est surtout consacrée à l'évaluation des institutions de recherche et des programmes de recherche financés sur fonds publics. Une organisation de l'évaluation des politiques de communication a également été mise en place à l'Agence fédérale sur les relations avec la presse et à l'Agence fédérale pour l'information de santé.

Le Parlement a joué un rôle important dans la mise en place de dispositifs d'évaluation. Ainsi qu'il a été constaté dans le cas des États-Unis, les doutes d'une large partie de l'opinion, reflétés par les médias, sur le bien-fondé de nombreux programmes d'intervention étatique sont à l'origine d'une exigence croissante du Parlement quant à l'évaluation des effets de ces programmes. Depuis 1970, de très nombreux votes de lois et de programmes dans les domaines économiques, sociaux et éducatifs ont été assortis d'une obligation pour le gouvernement fédéral d'organiser une évaluation de la mise en œuvre et des effets obtenus.

Des expérimentations ont simultanément été conduites dans les domaines de l'éducation et de la santé. Le choix de ces domaines n'est pas étranger à des raisons politiques. En effet, l'éducation et la santé sont constitutionnellement attribuées aux régions et ne relèvent pas du gouvernement fédéral qui n'est autorisé dans ces domaines qu'à définir un cadre général. L'expérimentation a constitué pour le gouvernement fédéral un moyen de lancer des réformes sur des territoires hors de sa compétence, en passant des accords administratifs entre ses services et ceux des gouvernements locaux. La justification d'une telle procédure étant de multiplier les expériences afin de clarifier les bases d'une réglementation générale qui, elle, est du ressort du gouvernement fédéral.

La répartition des tâches en matière d'évaluation des effets des politiques est très nette quant à la réalisation des études, plus complexe quant à leur prescription et leur utilisation. La réalisation des travaux d'évaluation est considérée comme une tâche spécifique qui ne peut être accomplie, sauf dans les cas simples, par les institutions traditionnelles de contrôle ou par les services de l'administration. Lorsque certains ministères disposent d'unités spécialisées dans l'évaluation, ces unités, en général, ne conduisent pas elles-mêmes les recherches. L'évaluation est fondamentalement externe sous forme de recherche sur contrat avec des centres de recherche universitaires ou des sociétés d'études.

L'initiative de l'évaluation est généralement le fait du Parlement. Elle peut aussi venir de la Cour des comptes ou de la Chancellerie mais rarement du ministère chargé de la politique. La fonction des unités spécialisées dans les ministères est plutôt d'administrer l'évaluation externe et de faire un rapport au Parlement sur la base des études d'évaluation en proposant des modifications à la politique évaluée.

La fonction d'évaluation s'inscrit donc dans un dispositif institutionnel qui reste peu centralisé. Les services de la Chancellerie se limitant à une fonction de coordination, et le ministère des Finances n'ayant pas comme en France un rôle intégrateur, chaque ministre a, surtout dans les gouvernements de coalition, une grande autonomie, y compris dans ses rapports avec le Parlement.

Le fait que le Parlement ait joué un grand rôle dans ce processus peut expliquer que l'évaluation n'est pas très populaire dans l'administration fédérale où elle est souvent présentée sous la même terminologie que les contrôles traditionnels. Elle est ainsi parfois associée dans les esprits aux inspections ayant pour objet le respect des règles et la mesure des performances individuelles. Dans ces conditions, la résistance à l'évaluation peut exister et ne sera surmontée que s'il existe une base légale à sa réalisation ou si les critiques portées contre la politique sont très fortes. Dans ces cas où la probabilité d'une évaluation négative est élevée, on observe que les problèmes méthodologiques s'accroissent pour l'évaluateur, par défaut de coopération dans la fourniture des données et l'ouverture des dossiers. La tentation est forte ensuite pour les services de contester la validité de telles évaluations surtout si celles-ci sont globales et négatives. Les évaluations comportant une analyse approfondie de la mise en œuvre de la politique et des facteurs explicatifs des effets mesurés paraissent mieux acceptée par les administrations.

Enfin, la tendance semble être actuellement en faveur d'évaluations très sélectives, au besoin en recourant à des études préalables de faisabilité, utilisant des techniques robustes et non plus sophistiquées, faisant une large part à l'analyse de la mise en œuvre de la politique.

2) Les années 1970 ont vu un développement assez considérable des travaux de type PPBS aux Pays-Bas⁷. Toutefois, ce sont surtout les nouvelles politiques qui ont été examinées, conformément à l'attrait que suscitait à l'époque les tâches de conception. La troisième étape de la boucle évaluation→formulation→mise en œuvre était donc assez largement négligée.

Le souci de juger des actions publiques effectivement menées s'est installé progressivement à la fin des années 1970 et au début des années 1980.

En réponse à des critiques visant certaines actions ou bien parce que l'obligation de réaliser une évaluation était prévue légalement, des travaux ont été lancés dans le domaine de l'environnement, sur la réforme universitaire, sur la mise en œuvre par les administrations locales des politiques décidées centralement (entre autres).

7. Cette analyse s'appuie sur les documents suivants :

— E. Sorber, *Some experiences with program evaluation in the Dutch central government*, Workshop on Policy Evaluation, Louvain, septembre 1985.

— Th. Van der Tak, *Policy evaluation*, Workshop on Policy evaluation, Louvain, septembre 1985.

— H.P. Potman, *Evaluation of the Noise Nuisance Act in the Netherlands : the use of Acceptance of policy in evaluation research*, Workshop on Policy Evaluation, Louvain, septembre 1985.

— M. Horrevoets et F. Vonks, *Communication*, Workshop on Policy Evaluation, Louvain, septembre 1985.

— J. Van Geest et A. Ringeling, *Administrative Impact Assessments*, Workshop on Policy Evaluation, Louvain, septembre 1985.

M. Sorber a par ailleurs présenté une communication au groupe de travail.

Jusqu'à une période récente, ces démarches étaient restées assez partielles. Elles étaient souvent l'œuvre de chercheurs universitaires dont les préoccupations ne répondaient pas toujours à celles des décideurs. Une certaine tradition s'était ainsi dégagée mais sans qu'aucune organisation spécifique fut créée.

La situation a évolué récemment avec la création en 1982 au sein de la direction du Budget d'une cellule de coordination méthodologique et de diffusion de travaux en matière d'analyse des politiques publiques. Une nouvelle phase de la recherche est ouverte. Après des évaluations spécifiques employant des techniques lourdes, fortement orientées vers la quantification, deux axes de réflexion se dessinent : l'un met l'accent sur la mesure des effets, l'autre se concentre sur l'analyse des organisations publiques et des procédures d'action. Par ailleurs, les analystes essaient de délivrer des messages plus directement utilisables par les décideurs ; quant aux services gouvernementaux, leur approche de l'évaluation se modifie en raison :

- de la tendance à la privatisation et à la déréglementation ;
- de clauses prévoyant pour certaines lois soit une durée limitée (sunset), soit une évaluation périodique ;
- d'une attitude plus tournée vers la production de résultats.

Ces services tentent de développer leur propre activité évaluative, le développement créé à la direction du Budget ayant un rôle de soutien et de diffusion des méthodes et des travaux. Sont ainsi menées des actions de formation, diffusés des ouvrages de méthodes et un périodique : *Belei Analyse* (Analyse de Politiques).

Un aspect de l'évaluation aux Pays-Bas est l'échange qui est organisé entre les experts du gouvernement central, les autorités locales, les chercheurs universitaires et les sociétés d'études privées. Pour promouvoir cette idée, le ministre des Finances a invité à une conférence sur l'état de l'art en matière d'analyse de politique des décideurs et des analystes représentant les quatre groupes précités. Cette réunion a donné lieu à la création d'un groupe permanent et d'un programme de développement et d'animation du mouvement évaluatif.

3) A certains égards, le contexte dans lequel s'est développée l'évaluation au Royaume-Uni⁸ présente des similitudes avec le cas français. Du côté de la demande d'évaluation il y a, à la différence des États-Unis, une tradition de « dédain culturel » de la part des gouvernants et des hauts fonctionnaires pour les sciences sociales et leur contribution potentielle à la décision publique. D'autre part, contrairement à bien des idées reçues, les politiques gouvernementales sont préparées dans un cercle décisionnel très centralisé, constitué de hauts fonctionnaires des administrations centrales et des groupes d'intérêt qui sont leurs partenaires permanents. Le développement de l'évaluation trouve ses racines dans le désenchantement qui résulte des espoirs excessifs mis, dans les années 1960-1970, dans la rationalisation des choix par l'usage du calcul économique et des systèmes budgétaires intégrés. L'idée d'évaluation a également pénétré dans les collectivités territoriales et le secteur des organisations autonomes de santé, mais ces secteurs ont des moyens limités à y consacrer.

8. L'exposé de l'expérience britannique est principalement issu de la présentation de M.J. Spackman, *L'évaluation au Royaume-Uni*, avec un commentaire de J.C. Thoening, in EPP, chap. 3, ainsi que de B.W. Hogwood L.A. Gunn, *Pc. Analysis for the real world*, Oxford University Press, 1984.

Du côté de l'offre, la tradition britannique est celle de sciences sociales relativement cloisonnées et centrées sur les problèmes théoriques. L'intérêt pour l'application des sciences sociales aux problèmes pratiques d'administration est récent et apparaît à une partie de la communauté scientifique comme le moyen de développer une approche de la recherche et de l'enseignement plus appliquée et plus interdisciplinaire. L'analyse de politique en tant que discipline a mis de nombreuses années à être reconnue dans les milieux scientifiques et à être enseignée dans quelques universités, instituts polytechniques et écoles de gestion. Elle fait également l'objet aujourd'hui de programmes de formation pour fonctionnaires dans ces institutions et au « Civil Service College ». Dans le cas anglais, il est commode de classer les procédures d'évaluation selon leur objectif.

a) **Les procédures d'évaluation comme moyen de promouvoir ou de faciliter la mise en œuvre d'une politique nouvelle**

Il n'existe pas actuellement de procédure centralisée prévoyant l'évaluation de politique. Il n'en n'a pas toujours été ainsi. Le « **Program Analysis and Review** » (PAR) a fonctionné pendant plusieurs années, jusqu'à une date récente. Chaque année, les ministères devaient définir avec le Trésor quelles étaient les politiques publiques qu'il convenait de réexaminer. Pour chaque politique étudiée, une commission était constituée sous l'égide du ministère concerné avec la participation du Trésor et des autres ministères intéressés. La procédure était secrète : les PAR n'étaient pas annoncés officiellement et les rapports n'étaient pas publics. La procédure des PAR s'est avérée très utile. Il y a été mis fin quand les ministères ont été dotés de capacités d'évaluer ou de faire évaluer eux-mêmes leurs politiques. Les capacités d'évaluation sont très variables d'un ministère à l'autre. Certains disposent de services spécialisés qui participent à l'évaluation. La plupart des ministères supervisent des évaluations faites par les universités ou les institutions extérieures.

Il n'existe plus de service central chargé de faire des évaluations pour contribuer à l'élaboration de nouvelles politiques. De 1973 à 1983, un tel organisme, le « **Central Policy Review Staff** », faisait partie du secrétariat général du gouvernement et était composé pour moitié de fonctionnaires détachés pour deux ou trois ans et pour moitié de personnalités scientifiques venant de diverses origines. La disparition de cet organisme unique a laissé place à des divisions plus petites, proches du Premier ministre, qui étudient l'efficacité des ministères et leur fonctionnement et pas seulement les grands problèmes de politiques nationales. Une tradition particulière au Royaume-Uni est constituée par les Commissions spéciales et les Comités d'enquêtes qui mêlent fonctionnaires et personnalités compétentes extérieures à l'administration.

Outre les travaux effectués ou financés par l'administration, une grande partie de l'évaluation visant à promouvoir de nouvelles politiques est bien entendu réalisée en toute indépendance par des instituts de recherche et des groupes de pression.

b) **Les procédures d'évaluation dans le cadre d'un réexamen périodique**

Ce type de procédure n'est pas systématique mais certaines mesures sont assorties d'une décision de réexamen au terme d'une période déterminée ou de façon

récurrente. Il ne s'agit pas seulement d'un réexamen lié à la procédure budgétaire (bien que celle-ci puisse conduire à des études plus approfondies) mais d'une évaluation spécifique à échéance fixe. Il s'agit par exemple de la décision de privatiser les procédures de règlement de contentieux d'assurance ou de la politique de fixation de la marge acceptable dans les marchés publics de défense nationale n'ayant pas fait l'objet d'un appel d'offres. Dans ce dernier cas, le réexamen du taux de marge et son mode de calcul est effectué tous les deux ans par une commission indépendante.

c) Les procédures d'évaluation comme moyen d'améliorer les processus de décision et de gestion des dépenses

Les évaluations de ce type sont, si elles portent sur des objectifs limités, effectuées par le ministère chargé de la politique, ou, si elles portent sur des objectifs plus larges, supervisées ou réalisées par un organisme distinct de celui qui est responsable de la politique.

Comme exemple du premier type, on peut citer les évaluations (assez classiques) réalisées dans le domaine des transports (étude d'axes nouveaux ou aménagés, de l'effet des péages ou de nouvelles voies de circulation sur l'activité industrielle...). Des réexamens plus complets sont effectués par des services de recherche internes à l'administration mais externes au ministère décideur et par des centres de recherche universitaires ou des sociétés d'études ou d'audits privés. **L'appel à ces experts extérieurs s'est beaucoup renforcé dans les dernières années.** Ces études portent sur des domaines très variés tels que l'évaluation des établissements scolaires, la politique pénale, l'efficacité des organisations de santé ou les aides à l'industrie.

L'audit externe de l'administration a évolué d'autre part en direction de l'évaluation. L'audit externe de l'administration centrale est effectué par l'Exchequer and Audit Department (E & AD) sous l'autorité du Comptroller and Auditor General, qui est l'homologue du Premier président de la Cour des comptes. Traditionnellement consacrés à la vérification des comptes et à la dénonciation des irrégularités et des gaspillages manifestes, ces contrôles sont étendus depuis le National Audit Act de 1983 à des missions d'études « portant sur les économies réalisées, l'efficiences et l'efficacité ». L'Exchequer and Audit Department a vu ses moyens renforcés quantitativement et qualitativement notamment par l'apport d'experts extérieurs. Les rapports de l'E & AD sont examinés par la Commission des comptes publics de la Chambre des communes qui généralement les publie ainsi que ses propres recommandations après avoir, le cas échéant, entendu publiquement les hauts fonctionnaires concernés par ces rapports.

Une telle procédure et la publicité qui l'accompagne donnent beaucoup de poids aux évaluations. Il est intéressant de signaler que la Commission des comptes publics regroupe des membres de tous les partis composant le Parlement et qu'elle est normalement présidée par un ancien ministre des Finances appartenant à un parti de l'opposition.

d) Les procédures d'évaluation comme moyen de faire connaître les résultats obtenus à l'opinion publique

Outre la Commission des comptes publics, d'autres commissions parlementaires suivent l'action de chaque département ministériel. Elles peuvent étudier spécifiquement telle ou telle politique et entendre les représentants de ministères concernés. Le Livre blanc sur les dépenses publiques publié chaque année est utilisé de plus en plus comme outil d'évaluation et comporte de plus en plus de mesures de résultats.

Au total, l'expérience britannique se caractérise par un apprentissage de la démarche évaluative qui l'a fait passer d'une activité centrale et globalisante à une activité polycentrique et liée directement à l'action concrète. L'évolution récente s'est traduite par l'élimination des risques de monopole et par l'ouverture du système d'évaluation à de multiples instances : auprès du Premier ministre, dans les ministères, à l'E & AD, dans les centres de recherche universitaires et dans les sociétés d'études et d'audit privées. D'une certaine façon l'évaluation est en voie de se banaliser. Ce n'est plus une activité nouvelle dotée de vertus magiques mais un acte de bonne gestion parmi d'autres.

e) La Suède⁹ est l'un des pays à économie de marché dans lequel la part des dépenses publiques a le plus augmenté dans les trente dernières années. C'est aussi l'un de ceux dans lesquels la remise en cause de cette croissance s'est manifestée le plus tardivement.

Homogène, de petite dimension, urbanisée et industrialisée, la société suédoise a manifesté jusqu'à une période récente un consensus fort quant au développement du secteur public, essentiellement dans les domaines sociaux et de l'éducation. Ce secteur public est largement décentralisé puisque les deux tiers des budgets publics relèvent des comtés et des municipalités. Dans les années 1960 ont été mis en place des systèmes de planification et de contrôle de type PPBS qui ont largement été utilisés comme moyens de décentralisation de l'administration. L'application de ces méthodes n'a pas freiné l'accroissement des budgets publics mais a contribué à l'efficacité de la gestion, peut-être parce que, à la différence de nombreux autres pays, ces méthodes n'ont pas été appliquées à de vastes administrations centrales mais à des organismes autonomes et de taille restreinte, et avec un accent mis sur les systèmes d'information et non sur les études analytiques qui ont connu les mêmes difficultés de réalisation que dans les autres pays.

Ce consensus sur le développement des politiques publiques, notamment dans le domaine social, et le sentiment partagé d'une assez bonne gestion des organismes publics, peuvent expliquer la relative lenteur du développement des travaux d'évaluation en Suède. D'autres facteurs peuvent également être évoqués. Le premier est que les petits pays sont relativement désavantagés. Bien que le nombre de questions sur lesquelles une expertise est utile à la décision soit équivalent à celui d'un grand pays, les ressources disponibles en matière de recherches économiques et sociales dans un pays comme la Suède sont beaucoup plus limitées.

En second lieu, le caractère très décentralisé de la décision publique constitue une forme de substitut à l'évaluation. Les citoyens étant proches des décideurs, la

9. Voir : — B. Ch. Ysander, *L'évaluation des politiques publiques en Suède*, in EPP, chap. 4.
— R. Fremfors, *Social research and governmental commissions in Sweden*, American Behavioural Scientist, vol. 26, n° 5, 1983.

pression exercée sur le gouvernement central pour qu'il évalue les politiques a été relativement moins forte que dans d'autres pays, compte tenu également du courant redonnant davantage de décentralisation et un allègement des réglementations nationales. Bon nombre des études d'évaluation effectuées au cours de la période récente ont en effet été centrées sur des modifications de la structure des pouvoirs, à côté des études portant sur la réglementation du marché du travail, les aides sociales, les traitements médicaux, les systèmes de retraite ou l'éducation qui constituent avec les études sur les politiques de l'énergie, les gros bataillons des travaux suédois.

L'organisation des travaux d'évaluation explique en effet la lenteur des progrès de l'évaluation de l'efficacité des politiques et surtout leur caractère peu apparent.

Le principal véhicule de l'évaluation des politiques en Suède est le système des **commissions nationales**. Les lois importantes et les modifications majeures de politiques sont généralement élaborées par une commission nationale nommée par le gouvernement, composée de parlementaires issus des grands partis, de représentants des groupes sociaux (syndicats, fédérations patronales), de fonctionnaires qui assurent le secrétariat et d'experts extérieurs.

Étant donné qu'il y a toujours deux ou trois cents commissions qui se réunissent simultanément, et ce en moyenne pendant trois ans, la quasi-totalité des experts du pays apportent leur contribution aux travaux des commissions. La tâche des experts porte essentiellement sur l'évaluation des politiques passées et sur l'analyse des nouvelles options. Mais, d'une part, les directives du gouvernement limitent la portée des évaluations réalisées et, d'autre part, la principale fonction de ces commissions est de préparer la voie à un consensus parlementaire. Les travaux des experts peuvent être ainsi utilisés plus comme des armes politiques que comme une base de décision objective et avoir une faible visibilité en tant qu'information indépendante.

Une autre source importante d'évaluation est constituée par les **Commissions temporaires de recherche et de développement** créées par les principaux ministères mais jouissant d'un statut semi-indépendant et composées de représentants du ministère et des agences nationales concernées ainsi que d'experts extérieurs. Les ministères peuvent également faire appel directement à une prestation extérieure pour une évaluation, procédure actuellement en développement.

La troisième source d'évaluation est le **Bureau national de contrôle** dont la vocation est de contrôler l'efficacité des agences nationales qui sont les instruments d'action des ministères, ceux-ci étant réduits à un état-major de 40 à 100 personnes. Le Bureau national de contrôle a évolué progressivement d'un contrôle externe de l'efficacité de chaque agence, celles-ci ayant par ailleurs leurs propres systèmes de contrôle interne, vers une évaluation directe des politiques en débordant du cadre d'activité de chaque agence.

Enfin, le **niveau local** qui pèse d'un poids considérable, mais dont l'activité est pour une part réglementée au niveau national, est très surveillé au plan de l'efficacité de la gestion. Il ne fait pas l'objet d'évaluations de politiques très fréquentes sauf dans les grandes municipalités. De tels travaux sont encouragés par les deux fédérations nationales des collectivités locales qui disposent de services de recherche qui leur sont propres.

Cette organisation très fragmentée et l'affaiblissement du rôle du Parlement ne sont pas favorables à un développement rapide de l'évaluation ni à des évaluations approfondies et de grande ampleur. Seul un examen plus détaillé permettrait de comprendre si cet état de choses tient à la faiblesse de la demande sociale pour l'évaluation ou au fait que cette demande est au moins partiellement satisfaite par un substitut constitué d'une part par l'usage systématique du contrôle de gestion dans les agences nationales et collectivités territoriales, d'autre part, et surtout, par une forte participation des experts ou représentants de l'opinion dans le cadre des commissions nationales et des organisations décentralisées.

f) En Suisse¹⁰, l'évaluation des politiques publiques à l'aide des instruments de la recherche scientifique a commencé durant les années 1970. Au début, l'initiative d'évaluer était prise par le Gouvernement, le Parlement, l'administration ou encore par les chercheurs indépendants. Le mouvement a été notablement renforcé grâce aux programmes nationaux de recherche de la **Fondation nationale suisse**. Cette institution de recherche subventionnée par le gouvernement a fait réaliser un nombre élevé d'études d'évaluation dans différents domaines. Un accent particulier a été mis sur les effets régionaux de mesures budgétaires et de programmes de régulation ainsi que sur l'analyse des procédures de décision.

Ces travaux n'ont pas encore répondu complètement aux espoirs qu'ils avaient suscités en ce qui concerne leur traduction politique. Ceci tient en partie au fait que de nombreux programmes de recherche sont encore en cours de réalisation.

Pour s'en tenir aux études menées et terminées durant la période 1975-1983, en dehors des programmes de recherche, on peut citer notamment :

- la réduction de la vitesse limitée dans les agglomérations de 60 à 50 km/h (cf. partie II) ;
- la ceinture de sécurité ;
- les effets des autoroutes sur les régions concernées ;
- l'équité du financement des routes soit par la fiscalité, soit par la tarification.

Les fonctions des évaluations sont multiples : production d'informations nouvelles, préparation de décisions à venir, renforcement du consensus, légitimisation de la politique suivie. Les méthodes employées sont diverses, analyse coûts-avantages, analyse d'efficacité, expérimentation, simulation, construction d'indicateurs d'impact qualitatifs et quantitatifs.

Les évaluations sont menées tant par l'administration fédérale, cantonale,... elle-même que par des experts extérieurs et conduisent à des échanges fructueux.

10. Dr. Armin Jans, *Some Notes on Policy Evaluation in Switzerland*, Workshop on Policy Evaluation, Louvain, septembre 1985.

V - 4 Le cas de la France

1) La recherche appliquée et les techniques d'aide à la décision sont, en France également, à l'origine de l'évaluation mais dans des proportions et dans un contexte très différents.

La recherche dans les sciences sociales s'est développée en France dans un rapport distant avec le pouvoir. La recherche économique elle-même ne s'est implantée dans les circuits de décision publics qu'à partir des années 1950¹¹. Ce lien entre la recherche empirique et la décision politique et administrative reste jusqu'aux années 1960 épisodique et propre à quelques secteurs privilégiés : contributions de psychologues et sociologues dans le domaine de l'éducation, du Centre de sociologie des organisations en matière de réformes administratives par exemple. Des travaux que l'on pourrait qualifier aujourd'hui d'évaluations apparaissent mais ne sont pas élaborés ni utilisés dans cette perspective.

En parallèle, le courant du calcul économique¹² et de l'aide à la décision se développe dans les années 1960 sous l'impulsion des ingénieurs-économistes de l'administration, notamment dans les secteurs de la Défense, de l'Équipement, et de l'Énergie. Ces secteurs font alors l'objet d'une forte demande et il y a un relatif consensus sur les efforts d'investissement qu'ils nécessitent.

Le lancement de la rationalisation des choix budgétaires, en 1968, marque l'extension de l'analyse économique *a priori* dans de nombreux domaines en même temps que la multiplication des services d'études économiques et de programmation.

Cette nouvelle phase se déroule toujours en période de croissance mais l'ampleur et la diversification de la demande des services publics ne permet plus d'obtenir un consensus quant aux projets à réaliser. Les divergences sur les estimations des effets attendus des politiques étudiées commencent à s'accroître, conduisant à l'utilisation de méthodes multicritères qui sont davantage utilisées comme schémas de négociation que comme instruments de calcul.

A partir de 1975, les conséquences de la crise vont se faire sentir sur les modalités même de prise de décisions. Les études *a priori* s'avèrent de plus en plus insérées dans les processus de négociation entre organisations et groupes sociaux. Elles se diversifient dans leur nature et certaines incluent des évaluations *ex post* de politiques¹³. L'éclatement du mouvement RCB, comme celui du PPBS aux États-Unis, conduit à une certaine prise de conscience de l'insuffisance des données et de la connaissance des processus par lesquels les programmes publics produisent leurs effets dans le public. Là aussi un certain rapprochement se réalise entre les deux courants avec le développement limité, mais incontestable, dans les années 1975/1985, de travaux d'évaluation *ex post* menés principalement à l'initiative de

11. F. Fourquet, *Les comptes de la puissance*, Encres, 1980.

12. R. Poinard et B. Walliser, *Évaluation et évolution des études d'aide à la décision*, Bulletin RCB, n° 50, septembre 1982.

13. R. Poinard et B. Walliser, *Rapport d'évaluation des études prioritaires*, Bulletin RCB, n° 43, décembre 1980.

l'administration mais auxquels contribue beaucoup plus largement que par le passé la communauté scientifique en économie, sociologie, science politique et gestion¹⁴. En 1986, l'évaluation est en France une activité qui progresse faiblement et n'est pratiquement pas organisée ni institutionnalisée au sein du système politico-administratif. La production qu'on peut estimer grossièrement à quelques dizaines d'études par an, avec d'assez grandes différences entre acteurs, est à une échelle encore artisanale tant au sein des administrations qui en sont les premiers prescripteurs et réalisateurs que dans les centres de recherche. Les évaluations réalisées sont peu utilisées par les mass médias et les hommes politiques, elles ne constituent pas un élément majeur de préparation des décisions publiques. Un certain rapprochement entre les deux courants a été réalisé du côté de la recherche universitaire par le progrès du travail empirique et une moindre acuité des luttes entre écoles de pensée (néo-classiques contre marxistes en économie, structuro-fonctionnalistes contre marxistes en sociologie, comportementalistes contre freudiens en psychologie, etc.). D'autre part, certaines parties de l'administration, notamment les comités de recherche du Commissariat général du plan, ont joué un rôle pilote pour la mise en rapport des services d'études de l'administration et des centres de recherche *extra-muros*.

Une troisième composante qui contribue à l'originalité du mouvement évaluatif en France est constituée par le rôle des corps de contrôle (voir partie I). La préoccupation de l'évaluation s'est fait jour à la Cour des comptes et dans les inspections générales où elle représente une activité encore réduite mais significative. Malgré ces limitations, ce courant d'évaluation joue un rôle important par son ambivalence. D'une part les contrôles ayant un objectif d'évaluation contribuent au développement de celle-ci, mais d'autre part l'existence même et le mode de fonctionnement des corps de contrôle peuvent constituer un frein à l'institutionnalisation d'une évaluation de type scientifique. En effet, non seulement de larges parties de l'opinion des milieux administratifs et politiques ne perçoivent pas clairement la différence entre l'évaluation et les contrôles traditionnels, mais nombreux sont ceux qui considèrent plus ou moins implicitement que si une fonction d'évaluation doit exister, elle doit rester interne à l'administration, ce qui conduit plus ou moins à assimiler évaluation et modernisation des contrôles administratifs.

III apparaît donc que l'évaluation en France existe de façon significative depuis le début des années 1980 mais n'est pas encore reconnue au plan institutionnel comme une activité normale, régulière et organisée. Elle reste une activité sporadique même si de nombreux signes manifestent qu'elle est l'objet d'attentes diverses, éventuellement excessives. Cet état de fait s'explique par le jeu d'un certain nombre de freins spécifiques qui limitent son développement en France.

2) A cet égard, les difficultés méthodologiques mises en avant pour expliquer le développement limité de l'évaluation en France doivent être appréciées à leur juste valeur. Pour l'essentiel, les obstacles de type méthodologique naissent de contradictions opposant les conditions de mise en œuvre des politiques publiques et les exigences de la recherche scientifique. Sous ce dernier aspect, évaluer une

14. Sur le développement et la nature des études d'évaluation en France, voir J.-P. Nioche et R. Poinard, *L'évaluation des politiques publiques, en France*, in EPP, op. cit.

politique revient à établir une causalité entre les effets (attendus ou non) et la cause que constitue la politique. C'est certainement la procédure expérimentale qui offre les meilleures garanties pour établir la relation causale dans de bonnes conditions. Malheureusement, il est vrai que les plans de recherche de ce type supposent satisfaire de nombreuses conditions (cf. II ci-dessus).

Toutefois, les évaluations menées depuis des années dans plusieurs pays montrent que les sciences sociales offrent de nombreuses solutions techniques qui, pour peu que l'on ne tombe pas dans le fétichisme méthodologique, permettent d'apporter des éléments de réponse sinon dans tous les cas et de façon totalement fiable, du moins de manière satisfaisante et avec un degré de clarté très supérieur à celui du débat politique usuel.

Les difficultés méthodologiques n'expliquent pas vraiment le faible développement de l'évaluation en France. Celui-ci est plutôt le fruit de l'environnement administratif, politique et culturel qui limite le recours à l'évaluation des politiques publiques et rend difficile l'utilisation de certaines méthodes.

3) **Les obstacles administratifs paraissent importants mais peuvent être surmontés par certaines réformes.** Le premier de ces obstacles est l'absence de gestion par programme. Pourtant des tentatives ont eu lieu en ce sens au cours des vingt dernières années : les « programmes prioritaires » du plan, les « budgets de programmes » mis en place dans le cadre budgétaire au niveau de chaque département ministériel, les « plans d'action » élaborés par certains services administratifs. Sauf en ce qui concerne les programmes prioritaires, l'insertion de ces instruments dans les procédures de négociation et leur impact sur les décisions sont restés limités¹⁵. S'ils ont pu contribuer d'une part à rendre plus habituel le raisonnement en terme de moyens et de résultats de l'action administrative et, d'autre part, à produire une information nouvelle, les outils de ce type n'ont pas empêché que la préparation des budgets, les arbitrages gouvernementaux, les débats parlementaires, la gestion des services, les contrôles, s'exercent toujours selon la logique du budget de moyens. Le degré de précision des politiques, de leurs objectifs, des moyens qui leur sont attribués, de leur enjeu, est donc faible en moyenne, ce qui complique la tâche des évaluateurs. Il s'agit là d'une des grandes différences entre la France et les États-Unis où la gestion par programme est la règle.

Par ailleurs, le fonctionnement des structures administratives réalise un équilibre « bas » en matière d'évaluation. La rareté des évaluations et la faiblesse de leur impact décisionnel ne conduisent pas à la création d'équipes spécialisées. En conséquence, le partage des tâches d'évaluation entre divers services opérationnels ou d'études, inspections et corps de contrôle, qui pourrait être une source de richesse, a du mal à dépasser le stade de l'éparpillement et à opérer des synergies effectives. Le même processus est observable au plan financier puisque les cas de politiques prévoyant les moyens de leur évaluation dès leur lancement sont rarissimes. Or, l'évaluation sans moyens ne peut montrer qu'elle peut être efficace et constituer, par là, un facteur d'économie.

15. P. Poinard, *Les budgets de programmes, quinze ans après*, Economie et Prévision, n° 71, 1985/5.

D'une façon générale, l'administration française souffre sans doute de certaines lacunes dans son système de formation. A l'exception de l'économie (et de manière non systématique) les futurs hauts fonctionnaires sont peu mis en contact avec la recherche contemporaine dans les différentes sciences sociales. La formation permanente ne venant pas compenser les handicaps de la formation initiale, les méthodes de recherche en sciences sociales et leurs développements actuels, donc les ressources qu'elles peuvent constituer dans l'action administrative, sont en définitive peu connues.

Enfin, l'évaluation peut apparaître à certains comme redondante avec les multiples formes de contrôles administratifs, politiques et juridictionnels dont font l'objet leurs actions.

4) **les obstacles politiques et culturels** sont par nature plus insaisissables mais jouent sans doute un rôle majeur dans les réticences que rencontre l'évaluation. Peu développée dans l'administration, l'évaluation n'a pas trouvé de conditions favorables à son implantation dans d'autres sphères, notamment à cause de la prééminence de l'exécutif sur le législatif. Il s'agit là d'une différence fondamentale avec les États-Unis où la compétition entre exécutif et législatif a conduit à une sophistication croissante des mécanismes d'évaluation de part et d'autre.

En France, la situation privilégiée de l'administration en matière d'informations économiques et sociales ainsi que la relative faiblesse des mass medias dans leur rôle d'évaluateur de l'action des pouvoirs publics ont freiné le développement de l'évaluation.

De façon structurelle, la nature unitaire et très centralisée de l'État ainsi que l'existence d'un droit administratif très cohérent sont interprétés comme limitant les possibilités de modulation des politiques en fonction de conditions locales (au sens large). Les occasions d'évaluation dans des conditions d'expérimentation « naturelle » par comparaison entre différentes options de politiques sont donc réduites. A cet égard, la décentralisation peut constituer une chance pour l'évaluation et notamment l'expérimentation de politiques¹⁶.

Plus profondément encore, certaines représentations mentales, qui conditionnent les comportements administratifs et politiques, sont des freins puissants à un développement de l'évaluation.

Notre culture intellectuelle conduit à donner l'avantage à la déduction sur l'induction, au jugement normatif sur la recherche empirique, aux tâches de conception sur les fonctions opérationnelles. L'évaluation, qui repose sur l'observation, le doute et la mesure sur le terrain se heurte à ces valeurs et à ces traditions. Elle s'est heurtée à **une représentation dominante, bien que déclinante, de l'État comme promoteur infallible de l'intérêt général**. Cette idéologie consiste à regarder ou proclamer une politique comme satisfaisante, c'est-à-dire conforme à l'intérêt public, du moment qu'elle a été décidée régulièrement par une autorité légitime, ce qui exclut implicitement tout besoin d'évaluation.

16. A condition qu'il n'y ait pas une course à l'appropriation de l'évaluation.

Naturellement, et pour le même type de raisons, les méthodes d'évaluation les plus satisfaisantes scientifiquement se heurtent dans notre pays aux obstacles les plus nombreux. Ainsi, l'expérimentation est peu répandue, non seulement à cause des difficultés techniques qu'elle présente, qui ne sont pas propres à la France, mais parce qu'elle oblige ceux qui veulent y recourir à admettre publiquement qu'ils peuvent avoir des doutes sur l'efficacité des mesures qu'ils proposent, et à affronter le risque d'être accusé d'injustice, d'immoralité ou d'illégalité pour non respect du principe d'égalité de tous devant la loi.

V - 5 Les limites d'un développement spontané de l'évaluation dans notre pays

L'évaluation est donc, en France, une activité encore endémique et peu insérée dans les processus de décision. On a dit qu'elle faisait l'objet de nombreuses attentes dont certaines reposent sur des conceptions floues, hétérogènes et parfois contradictoires. Le développement modeste mais incontestable de l'évaluation dans la période récente, malgré l'ampleur des obstacles administratifs, politiques et sociologiques laisse toutefois supposer que des mécanismes plus profonds sont en jeu, comme le montrent certains exemples étrangers.

La question fondamentale qui se pose est de savoir si ces mécanismes sont susceptibles d'engendrer, spontanément et sans intervention volontariste de la puissance publique, un développement continu et solide de l'évaluation des politiques publiques. Pour prendre une image économique, peut-on compter sur le « libre jeu de l'offre et de la demande » pour assurer le développement de l'évaluation, sa qualité et sa bonne insertion dans les débats et les choix politiques¹⁷ ? En quelque sorte, l'évaluation a-t-elle un marché ?

1) L'« offre » d'évaluation est diversifiée mais cette abondance apparente masque une mauvaise distribution des capacités techniques et organisationnelles nécessaires à une évaluation de qualité. Il n'existe en France aucune institution ayant pour vocation principale de faire de l'évaluation de politique. Il y a très peu d'organismes dans lesquels l'évaluation atteint une masse critique et une permanence suffisantes pour accumuler l'expérience et faire progresser l'expertise. Le pouvoir législatif ne dispose pas des moyens requis pour faire des évaluations de qualité. Il est en effet caractérisé en France par la faiblesse de ses capacités d'expertise propres¹⁸. Les services de l'Assemblée nationale sont dotés de moyens d'études encore limités dont les efforts actuels de modernisation sont orientés

17. Sur cette analyse du « marché » de l'évaluation, voir J.-P. Nioche et R. Poinard, *L'évaluation des politiques publiques en France*, in EPP, op cit.

18. J.-L. Quermonne et L. Rouban, *L'évaluation et le contrôle politique*, in EPP, op. cit.

prioritairement vers la collecte et l'utilisation des informations produites par l'administration. Le Parlement se trouve donc dans une situation de dépendance de fait vis-à-vis de l'administration quant aux données sur les effets des politiques gouvernementales. La création de l'Office parlementaire d'évaluation des choix scientifiques et technologiques (loi du 8 juillet 1983) a introduit une innovation notable, mais celui-ci a pour objet l'évaluation *a priori* et non *a posteriori*.

Le Conseil économique et social se trouve dans une situation analogue malgré l'apparition notable d'évaluations dans ses travaux récents.

L'évaluation à la Cour des comptes est le résultat des programmes de travail que se donne cette juridiction. La part qu'elle fait à l'évaluation est encore limitée compte tenu du poids de ses métiers traditionnels.

Les inspections générales ministérielles¹⁹ se trouvent dans une position similaire bien qu'il existe des différences importantes entre elles. Les contrôles débordant sur l'évaluation de politique existent dans presque toutes les inspections mais avec des fréquences variables, qui sont assez élevées, tout en étant minoritaires, à l'Inspection générale des finances, ou à l'Inspection générale de l'industrie.

D'une façon générale, les évaluations réalisées dans les corps de contrôle restent d'ampleur limitée et leurs méthodes sont fondamentalement marquées par la tradition de vérification de régularité et d'audit financier.

Lorsque des méthodes nouvelles, telles que les enquêtes, sont utilisées, leurs modalités restent assez éloignées des normes professionnelles des milieux scientifiques, notamment compte tenu des contraintes de délais.

Ces normes ont davantage pénétré les services d'études des ministères. Mais, là aussi, l'évaluation constitue une activité marginale par rapport aux études de surveillance de l'environnement et d'éclairage des choix. Elle semble relativement plus implantée dans les secteurs correspondant à des interventions de l'État plus récentes (environnement, culture) ou plus conflictuelles (affaires sociales) et dans les services d'analyse à vocation interministérielle (Direction de la prévision, Commissariat général du plan).

Les services opérationnels effectuent eux-mêmes un certain nombre d'évaluations qui semblent plus souvent limitées à l'analyse de leur propre fonctionnement et des productions qui en découlent.

L'orientation est généralement opposée dans les évaluations réalisées par des évaluateurs extérieurs à l'administration. La presse ou les sociétés d'études mettent généralement l'accent sur la collecte d'informations par enquête dans le public sur les effets des politiques.

Enfin les centres de recherche (universités, grandes écoles, centres indépendants) contribuent de manière significative à la production d'évaluations mais, pour chacun d'eux, il s'agit le plus souvent d'une activité occasionnelle ou secondaire. Elle est, pour les laboratoires spécialisés dans une discipline, l'occasion de tester une théorie et, pour les centres spécialisés dans un secteur, une forme, parmi d'autres, de connaissance du terrain.

19. Voir P. Milloz, op. cit.

Non seulement l'offre d'évaluation est le fait d'évaluateurs pour lesquels elle constitue une activité marginale, mais leurs capacités d'évaluation sont contraintes par leur activité principale. Ainsi les points forts des corps de contrôle sont leur crédibilité auprès des décideurs, leur disponibilité, le caractère opérationnel de leurs recommandations et leur facilité d'accès aux informations internes à l'administration. De manière opposée, les centres de recherche de type universitaire extérieurs à l'administration disposeront en général d'une plus grande indépendance²⁰, seront plus aptes à opérer des remises en cause fondamentales, à collecter l'information externe à l'administration et à mettre en œuvre des méthodes de type scientifique.

2) **La demande d'évaluation** qui en principe pourrait être aussi diverse que l'offre est en fait très concentrée : l'administration y joue un rôle majeur alors que le Parlement, la presse ou les publics concernés sont rarement à l'origine d'évaluations. Bien qu'il soit parfois difficile de départager, dans les évaluations lancées par l'exécutif, celles qui l'ont été par l'administration et celles qui l'ont été par les responsables politiques, il semble que les services opérationnels soient les premiers prescripteurs. Si l'administration joue un rôle prépondérant dans la demande d'évaluation, celle-ci s'exprime de façon peu organisée. Il existe de vastes zones d'ombre, à côté de secteurs faisant l'objet d'évaluations multiples. Ces demandes multiples n'apparaissent pas coordonnées à des fins de vérification ou de complémentarité mais plutôt comme le reflet du souci d'acteurs désireux d'avoir leur information propre sur un sujet ou de répondre à une évaluation par une autre. L'essentiel de la demande se situant au niveau de ministères ou de parties de ministères, il n'est pas étonnant que les politiques interministérielles soient peu évaluées. L'organisation de la Cour des comptes par chambres sectorielles ne l'incitait pas à faire exception à cette tendance. Un décret du 11 février 1985 permet à la Cour de conduire des travaux relevant de plusieurs chambres et facilitera donc l'étude de politiques inter-organismes.

Les centres de recherche ont eux aussi un rôle de prescripteur pour les évaluations qu'ils font à leur propre initiative, mais dans le cadre de moyens généralement limités. **D'une façon générale, les acteurs sociaux autres que l'administration participent faiblement à la demande réelle d'évaluation.** Cela est probablement dû à la faiblesse de leurs moyens financiers et aux difficultés qu'ils rencontrent pour accéder à l'information.

3) Les mécanismes d'ajustement entre l'offre et la demande ne favorisent ni l'insertion de l'évaluation dans les processus de décision ni l'amélioration de la qualité des études ni une contribution croissante au débat public sur les politiques.

Les pratiques actuelles d'évaluation sont caractérisées par l'importance de l'auto-saisine et de l'auto-évaluation. L'auto-saisine correspond aux cas dans lesquels un évaluateur définit lui-même le champ et les objectifs d'une évaluation. L'auto-saisine n'est pas en soi un obstacle à la qualité mais peut fréquemment correspondre à des situations dans lesquelles soit il n'y a pas de demande, soit il n'y a pas coïncidence entre les termes de l'offre et ceux de la demande potentielle, donc une faible chance de voir l'évaluation servir à la décision. L'auto-évaluation,

20. Mises à part les contraintes liées aux activités sous contrat.

c'est-à-dire les cas d'évaluation menée à leur initiative, par les services opérationnels des politiques qu'ils mettent en œuvre, est un cas relativement fréquent. Il s'agit là de travaux qui peuvent servir à la gestion des services et à une responsabilisation de leur personnel mais qui comportent un risque de biais important qui a déjà été souligné.

L'importance de ces phénomènes relativise l'usage que l'on peut faire de la notion de marché pour décrire le mouvement évaluatif en France. Il y a une offre et une demande mais la plupart des évaluations ne font pas l'objet d'un échange entre un demandeur d'information et un offreur d'expertise. En outre, le débat public correspondant aux politiques évaluées bénéficie assez peu de ces travaux. D'une part parce que l'administration, qui joue un rôle prépondérant dans la production d'évaluation, ne publie qu'une part très faible des travaux correspondants. D'autre part parce que les évaluateurs qui publient leurs travaux, tels que les universitaires, le font souvent dans des termes impropres à un débat ouvert aux non-spécialistes.

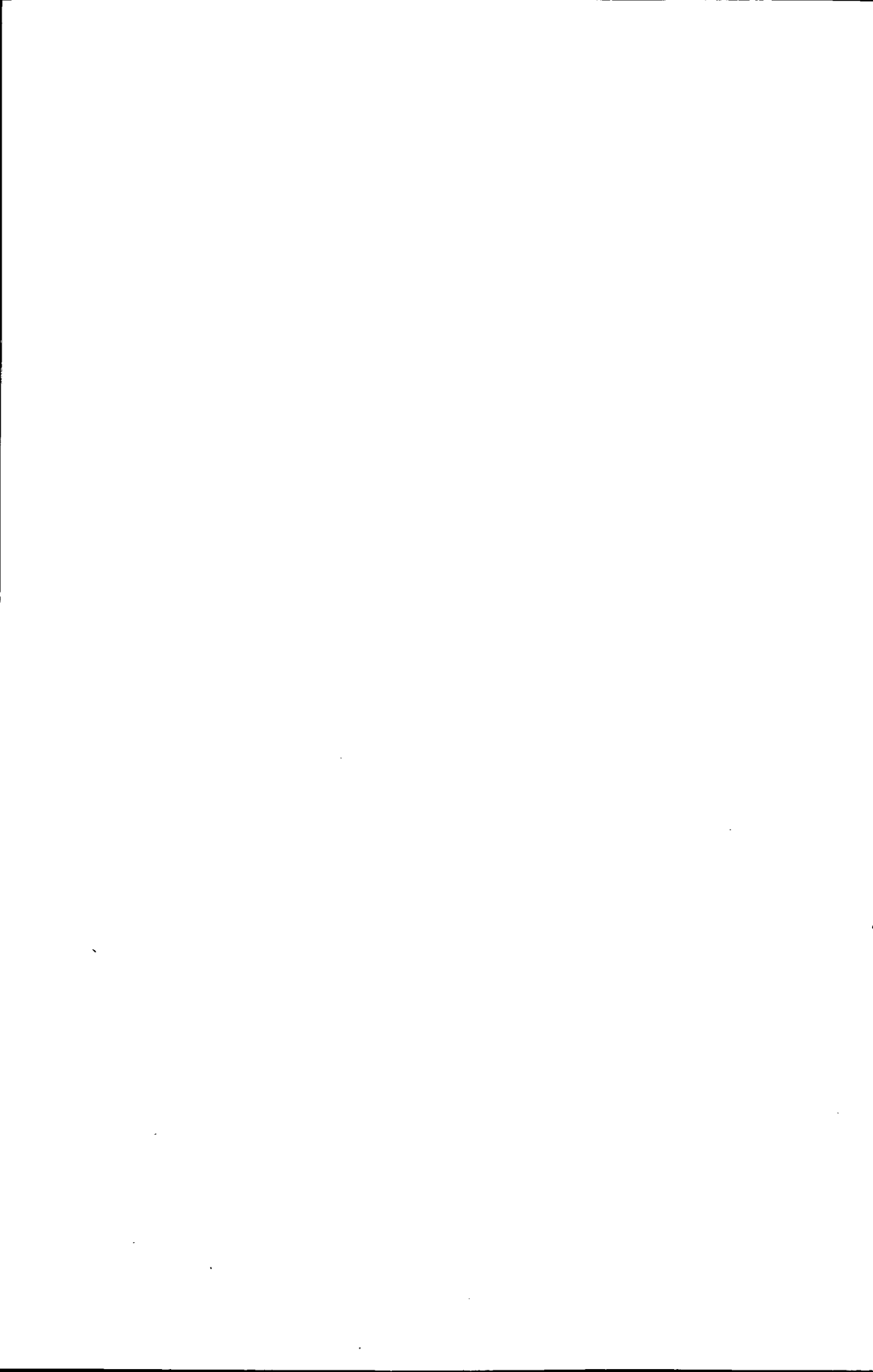
4) L'évaluation n'est ni une activité réglementée ni un véritable marché. Aucune règle ne rend l'évaluation obligatoire²¹, ne définit la manière de choisir quelles politiques doivent être évaluées en priorité, ne suggère des méthodes ou ne désigne des évaluateurs potentiels. A vrai dire, l'évaluation n'est pas une catégorie : les objectifs poursuivis, la nature des rôles, les méthodes mises en œuvre, l'utilisation des résultats ne sont pas codifiés par la règle ni par les usages.

Inversement, l'évaluation n'est pas un véritable marché. L'offre est peu structurée et insuffisante, la demande est excessivement concentrée, les échanges sont faibles et fragmentés. Il ne s'agit d'ailleurs pas d'une activité marchande par essence, même si elle ne peut se développer sans modalités de financement appropriées. On ne peut guère compter sur un élargissement naturel de la demande solvable pour « tirer » le mouvement évaluatif. Il est peu réaliste d'attendre un accroissement des flux d'évaluation d'une multiplication de demandes exprimées cas par cas par les hommes politiques. Ceux-ci sont plus tournés vers l'avenir que vers l'examen du passé, leur horizon temporel est souvent plus bref que celui exigé par l'évaluateur, etc. Mais les responsables politiques peuvent trouver un intérêt à des mesures plus générales permettant de mettre en place un dispositif d'évaluation assurant un flux régulier d'évaluations et mobilisable à la demande sur des projets spécifiques. Les médias sont potentiellement intéressés par l'évaluation mais assez rarement au point de financer eux-mêmes sa réalisation. Les informations relatives aux résultats des politiques gouvernementales sont exploitées le plus souvent de façon ponctuelle sous forme de « scoop moral » comme pour l'affaire des avions renifleurs. Il est probablement possible d'amener les médias à passer d'une recherche stricte d'un spectaculaire sans lendemain, comme l'exploitation annuelle du rapport de la Cour des comptes, à une surveillance permanente de l'action gouvernementale. Mais là aussi il est peu réaliste d'attendre cette transformation d'une évolution naturelle des médias. De la même manière, il ne faut pas espérer un accroissement spontané important de la demande d'évaluation par les administrations et, de toute façon, ces demandes resteront limitées dans leurs

21. Mis à part quelques cas de lois prévoyant explicitement qu'elles devraient faire l'objet d'une évaluation (ex : la loi sur le prix unique pour le livre).

objectifs : aucune administration ne réalisera ou ne fera réaliser une évaluation susceptible de compromettre son « fond de commerce », c'est-à-dire de remettre en cause radicalement la ou les politiques qui constituent son portefeuille d'activités et justifient les moyens dont elle dispose.

Les forces en présence ainsi que les logiques d'action des parties prenantes ne permettent donc pas d'obtenir de manière spontanée un développement rapide et satisfaisant de l'évaluation. Par le mot satisfaisant, nous voulons dire deux choses. D'une part que le respect de normes méthodologiques et déontologiques garantisse la qualité et la crédibilité des évaluations auprès des décideurs et du public afin de conduire à leur prise en compte dans les choix publics. D'autre part que le dispositif global d'évaluation soit « ouvert », c'est-à-dire qu'il assure aux différentes forces sociales la possibilité de faire ou de faire faire des évaluations ; aux évaluateurs qualifiés qu'il permette de répondre à ces demandes ; enfin que les citoyens et leurs représentants aient accès aux évaluations réalisées. Au stade actuel, le développement de l'évaluation relève donc pour partie d'une intervention volontariste de la puissance publique. Une telle orientation ne signifie ni une prise en charge par l'État de l'activité d'évaluation ni une réglementation détaillée, mais suppose la mise en œuvre d'un nombre relativement restreint de mesures visant à faciliter l'émergence d'un dispositif pluraliste de production et d'utilisation de l'évaluation.



VI

Développer l'évaluation des politiques publiques en France Cinq propositions



VI

Développer l'évaluation des politiques publiques en France

Cinq propositions

Le développement de l'évaluation des politiques dans notre pays apparaît hautement souhaitable, à la fois pour affermir la maîtrise et l'efficacité des actions de la puissance publique et pour éclairer le dialogue social.

Il est susceptible, en particulier, de provoquer ou d'accompagner de manière utile l'adaptation de ces politiques, de leurs domaines ou de leurs modalités, à l'évolution de l'environnement économique et social.

Le mouvement des esprits et des structures est propice mais les chances d'un développement spontané, immédiat, de l'évaluation, hors de toute intervention volontariste de la puissance publique, paraissent limitées. La dispersion et l'hétérogénéité des compétences comme des intérêts rendent improbable l'enclenchement automatique d'une dynamique cohérente. Ce constat ne signifie pas pour autant qu'il faille viser une prise en charge colbertiste de l'activité d'évaluation. Il implique, par contre, que soient engagées quelques actions résolues, propres à créer les conditions structurelles nécessaires pour un développement dans lequel l'État ne serait plus, à terme, qu'un acteur parmi d'autres.

L'analyse faite par le groupe, conformément au mandat qui lui avait été fixé, conduit à formuler cinq propositions principales :

- rendre les politiques publiques évaluables ;
- organiser les compétences ;
- renforcer l'évaluation des politiques au sein de l'administration ;
- accroître la capacité d'évaluation externe à l'administration et favoriser sa diffusion dans la société ;
- identifier les thèmes prioritaires.

Pour chacune de ces propositions, un certain nombre d'orientations spécifiques sont définies. Leur mise en œuvre détaillée appelle des dispositions et décisions complémentaires, notamment institutionnelles, qu'il n'appartenait pas au groupe de travail de fixer. Un souci est présent dans l'ensemble de ces propositions : celui d'aboutir à un développement pluraliste de l'évaluation, assis sur la production de travaux de qualité.

VI - 1 Rendre les politiques publiques évaluables

Il a été souligné à plusieurs reprises que les difficultés propres à l'évaluation sont très loin d'être uniquement techniques. Des obstacles plus fondamentaux se rencontrent à l'amont, pour la reconnaissance même de « l'objet » à évaluer. Si certaines conditions minimales ne sont pas vérifiées au stade de la formulation de la politique, l'activité d'évaluation peut être très malaisée voire impossible à conduire. Par rapport à la situation française actuelle, des progrès doivent être recherchés dans cinq directions complémentaires.

1) Il convient tout d'abord de **formuler aussi systématiquement que possible les politiques de manière à en favoriser l'évaluation**. Les pratiques « sédimentaires », par développement à la marge d'un système, l'absence d'objectifs précis et quantifiés, le caractère implicite des hypothèses d'environnement, peuvent engendrer de graves difficultés pour la conduite et l'évaluation des politiques concernées. Il serait donc fort utile, *a contrario*, de rechercher une organisation des politiques par programme, une spécification des objectifs opérationnels, un affichage des hypothèses d'environnement.

2) Il existe une **forte complémentarité entre les études d'aide à la décision, préalables à la formulation et à l'exécution de la politique, et d'évaluation des résultats ex post**. Ces études sont loin d'être systématiques. L'absence d'une connaissance satisfaisante des caractéristiques des groupes concernés par la politique, des mécanismes de mise en œuvre, des effets attendus, est fortement préjudiciable à la validité d'une évaluation *a posteriori*. L'argument selon lequel l'étude *a priori* ralentit l'action ne doit pas être surestimé. Les cas où les politiques doivent être appliquées en véritable urgence pour des raisons techniques sont en fait minoritaires.

3) **L'évaluation se déroulera de manière d'autant plus efficace que son organisation est prévue très tôt, dès la formulation de la politique**. Cette observation concerne, tout d'abord, le dispositif de collecte des informations nécessaires à l'évaluation, soit par recours à des sources existantes (comptabilité, systèmes statistiques), soit par création d'instruments *ad hoc* (enquêtes). Dans ce domaine l'improvisation *ex post*, si elle peut être nécessaire, est souvent coûteuse d'un point de vue financier, mais aussi pour la validité des résultats. Il convient par ailleurs de prévoir à l'origine les moyens de l'évaluation, financiers et humains. Cette pratique est courante aux États-Unis pour tout nouveau programme.

4) L'évaluation d'une politique publique doit faire l'objet d'un **engagement**. Ceci implique que l'on définisse des **échéances d'évaluation**, appelant éventuellement une **procédure d'extinction de la politique**. Une telle discipline limiterait les risques liés à une pratique excessivement discrétionnaire de l'évaluation. Elle constituerait un frein à l'empilement, souvent dénoncé dans ce rapport, de mesures prises successivement par les pouvoirs publics sur un même domaine et dont la cohérence et l'efficacité d'ensemble peuvent être mises en doute.

5) Le recours à l'**expérimentation** constitue une voie privilégiée pour l'évaluation des politiques publiques, notamment les politiques incitatives et réglementaires. Sa pratique peut, dans certaines situations, soulever des difficultés juridiques liées

notamment à la notion d'égalité devant le service public. Il conviendrait de reconnaître précisément ces difficultés et les cas où elles sont réellement insurmontables. En outre, la décentralisation des pouvoirs peut être un facteur favorable au développement de l'expérimentation, en facilitant la mise en place, dans le cadre réglementaire et législatif existant, de dispositifs différenciés suivant les régions. Le cas échéant, la réalisation d'expérimentations pourrait justifier un aménagement *ad hoc* de ce cadre.

6) De ce point de vue, l'évaluation des politiques publiques doit prendre explicitement en compte le fait de la **décentralisation** et le partage des responsabilités entre pouvoirs centraux et locaux qui en résulte. Toutefois, si la décentralisation peut être une chance pour l'évaluation en multipliant les cas de figure, elle comporte le risque d'une segmentation, avec un développement de sources d'information et d'outils méthodologiques locaux inutilement spécifiques, ce qui réduirait la comparaison et la généralisation des résultats obtenus. **Une réflexion sur l'évaluation des politiques menées par les autorités décentralisées est donc nécessaire.** Une première recommandation peut être formulée, qui consiste à rechercher un accord entre les différents acteurs, nationaux et locaux, sur les méthodes, tout en favorisant le pluralisme des initiatives.

VI - 2 Organiser les compétences

La partie proprement technique de l'évaluation appelle des compétences spécifiques. Des efforts sont nécessaires pour les développer en France. Ils doivent prendre en compte l'existant et éviter dispersion et double emploi.

1) Le développement ou la mise en place des **formations** nécessaires, initiales ou permanentes, devrait en priorité **s'appuyer sur les organismes compétents existants**, sans nécessairement chercher à créer des structures *ad hoc* dont la pérennité ne serait plus assurée en régime de croisière. Il n'est pas dans l'objet du rapport de dresser une liste exhaustive de ces points d'appui dont on se bornera à mentionner quelques uns, tels que l'École nationale de la statistique et de l'administration économique, le Centre d'enseignement de la statistique appliquée à la médecine et à la biologie générale, l'Institut de management public, le Centre HEC-ISA. C'est par leur intermédiaire que devraient se diffuser les techniques insuffisamment utilisées en France, en matière d'expérimentation, d'analyse statistique ou de modélisation microéconométrique.

2) La constitution de **pôles d'étude de dimension critique, susceptibles de capitaliser compétences et expériences**, notamment à l'extérieur de l'administration, est une priorité. Elle suppose une mobilisation des instruments existants à la disposition des pouvoirs publics (financement de la recherche, contrats, appels d'offres) en évitant toute dispersion. On ne peut que déplorer le manque de rigueur

méthodologique qui caractérise souvent les appels d'offres de recherche, par exemple en matière de politiques sociales, et qui conduisent à des programmes fourre-tout. De même, l'attribution des moyens financiers participe trop souvent de préoccupations distributives. Le développement de l'évaluation ne requiert pas nécessairement de dégager des moyens nouveaux, si l'on veut bien mobiliser de manière efficace l'existant. Dans le même ordre d'idées, la puissance publique devrait inciter les centres de recherche publics et parapublics, ainsi évidemment que ses propres services d'études et de statistiques, à rééquilibrer leurs activités en faveur des tâches d'évaluation.

3) Enfin, les procédures de collaboration entre prescripteurs administratifs et évaluateurs, externes ou internes, doivent être améliorées. Trop souvent, les objectifs poursuivis par le prescripteur administratif sont insuffisamment précisés. De même, les conditions d'échange d'informations sont mal définies et une faible attention est portée à la lisibilité des travaux. Il en résulte une insatisfaction mutuelle et la tentation pour chaque partie de camper sur son propre territoire, celui de l'opérationnel ou celui de la connaissance. L'établissement de « cahiers des charges » précis, entre prescripteurs et évaluateurs, permettant notamment un suivi attentif des travaux, permettrait de pallier ces difficultés.

VI — 3 Renforcer l'évaluation des politiques publiques au sein de l'administration

Un développement de l'évaluation purement externe à l'administration aurait toute chance de n'avoir qu'un impact assez faible sur l'action de la puissance publique, directement responsable de la conception et de l'exécution des politiques et détentrice d'avantages d'information importants. Un progrès équilibré de l'évaluation dans notre pays, porteur de changements effectifs, suppose donc qu'un effort particulier soit fait au sein de l'administration elle-même. Il doit s'appuyer sur plusieurs principes.

1) Le premier concerne l'initiative de l'évaluation des politiques publiques réalisée par l'administration. La prise en compte du principe de subordination hiérarchique des services administratifs à l'égard de l'autorité ministérielle quant à l'initiative de décision, conduit naturellement à affirmer également la primauté de celle-ci en matière de prescription d'évaluation et à rejeter l'idée d'une seule auto-évaluation administrative par auto-saisine des services eux-mêmes. Pour éviter toutefois des pratiques excessivement discrétionnaires, il conviendrait de privilégier des procédures d'engagement, conduisant à une réalisation irréversible une fois le principe de l'évaluation retenu, et d'améliorer les conditions de diffusion à l'intérieur comme à l'extérieur de l'administration.

2) Les évaluations réalisées au sein de l'administration doivent se conformer aux deux règles déontologiques de **reproductibilité** des protocoles et de **séparation** entre les responsabilités d'évaluation et d'exécution.

La première règle a pour objet de rendre les évaluations vérifiables, en garantissant des conditions suffisantes de transparence. Ces conditions supposent un affichage des protocoles suivis comme une maîtrise des techniques utilisées. Pour atteindre ce dernier objectif, il sera le plus souvent nécessaire d'associer diverses compétences, par appel aux spécialistes concernés de l'administration et à des experts extérieurs, étant entendu que les institutions exerçant actuellement des responsabilités de contrôle (Cour des comptes, Inspection générale des finances, Inspections ministérielles) devraient jouer un rôle particulièrement moteur dans cette dynamique.

La règle de séparation constitue également un impératif, même si son application prend en compte les conditions particulières propres à chaque évaluation comme cela a été souligné dans le rapport. Une attitude *a priori* trop favorable à l'auto-évaluation est susceptible de conduire à une indépendance insuffisante du responsable de l'évaluation par rapport aux services opérationnels chargés de la mise en œuvre de la politique. Il faut réaffirmer qu'une fois la mission confiée à l'évaluateur, celui-ci doit être indépendant dans le choix de ses méthodes et la formulation de son jugement sur les résultats de la politique.

3) Rechercher des règles systématiques sur la diffusion des résultats de l'évaluation serait illusoire. On peut par contre recommander des **schémas gradués** de diffusion assurant au minimum, en cas d'obligation de confidentialité sur l'information de synthèse ou les conclusions, une possibilité d'accès à l'information de base collectée, en s'inspirant de règles suivies en matière d'information statistique ou de déontologie des sondages.

4) Enfin, il paraît important d'assurer un **caractère interministériel** au développement de l'évaluation au sein de l'administration. Les modalités retenues à cette fin peuvent être variables. **Le groupe a reconnu l'intérêt d'un rattachement particulier du développement de l'évaluation au sein de l'administration auprès du Premier ministre, éventuellement par le biais d'une instance spécifique.** Quelle que soit la solution institutionnelle retenue, placer l'évaluation à un niveau interministériel est une réponse, en termes techniques et de légitimité, au caractère souvent global des politiques et de leurs résultats. Le dépassement des seuls champs ministériels est en outre indispensable pour :

- améliorer la qualité des travaux en soumettant ceux-ci à une plus large confrontation et expertise ;
- favoriser la diffusion de leurs résultats au sein de l'administration et éviter la rétention d'informations ;
- permettre une capitalisation des expériences ;
- favoriser l'intégration des résultats des évaluations dans les procédures d'arbitrage interministérielles, notamment de caractère budgétaire.

VI — 4 Accroître la capacité d'évaluation externe à l'administration et favoriser sa diffusion dans la société

Le développement de l'évaluation au sein de l'administration est un point de passage obligé pour assurer sa diffusion dans la société. Les modalités évoquées précédemment participent donc de ce dernier objectif. Elles doivent être explicitement finalisées en ce sens et complétées par des dispositions spécifiques.

1) **Le renforcement des pôles de compétences externes à l'administration comme une publicité accrue des évaluations faites au sein de l'administration sont évidemment des moyens importants pour fournir des outils et des méthodes aux évaluateurs externes.** Un parallèle peut être fait avec la politique menée depuis le rapport Lenoir en matière de diffusion de l'expertise économique. La constitution volontariste de centres extérieurs à l'administration (OFCE, IPECODE) a effectivement permis un transfert de technologie et autorisé une remise en cause du monopole administratif traditionnel dans ce domaine. Des voies similaires peuvent être suivies en matière d'évaluation, à cette différence près que la capacité d'expertise de l'administration en matière d'évaluation se situe probablement à l'heure actuelle en deçà de celle-ci, qui existait en matière d'analyse macroéconomique lorsque ces centres externes ont été constitués.

2) **Le potentiel d'évaluation paraît actuellement très déséquilibré au préjudice du législatif, contrairement à la situation d'autres pays.** Trois leviers sont susceptibles d'être utilisés pour corriger cette carence :

— développer les moyens propres à la disposition de l'Assemblée nationale et du Sénat en termes de personnels, de mode de traitement de l'information, de crédits, ce qui permettrait en particulier l'appel à des experts extérieurs ;

— permettre un recours du législatif à des prestations fournies par les services administratifs à l'instar de ce qui existe en matière d'analyse macroéconomique (prestations de l'INSEE et de la Direction de la prévision pour le Sénat et le Conseil économique et social) ;

— rendre plus systématiques les procédures prévues de saisine de la Cour des comptes par l'Assemblée nationale et le Sénat tout en assurant mieux les fonctions de conseil que la Cour peut remplir auprès du gouvernement.

3) **Le Commissariat général du plan pourrait assurer une mission d'animation et d'information permanente dans le domaine de l'évaluation, en raison de sa vocation interministérielle et de sa position charnière vis-à-vis des acteurs économiques et sociaux.** Il peut tout d'abord remplir des tâches de synthèse et de diffusion méthodologique à l'instar du rôle qu'il tient en matière de calcul économique (production de circulaires de méthodes, comparaisons internationales, etc.). Il peut être un endroit privilégié pour réaliser une confrontation dans le cas de multiévaluations. Par ailleurs, le Commissariat général du plan pourrait se voir, assez naturellement, confier une responsabilité particulière dans l'évaluation des politiques dépassant le cadre habituel, notamment budgétaire, des procédures de décisions ministérielles. Il s'agit d'une part des politiques de moyen terme, maintenues sur longue période, dont les effets sont cumulatifs et globaux (par

exemple sécurité routière, politiques de l'environnement et de l'espace, politiques de restructuration industrielle, politiques de la concurrence). Le deuxième type de politiques dont l'évaluation pourrait être confiée au Commissariat général du plan correspond aux politiques générales pour lesquelles les conditions d'application locales sont importantes. Ce rôle d'évaluation suppose que ne soit pas privilégiée la recherche d'un consensus, traditionnelle dans les travaux du plan, au détriment de l'exercice sans partage d'une responsabilité de jugement sur les résultats.

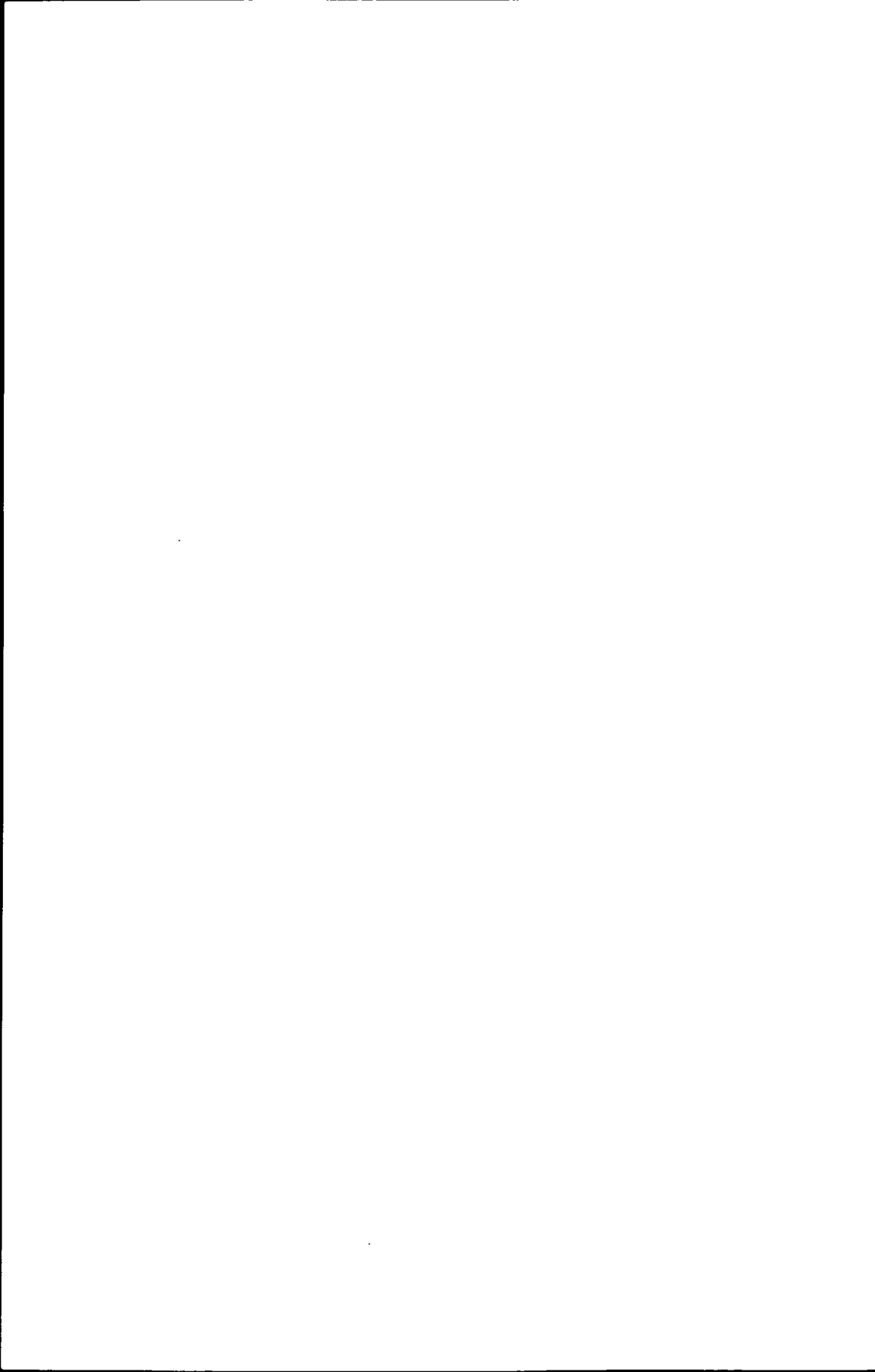
4) Enfin, la diffusion des évaluations par les médias doit être favorisée à partir de pratiques s'inspirant de celles suivies aux États-Unis par le General Accounting Office (conférences de presse, communication de résumés non techniques directement utilisables, etc.).

VI - 5 Identifier les thèmes prioritaires

La relative faiblesse des moyens rapidement utilisables pour réaliser des évaluations solides impose de canaliser l'effort sur quelques thèmes bien choisis importants en termes économiques ou sociaux, et offrant des possibilités à l'évaluateur.

1) L'évaluation doit tout d'abord être utilisée pour sa fonction de remise en cause. Une attention particulière doit être portée aux « politiques silencieuses » qui n'ont pas d'implication budgétaire directe et tombent plus facilement dans l'oubli sans cesser d'agir. Cette observation concerne prioritairement les politiques réglementaires. Dans le même ordre d'idées, il convient de s'intéresser aux « politiques enracinées » faisant l'objet, pour ainsi dire, d'une perpétuelle reconduction tacite. Sans aller jusqu'à prôner des pratiques de révision systématique des crédits votés, il convient de s'intéresser aux politiques résultant d'un processus d'empilement long, coûteuses en termes budgétaires. Cette observation concerne tout particulièrement les diverses aides aux entreprises, agricoles comme industrielles, dont l'efficacité sur l'orientation des choix est loin d'être assurée, ainsi que de nombreux transferts sociaux, dont, au contraire, les apparentes vertus redistributrices sont susceptibles de dissimuler des effets incitatifs néfastes.

2) Par ailleurs, au delà de cette problématique d'efficacité et de réduction de coûts implicites ou explicites, l'évaluation doit porter sur les politiques apparaissant particulièrement déterminantes pour le développement sur moyenne période, de l'économie. Tel est le cas en particulier de l'enseignement, de la recherche et de la santé.



Glossaire



Glossaire¹

Analyse des données :

Ensemble des techniques statistiques permettant de traiter de l'information brute (généralement issue d'enquêtes statistiques ou de fichiers de gestion) afin d'en tirer les enseignements utiles.

Analyse multivariée :

Ensemble de techniques statistiques reposant sur des relations linéaires et permettant de juger de la présence ou de l'absence d'un effet.

Dans le cas de l'analyse de la variance, les variables explicatives sont des variables artificielles prenant les valeurs 0 ou 1, qui indiquent la présence ou l'absence d'un effet sur une variable.

Dans le cas de l'analyse de la covariance, les variables explicatives sont d'une part des variables numériques, d'autre part des variables 0 et 1.

Données en coupes transversales :

Observations relatives à une même unité de temps (mois, trimestre, année) et collectées sur des sujets ou des événements différents.

Données en séries temporelles :

Observations collectées sur un nombre plus ou moins grand d'unités de temps successives, mais concernant constamment le même sujet (ou le même ensemble de sujets).

Données sur panel :

Mélange d'observations en coupes transversales et en séries temporelles (on repère l'évolution au cours du temps des caractéristiques de chacun des sujets observés).

Econométrie :

Ensemble des techniques statistiques permettant d'estimer numériquement les paramètres d'une relation analytique entre une variable expliquée et des variables explicatives, cette relation étant postulée en général à partir d'une théorie *a priori* (notamment économique).

Etude de cas :

Description analytique d'un événement, d'un organisme, d'un processus ou d'un programme.

Evaluateur :

Personne(s), groupe de personnes ou d'organismes ayant la charge de réaliser une évaluation.

1. Ce glossaire constitue une aide à la lecture du rapport en recensant les principaux termes techniques utilisés. Les définitions données se veulent pédagogiques plus que totalement rigoureuses.

Evaluation externe :

Cas où l'évaluateur est indépendant du (ou des) organisme(s) qui a (ont) décidé ou mis en oeuvre la politique.

Evaluateur interne :

Cas où l'évaluateur dépend du (ou des) organisme(s) qui a (ont) décidé ou mis en oeuvre la politique.

Expérimentation vraie :

Attribution de traitements (ou de politiques) différents(es) à des groupes représentatifs de la politique étudiée et équivalents (ce qui implique qu'ils résultent d'un tirage au sort).

Groupes équivalents :

Groupes (d'individus, d'entreprises, etc.) ayant la même distribution de caractéristiques.

Groupes non équivalents :

Groupes (d'individus, d'entreprises, etc.) dont les distributions de caractéristiques sont différentes.

Groupes non équivalents issus de l'observation :

Groupes dont la constitution et le traitement ne résultent pas de la volonté de l'évaluateur, mais des conditions effectives d'application de la politique à évaluer.

Groupes non équivalents issus d'une quasi-expérimentation :

Groupes dont le traitement résulte de la volonté de l'évaluateur mais qui, n'étant pas tirés au sort, ne sont pas strictement équivalents.

Groupe témoin :

Groupe (d'individus, d'entreprises, etc.) n'ayant pas subi la politique que l'on cherche à évaluer.

Groupe de traitement :

Groupe (d'individus, d'entreprises, etc.) auquel l'on applique la politique que l'on veut évaluer.

Mesure :

Procédure consistant à assigner un nombre à une caractéristique donnée d'un objet ou d'un événement.

Modélisation :

Construction d'une représentation schématique formalisée de comportements et de mécanismes à partir d'une théorie *a priori* explicative, et susceptible d'estimation numérique.

Plan de recherche (ou protocole) :

Organisation méthodologique générale de l'évaluation en fonction de l'information empirique (observations) disponible ou à collecter.

Prescripteur (d'une évaluation) :

Personne(s), groupe de personnes ou d'organismes demandant que la politique publique soit évaluée, et en confiant la mission à l'évaluateur.

Quasi-expérimentation :

Constitution, par l'évaluateur, de groupes de comparaison par un procédé autre que le tirage au sort, mais qui vise à limiter autant que possible les différences de caractéristiques entre ces groupes.

Typologie :

Regroupement d'observations en classes possédant des caractéristiques aussi homogènes que possible.

Technique d'analyse :

Outil particulier de traitement d'une information ou d'un ensemble d'informations.

Validité conceptuelle :

Légitimité de généraliser les conclusions de l'évaluation à d'autres variantes de la même politique.

Validité externe :

Légitimité de généraliser les conclusions de l'évaluation à la même politique visant d'autres cibles ou opérant dans un autre contexte.

Validité interne :

D'une part, possibilité de déceler des corrélations entre la politique évaluée et certains phénomènes. D'autre part, légitimité de considérer ces phénomènes comme des effets de la politique.

Variance :

Mesure de la dispersion des caractéristiques d'une population.

Créer de la variance : multiplier les observations d'unités suffisamment dissemblables au sein d'une population et mesurer leurs caractéristiques.



Annexes



Annexes

- 1 — Composition du groupe de travail au 30/12/85
- 2 — Communications présentées au groupe

Composition du groupe de travail « Méthodes d'évaluation des politiques publiques » au 30/12/85

Président

M. Michel Deleau Sous-directeur à la Direction de la prévision. Ministère de l'Économie, des Finances et du Budget

Secrétariat du groupe :

M. Philippe Penz Chargé de mission au Service économique du Commissariat général du plan

Membres de la cellule d'animation

M. Jean-Pierre Nioche Professeur à l'École des hautes études commerciales
M. Robert Poinsard Secrétaire de la Commission RCB - Direction de la prévision - Ministère de l'Économie, des Finances et du Budget

Membres permanents :

M. Thierry Bert (*) Inspecteur des finances, chargé de mission auprès du chef de l'Inspection des finances

M. Bernard Bobe Conseiller scientifique au Commissariat général du plan

Mme Aude-Marie Debarle Chargée de mission au Service des études et de la recherche du Commissariat général du plan

M. Bernard Dechambre Chargé de mission à la Mission d'analyse et d'évaluation - Direction des Affaires économiques et financières - Ministère de l'Agriculture

M. Bernard Devev Directeur à la COFREMCA

Mme Elvyne Février Chargée de mission auprès du Sous-directeur de la sécurité et de l'exploitation de la route - Direction de la sécurité et de la circulation routière - Ministère de l'Urbanisme, du Logement et des Transports

(*) M. Thierry Bert a succédé à M. Laurent Perpère, inspecteur des Finances.

M. Thierry Gaudin	Directeur du Centre de prospective et d'évaluation - Ministère de la Recherche et de l'Industrie
M. Daniel Goudard	Chargé de mission au Commissariat général du plan - Adjoint au chef du Service économique
M. Pierre Greiner (*)	Chef de la Mission d'analyse et d'évaluation - Direction des affaires économiques et financières - Ministère de l'Agriculture
M. Christian Join- Lambert(**)	Conseiller référendaire à la Cour des comptes
M. Pierre Mondon	Chef du Service de la prévision, des statistiques et de l'évaluation (SPRESE) - Ministère de l'Éducation Nationale
M. Henry Ollagnon	Chargé de mission à la Mission d'analyse et d'évaluation - Direction des affaires économiques et financières - Ministère de l'Agriculture
Mme Elisabeth Pascaud (***)	Responsable d'Études au Centre de recherche pour l'étude et l'observation des conditions de vie (CREDOC)
Mme Brigitte Roguet	Responsable du secrétariat des comptes de l'emploi - Service des études et de la statistique - Ministère du Travail, de l'Emploi et de la Formation professionnelle
M. Daniel Schwartz	Directeur de l'unité de recherches statistiques de l'Institut national de la santé et de la recherche médicale (INSERM)
M. Olivier Villey	Chef de la division de l'emploi et des politiques de l'emploi - Service des études et de la statistique - Ministère du Travail, de l'Emploi et de la Formation professionnelle
M. Michel Volle	Chef de la Mission d'études économiques - Centre national d'études des télécommunications

(*) M. Pierre Greiner a succédé à M. Dominique Gagey.

(**) M. Christian Join-Lambert a succédé à M. Jean-Marie Bertrand, conseiller référendaire à la Cour des comptes.

(***) Mme Elisabeth Pascaud a succédé à Mme Catherine Blum, directeur du CREDOC.

Communications présentées au groupe :

Les méthodes de l'évaluation (31/01/1985)

- Repérage de l'évaluation par rapport à des genres voisins. Aperçu des problèmes méthodologiques de l'évaluation (M. Nioche, HEC).
- Les techniques d'analyse causale et de mesure des effets globaux (M. Volle, Centre national d'études sur les télécommunications).
- Évaluation et expérimentation : le cas de l'épidémiologie (M. Schwartz, INSERM).
- Évaluation et calcul économique (M. Henry, Laboratoire d'économétrie de l'École polytechnique).

Les méthodes d'évaluation en matière de politique d'éducation (14/02/1985)

- L'observatoire des entrées dans la vie active (M. Germe, Centre d'études et de recherches sur l'emploi et les qualifications).
- L'Évaluation en matière pédagogique au ministère de l'Éducation nationale (M. Tournardre, Service de la prévision, des statistiques et de l'évaluation, ministère de l'Éducation nationale).
- Une évaluation internationale de l'enseignement des mathématiques (M. Robin, Institut national de recherche pédagogique, ministère de l'Éducation nationale).

Les méthodes d'évaluation en matière de politiques sociales (07/03/1985)

- La problématique générale de l'évaluation en matière de politiques sociales (Mme Catherine Blum, CREDOC).
- Mères isolées, allocation de parent isolé, et travail. Essais de mesure de l'efficacité économique et sociale de l'API (M. Ray, université de Nancy II, LASARE).

Les méthodes d'évaluation en matière de politiques de transports et d'environnement (21/03/1985)

- La problématique générale de l'évaluation en matière de transports (M. Gabet, Direction des transports terrestres, ministère de l'Urbanisme, du Logement et des Transports).
- L'évaluation dans le domaine de la sécurité routière (Mme Février, Direction de la sécurité et de la circulation routières, ministère de l'Urbanisme, du Logement et des Transports).
- L'inventaire permanent du littoral, outil d'évaluation en matière de politique d'aménagement (M. Ecobichon, Direction de l'urbanisme et des paysages, ministère de l'Urbanisme, du Logement et des Transports).

L'évaluation du plan textile (30/04/1985)

- Exposé de M. Tuloup (Direction des industries chimiques, textiles et diverses, ministère du Redéploiement industriel et du Commerce extérieur).

- Exposé de MM. Boaretto et du Buisson (Direction de la prévision et Direction du budget, ministère de l'Économie, des Finances et du Budget).
- Les méthodes utilisées pour l'évaluation du plan textile (M. Nioche, HEC).

Les méthodes de l'évaluation en matière de politique de santé (04/06/1985)

- La problématique générale de l'évaluation en matière de politique de santé (Mme Weill, Direction des hôpitaux, ministère des Affaires sociales et de la Santé).
- L'évaluation des innovations technologiques (Mme Claudine Blum, Direction du plan de l'Assistance publique de Paris).

Les méthodes d'évaluation des politiques de la recherche (18/09/1985)

- L'évaluation des Directions régionales de l'industrie et de la recherche (M. Gaudin, Centre de prospective et d'évaluation, ministère de la Recherche et de l'Industrie).
- Le Comité national d'évaluation des universités (M. Staropoli, Comité national d'évaluation des universités, ministère de l'Éducation nationale).

L'évaluation par la société civile (03/10/1985)

- Le rôle des médias dans le déclenchement et dans la publicité de l'évaluation (M. de La Taille, TF1).
- Évaluation par les citoyens : apport des systèmes d'observation et de mesure des sensibilités et des mentalités du public (M. Deveze, Cofremca).

Les méthodes d'évaluation des politiques d'emploi (17/10/1985)

- Les méthodes d'évaluation en matière de politiques d'emploi (Mme Elbaum, délégation à l'Emploi).
- La mise en place de l'évaluation des travaux d'utilité collective (M. Villey, Direction des études statistiques, ministère du Travail, de l'Emploi et de la Formation professionnelle).
- Les méthodes d'évaluation utilisées dans l'étude prioritaire RCB sur les aides à l'emploi (M. Hara, Direction de la prévision, ministère de l'Économie, des Finances et du Budget).
- L'apport des enquêtes auprès des entreprises dans l'évaluation des aides à l'emploi (M. Eymard-Duvernet, directeur du Centre d'études de l'emploi).
- Les modèles micro-analytiques de fonctionnement du marché du travail et leur utilité pour l'évaluation des politiques d'emploi (M. Ballot, université Paris II).

Les méthodes d'évaluation des politiques agricoles (24/10/1985)

- La pratique de l'évaluation au ministère de l'Agriculture (M. Gagey, Mission d'analyse et d'évaluation, ministère de l'Agriculture).
- Présentation méthodologique de l'étude d'évaluation de la politique publique de l'élevage bovin (M. Le Pen, Legos, université Paris IX).

- L'approche « Système-acteurs » (M. Ollagnon, Mission d'analyse et d'évaluation, ministère de l'Agriculture).

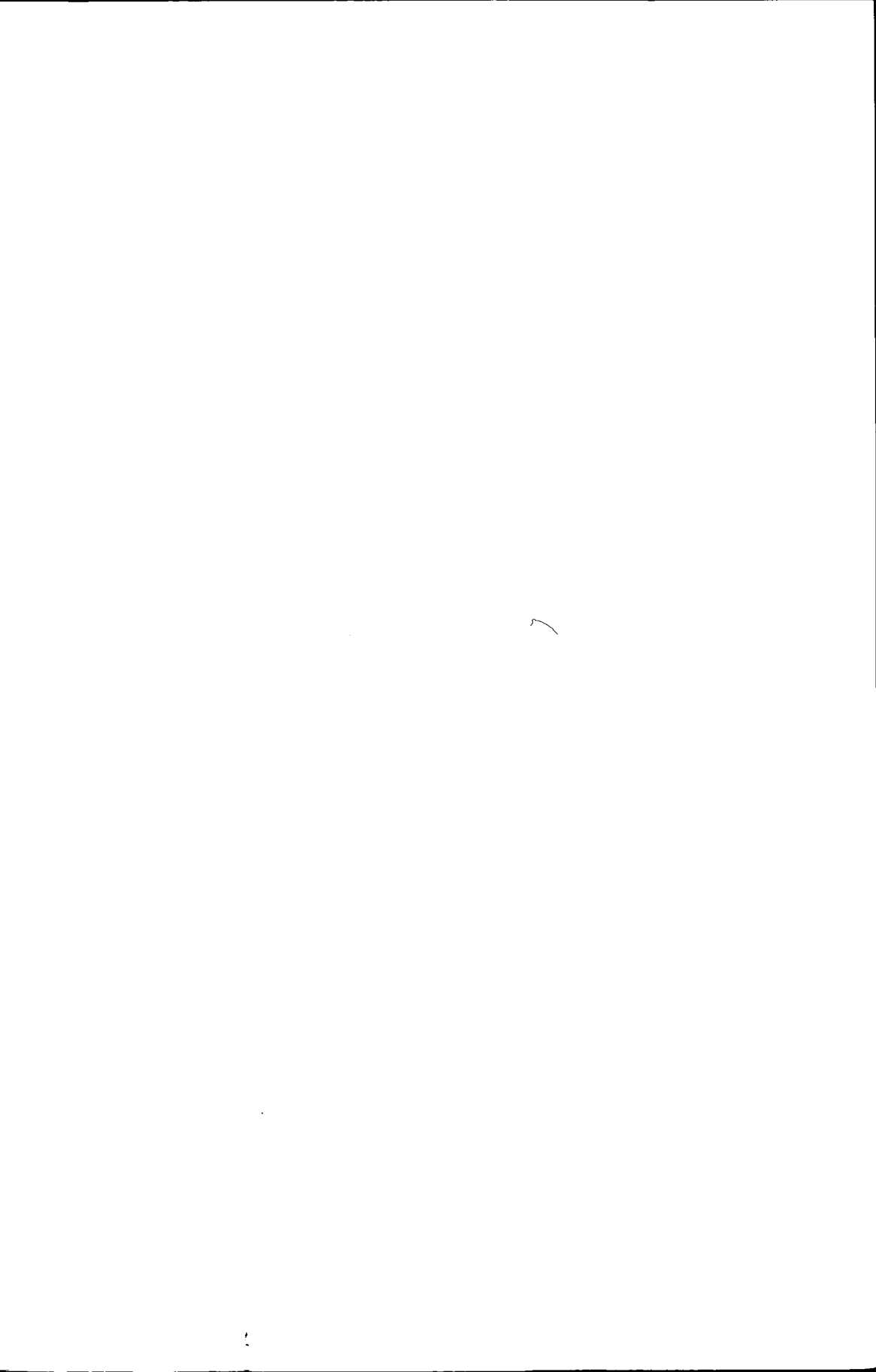
Contrôle et évaluation (07/11/1985)

- Les problèmes d'organisation générale des rapports entre contrôle et évaluation (M. Joint-Lambert, Cour des comptes).
- Le rôle de l'Inspection des finances (M. Bert, Inspection des finances).
- Le rôle des inspections générales ministérielles (M. Milloz, Inspection générale de l'industrie).
- Le rôle de la Cour des comptes (M. Logerot, secrétaire général de la Cour des comptes).
- Les corps de contrôle et l'évaluation (M. Gibert, CESMAP).

Les expériences étrangères en matière d'évaluation (11/12/1985)

- L'évaluation des politiques publiques aux Pays-Bas (M. Sorber, ministère des Finances des Pays-Bas).
- L'évaluation des politiques publiques aux États-Unis (M. Rouban, Centre de recherches administratives et politiques de l'université de Rennes I).



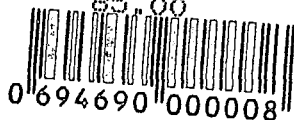




EVALUER POLIT PUBLI

85,00

504



LA DOCUMENTATION FRANÇAISE
29-31, quai Voltaire
75340 Paris Cedex 07
Tél. : (1) 42.61.50.10
Télex : 204826 Docfran Paris

Imprimé en France
ISBN : 2-11-001702-3
DF 1345