

Werner Bussmann

**Accompagner et mettre à profit avec succès
les évaluations des mesures étatiques:**

Guide de réflexion

Traduit de l'allemand
par Françoise Nagy

Paru chez Georg Éditeur, 1995
(la mise en page du présent document électronique ne correspond pas
avec la mise en page de la version imprimée)

Table des matières

| | |
|--|-----------|
| Avant-propos | 4 |
| 0. Aperçu | 5 |
| 1. Six étapes pour commander, accompagner et mettre à profit avec succès un projet d'évaluation | 6 |
| 1.1 Distinguer le problème | 6 |
| 1.2 Cerner l'environnement et les questions | 7 |
| 1.21 Définir la question de recherche | 7 |
| 1.22 Déterminer approximativement le groupe des parties prenantes..... | 10 |
| 1.23 Evaluation interne ou externe? | 10 |
| <i>Evaluation interne</i> | 10 |
| <i>Evaluation externe</i> | 11 |
| 1.24 Déterminer l'objectif d'utilisation de l'évaluation | 12 |
| 1.25 Fixer le cadre temporel et estimer les moyens nécessaires, en finances et en personnel | 12 |
| 1.26 Décider de la poursuite de l'évaluation..... | 13 |
| 1.3 Demander des offres et adjuger le mandat | 13 |
| 1.31 Faut-il donner mandat d'une étude d'évaluabilité? | 13 |
| 1.32 Préparer la présentation de l'offre | 14 |
| 1.33 Examen des offres sous l'angle de leurs lacunes | 14 |
| 1.34 Choisir | 15 |
| 1.35 Adjudication du mandat en tant que négociation | 18 |
| 1.4 Accompagner utilement l'évaluation | 19 |
| 1.41 Qui accompagne l'évaluation? | 19 |
| 1.42 Principes de l'accompagnement..... | 19 |
| 1.5 Mettre à profit de manière optimale les résultats de l'évaluation..... | 21 |
| 1.51 Rendre compte | 21 |
| 1.52 Améliorer l'exécution | 21 |
| 1.53 Economiser des frais..... | 22 |
| 1.54 Supprimer des mesures | 22 |
| 1.55 Changer ou remanier des mesures | 23 |
| 1.56 Les processus d'apprentissage en tant qu'objectif de l'évaluation..... | 24 |
| 1.57 Digression: Abus possible des évaluations..... | 25 |
| 1.6 Prendre des dispositions en vue de mieux saisir à l'avenir l'efficacité des mesures | 26 |
| 2. Qu'entend-on par évaluations? | 30 |
| 2.1 La notion d'évaluation | 30 |
| 2.2 Que peut-on, en fait, évaluer? | 30 |
| 2.3 Quels aspects et critères les évaluations peuvent-elles contenir? | 31 |
| 2.4 Quel horizon temporel couvrent les évaluations? | 32 |
| 2.5 Que coûtent les évaluations et combien de temps durent-elles? | 33 |
| 3. De quelles stratégies et techniques de recherche les évaluations se servent-elles? | 35 |

| | | |
|-----------------|--|-----------|
| 3.1 | Possibilités de classification des stratégies d'évaluation..... | 35 |
| 3.2 | Notions de stratégie d'évaluation et de technique de recherche..... | 35 |
| 3.3 | Aperçu des stratégies d'évaluation | 36 |
| 3.31 | Questions descriptives, normatives et à orientation causale..... | 36 |
| 3.32 | Genres de questions et stratégies d'évaluation | 37 |
| 3.4 | Stratégies d'évaluation pour les questions descriptives et normatives | 38 |
| 3.5 | Stratégies d'évaluation pour les questions à orientation causale | 40 |
| 3.51 | Stratégies d'évaluation explicitement et implicitement comparatives | 40 |
| 3.52 | Comparaison avec groupe de contrôle | 43 |
| 3.53 | Comparaisons transversales | 44 |
| 3.54 | Comparaison longitudinale | 46 |
| 3.6 | Techniques de recherche | 48 |
| 3.61 | Aperçu/présentation de techniques de recherche choisies | 48 |
| 3.62 | Etudes de cas | 49 |
| 3.63 | Sondages..... | 51 |
| 3.64 | Analyse de données existantes | 52 |
| 3.7 | Digression: Les modèles en tant que point de départ et résultat des évaluations | 53 |
| 3.8 | Mélange de méthodes et triangulation | 56 |
| 3.9 | Indications sur le choix de la stratégie d'évaluation et de la technique de recherche | 57 |
| Annexe 1 | Explication des principales notions..... | 60 |
| Annexe 2 | Glossaire des principales notions en quatre langues..... | 64 |
| Annexe 3 | Bibliographie..... | 65 |

Avant-propos

Le développement considérable des activités de l'Etat dans la période d'après-guerre a fait douter, depuis les années 70, de plus en plus que les mesures étatiques soient correctement exécutées, qu'elles produisent les effets souhaités et qu'elles puissent être justifiées économiquement. Les évaluations sont l'une des réponses données à ces questions brûlantes. Développées aux Etats-Unis depuis les années 60, elles se sont répandues de plus en plus largement en Europe; en Suisse, elles sont engagées toujours plus fréquemment au niveau de la Confédération, des cantons et des communes.

Le présent guide est destiné à faciliter l'approche de l'instrument "évaluations", qui peut être engagé de manière très diversifiée et souple. Il s'adresse aux personnes qui aimeraient commander, accompagner et mettre à profit des évaluations. L'expérience montre que, dans bien des cas, c'est l'administration qui demande des évaluations; c'est pourquoi beaucoup d'exemples seront tirés de ce contexte. Mais les partis, les associations et d'autres groupes auront toujours plus fréquemment recours à cet instrument. A eux aussi, ce guide devrait apporter une aide utile.

Cet ouvrage constitue la première partie du résultat final du Programme national de recherche sur "les effets des mesures étatiques" (PNR 27). Une seconde partie paraîtra dans un an environ sous la forme d'un "manuel sur les méthodes de recherche évaluative"; il sera axé avant tout sur la science et la pratique des évaluations.

L'impulsion à l'élaboration de ce guide m'a été donnée en avril 1993, alors que j'effectuais un séjour de recherche de trois semaines aux Etats-Unis. A cette occasion, j'ai pu tirer profit, en particulier, des contacts noués avec le GAO (General Accounting Office) et de plusieurs conversations que j'ai eues avec Eleanor Chelimsky et Ray Rist. La suite de mon travail s'est fondée sur de nombreux entretiens intervenus dans le cercle du groupe d'experts et des collaborateurs de projets participant au PNR 27. Andreas Balthasar, Frohmut Gerheuser, Leo Meyer et Ruedi Meier, notamment, m'ont communiqué des commentaires détaillés et précieux à propos des premiers projets du guide. Madame Madeleine Rieke a donné au texte une forme imprimable et fait preuve, dans cette entreprise, d'une persévérance et d'une patience infinies. Françoise Nagy a mis beaucoup de soin dans la traduction du texte. Heinrich Koller, directeur de l'Office fédéral de la justice, a mis à disposition son infrastructure de bureau. Ernst Buschor, président du groupe d'experts du PNR 27, m'a donné beaucoup de confiance et de liberté dans l'élaboration des rapports finaux. Je tiens à remercier ici chaleureusement toutes les personnes mentionnées. Je dois en outre des remerciements particuliers à ma femme et à mes deux enfants, dont la vitalité m'a aidé à surmonter sans dommage ma charge de travail parfois lourde.

Berne, octobre 1994

0. Aperçu

Ce guide s'adresse aux personnes qui souhaitent commander, accompagner et mettre à profit une évaluation, ou procéder à des éclaircissements à ce propos. Ce n'est pas un guide d'auto-évaluation, mais il peut donner des idées sur les aspects à considérer lors de l'appréciation propre de l'exécution ou de l'efficacité.

L'ouvrage se divise en trois parties. La première traite du déroulement d'une évaluation - du point de vue du mandant -, de la formulation du problème à l'exploitation du résultat. La seconde expose ce que sont les évaluations, et la troisième a pour thème les stratégies et les techniques de recherche engagées pour obtenir des résultats significatifs.

La première partie explique en six étapes la procédure concrète d'adjudication, d'accompagnement et de mise à profit des évaluations. Le chapitre 1.1 énumère, en donnant des exemples, les déclencheurs possibles d'évaluations. Le chapitre 1.2 vous montre quels points vous devriez prendre en considération quand vous déterminez le cadre général de l'évaluation. Le chapitre 1.3 vous donne des indications sur la demande d'offres, le choix à faire entre celles-ci, et la conclusion du contrat. Le chapitre 1.4 donne des informations sur l'accompagnement de l'évaluation et le temps qu'il faut y consacrer. Le chapitre 1.5 vous montre à quelles fins vous pouvez mettre à profit les résultats d'une évaluation. Les évaluations peuvent avant tout contribuer à améliorer l'exécution, à étudier et à représenter de manière fondée les effets des mesures en discussion, ainsi qu'à prendre en mains la modification ou la nouvelle conception d'une mesure. Le chapitre 1.6 vous donne des indications sur la manière dont vous pouvez vous assurer, lors du remaniement d'une mesure, que son efficacité pourra être par la suite contrôlée et attestée.

Les deuxième et troisième parties résument les connaissances élémentaires qui vous seront utiles en matière d'évaluations. La deuxième partie se penche sur les principaux genres d'évaluations, les critères qui leur servent de base, et mentionne également le temps nécessaire et les coûts des évaluations. La troisième partie examine les stratégies et méthodes de recherche les plus fréquemment utilisées dans la recherche évaluative. On insiste particulièrement sur les stratégies d'évaluation visant à distinguer les effets.

L'annexe 1 définit brièvement les principales notions utilisées, l'annexe 2 contient un glossaire des termes importants en quatre langues, l'annexe 3 donne des indications bibliographiques et l'annexe 4 propose un index des matières.

Quelques remarques encore sur ce que vous ne trouverez pas dans les chapitres ci-après. Il ne s'agit pas d'un cours de gestion administrative réussie. Le thème serait trop large. Nous nous concentrerons sur la question de savoir comment on peut se procurer de meilleures informations sur l'exécution et les effets. Nous savons que ces efforts pour distinguer l'efficacité et l'efficacité des mesures étatiques sont parfois vus d'un mauvais oeil, mais nous avons la conviction qu'ils sont dans l'intérêt des citoyennes et des citoyens et qu'à long terme ils seront reconnus comme tels. Dans les pages qui suivent, nous nous limiterons aussi aux évaluations de mesures étatiques fondées sur les sciences sociales. C'est là une voie parmi les nombreuses qu'il faudra parcourir. Une administration novatrice et coopérative a de très nombreuses autres possibilités, chacune avec ses avantages et ses inconvénients, d'obtenir des informations sur l'efficacité des mesures étatiques: contacts administratifs, réactions de citoyennes et de citoyens, contacts avec des associations, interventions politiques, communications dans les journaux, indicateurs de performance, statistiques officielles, etc. Ces possibilités ne seront pas traitées ici. Nous aimerions nous concentrer sur l'instrument "évaluations", de manière à vous donner à son sujet des informations significatives. Toutefois, nous tenons à insister sur le fait que les évaluations doivent compléter et non remplacer les autres mécanismes existants de contrôle des résultats, et qu'elles peuvent contribuer à ce que l'action étatique réponde encore mieux que jusqu'ici aux besoins des citoyennes et des citoyens.

1. Six étapes pour commander, accompagner et mettre à profit avec succès un projet d'évaluation

1.1 Distinguer le problème

Au début de toute évaluation, il y a un problème dont la solution ou la maîtrise judicieuse est entravée par le manque d'informations essentielles. Il peut s'agir par exemple d'un des problèmes suivants:

Exemple 1: A la Direction des travaux publics d'un petit canton montagnard de Suisse centrale, le Conseil exécutif et les fonctionnaires se demandent quel serait le meilleur moment pour dégager les trois grands cols de la neige. Les frais seraient-ils moindres si le déblaiement de la neige avait lieu plus tard? Mais n'y aurait-il pas alors des effets négatifs sur le tourisme et l'hôtellerie? (situation de départ de l'étude Ecoplan 1993b dans le canton d'Uri)

Exemple 2: L'Office fédéral de l'environnement, des forêts et du paysage s'inquiète de plus en plus de savoir si les prescriptions de protection de la nature et du paysage valables pour les pistes de ski sont effectivement respectées. Il doute de plus en plus que ses directives et instructions soient véritablement suivies dans la pratique et aimerait savoir comment contribuer à une meilleure application de ses prescriptions. (situation de départ de l'étude Hirsch et al. 1995)

Exemple 3: Une grande ville suisse est forcée de réduire ses dépenses également dans le domaine social. Les centres communautaires de la ville, gérés par une organisation d'utilité publique, doivent économiser plus d'un million de francs. Mais où va-t-on économiser? De manière linéaire ou de manière ciblée dans les domaines où le dommage n'est pas négligeable? La directrice du Bureau d'aide sociale souhaite qu'une stratégie d'économies raisonnable soit élaborée en commun avec les responsables des centres communautaires. (situation de départ de l'étude Price Waterhouse 1993)

Exemple 4: Un vice-directeur de l'Office fédéral de la justice aimerait faire examiner l'avant-projet de loi fédérale sur l'égalité entre la femme et l'homme sous l'angle de ses conséquences présumées dans le domaine économique. Quels coûts économiques une interdiction de la discrimination entraînerait-elle? Le bilan pourrait-il même se révéler positif? (situation de départ de l'étude Littmann-Wernli 1991)

Exemple 5: En 1984, le canton du Tessin a fait oeuvre de pionnier en adoptant une loi sur les administrateurs de biens. Dix ans plus tard, il est temps de faire le bilan. La loi a-t-elle fait ses preuves? Faut-il la conserver, la développer ou la supprimer? (situation de départ d'une étude encore en cours en 1994)

Exemple 6: Le service de l'OFIAMT compétent pour la promotion de l'hôtellerie se préoccupe depuis assez longtemps de l'efficacité et de l'efficacité des mesures d'encouragement. On sait combien d'hôtels reçoivent des cautionnements et des prêts. En revanche, on ne sait pas très bien si les hôtels bénéficiant de la promotion s'affirment ensuite véritablement par rapport à la concurrence et si les investissements renforcent durablement leur capacité d'investissement. (situation de départ de l'étude Hanser/Kuster/Cavelti 1994.)

Les évaluations sont un instrument qui peut être engagé de manière souple et diverse. On peut s'en servir pour rendre compte de mesures étatiques (exemples 5 et 6), pour améliorer l'exécution (exemple 2), pour économiser des frais ou améliorer l'efficacité (exemples 1 et 3), pour supprimer (év. exemple 5) ou pour changer des mesures (exemples 2 et 5), ou pour en élaborer de nouvelles (exemple 4). Les évaluations deviennent de plus en plus un instrument indispensable pour rendre l'action étatique plus transparente, plus efficace et plus efficiente. Elles sont utilisées par le gouvernement dans sa fonction de surveillance, par l'administration dans le cadre de sa responsabilité de préparation et d'exécution des mesures étatiques, par le Parlement dans l'exercice de sa mission de haute surveillance, mais aussi par les partis, les associations et les groupes de citoyens.

Pour élaborer des évaluations significatives, on utilise toutes les méthodes scientifiques (sociales) courantes de collecte et d'exploitation des données, ainsi que les procédures de décision et de solution des problèmes. Depuis quelque temps, on accorde en outre une importance accrue à la mise en valeur appropriée des résultats des évaluations.

La principale contribution de la recherche évaluative par rapport à d'autres instruments de collecte d'informations (instruments de gestion d'entreprise, controlling, etc.) réside avant tout dans la mise en évidence des possibilités et des limites de l'identification des effets des mesures étatiques dans le domaine social et économique. Le principal problème à cet égard est celui d'isoler l'influence des mesures étatiques d'autres facteurs d'influence (conjoncture, changement des valeurs, etc.) affectant les rouages complexes de l'évolution sociale.

→ Le premier chapitre de la deuxième partie donne d'autres éclaircissements sur ce que sont les évaluations, ce qu'elles peuvent faire et ce qu'elles coûtent. Mais on trouvera aussi dans les exemples ci-après des indications sur les possibilités d'engagement des évaluations. Le chap. 3.5 donne des détails sur les stratégies d'évaluation pour la détermination des effets des mesures étatiques.

Ainsi, les évaluations ne sont pas un instrument miracle. Elles coûtent cher en argent et en temps (cf. chap. 2.5). C'est pourquoi elles présupposent une planification attentive dans la perspective de ces deux éléments. Elles exigent tant du prescripteur que des destinataires qu'ils comprennent la situation, qu'ils s'engagent vigoureusement, qu'ils soient capables de se concentrer sur l'essentiel, capables aussi d'apprendre et prêts à coopérer. Dans les chapitres ci-après, nous aimerions montrer quelles étapes concrètes constituent la condition préalable à l'exécution et à la mise à profit d'évaluations significatives¹.

1.2 Cerner l'environnement et les questions

La première étape consiste à éclaircir les questions et à cerner de manière plus précise le contexte d'une évaluation possible. En d'autres termes, il faut délimiter le cadre général de la recherche. Concrètement, cela revient à décider

- si l'étude du problème doit être poursuivie,
- qui il faut engager dans la suite des travaux,
- si le travail sur le problème doit se poursuivre dans le cadre interne ou par le biais d'un mandat externe,
- à quoi un éventuel mandat doit servir,
- quel est le cadre temporel,
- quelles ressources financières et humaines doivent être engagées.

1.21 Définir la question de recherche

Si vous songez à confier un mandat d'évaluation (interne ou externe), c'est la plupart du temps dans l'espoir de pouvoir répondre à des questions très vastes, telles par exemple celles-ci:

- Des subventions déterminées produisent-elles des effets durables?
- Quels sont les points faibles de l'exécution de la loi?
- Atteignons-nous, par notre campagne d'information, véritablement ceux qui ont le plus besoin de cette information?

En règle générale, les évaluations ne peuvent donner des réponses fondées qu'à des questions précises. Elles ne sont pas en mesure de répondre à des coûts supportables à tous les besoins d'information. Selon les besoins de connaissance et les questions, on en vient à se concentrer sur un ou plusieurs des axes suivants d'évaluation:

¹ Dans la formulation des chapitres ci-après, j'ai fait des emprunts importants aux documents préparés par Andreas Balthasar pour le cours "Engagement d'experte(s) extérieur(e)s".

1. Evaluation de l'exécution (analyses, contrôles et études d'exécution, analyses de procédure): Une évaluation de l'exécution porte sur les activités de l'administration ainsi que sur le comportement des particuliers qui y est lié. Souvent, les instances d'exécution ont de bonnes connaissances de certains aspects de l'exécution (contacts personnels, conférences spécialisées, données statistiques, etc.). Toutefois, un examen de l'exécution par une équipe d'évaluation peut être utile dans différents cas, p.ex. dans la perspective d'un inventaire complet en même temps que critique, pour l'approfondissement de questions d'exécution importantes (p.ex. fixation de priorités lors de l'engagement de moyens limités) ou pour une comparaison transversale entre organisations d'exécution cantonales ou communales.
Exemple: Une enquête sur l'exécution déjà presque classique est l'étude de Delley et al. (1982) sur l'exécution de la lex Furgler de l'époque. Le résultat étonnant de l'enquête fut que les cantons tirèrent parti de cette loi de manière très novatrice pour leurs besoins respectifs.
2. Evaluation des effets (analyses et contrôles des effets, analyses d'impact): La détermination des effets des mesures étatiques se trouve aujourd'hui au centre de nombreux efforts, car elle représente la condition préalable à une appréciation fondée. La question est la suivante: Qu'est-ce qui serait différent si la mesure n'était pas exécutée? Il s'agit donc d'isoler l'influence exercée par la mesure sur la complexité des événements sociaux et économiques. Il ne faut pas oublier que la fiabilité des évaluations d'effets dépend en partie de la possibilité d'établir des comparaisons judicieuses (comparaisons avec groupe de contrôle, comparaisons transversales et longitudinales, cf. chap. 3.5). En outre, il est nécessaire de se concentrer sur certains aspects des effets pour obtenir des résultats significatifs.
Exemple: L'étude Hanser/Kuster/Cavelti 1994 cherchait à savoir quels effets la promotion de l'hôtellerie avait sur les activités d'investissement et le résultat d'exploitation des hôtels qui en bénéficiaient. Les collaborateurs du projet reconstruisirent dans chaque cas la décision d'investissement, analysèrent les données relatives à la situation des hôtels avant et après l'investissement et firent aussi des comparaisons avec des hôtels qui finançaient leurs investissements sans aide de l'Etat ou n'avaient pas investi depuis longtemps.
3. Evaluation de l'efficacité (analyses d'efficacité ou de rentabilité): De telles études servent à répondre à la question de savoir si l'output ou les effets pourraient aussi être réalisés avec moins de ressources ou si les dépenses (coûts) d'une mesure déterminée sont justifiées par ses effets. Les analyses d'efficacité, qu'elles soient menées de manière impressionniste ou systématique, correspondent ainsi à un désir croissant du public de plus de transparence concernant les avantages et les coûts de l'action étatique. L'exécution d'analyses coûts-avantages proprement dites exige une appréciation des effets en termes monétaires et présuppose des connaissances approfondies de la matière. Les analyses coût-efficacité posent des exigences un peu moins grandes en matière de quantification monétaire. Toutes ces études peuvent être entreprises à titre prospectif, avant la prise d'une mesure ou la mise en oeuvre d'un investissement, ou à titre rétrospectif en vue de procéder à d'éventuelles corrections.
Exemple: Dans une enquête sur le déblaiement de la neige sur les cols uranais (Ecoplan 1993 b), les avantages financiers d'une ouverture retardée des cols furent confrontés aux inconvénients qui en résulteraient (effets sur le secteur touristique). Il fut recommandé de repousser légèrement l'ouverture des cols.
4. Monitoring (contrôle et pilotage de l'exécution): Le monitoring, ou suivi, consiste à mettre à disposition en permanence des informations sur des aspects de l'exécution et des effets. Les données recueillies peuvent servir à surveiller l'exécution, p.ex. par la comparaison entre les valeurs idéales et les valeurs existantes ou la confrontation aux chiffres d'autres cantons. Le monitoring peut se référer à des grandeurs importantes pour l'objectif et donner une idée de leur développement. A ce propos, on mentionnera les efforts

entrepris pour élaborer des indicateurs de performance, qui doivent donner des éclaircissements sur l'output de l'administration et les effets obtenus. Mais, la plupart du temps, on ne peut pas certifier, sur la base des indicateurs de performance, que le résultat a été obtenu exclusivement sur la base des activités étatiques. Le monitoring et les indicateurs de performance sont des instruments ponctuels, mais permanents, alors que les évaluations des effets donnent des résultats plus approfondis, mais souvent limités à un moment déterminé. Ces instruments sont complémentaires.

Exemple: Les possibilités offertes dans le cadre des mesures dites préventives contre le chômage (avant tout les cours de perfectionnement) servent à améliorer les chances des chômeurs de trouver un emploi. Dans quelle mesure ces cours atteignent-ils leur but à court, à moyen et à long terme? Dans une étude approfondie, Freiburghaus (1987) a montré comment un contrôle permanent des résultats des mesures préventives peut être organisé. Pour contrôler le résultat de telles mesures de perfectionnement, des indicateurs (fréquentation des cours, appréciation au terme des cours, etc.) devraient être saisis.

5. Evaluation des effets futurs présumés d'une mesure prévue (évaluation ex ante, évaluation prospective): Cette perspective est centrale lors de la prise de nouvelles mesures ou de la révision de mesures existantes. Des méthodes très diverses sont utilisées pour répondre aux questions prospectives: évaluation impressionniste de déroulements d'effets possibles, élaboration de scénarios (p. ex. à l'aide de la méthode Delphi), appréciation des projets de mesures dans différentes conditions cadres, élaboration de modèles quantitatifs et simulation des effets. La collaboration d'une équipe d'évaluation peut servir différents objectifs: aide dans l'exploration de l'avenir (p.ex. exécution d'un sondage Delphi sur les scénarios possibles de l'évolution dans l'avenir), élaboration d'hypothèses fondées concernant la structure d'effets des mesures, élaboration de propositions en vue d'une meilleure évaluabilité des mesures prévues, etc.

Exemple: Après l'accident du réacteur de Tchernobyl, les voix en faveur de l'abandon de l'énergie nucléaire se multiplièrent. Le Conseil fédéral chargea un groupe de travail d'imaginer des scénarios énergétiques, afin d'étudier les possibilités et les conséquences d'un abandon de l'énergie nucléaire par la Suisse. Le groupe de travail examina au total 18 scénarios, dont 5 étaient particulièrement importants. Des divergences d'opinion au sein du groupe se manifestèrent notamment à propos des effets d'une politique sévère d'économies d'énergie sur le confort, ainsi que des effets économiques d'un abandon (rapport du groupe d'experts sur les scénarios énergétiques 1988).

La définition de la question de recherche est une démarche capitale de toute évaluation. Elle implique qu'on reformule le problème politique ou administratif qui se pose pour en faire une question de recherche utilisable.

Plus vous pourrez limiter et spécifier vos questions et vous faire une idée claire de la contribution de l'équipe d'évaluation à la solution de vos problèmes, plus l'évaluation sera précise et avantageuse. Bien qu'il incombe à une équipe d'évaluation compétente d'établir avec vous une formulation judicieuse des questions faisant l'objet de l'évaluation, la préparation sera facilitée si vous avez, à l'avance, des idées aussi réalistes que possible à ce sujet. Voici un exemple de la manière dont un ensemble de questions assez large au départ peut être réduit et précisé.

Le Conseil exécutif du canton X, compétent pour la promotion économique aimerait savoir si la politique de promotion économique du canton "fonctionne bien". Après plusieurs discussions avec les fonctionnaires chargés de l'exécution et un expert en évaluation appelé à titre informel, les questions de recherche possibles suivantes se dégagent:

- Les services de l'instance cantonale chargée de la promotion économique sont-ils suffisamment connus des petites et moyennes entreprises?
- Comment la collaboration entre les instances cantonales qui s'occupent de promotion économique pourrait-elle être améliorée?

- Est-il judicieux de soutenir en priorité les entreprises high-tech?

Après des entretiens avec le Conseil exécutif, il apparaît que les deux dernières questions sont jugées prioritaires. La question de la coordination est traitée dans le cadre interne, en collaboration avec les experts en organisation du département du personnel; pour la question des entreprises high-tech, on demande ensuite des offres à des équipes d'évaluation connaissant bien les questions régionales et économiques.

Le problème de savoir quelles sont, parmi les questions posées au départ sur une large base, celles auxquelles une évaluation pourrait répondre utilement doit être éclairci de cas en cas, et intégré par la suite également dans les négociations avec l'équipe d'évaluation.

1.22 Déterminer approximativement le groupe des parties prenantes

La plupart du temps, plusieurs services et responsables sont intéressés à une évaluation. Il s'agit donc de délimiter approximativement le groupe des personnes et instances concernées par l'évaluation. Votre service est-il le seul impliqué? D'autres services s'intéressent-ils à ces questions, ou leur participation est-elle même nécessaire dans la perspective d'une évaluation réussie? Est-il judicieux d'impliquer dès maintenant d'autres intéressés (associations, partis, etc.)? Qui doit signer officiellement un éventuel mandat d'évaluation? (Un éclaircissement préalable visant à connaître les idées de cette personne est en règle générale utile.) Quelles sont les personnes et institutions qui s'intéressent à une évaluation? De qui peut-on attendre un appui ou, selon les cas, des résistances?

1.23 Evaluation interne ou externe?

Dans ce guide, nous partons de l'idée que vous confiez un mandat d'évaluation à l'extérieur de votre propre service, que ce soit à un établissement proche de l'administration (p.ex. station fédérale de recherches agronomiques) ou à une instance indépendante (institut universitaire, bureau privé). Cela ne signifie pas que nous donnions la préférence à une évaluation externe. Comme on le verra dans les remarques suivantes, il y a beaucoup de bonnes raisons de confier une évaluation à une instance interne de l'administration, ou même de s'en charger soi-même. Mais, dans les deux derniers cas, le service dispose déjà en règle générale de contacts de recherche établis, ou même parfois d'un know-how en matière d'évaluation. Ce sont avant tout les personnes et les services qui n'ont pas de contacts de recherche réguliers qui devraient avoir le plus grand besoin d'un guide tel que celui-ci, c'est pourquoi nous en avons orienté le contenu sur ce cas hypothétique.

A propos de la question de l'opportunité d'une évaluation interne ou externe, il ne s'agit pas de choisir l'une aux dépens de l'autre, mais de combiner au mieux les deux possibilités. Il s'agit de méthodes d'action qui sont complémentaires et ne s'excluent pas. Il conviendra de garder cela à l'esprit quand nous examinerons ci-après quelques-uns des avantages et des inconvénients des évaluations interne et externe.

Evaluation interne

Dans une certaine mesure, tout service pratique l'évaluation interne. Il s'agit de représenter et d'apprécier les activités propres et leurs effets. La saisie des effets intervient dans la plupart des cas sous forme intuitive, sans que l'action étatique et la transformation des grandeurs cibles se trouvent dans un rapport de corrélation déterminé. L'accent principal de l'évaluation interne réside généralement dans l'évaluation de l'exécution. Dans la mesure où les connaissances scientifiques nécessaires existent dans le service en cause ou dans un département de recherche de l'institution correspondante, une évaluation interne des effets est certainement aussi possible.

- **Avantages**
En règle générale, une évaluation interne peut s'appuyer sur une bonne compréhension de la matière ainsi que sur la connaissance du domaine politique et de la discipline scientifique concernée. De bonnes relations personnelles existent avec les services administratifs, les instances d'exécution et les destinataires de la politique. Les données et statistiques administratives internes sont connues. Pendant l'évaluation, on peut transmettre en permanence les résultats intermédiaires. Généralement, les équipes internes d'évaluation bénéficient d'une grande crédibilité au sein de l'administration, et les services visés par les résultats et les recommandations acceptent volontiers de les mettre en pratique.
- **Inconvénients**
L'évaluation interne peut être liée à un aveuglement d'entreprise. Du point de vue méthodologique, les instances de recherche internes ne disposent pas obligatoirement de toutes les connaissances nécessaires à une évaluation. La crédibilité des résultats et des recommandations pour les institutions et personnes extérieures à l'administration (personnalités politiques, associations, partis, etc.) n'est généralement pas très élevée.

Evaluation externe

Pour une évaluation externe, vous avez la possibilité de choisir entre les instituts universitaires et les bureaux privés. Malgré les différences qui existent dans le profil de telles institutions (cf. AGREB 1990, p. 13), nous présenterons ensemble les avantages et les inconvénients des uns et des autres.

- **Avantages**
Les services d'évaluation externes disposent généralement de bonnes connaissances de la procédure (stratégies d'évaluation, techniques de recherche). Ils ont l'expérience du domaine en cause ou d'autres domaines politiques. Parfois, ils offrent une large palette de capacités (connaissances scientifiques, pratique des conseils). Ils ont la distance nécessaire par rapport à l'objet de la recherche et peuvent apporter des perspectives et des idées nouvelles. Cela accroît la crédibilité des résultats aux yeux des personnes extérieures.
- **Inconvénients**
Les équipes d'évaluation externes connaissent souvent moins bien l'objet de l'enquête que les praticiens de l'administration, et il leur faut parfois un certain temps pour démarrer. Les évaluateurs ne sont pas obligatoirement au courant des préoccupations concrètes de l'administration et ils doivent faire des efforts particuliers pour que les résultats correspondent aux besoins.

A notre avis, il n'y a pas de différences marquées entre les coûts et l'objectivité de l'enquête. En règle générale, les évaluations externes sont liées à des frais budgétaires déclarés comparativement plus élevés, étant donné que divers frais d'évaluation internes (frais de personnel, frais de bureau, frais de formation) demeurent cachés. Mais, du point de vue du coût global, elles ne sont pas obligatoirement plus chères. Inversement, les évaluations internes ne sont pas nécessairement plus optimistes, ou moins critiques par rapport au mandant. Il y a des équipes d'évaluation externes qui sont très soucieuses de leur réputation d'indépendance et de sérieux, mais aussi d'autres qui, dans la perspective de l'adjudication d'autres mandats de recherche, tiennent à maintenir une bonne entente avec le prescripteur de l'évaluation.

Les évaluations internes et externes peuvent être combinées de manière appropriée de cas en cas. En voici deux exemples.

Exemple 1: Un service aimerait savoir comment il pourrait préparer des données d'exécution appropriées pour saisir les effets d'une mesure déterminée. Il confie un mandat d'étude d'évaluabilité à un scientifique versé dans les questions d'évaluation. L'expertise fournie au service intéressé traite en détail des objectifs et des mécanismes d'effets de la mesure étudiée et indique les moyens de saisir en permanence ces effets (monitoring, indicateurs de

performance, indications pour évaluations futures des effets). Le service décide de recueillir auprès des instances d'exécution et des destinataires de la mesure différentes données qui seront utiles pour l'appréciation de l'efficacité. Par manque de personnel, différentes recommandations de plus longue portée concernant la collecte de données ne peuvent être mises en oeuvre (étude Freiburghaus 1987).

Exemple 2: Une évaluation de l'exécution et des effets est réalisée par un bureau privé. L'évaluation concerne aussi une institution semi-privée qui assure la plus grande partie de l'exécution. Les bases statistiques nécessaires à l'analyse des effets sont fournies par cette institution elle-même. Elle est tenue constamment au courant des résultats intermédiaires et met en oeuvre différentes propositions déjà avant la présentation du rapport final au public.

1.24 Déterminer l'objectif d'utilisation de l'évaluation

Il vaut la peine de réfléchir assez tôt à l'utilisation possible de l'évaluation. La délimitation du groupe des parties prenantes vous a peut-être déjà donné certaines indications. Vous savez qui est intéressé à l'évaluation et à ses résultats et comment la personne ou institution correspondante pourrait utiliser les éventuels résultats.

Maintenant déjà, il est probablement possible de savoir si l'évaluation sera engagée "simplement" au sens d'un compte rendu ou de manière ciblée, par exemple dans la perspective de réductions de frais ou de changements de mesures.

→ Vous trouverez des indications sur les objectifs d'utilisation possibles des évaluations au chap. 1.5.

En rapport avec l'objectif d'utilisation, la question se pose aussi de savoir si une publication va suivre l'étude d'évaluation. A qui faut-il communiquer les résultats? Au service intéressé et à ses responsables, aux autres parties prenantes, au grand public? Notamment pour faciliter les processus d'apprentissage (cf. chap. 1.56), nous préconisons la plus grande transparence et ouverture possible dans le domaine des évaluations. Pourtant, il sera parfois indiqué de concevoir les évaluations au sens strict de processus de conseils et de ne pas communiquer les résultats à d'autres intéressés ou au grand public. Cette procédure est conseillée, en particulier, quand une évaluation porte aussi sur le développement de l'organisation ainsi que sur l'amélioration des structures de direction et du personnel.

1.25 Fixer le cadre temporel et estimer les moyens nécessaires, en finances et en personnel

La plupart du temps, les exigences posées au processus d'évaluation découlent de l'objectif d'utilisation. Il s'agit de déterminer quand les résultats intermédiaires et finaux devront être produits.

En outre, il importe de fixer approximativement le cadre des ressources humaines et financières. Si vous exécutez l'évaluation au niveau interne, vous devez libérer le personnel nécessaire à l'exécution de l'évaluation. Pour une évaluation externe, vous devez réserver des moyens budgétaires et prévoir le personnel nécessaire à l'accompagnement et à la mise à profit de l'étude d'évaluation.

→ Il est important que vous partiez d'hypothèses réalistes concernant le temps et les frais. Le chap. 2.5 vous donne des indications correspondantes. Vous trouverez au chap. 1.4 des informations sur le temps nécessaire à l'accompagnement et à la mise à profit.

1.26 Décider de la poursuite de l'évaluation

Une évaluation n'a de sens que si le profit qui en est attendu dépasse les ressources en personnel et en finances qu'il faudra vraisemblablement y consacrer.

Dans de nombreux cas, des variantes de procédure sont aussi possibles, et il s'agit de choisir parmi elles celle qu'on présume la meilleure.

Exemple: L'OFEFP voulait savoir si les prescriptions de protection de la nature et du paysage étaient respectées dans le cas d'atteintes au paysage conditionnées par la pratique du ski. Il envisagea comme première variante une observation de l'environnement, sous forme de prise et d'interprétation de photographies aériennes de nombreux domaines skiables, afin d'en suivre l'évolution dans le temps. La seconde variante consistait à étudier très intensivement un certain nombre de domaines skiables selon des perspectives diverses (scientifique, juridique, sociologique et économique), afin d'observer le poids des impératifs écologiques dans la totalité des intérêts en jeu. (L'OFEFP s'est décidé pour la première variante; par la suite, le PNR 27 a financé la seconde, cf. Hirsch et al 1995.)

Documentation complémentaire

Pour de plus amples informations, on se référera avant tout au guide pour le management des projets de recherche de l'administration fédérale (AGREB 1990) et à ses annexes.

1.3 Demander des offres et adjuger le mandat

Quand vous êtes au clair sur le cadre général et les principales questions de l'évaluation, vous pouvez demander des offres². Même si vous entretenez une bonne et étroite collaboration de recherche avec un institut universitaire ou une équipe d'évaluation/un bureau de conseils, nous vous recommandons de demander plus d'une offre, et cela pour les raisons suivantes:

- Les offres concurrentes peuvent ouvrir des perspectives intéressantes sur la formulation des questions et la conception de stratégies d'évaluation nouvelles et prometteuses.
- Vous avez une meilleure position de négociation, avant tout du point de vue des coûts.
- Vous pouvez justifier plus valablement votre adjudication lors de contrôles ultérieurs.
- Il se peut que la meilleure approche du projet consiste à en faire une collaboration entre différentes équipes d'évaluation, ou à le diviser en mandats partiels.

Après une première sélection restreignant le choix à deux ou trois mandataires possibles, il est dans la plupart des cas utile de discuter des propositions soumises dans les offres avec les équipes d'évaluation correspondantes. On éclaircit à cette occasion les questions de compréhension, et on concrétise mieux la formulation des questions, la stratégie d'évaluation et la technique de recherche.

1.31 Faut-il donner mandat d'une étude d'évaluabilité?

Lors de projets d'évaluation relativement importants (plus de 200'000 francs), il peut être opportun de demander une étude d'évaluabilité (cf. aussi Balthasar 1993, pp. 6 ss). On entend par là des éclaircissements et des réflexions sur la planification et la praticabilité d'une enquête, en d'autres termes une étude de faisabilité concernant une évaluation. Cette étude doit notamment donner des éclaircissements sur la mesure à évaluer, les critères qui doivent être à la base de l'évaluation, les stratégies d'évaluation possibles, les techniques de recherche possibles et la situation des données. Dans le cadre d'une étude d'évaluabilité, on abordera par exemple les questions suivantes: La mesure à évaluer peut-elle être suffisamment cernée? Sur quels objectifs (objectifs formulés dans la loi, déclarations d'associations, besoins

² Nous partons ici du cas d'une évaluation externe; toutefois, certaines des étapes mentionnées ci-après peuvent également être utiles lors d'une évaluation interne.

des groupes cibles, etc.) l'évaluation pourrait-elle s'appuyer? Y a-t-il des sources d'information appropriée sur l'exécution ou les effets? Y a-t-il des résistances à une évaluation? Quel devrait être le rôle de l'équipe d'évaluation (fournir des informations, conseiller, arbitrer, etc.)?

Les études d'évaluabilité accroissent les chances de succès d'une évaluation. Elles peuvent aussi amener à la conclusion que les questions posées ne peuvent recevoir de réponse ou ne le pourraient qu'à un coût très élevé, en moyens financiers et en temps. Le cas échéant, une étude d'évaluabilité peut aussi indiquer à un service comment il pourrait trouver lui-même des informations appropriées sur l'efficacité des mesures mises en oeuvre par lui, et constituer une solution de rechange à une évaluation proprement dite.

1.32 Préparer la présentation de l'offre

La présentation de l'offre est facilitée, et le résultat amélioré, si vous fournissez des indications aussi précises que possible à propos du mandat. Vous devez en indiquer l'objectif et les principes; il appartiendra alors aux offrants de définir la voie par laquelle les buts prescrits seront atteints.

C'est pourquoi un cahier des charges que vous remettez aux intéressés pourrait contenir des indications sur les points suivants:

- motif, problème
- questions
- objectif d'utilisation des résultats
- état de fait à évaluer, mesure étatique à étudier
- questions spécifiques à creuser
- éventuels travaux préparatoires, études préalables
- mandant, personne compétente/personne de contact
- autres services concernés par l'évaluation
- exigences de qualifications posées à l'équipe d'évaluation
- impératifs temporels pour les résultats intermédiaires et finaux
- forme des produits intermédiaires et finaux (informations orales, écrites)
- cadre des coûts, éventuels impératifs concernant le niveau des honoraires
- date de remise des offres, ampleur attendue

1.33 Examen des offres sous l'angle de leurs lacunes

Avant de faire un choix entre les offres, étudiez-les sous l'angle de leurs éventuelles lacunes. Vous pourrez ainsi déterminer si elles donnent des réponses à toutes les questions importantes. La liste de contrôle ci-après énumère les points qui, sur la base des expériences du PNR 27, devraient être éclaircis avant la mise en oeuvre de l'évaluation. Selon le type de questions toutefois, il y aura d'autres points à régler; de même, certaines des questions figurant sur la liste de contrôle pourront ne pas être pertinentes.

Liste de contrôle pour l'examen de l'intégralité des offres

Les questions sur lesquelles il faut être au clair sont les suivantes:

- Quelle est la mesure à évaluer (p.ex. acte normatif complet, réglementation juridique particulière, subventions particulières, etc.)?
- Sur quelles questions (exécution, effets, monitoring, autres) l'évaluation se concentrera-t-elle?
- Selon quels critères (effectivité, efficacité, efficience) la mesure doit-elle être appréciée?

- Sur quoi l'analyse des buts s'appuiera-t-elle (buts formulés juridiquement, buts ou déclarations informels des protagonistes, besoins du groupe cible, etc.)?
- Quel modèle d'effets sera à la base de l'évaluation?
- Quelles stratégies d'évaluation va-t-on utiliser?
- Quelles techniques de recherche va-t-on utiliser?
- Quelles sont les personnes qui vont mener l'évaluation?
- Quelles étapes de travail seront entreprises, quand (év. par qui)?
- Qui accompagnera le travail de l'équipe d'évaluation?
- Quand des inputs (p.ex. documents statistiques, commentaires, communications à ceux qui sont concernés par l'évaluation) sont-ils attendus de vous ou d'autres services?
- L'accès aux données est-il assuré?
- De quoi se composeront les produits intermédiaires et finaux (rapports, ampleur de ces rapports, présentations orales, traductions de rapports, etc.)?
- Quand recevrez-vous ces produits intermédiaires et finaux?
- Des mesures supplémentaires de mise en valeur des résultats sont-elles prévues (p.ex. organisation de séminaires, présentations à des instances de décision, articles dans des revues)?
- Combien l'évaluation coûtera-t-elle (niveaux des prix) et de quels éléments se composeront les coûts?

1.34 Choisir

Vous êtes maintenant en mesure de prendre une décision fondée. Les critères déterminants pour cela dépendent du projet d'évaluation concret. Nous vous donnerons d'abord quelques indications sur les facteurs qui, selon notre expérience, sont importants pour le choix, et nous vous proposerons ensuite une liste de contrôle pour l'appréciation des offres.

Facteurs touchant le contenu

- Stratégie d'évaluation solide: Il s'agit de garantir que les difficultés rencontrées lors de la réalisation de certaines étapes de recherche ne mettront pas en danger l'ensemble du projet. Le meilleur moyen pour cela consiste à combiner les stratégies d'évaluation et à mélanger les méthodes.
- Flexibilité de la procédure: Si, dans le cadre d'une évaluation, on explore un terrain nouveau, on ne peut guère exclure les surprises (p.ex. concernant la disponibilité des informations). Dans ce cas, il est souhaitable que la structure de l'enquête permette des changements. Dans cette perspective, une procédure souple est plus facilement possible, par exemple, lors d'études de cas et d'interviews orales que dans le cadre d'une vaste interrogation écrite ou d'exploitations statistiques très coûteuses.
- Eclaircissement de la situation des données: Il faut éclaircir avec soin la question de savoir si les données nécessaires existent et peuvent être exploitées (p.ex. avec soutien informatique). La limitation à une stratégie d'évaluation ou à des techniques de recherche déterminées présuppose un éclaircissement comparativement plus approfondi de la situation des données. Cela peut nécessiter p.ex. que, lors d'un sondage dans un domaine délicat (p.ex. situation financière) on effectue des tests pilotes, que lors d'exploitations statistiques secondaires on examine les conditions préalables à une exploitation assistée par l'informatique ou que, lors d'une comparaison transversale, on procède à des éclaircissements de base sur les données (analyse élémentaire des mesures, comparabilité de l'environnement).
- Mesures de mise en valeur des résultats: Elles peuvent constituer un facteur de coût déterminant, dont l'importance est parfois sous-estimée lors de l'adjudication du mandat. Une offre qui ne contient pas de mesures de mise en valeur des résultats peut, au premier abord, paraître plus avantageuse qu'une offre qui prévoit les mesures correspondantes (y compris les frais!); mais, en fin de compte, elle sera peut-être plus coûteuse. Avant

tout si vous souhaitez déclencher des processus d'apprentissage dans l'administration (cf. chap. 1.56), la communication est une tâche permanente à laquelle il convient de prêter toute l'attention voulue. Une procédure participative, qui essaie d'inclure constamment les préoccupations des mandataires dans le processus d'évaluation et permet en même temps des réactions permanentes peut être très utile aux processus d'apprentissage (toutefois, elle prend plus de temps et suscite des frais). Les mesures de mise en valeur des résultats peuvent être fixées avant le début des travaux. Si trop d'incertitudes demeurent sur leur forme, on peut également les déterminer par la suite. Dans ce cas, il faudrait prévoir pour elles une réserve qui ne soit pas trop chichement mesurée, afin que, plus tard, il ne manque pas les moyens financiers nécessaires pour diffuser utilement les résultats de l'étude d'évaluation dans l'administration et à l'extérieur.

Facteurs personnels

- Qualité de l'équipe d'évaluation: Elle peut être d'une importance déterminante pour la réussite de l'évaluation. La qualité de l'équipe d'évaluation repose sur quatre facteurs qu'il vous appartiendra de pondérer.
 - L'équipe d'évaluation devrait bien connaître la stratégie d'évaluation correspondante et les techniques de recherche à appliquer. Cela ne concerne pas seulement les techniques difficilement compréhensibles pour le profane telles que les analyses multivariées ou les procédures économétriques. Des techniques apparemment (!) plus simples telles que les sondages ou les interviews en profondeur présupposent aussi des connaissances et de l'expérience si on veut qu'elles conduisent à des résultats significatifs.
 - L'équipe d'évaluation devrait avoir de bonnes connaissances du domaine qui fait l'objet de la recherche (p.ex. politique agricole, sécurité sociale, politique technologique), ou de la discipline déterminante (économie, sociologie, science politique, etc.). Cela permet une procédure rapide et précise, et facilite l'accès de l'équipe d'évaluation aux données (p.ex. disponibilité des personnes clés pour les interviews). En outre, les résultats d'une équipe d'évaluation qui compte parmi les connaisseurs d'un domaine déterminé sont en règle générale plus crédibles.
 - Si plusieurs personnes ou équipes d'évaluation indépendantes participent au projet, une bonne collaboration devrait être garantie (p.ex. par des formes déjà existantes de collaboration).
 - Last but not least, l'équipe d'évaluation devrait disposer de bonnes capacités de communication (contacts avec vous, avec les personnes concernées par l'évaluation, présentations écrites et orales).

La liste de contrôle ci-après pourra vous aider dans l'appréciation des offres entre lesquelles vous aurez à faire un choix. C'est à vous qu'il appartiendra de donner un poids plus ou moins grand aux différents facteurs. Nous vous recommandons de juger les offres avec réalisme. Ne vous laissez pas aveugler par les belles expressions techniques ou les graphiques perfectionnés. Il est important que vous compreniez la procédure proposée. Prenez garde quand l'équipe d'évaluation vous promet des réponses à toutes vos questions. Il se pourrait qu'elle n'ait pas du tout conscience de la portée de vos questions et que, par exemple, elle croie pouvoir répondre aux questions sur les effets d'une mesure par une simple énumération des indicateurs de buts. Faites preuve de prudence également quand l'équipe d'évaluation vous promet des réponses à des questions que vous n'avez pas du tout posées. Il est possible qu'elle soit spécialisée dans d'autres questions et d'autres méthodes que ce qui serait nécessaire dans le cas concret.

- Si vous souhaitez confier un mandat d'évaluation des effets et que les offres reçues ne vous convainquent qu'en partie, nous vous recommandons de vous référer au chap. 3.5.

Liste de contrôle pour l'appréciation d'offres

Offrant(e)::.....

| Critères d'appréciation | oui/non | commentaire |
|--|---|-------------|
| Critères touchant le contenu | | |
| La définition des tâches a été correctement saisie. | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | |
| Toutes les questions ont été examinées. | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | |
| La manière dont les questions seront abordées est représentée de manière compréhensible. | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | |
| La procédure et la perspective méthodologique choisies sont convaincantes. | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | |
| Dans le cas d'analyses d'effets, la stratégie d'évaluation et les méthodes de recherche sont appropriées pour saisir les effets des mesures. | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | |
| Le calendrier de travail apparaît réaliste, il n'y a pas de goulots d'étranglement. | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | |
| Du point de vue des résultats intermédiaires et finaux, les étapes du projet sont correctement réparties. | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | |
| Les propositions de mise en valeur des résultats sont judicieuses. | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | |
| Critères personnels | | |
| L'équipe d'évaluation a des connaissances méthodologiques (sur les stratégies d'évaluation, les techniques de recherche), spécialisées (p.ex. en économie) et pratiques suffisantes. | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | |
| L'organisation du projet prévue semble appropriée à l'accomplissement du mandat. | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | |
| La forme de la collaboration entre mandant et mandataires est clairement définie. | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | |
| Critères financiers | | |
| Le cadre financier proposé est adapté aux travaux prévus. | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | |
| Les modalités de décompte proposées permettent un contrôle permanent du rapport prix/prestations. | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | |
| Les éventuels impératifs concernant les honoraires sont respectés. | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | |

Autres remarques

Source: Andreas Balthasar, documents destinés au cours "Engagement d'expert(e)s extérieur(e)s", dans le cadre du programme de perfectionnement "La commune et l'administration urbaine en mutation" du Centre de recherche sur la politique suisse (avec petits compléments de l'auteur).

1.35 Adjudication du mandat en tant que négociation

Il est rare que, parmi les offres qui vous sont soumises, vous trouviez une proposition parfaitement adaptée à vos besoins. En règle générale, une phase de négociation est nécessaire avant l'adjudication du mandat, au cours de laquelle vous concrétisez encore vos désirs et vos idées alors qu'inversement l'équipe d'évaluation précise ses options méthodologiques. Cette phase peut être très importante pour le succès de l'évaluation. On pourra peut-être adapter encore mieux à vos besoins la procédure, le calendrier et les coûts. Dans la mesure où vous excluez délibérément certaines méthodes (p.ex. pour des raisons de fiabilité, de temps ou de frais), vous accepterez d'autant mieux la procédure choisie et ses résultats. La phase de négociation vous permet d'évaluer le niveau de qualification méthodologique de l'équipe d'évaluation, et de vérifier si elle est en mesure de répondre à vos questions critiques sur la stratégie d'évaluation et les méthodes de recherche.

A la fin du processus de négociation, vous constaterez probablement qu'en matière d'évaluation il n'y a pas de voie "idéale" ou "unique". La définition d'un mandat d'évaluation est beaucoup plus un processus d'optimisation dans lequel un rapport d'équilibre doit être établi entre, d'une part, les besoins d'information et d'apprentissage et, d'autre part, le temps et les coûts.

Il nous paraît important qu'avant la signature du contrat vous fassiez une courte pause pour considérer objectivement l'esquisse du projet. En règle générale, une évaluation ne donnera de bons résultats que dans la mesure où on peut s'y attendre sur la base du plan de recherche. Il ne faut donc pas espérer des miracles. Avant l'adjudication du mandat, il peut valoir la peine de considérer une dernière fois les questions suivantes:

- Est-ce que je comprends la stratégie d'évaluation? A-t-elle un sens? Est-elle appropriée pour donner une réponse aux questions posées? Peut-on prévoir que les résultats seront utilisés et appliqués³? N'est-il plus possible d'apporter d'autres améliorations au plan de recherche?
- L'équipe d'évaluation mérite-t-elle notre confiance dans la perspective du traitement de questions difficiles?
- Le calendrier de travail est-il réaliste? N'y a-t-il pas de goulots d'étranglement (délais trop courts)? Les résultats arriveront-ils à temps pour la prise des décisions qui en dépendent?
- Les frais de l'évaluation peuvent-ils être justifiés par le profit qu'on pourra vraisemblablement en tirer?

En cas de doute, il vaut la peine de revoir encore une fois la question et de discuter des points correspondants avec l'équipe d'évaluation, voire de renoncer à l'évaluation. Vous avez le droit d'être au clair sur le plan de recherche. En revanche, si vous pouvez répondre affirmativement à toutes les questions ci-dessus, plus rien ne devrait s'opposer à la conclusion du contrat.

La dernière étape de l'adjudication du mandat est la formulation et la signature du texte du contrat. A cet égard, on peut en règle générale partir de l'offre, complétée et précisée sur la base du résultat obtenu dans les négociations.

Documentation complémentaire

Comme autres instruments de travail de nature générale, vous pouvez utiliser AGREB 1990 et l'Office fédéral du personnel 1993. En outre, Wholey 1983 donne une bonne introduction à la procédure en matière d'évaluations.

³ Si vous souhaitez approfondir ce point, vous trouverez d'autres données au chap. 1.5.

1.4 Accompagner utilement l'évaluation

Dans la perspective de l'accompagnement et du pilotage du projet d'évaluation, deux questions sont au premier plan: Premièrement, qui doit accompagner l'évaluation? Deuxièmement, comment, c'est-à-dire selon quels principes l'accompagnement doit-il intervenir?

1.41 Qui accompagne l'évaluation?

La définition de l'organisation du projet dans la perspective de l'accompagnement a des effets sur la mise en valeur des résultats, mais aussi sur les frais. Les questions suivantes doivent recevoir une réponse: (1) Quelle(s) personne(s) du service qui commande (et finance) l'évaluation doit/doivent assumer la responsabilité face à l'équipe d'évaluation? (2) Faut-il éventuellement constituer un groupe chargé de l'accompagnement et du pilotage de l'évaluation? Quelle sera la tâche de ce groupe d'accompagnement? De qui doit-il se composer?

Sur la base des expériences acquises dans le cadre du PNR 27, les groupes d'accompagnement peuvent représenter un instrument utile pour accompagner et contribuer à façonner le processus d'évaluation. Ils devraient se composer des principales personnes intéressées au processus d'évaluation et à ses résultats, et en même temps ne pas dépasser une certaine grandeur (8-12 personnes). Le groupe d'accompagnement peut, d'une part, fournir des informations précieuses à l'équipe d'évaluation. De l'autre, par son accès privilégié aux résultats de l'évaluation, il est très souvent un instrument utile pour la mise en valeur des résultats. On peut lui confier par exemple les tâches suivantes:

- collaboration à la définition de la question de recherche et du plan de recherche
- détermination des personnes ou groupes à inclure dans l'évaluation (p.ex. sous forme d'interviews)
- appréciation des stratégies d'évaluation et des techniques de recherche (p.ex. du point de vue de la plausibilité et de la compréhensibilité)
- appréciation des résultats de l'évaluation (présentations orales, rapports intermédiaires, projet de rapport final)
- collaboration à la détermination des changements qui doivent être décidés à la suite de l'évaluation.

L'aspect négatif est le temps nécessaire à l'action du groupe d'accompagnement. Les délais doivent être fixés assez tôt et comporter une réserve pour les difficultés qui pourraient éventuellement survenir lors de l'évaluation. Les groupes d'accompagnement suscitent également des dépenses, à savoir des frais administratifs (p.ex. locations de salles, repas, év. indemnités de déplacement aux participant/e/s) ainsi que du temps pour la préparation des séances. Le temps étant un bien coûteux, les séances du groupe d'accompagnement doivent être préparées très sérieusement par l'équipe d'évaluation (ce qui se répercute sur le budget).

Nous déconseillons d'établir des groupes d'accompagnement en tant que simple instrument alibi. Ils doivent permettre au contraire une authentique collaboration d'autres personnes au processus d'évaluation. Le mandant formel n'a pas à craindre cette collaboration si, de son côté, il peut faire en sorte que le groupe d'accompagnement ait suffisamment de connaissance de la matière.

1.42 Principes de l'accompagnement

L'accompagnement de l'évaluation exige un bon dosage de confiance et de contrôle. Les équipes d'évaluation méritent généralement la confiance qui leur est accordée. Elles apportent un know-how professionnel et de l'expérience dans la gestion de projets de recherche, et elles ont intérêt à ce que l'évaluation soit un succès.

En même temps, il est important que vous montriez à l'équipe d'évaluation que vous êtes soucieux de la communication ponctuelle des résultats intermédiaires et finaux. Plus vous saurez marquer votre présence et mettre en évidence l'importance de résultats conformes au but et à la matière, mieux le résultat correspondra à vos idées. Un accompagnement attentif de l'évaluation peut être très productif pour la mise en valeur des résultats. Plus vous suivrez l'évaluation avec compétence, plus les décisions à prendre encore pendant le processus d'évaluation (p.ex. ampleur de l'échantillon du sondage, détermination des interviews à faire lors de la deuxième étape, ampleur et forme du rapport intermédiaire) correspondront à vos besoins. Vous tirerez profit des dépenses et des efforts d'accompagnement au moment de la mise en valeur des résultats.

Prévoyez donc assez de temps pour l'accompagnement de l'évaluation. L'expérience nous a montré que, pour l'accompagnement d'une évaluation relativement importante - p.ex. de dix mois -, il faudrait prévoir au moins 1 à 2 jours par mois. Vous aurez besoin de ce temps pour:

- les contacts téléphoniques (p.ex. toutes les 2 semaines); la discussion des questions de détail qui surgissent;
- les rencontres avec l'équipe d'évaluation (environ une fois par mois) ou avec un éventuel groupe d'accompagnement (tous les 2-4 mois); la préparation des décisions importantes touchant le projet de recherche;
- l'information d'autres parties prenantes, au sein et en dehors de votre service, concernant la progression du projet et les résultats (intermédiaires);
- l'analyse des résultats intermédiaires et finaux;
- la mise en valeur des résultats.

Il est tout à fait possible que l'équipe d'évaluation, au cours de l'opération, interprète le plan de recherche différemment de vous ou que, par exemple, des changements doivent être introduits en raison de la situation des données. Dans ce cas, il est important de mettre les problèmes sur le tapis et d'en discuter de manière approfondie. Souvent, il sera possible de trouver une solution appropriée prenant en compte tant vos besoins d'information que les limites fixées par le budget.

Nous ne saurions taire le fait que, dans de rares cas, des difficultés sérieuses peuvent surgir lors d'évaluations. Les raisons peuvent en être très diverses:

- résistance des services qui sont partie prenante et/ou des groupes concernés;
- changement des conditions cadres (rejet d'un projet en votation populaire, nouvelles décisions d'orientation des autorités politiques, problèmes budgétaires, etc.);
- plan de recherche irréalisable, ou seulement difficilement réalisable (manque de qualification d'un groupe de contrôle pour la comparaison, séries de nombres insuffisantes lors d'une comparaison longitudinale, problèmes de collecte des données, difficultés dans l'exploitation assistée par ordinateur);
- problèmes personnels au sein de l'équipe d'évaluation (remise de tâches de recherche à une personne non qualifiée, mauvaise collaboration au sein de l'équipe d'évaluation, maladie ou indisposition prolongée de membres clés de l'équipe d'évaluation, départ d'une personne très capable de l'équipe d'évaluation).

Il n'y a pas de recettes passe-partout pour faire face à ces situations délicates. Vous pouvez essayer de remédier quelque peu aux difficultés et de trouver des solutions de remplacement (p.ex. remplacement d'un collaborateur de projet non qualifié, échantillon composé un peu différemment) afin d'arriver néanmoins à des résultats utilisables. Lors de problèmes dont l'équipe d'évaluation n'est pas responsable (p.ex. insuffisance des données fournies par l'administration), un mandat complémentaire sera peut-être nécessaire. Lorsqu'une conclusion prometteuse de l'évaluation n'est pas en vue, on devra envisager de limiter, voire d'abandonner le projet. Dans tous les cas, nous vous déconseillons, face à des résultats (intermédiaires) insuffisants dus à une planification lacunaire et à une gestion de projet insatisfaisante de l'équipe d'évaluation, d'investir des moyens supplémentaires dans le projet d'évaluation. Inversement, il peut être très profitable de donner des mandats supplémentaires

quand l'équipe d'évaluation a fait du bon travail et que d'intéressantes questions complémentaires se posent.

1.5 Mettre à profit de manière optimale les résultats de l'évaluation

Vous vous trouvez en face d'un problème et vous avez commandé et accompagné une évaluation pour y trouver des réponses. Déjà lors de la préparation du projet d'évaluation, vous avez peut-être rencontré des questions que vous ne vous étiez encore jamais posées sous cette forme. Par exemple: Quels sont nos principaux objectifs? Quelles personnes, entreprises ou régions nos mesures doivent-elles atteindre? La mesure a-t-elle suscité les effets secondaires redoutés? Y a-t-il un potentiel d'économies qui ne mette pas en danger les effets principaux?

Par le contact avec votre mandataire, vous avez eu connaissance des résultats intermédiaires. Peut-être ce contact étroit a-t-il déjà déclenché des processus d'apprentissage (souvent, les processus d'apprentissage les plus importants se déroulent pendant l'évaluation). Maintenant, les résultats sont là (par écrit ou sous forme de présentation orale). Vous ou d'autres personnes pouvez les mettre à profit à différentes fins.

1.51 Rendre compte

Peut-être les résultats ne sont-ils pas nouveaux. Ils mettent en évidence des faits que vous connaissiez déjà. Peut-être avez-vous maintenant une connaissance plus fondée de l'exécution, de l'efficacité ou de l'efficience. Vous pouvez attester que le public cible est atteint. Il est apparu que les effets secondaires étaient moins graves qu'on l'avait prétendu. C'est bien. Il n'est pas obligatoire qu'une évaluation fasse apparaître la nécessité de changements et d'améliorations. Le public a le droit de savoir si des mesures et des lois qui sont coûteuses ou qui imposent des devoirs particuliers aux privés atteignent les objectifs fixés. La transparence des mesures étatiques aide à créer ou à accroître la confiance des citoyennes et des citoyens en l'Etat. Par une évaluation, vous pouvez aussi, le cas échéant, prévenir une "fausse alarme" et vous procurer des arguments permettant de faire face à l'appel à une réduction de subventions ou à une suppression de prescriptions (ou le contraire). La démocratie a besoin d'arguments fondés.

Nous aimerions vous déconseiller expressément d'utiliser les évaluations comme un simple instrument alibi propre à vous donner une légitimation. A court terme, vous pouvez en tirer un bénéfice. Mais nous avons la conviction qu'à l'occasion les organes de contrôle, au niveau de la Confédération et des cantons, vérifieront si l'adjudication des évaluations correspond à un véritable besoin et est menée avec soin.

1.52 Améliorer l'exécution

Très souvent, les évaluations peuvent donner des indications sur la manière dont vous pourriez améliorer l'exécution. Les évaluations de l'exécution, en particulier, poursuivent cet objectif. Les évaluations des effets peuvent aussi vous donner des indications sur les améliorations possibles de l'exécution quand elles ne sont pas conçues comme une évaluation "boîte noire" et qu'elles reconstruisent soigneusement la piste d'effets.

Exemple 1: Vous avez appris par l'évaluation que les services que vous offrez sont connus de 90 % des grandes entreprises, mais de 40 % seulement des petites. Dans les années à venir, vous allez renforcer sensiblement vos efforts d'information auprès des petites entreprises et vérifier ensuite si le degré de connaissance et l'utilisation des services se sont améliorés dans ces entreprises.

Exemple 2: Une évaluation de l'exécution a essayé avant tout de déterminer à quoi sont imputables les différences d'exécution qui existent entre les communes de votre canton.

L'enquête a montré que les responsables de cet état de choses ne sont pas la structure politique ou la situation financière des communes, mais la sensibilité aux problèmes et le niveau de formation du personnel d'exécution. C'est pourquoi, dans les mois et les années à venir, vous allez centrer vos efforts principalement sur la formation et le perfectionnement (cf. étude Balthasar/Knöpfel 1994).

Exemple 3: L'étude commandée montre que l'amélioration de la qualité de l'environnement visée par la mesure est anéantie par la multiplication des émetteurs (p.ex. véhicules, bâtiments). Mais, en même temps, un potentiel considérable est découvert encore au niveau de l'application systématique des prescriptions. C'est pourquoi vous allez concentrer vos efforts en premier lieu sur une exécution conforme aux normes, et patienter en ce qui concerne l'aggravation des normes juridiques.

1.53 Economiser des frais

En tant qu'instrument polyvalent, les évaluations peuvent contribuer à économiser des frais. Des propositions d'économies possibles peuvent résulter d'évaluations de l'exécution et des effets. Mais les évaluations peuvent aussi porter spécifiquement sur la recherche de possibilités de réduire les frais. Dans ce cas, on attribue la plus grande priorité au critère de l'efficacité. Il s'agit de déterminer quelles sont les économies qui auront le moins d'influence sur les effets principaux visés par la mesure. On mentionnera spécifiquement à cet égard les analyses coûts-avantages et les analyses coût-effectivité, qui sont engagées généralement de manière prospective pour évaluer les solutions d'action possibles.

Exemple 1: Une reconstruction des décisions d'investissement du groupe cible visé par la subvention fait apparaître que 70% des bénéficiaires auraient procédé à l'investissement également sans aide de l'Etat. C'est là un effet d'aubaine très élevé. Par un rétrécissement du groupe des ayants droits et une meilleure formation des personnes chargées de l'examen des demandes, on va essayer de diminuer l'effet d'aubaine.

Exemple 2: Une enquête auprès d'un groupe cible de bénéficiaires de subventions étatiques fait apparaître que certaines personnes et entreprises n'avaient pas besoin de cette prestation et ont été très surprises d'en bénéficier. On essaie maintenant, en changeant les bases juridiques, d'exclure de tels cas du droit aux subventions.

Exemple 3: Le goulot d'étranglement financier d'une grande ville impose une réduction des dépenses d'un service social égale à 1,4 million de francs. On envisage d'abord une réduction linéaire des dépenses. On y renonce en faveur d'une évaluation. Pour la première fois, on crée ainsi la transparence sur les frais globaux (frais de personnel, frais de matériel, frais de locaux incorporables) de ce service réparti dans toute la ville, on classe les prestations fournies par ordre d'importance, en collaboration avec le personnel d'exécution, et on fixe ensuite des priorités et des postériorités pour les réductions. Grâce à l'évaluation, on suscite pour la première fois dans le service en cause la conscience des objectifs visés par le travail social et des frais que cela implique. On peut alors effectuer les réductions là où cela sera le plus supportable pour les personnes touchées. Une comparaison transversale des différents services améliore en outre la conscience de leurs particularités et des potentiels d'amélioration (cf. étude Price Waterhouse 1993).

Il est légitime que les évaluations soient mises en oeuvre en vue d'économiser des coûts, implicitement ou explicitement. Elles peuvent fournir des bases approfondies sur la manière dont des réductions déterminées pourront être réparties au mieux entre différents postes de dépenses.

1.54 Supprimer des mesures

Le cas échéant, une évaluation peut aussi montrer qu'une mesure pourrait être supprimée sans inconvénients majeurs.

Exemple: L'évaluation de l'impératif d'égalité, dans la loi sur le travail à domicile, entre travailleurs à domicile et travailleurs en usine (Gerheuser/Schmid 1993) a montré, premièrement, que le nombre des personnes soumises à cette disposition diminue régulièrement, deuxièmement, qu'il est hautement vraisemblable qu'il n'y a plus de cas d'exploitation criante et, troisièmement, qu'une application systématique de l'impératif d'égalité conduirait à la disparition de certaines formes de travail à domicile, ce qui n'est en fait pas souhaité par les travailleurs intéressés. Dans leurs conclusions, les auteurs considèrent la suppression de la prescription comme un scénario possible, en partant de l'idée que le travail à domicile, principalement exercé par les femmes, sera suffisamment protégé par l'impératif d'égalité selon la (future) loi fédérale sur l'égalité entre la femme et l'homme.

Avant tout, les évaluations de l'efficacité peuvent donner des indications sur la question de savoir si les coûts d'une mesure sont compensés par l'utilité des effets qu'elle permet d'obtenir. Dans certains cas, elles peuvent aussi aider à justifier la renonciation à des mesures supplémentaires demandées par les groupes politiques.

1.55 Changer ou remanier des mesures

Les évaluations, en tant qu'enquêtes empiriques, sont d'abord orientées sur la saisie de la situation existante (cf. chap. 2.4). Toutefois, leur objectif le plus important réside souvent dans la préparation des décisions futures. Pour savoir où on va, il faut commencer par bien savoir où on est. C'est seulement quand on sait comment l'exécution se déroule et quels effets la mesure produit qu'on peut déterminer le besoin de changement de manière quelque peu fiable.

Si l'avenir était un simple prolongement du passé, le remaniement d'une mesure serait une chose très simple. Il suffirait d'éliminer les défauts de la mesure existante. Les changements résulteraient pour ainsi dire automatiquement d'une évaluation (ex post/rétrospective).

Mais la réalité est complexe et souvent imprévisible. C'est pourquoi trois autres étapes doivent s'ajouter à une évaluation de l'exécution et des effets:

Analyse de l'évolution du problème:

Il s'agit d'analyser les développements intervenus jusqu'ici et de formuler des hypothèses aussi fondées que possible sur la manière dont les conditions cadres et la situation du problème évolueront à l'avenir. Il faut par exemple répondre aux questions suivantes: Comment le problème a-t-il évolué? A-t-on constaté une amélioration, une détérioration, une stabilisation? Comment la situation va-t-elle vraisemblablement évoluer? Quels nouveaux facteurs futurs faut-il prendre en compte? Si, notamment dans les questions touchant l'économie, les transports, l'énergie et l'environnement, on peut parfois établir une modélisation quantitative des développements passés ou à venir, dans d'autres domaines politiques, pour autant qu'on s'y livre à des considérations explicites, on travaille avec des scénarios. Il est concevable aussi de chercher à connaître les idées que se font de l'avenir les experts du domaine en cause, en appliquant à cela la technique du sondage Delphi. Pour maîtriser les techniques complexes d'exploration de l'avenir, il peut être tout à fait indiqué de faire appel à un conseiller scientifique.

Analyse des possibilités d'action:

Sur la base de l'analyse du problème, de l'évaluation de l'exécution et des effets, compte tenu des exemples (positifs et négatifs) tirés de domaines politiques apparentés, dans d'autres cantons ou d'autres pays, et, last but not least, sur la base de votre créativité et de celle d'autres partenaires intéressés au remaniement, il s'agit de déterminer des solutions de rechange pour l'avenir. Celles-ci peuvent consister en copies de mesures existantes (dans d'autres domaines politiques, d'autres cantons ou pays), en adaptation de mesures existantes à une situation changée, en un amalgame d'éléments de mesures existantes, en une synthèse

de tels éléments de mesures pour en faire quelque chose de nouveau, ou encore relever de l'invention.

Choix de la mesure appropriée:

Dans une analyse grossière, il s'agit d'abord d'éliminer les solutions d'action inappropriées. Ensuite, les solutions restantes (sur la base de vos expériences et connaissances acquises jusqu'ici) seront jugées approximativement, dans la perspective des critères "effectivité/apptitude à l'exécution", "efficacité" et "efficience". Il faudra répondre à des questions telles que celles-ci: Quels problèmes d'exécution pourraient surgir selon les différentes solutions d'action? Combien de temps nécessitera l'exécution? Quels effets principaux et secondaires pourraient intervenir? Quelles charges financières et autres faut-il en attendre pour les privés? Comme l'expose le chapitre 3.7, il est possible, dans certains domaines politiques, de mettre en place des modélisations et des simulations des effets de mesures futures. Dans d'autres domaines, on travaille avec des scénarios. En règle générale, il s'agit là d'une appréciation approximative des effets. Elle constitue la base du choix de l'action. Celui-ci implique qu'on classe les objectifs selon leur importance, qu'on détermine le degré probable de réalisation des objectifs des différentes mesures et qu'on choisisse selon le profit à en attendre. En soi, il existe des modèles et des logiciels pour un choix d'action aussi systématique que possible, mais qui ne sont toutefois pas applicables à la plupart des processus de décision dans le domaine public. Ce qui est décisif, c'est de faire un choix conscient, fondé sur des arguments valables. Les possibilités d'action nouvelles et peu conventionnelles ne doivent pas être exclues trop vite sur la base d'arguments émotionnels; leur contribution à la solution du problème devrait être appréciée sérieusement.

En règle générale, les étapes énumérées ci-dessus ne sont pas exécutées par une seule personne, mais dans le cadre de travaux de commission. La collaboration de plusieurs personnes encourage la créativité et facilite l'approche et la discussion de solutions d'action possibles. Toutefois, le danger existe qu'une opinion de groupe s'impose rapidement et sans examen sérieux des différentes solutions d'action. Dans la mesure où vous pouvez faire intervenir votre responsabilité dans le processus, vous devriez faire en sorte que toutes les solutions valables soient étudiées sérieusement et que le choix se fasse sur une base solide.

L'exécution systématique et résolue des étapes mentionnées ne vous donne pas encore la garantie que les mesures pourront être exécutées sans difficultés et qu'elles agiront dans la mesure attendue. L'avenir est toujours ouvert aux surprises. L'objectif devrait être de façonner la mesure, sur la base des éléments connus, de telle sorte qu'elle prenne en compte les développements prévisibles possibles. On ne peut pas faire plus. Toutefois, il est important que, dans la formulation de la mesure, vous garantissiez qu'à l'avenir vous recevrez périodiquement les informations qui vous permettront de savoir si l'exécution se déroule dans le cadre planifié et si on peut admettre que les effets prévus sont atteints (cf. chap. 1.6).

1.56 Les processus d'apprentissage en tant qu'objectif de l'évaluation

L'objectif principal des évaluations est, en résumé, de permettre et de favoriser les processus d'apprentissage. De meilleures connaissances doivent rendre possible une meilleure action. Les processus d'apprentissage ont une dimension individuelle et collective. Seuls les individus acquièrent la connaissance. Mais la transformation de la réalité par l'action intervient dans un contexte collectif, dans des organisations ou même dans des réseaux. Souvent, il est très difficile de prévoir au début qui il faut intégrer dans ce processus d'apprentissage: un service déterminé, les fonctionnaires au niveau de la Confédération, des cantons et des communes, un réseau relativement large incluant les associations des personnes touchées, ou même le grand public? Les processus d'apprentissage qui se dégagent des évaluations ne sont pas entièrement planifiables. Par la transparence, et en particulier par l'ouverture de l'accès aux résultats, on crée toutefois la possibilité d'élargir au maximum le cercle des personnes et institutions impliquées dans le processus d'apprentissage.

Jusqu'ici, la recherche évaluative a pris en compte la nécessité d'une évaluation favorisant l'apprentissage en développant des modes d'action répondant très concrètement aux besoins des mandants. Depuis quelque temps, on pratique de plus en plus les méthodes d'évaluation interactives. Celles-ci comprennent l'évaluation comme un processus ouvert, façonnable, dans lequel peuvent intervenir tant les préoccupations de l'équipe d'évaluation que celles des intéressés (cf. chap. 3.62). L'évaluation interactive peut contribuer à susciter des processus d'apprentissage dans un réseau de personnes ou d'institutions. L'évaluation se rapproche ainsi de la médiation, c'est-à-dire de l'utilisation des connaissances des sciences sociales pour mieux maîtriser les conflits politiques et sociaux. Certes, à côté de cela, on continuera à avoir besoin, grand besoin même, d'évaluations de nature indépendante et libre. De telles évaluations peuvent, dans la mesure où elles créent un savoir nouveau sur une base fiable, susciter des processus d'apprentissage sans pour autant afficher explicitement cet objectif.

1.57 Digression: Abus possible des évaluations

Dans le contexte des processus d'apprentissage et de l'utilisation des évaluations, il importe de dire un mot sur l'abus possible des évaluations. A cet égard, il convient de signaler d'emblée qu'entre l'utilisation légitime des évaluations et leur abus, il existe des zones d'ombre qui rendent les limites mouvantes.

Les dangers d'abus se glissent déjà dans l'adjudication du mandat d'évaluation. La question de recherche peut être formulée de telle sorte que les effets positifs d'une mesure sont saisis et que, par exemple, les effets secondaires négatifs possibles ne sont absolument pas étudiés. Ce caractère unilatéral peut se manifester aussi dans le design d'évaluation: par exemple, on ne choisit comme unités de comparaison que des cantons ou d'autres pays dont on est sûr qu'ils s'en tirent comparativement moins bien. Egalement lors de l'application de certaines techniques de recherche, on peut orienter un résultat positif: par exemple, dans les sondages, on n'inclut pas les groupes qui se montrent critiques à l'égard de la mesure. Les statistiques qui donnent une image défavorable de la situation sont classées comme non fiables et ne sont même pas analysées. Le résultat d'une telle manière d'agir est une évaluation alibi, forcée de manière à donner une image positive de la mesure. Inversement, une évaluation - voulue comme une évaluation "tueuse" - peut aussi être utilisée pour "descendre" une mesure déterminée. Là encore, il s'agit d'une évaluation alibi qui ne tend qu'à légitimer une décision déjà prise.

Il n'y a pas de recettes passe-partout pour lutter contre ces abus des évaluations. Nous plaçons notre espoir dans une opinion publique critique, des médias qui posent des questions et connaissent leur sujet, une science qui veille soigneusement sur ses critères, et des organes de contrôle de l'exécutif et du Parlement⁴ soucieux de vérifier le fondement méthodique des études d'évaluation en procédant à des contrôles au hasard. Certes, tous les abus ne pourront être empêchés de la sorte. Mais nous avons la conviction qu'avec le développement des évaluations, on prendra également des dispositions propres à empêcher les éventuels excès. A l'avenir, les instituts universitaires et les bureaux privés seront obligés, plus encore qu'aujourd'hui, de se conformer à des directives éthiques implicites ou explicites s'ils veulent maintenir à long terme leur crédibilité scientifique.

Des abus peuvent encore se produire après l'établissement des résultats de la recherche. On classera très clairement parmi les abus la falsification de données désagréables, le retrait de passages critiques et la non-publication de résultats d'évaluation élaborés avec sérieux, alors que pourtant la publication avait été convenue. A côté de cela, il y a de nombreux cas qu'on ne peut qualifier véritablement d'abus. Parmi ceux-ci, il y a l'utilisation unilatérale de

⁴ Ainsi, aux Etats-Unis, le General Accounting Office, organe de contrôle du parlement, a accumulé un ensemble de connaissances important dans le but d'examiner scientifiquement les évaluations demandées par l'exécutif.

résultats d'évaluation par les médias et les politiciens. Il peut très bien arriver que, dans les médias, à partir d'un rapport aboutissant à un jugement dans l'ensemble positif sur une mesure, on extraie les quelques passages critiques en les présentant comme l'opinion d'ensemble du rapport (ou inversement). Dans la mesure où il n'y a pas de citations fausses, on ne peut guère faire quoi que ce soit juridiquement (p.ex. par le recours au droit de réponse) contre une telle utilisation unilatérale. On sait aussi que les politiciens choisissent des arguments dans les analyses scientifiques et s'en servent comme munitions dans les débats politiques. Prescripteurs et exécutants d'évaluations doivent se résigner au fait que la discussion des mesures étatiques se déroule dans l'arène politique et que c'est là que sont fixées les règles du jeu.

Dans le champ de tension entre l'exploitation légitime et l'abus, des questions très délicates apparaissent notamment lors de la formulation des rapports finaux destinés à l'opinion publique. Du point de vue de l'équipe d'évaluation, l'administration réagit souvent de manière hypersensible à certaines remarques critiques figurant dans le projet de rapport final. De leur côté, les mandants se représentent déjà comment ces passages seront extraits de leur contexte, gonflés et jetés dans le combat politique. Nous voudrions ici, d'une part, faire remarquer aux prescripteurs (avant tout dans l'administration) que la fonction de contrôle de la presse et de la politique exige qu'on montre à l'occasion les faiblesses et que les conséquences qui en découlent pour l'administration sont rarement graves. Nous croyons que le gouvernement et l'administration doivent développer une meilleure approche de la critique. D'autre part, nous aimerions faire remarquer en même temps aux mandataires que les évaluations sont très souvent utilisées comme des munitions dans les discussions politiques et que, selon la devise "if anything can go wrong, it will go wrong", il y a de grandes chances pour que certaines déclarations soient tirées de leur contexte et gonflées. En conséquence, nous aimerions recommander à l'équipe d'évaluation, lors de la rédaction de rapports destinés au grand public, de respecter le principe "dur sur le fond, modéré dans le ton".

Documentation complémentaire

A propos des objectifs possibles d'utilisation des évaluations, on trouvera des renseignements avant tout dans Chelimsky 1987. Des indications plus détaillées concernant l'appréciation prospective des effets figurent dans Bussmann 1992. Böhret/Hugger 1980 traitent des modèles de décision. Rose 1993 se penche sur les possibilités de tirer la leçon des expériences d'autres secteurs et de celles acquises lors de mesures précédentes. Dans la perspective de l'encouragement des processus d'apprentissage, parmi les projets du PNR 27, avant tout Rey et al. 1993, Bierter/Binder 1993 et Hirsch et al. 1995 donnent une idée des méthodes interactives correspondantes. Les exposés figurant dans Bussmann 1994, et en particulier celui de Binder 1994, abordent la thématique d'apprentissage ou l'utilisation des évaluations dans ce but.

1.6 Prendre des dispositions en vue de mieux saisir à l'avenir l'efficacité des mesures

Quand une mesure nouvelle ou changée a été acceptée par les instances de décision politiques, on pourrait dire: tout est bien qui finit bien. Mais l'action étatique ne connaît pas de fin, elle est un processus permanent. Tôt ou tard, des questions seront posées sur l'effectivité, l'efficacité et l'efficience des mesures prises. Si vous vous souciez de pouvoir justifier plus tard l'efficacité des mesures, ce que nous supposons, alors c'est maintenant que vous devez prendre les dispositions appropriées. Sans cela, une évaluation ne pourra être exécutée qu'à grands frais ou, selon les circonstances, seulement par des méthodes discutables. Si vous souhaitez améliorer l'évaluabilité, vous pouvez commencer dans les domaines suivants:

Collecte des données d'exécution

Les données d'exécution se présentent lors de l'activité de l'administration. Elles jettent une lumière tant sur les prestations de l'administration (p.ex. le personnel engagé pour certaines activités, les mesures prises, les autorisations accordées, les subventions fournies) que sur le comportement des destinataires de la mesure (p.ex. nombre des questions ou demandes formulées, nombre des installations ayant besoin d'assainissement, ampleur des travaux de construction entrepris). En règle générale, les données d'exécution sont très faciles à recueillir. Les instances d'exécution peuvent être priées ou avoir l'obligation de fournir périodiquement certaines données importantes pour l'appréciation de l'exécution ou de son efficacité. Une situation particulièrement favorable à l'accès aux données est celle où une prestation étatique est fournie (subvention, octroi d'un permis de construire, etc.).

Dans la détermination des données à saisir, il s'agit d'apprécier aussi bien que possible non seulement les besoins actuels, mais aussi les besoins futurs. Quels problèmes pourraient surgir au niveau politique? Sur quels aspects des effets faudrait-il à l'avenir disposer d'informations fondées? Quelles mesures partielles pourraient donner lieu à des discussions quant à leur efficacité? Il vaut la peine de réfléchir assez tôt à ces questions et à d'autres. Naturellement, il n'est pas possible de prévoir tous les besoins d'information. De même, il est inutile de prévoir des obligations d'information pour tous les cas concevables. Là comme partout, le principe de la proportionnalité s'applique. En partie, on ne pourra éviter que, dans le cas de besoins d'information surgissant rapidement, on doive recueillir de nouvelles données à court terme. En même temps, il faut rappeler que les séries de chiffres à plus long terme peuvent être importantes pour l'appréciation de l'efficacité d'une mesure. Lors de tout changement de mesure, il convient d'éclaircir si, jusqu'ici, on a saisi des données qui ne sont plus que d'une faible utilité ou qui pourraient être saisies plus simplement et mieux d'une autre manière. Cela peut arriver par exemple quand vous avez recueilli jusqu'ici de grandes quantités de données et que vous pouvez désormais vous limiter à un échantillon de l'ensemble pour lequel on saisira des données sensiblement plus détaillées.

Traitement et exploitation des données d'exécution

La saisie des données d'exécution ne vaut la peine que si on en assure le traitement et l'exploitation. Prenez contact avec les services responsables de tâches statistiques et demandez au besoin un conseil méthodologique auprès de l'Office fédéral de la statistique. Pensez assez tôt à réserver le personnel et les appareils informatiques nécessaires. Mieux vous planifierez le traitement et l'exploitation, moins vous devrez engager par la suite de personnel pour cela (utilisation d'informations exploitables par machine). Veillez à ce que les données puissent être exploitées dans la perspective de différents aspects. Il existe aujourd'hui des systèmes de banques de données faciles à manier, qui rendent possibles de telles analyses. Prenez également contact en temps utile avec l'instance de protection des données, afin de vous assurer que les exploitations ne permettent pas de tirer de déductions concernant des personnes déterminées. Une exploitation significative peut aller de pair avec le respect des principes de la protection des données.

Rendre possibles des comparaisons en vue d'une meilleure évaluabilité de la mesure

Les identifications d'effets les plus fiables sont celles qui se fondent sur des comparaisons appropriées (cf. chap. 3.5).

Nous mentionnerons en premier lieu la comparaison longitudinale, dans laquelle on compare la situation avant et après l'entrée en vigueur d'une mesure. A cet égard toutefois, il est important que l'influence de facteurs qui fluctuent dans le temps (p.ex. situation conjoncturelle) soit maintenue constante. En règle générale, une comparaison longitudinale est d'autant plus opportune que le changement de la mesure apparaît marqué. Cela contredit en partie la tradition des petits changements graduels. Dans tous les cas, nous vous recommandons de saisir la situation réelle avant l'entrée en vigueur de la mesure de manière aussi détaillée que possible, afin de pouvoir par la suite établir une comparaison temporelle.

Comme deuxième et troisième possibilités, nous mentionnerons la comparaison avec groupe de contrôle et la comparaison transversale. L'identification des effets est facilitée quand la mesure prise en considération n'est pas étendue à tous les destinataires ou à tous les domaines possibles. Certains groupes de personnes, branches ou domaines sont ainsi exclus de la réglementation. De ce fait, il devient possible de comparer la situation des destinataires de la réglementation (groupe cible) à celle d'un groupe de personnes qui n'est pas soumis à cette réglementation (comparaison avec groupe de contrôle), ou de comparer la situation entre plusieurs domaines (comparaison transversale).

Exemple 1: L'étude de l'impact sur l'environnement n'est prescrite que pour les projets de construction à partir d'un certain ordre de grandeur, p.ex. à partir de 200 places de stationnement pour les parkings en étages. Grâce à cette valeur limite, une exploitation statistique des effets exercés par l'étude de l'impact sur l'environnement sur la durée de la procédure de construction peut être menée. Par exemple, on peut chercher à savoir si la valeur limite de 200 places de stationnement a une influence marquée sur la durée de la procédure de construction. L'analyse d'un certain nombre de procédures de constructions soumises et non soumises à l'obligation d'étude d'impact sur l'environnement (EIE) peut donner par exemple les images suivantes:

Figure 1

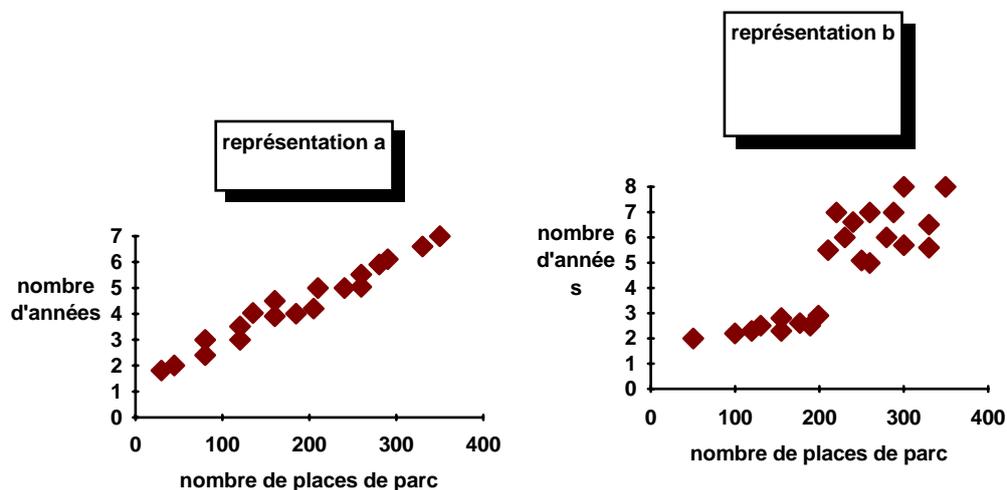


Figure 1: Durée de la procédure de construction en fonction du nombre de places de stationnement prévues dans le projet de construction (exemple hypothétique)

Dans la représentation a, la durée de la procédure de construction dépend du nombre de places de stationnement; l'obligation d'EIE n'exerce pas d'influence là-dessus. La représentation b fait apparaître une autre image: au-dessus de la valeur limite de 200 places de stationnement, on constate une augmentation sensible de la durée de la procédure de construction; dans ce cas, l'obligation d'EIE exerce une influence significative.

Exemple 2: Les aides mises à disposition par la loi fédérale sur l'encouragement du crédit à l'hôtellerie et aux stations de villégiature ne sont pas demandées par tous les hôtels qui pourraient bénéficier des prestations correspondantes. Différentes raisons expliquent cela (manque de connaissance des possibilités d'aide, base en capital propre suffisante, principes réglementaires). La comparaison entre les hôtels qui ont demandé des prestations (cautionnements, contributions aux frais d'intérêts) et ceux qui ne l'ont pas fait permet de tirer des conclusions sur la manière dont les hôteliers apprécient la promotion de l'hôtellerie et sur les effets des mesures (étude Hanser/Kuster/Cavelti 1994).

Exemple 3: Alors que la forme de la semi-détention pour les courtes peines privatives de liberté jusqu'à 3 mois est déjà possible depuis le 1er janvier 1974, les cantons peuvent, depuis le 1er janvier 1986, à titre expérimental, faire exécuter sous forme de semi-détention

également les peines d'emprisonnement et de détention de 3 à 6 mois. Pour les peines privatives de liberté jusqu'à une durée maximum de 30 jours, les cantons ont en outre, depuis le 1er mai 1990, la faculté de choisir comme forme d'exécution un travail d'utilité publique. Les cantons ont l'obligation d'analyser les expériences qu'ils font avec ces nouvelles formes d'exécution. Le fait que celles-ci ne soient pas appliquées systématiquement permet la comparaison avec les formes traditionnelles d'exécution des peines du point de vue des coûts, de l'aptitude à l'exécution et de la resocialisation.

Le fait que les destinataires "glissent à travers les mailles de la loi" ou que les cantons n'appliquent pas une prescription ou ne le fassent que de manière très inégale est peut-être contestable du point de vue éthique, car cela contrevient au postulat de l'égalité. Pourtant, l'inégalité de traitement, les réglementations d'exception et l'inégalité d'exécution offrent des possibilités d'en apprendre plus sur les effets de l'exécution. Notamment dans le cas de nouvelles mesures ou de conditions cadres incertaines, elles constituent une bonne occasion de créer de meilleures bases en vue de décisions ultérieures - conçues comme définitives. Au besoin, de telles mesures peuvent prendre juridiquement la forme de réglementations à titre d'essai; elles sont limitées dans le temps et ne sont maintenues que si elles font leurs preuves.

Exemple: La question de la remise d'héroïne aux toxicomanes est connue et contestée dans l'opinion publique. Sur la base de la loi sur les stupéfiants, 700 toxicomanes reçoivent dans les villes suisses, sous contrôle médical, de l'héroïne, de la morphine ou de la méthadone. Mais il y a à côté de cela différentes autres réglementations à titre d'essai, avant tout dans la circulation routière (p.ex. 50 km/heure à l'intérieur des localités), dans les médias (p.ex. essais de radios locales), dans la création et l'utilisation de banques de données électroniques et dans le domaine scolaire (p.ex. semaine de cinq jours, horaire continu).

Nous vous suggérons de considérer la question de l'évaluabilité dès la phase de formulation de la mesure et de l'inscrire à l'ordre du jour des questions à traiter. Si, dans votre service ou dans la commission qui traite de la nouvelle formulation de la mesure, on a trop peu de connaissances solides sur les stratégies d'évaluation, une personne connaissant bien ces questions pourra vous donner relativement facilement (et à peu de frais) des conseils en vue d'améliorer l'évaluabilité de la mesure.

Documentation complémentaire

Mader 1988 et 1993 et Mastronardi 1991 donnent un point de vue juridique sur les réglementations à titre d'essai.

2. Qu'entend-on par évaluations?

2.1 La notion d'évaluation

La notion d'évaluation se réfère à l'appréciation d'un objet ou d'un état de fait. Elle peut être utilisée dans un sens très large, qui ne limite ni l'état de fait à évaluer (produit, institution, mesure, domaine politique assez vaste), ni l'objectif de l'évaluation (décision d'achat, réorganisation, amélioration d'une mesure), ni la stratégie d'évaluation (méthodes de recherche relevant des sciences sociales, appréciation impressionniste), ni l'instance d'évaluation (autorité d'exécution, gouvernement, autorité de surveillance, recherche scientifique libre, etc.).

Dans notre contexte toutefois, nous nous limitons au domaine public (étant entendu que les stratégies d'évaluation peuvent parfaitement être appliquées à certaines questions qui se posent dans des entreprises, par exemple des programmes de formation interne ou des mesures de promotion culturelle propres à une entreprise).

Deuxièmement, nous nous concentrons sur les évaluations systématiques ou scientifiques, c'est-à-dire les évaluations qui

- se fondent sur des réflexions et des affirmations reproductibles,
- reposent sur une procédure systématique qui vise à saisir de façon complète les effets les plus significatifs,
- visent à acquérir des indications aussi objectives que possible, et
- sont présentées sous une forme qui atteste leur caractère méthodique.

Troisièmement, nous nous concentrons sur l'évaluation de mesures étatiques⁵. Il s'agit d'activités de l'Etat isolables analytiquement dans la perspective de buts déterminés, parfois formulés de manière vague, il est vrai, qui se réfèrent à des changements dans l'économie et la société. Elles comportent en règle générale des ressources en personnel et en moyens financiers.

2.2 Que peut-on, en fait, évaluer?

L'évaluation des mesures étatiques couvre un large domaine. Peuvent faire l'objet d'évaluations:

- des lois entières ou des dispositions particulières,
- des subventions (p.ex. contributions aux chemins forestiers, aux briseurs d'avalanches),
- institutions et établissements (p.ex. écoles, entreprises de transport, institutions soutenues par l'Etat telles que les offices du tourisme ou les associations d'utilité publique),
- installations (p.ex. routes, lignes de chemin de fer ou de bus).

La méthodologie utilisée pour tous ces objets d'évaluation est très semblable, de même que pour les programmes - paquets de mesures orientés sur des objectifs déterminés. Nous reviendrons en détail, dans la suite de ce guide (troisième partie), sur les stratégies d'évaluation et méthodes de recherche correspondantes; pour l'instant, nous nous limiterons à des considérations sur l'évaluation de domaines politiques entiers (p.ex. le système suisse de formation, la politique agricole, etc.). Les stratégies usuelles d'évaluation des mesures et programmes étatiques, comparatives ou non comparatives, ne conviennent ici que dans une mesure limitée, étant donné que différentes mesures agissent ensemble de telle manière qu'on ne peut distinguer avec précision leur contribution au résultat global. Les conditions préalables à une évaluation sont analogues à la situation d'un patient qui absorbe en même

⁵ Nous utilisons partout le terme de "mesure". Les différences par rapport à la notion de programme, qui domine dans la littérature d'évaluation anglo-saxonne, sont de nature graduelle. Aux Etats-Unis, le programme n'est pas simplement une notion analytique, mais aussi une notion de droit, ce qui fait que la reprise littérale de ce terme nous a paru problématique.

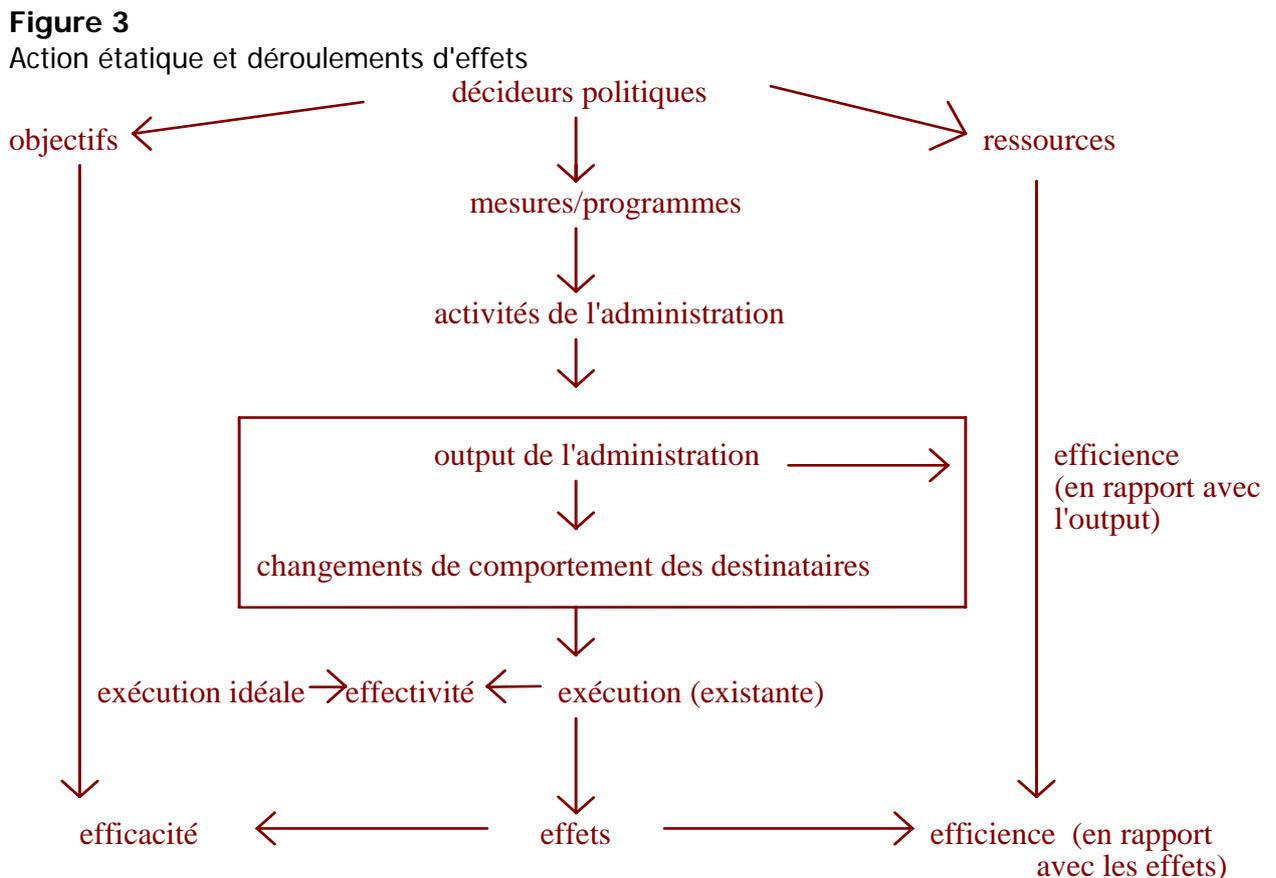
temps plusieurs médicaments; de ce fait, il va être difficilement possible d'estimer l'influence d'un médicament précis sur son état de santé. Un objet global - un domaine politique entier - demande aussi des méthodes d'évaluation globales. On peut se servir pour cela, notamment, des moyens suivants: analyse d'indicateurs de performance existants (p.ex., dans le système éducatif, du nombre de ceux qui arrivent au terme des études primaires, secondaires et tertiaires) et des frais; appréciation par des experts étrangers (peer review), en partie à l'oeuvre dans des organisations internationales (p.ex. OCDE, Banque mondiale, OMS).

Figure 2
 Méthodologie d'enquête pour les mesures étatiques, programmes et domaines politiques

| Objet | méthode d'enquête principale |
|---------------------|---|
| Mesures étatiques | stratégies d'évaluation comparatives/non comparatives |
| Programmes | dito |
| Domaines politiques | analyses quantitatives, appréciations globales, peer review |

2.3 Quels aspects et critères les évaluations peuvent-elles contenir?

Entre ce que font les fonctionnaires et ce qui se passe dans la réalité sociale, il existe des déroulements d'effets complexes. Il y a plusieurs possibilités de récapituler et de classer les divers activités et états de faits; les notions ne sont pas (encore) utilisées de manière uniforme. Nous partons de la représentation suivante:



Les ressources de l'administration comprennent les compétences juridiques, les moyens matériels, les moyens financiers et le personnel.

L'output correspond aux biens produits et aux services fournis par l'administration (imprimés, ordonnances, autorisations, amendes, subventions, etc.).

L'exécution comprend le jeu réciproque des activités de l'administration, d'une part, et des destinataires visés par les mesures étatiques, de l'autre. Le critère déterminant de la qualité de l'exécution est l'effectivité. On considère là le degré de concordance entre le comportement juridiquement exigé ou prévu des instances d'exécution et/ou des destinataires et leur comportement effectivement observé. Selon le type de mesure, la question de l'effectivité se pose de manière différente: degré de respect en cas d'injonction et d'interdictions, degré d'application s'il s'agit de normes de droit public nécessitant une mise en oeuvre par des instances administratives (p.ex. octroi d'autorisations), degré d'utilisation pour les règles de droit privé à caractère facultatif et pour les aides financières, degré d'observation pour les informations et recommandations. L'effectivité est une condition nécessaire mais non suffisante de l'efficacité.

Les effets sont les transformations imputables à une mesure qui interviennent dans l'économie et la société. Deux critères sont déterminants ici. Le premier est celui de l'efficacité, c'est-à-dire la relation entre les effets visés et les effets réellement intervenus. En termes de grandeur quantitative, on appelle cette relation le degré de réalisation de l'objectif. La constatation et l'appréciation de l'efficacité présupposent la détermination des objectifs des mesures étatiques.

Le second critère est celui de l'efficience; il se réfère au rapport entre l'output, ou les effets obtenus, et les ressources engagées (efficience en rapport avec l'output, ou efficience en rapport avec les effets). On atteint une meilleure efficience soit en engageant moins de ressources pour obtenir un certain output ou certains effets en rapport avec le but, soit en obtenant avec les mêmes ressources plus d'output ou plus d'effets en rapport avec le but. Par efficience, on entend souvent aussi, dans la recherche évaluative, le rapport, exprimé en grandeurs monétaires, entre les avantages/effets et les ressources/moyens investis.

2.4 Quel horizon temporel couvrent les évaluations?

Dans l'usage linguistique courant, on entend souvent par évaluation le choix d'un produit dans la perspective d'une utilisation ultérieure. Par exemple, un service étudie l'utilité pratique d'un nouvel équipement informatique, ou le Département militaire fédéral se demande quel type de vélo militaire pourrait convenir à la troupe durant les prochaines décennies. En règle générale, une telle évaluation intervient sur la base de modèles existants ou, le cas échéant, de prototypes. Ceux-ci sont comparés les uns aux autres au moyen de descriptions de produits (stations de travail), ou soumis à un test sur le terrain (vélos militaires).

A la différence des installations informatiques ou des vélos militaires, nous ne savons souvent que très approximativement, dans le cas des mesures étatiques, comment est fait le "produit", c'est-à-dire comment l'exécution se déroule et quels effets la mesure suscite. Ces informations ne peuvent être obtenues que par la saisie des processus qui se sont déroulés jusqu'ici. C'est pourquoi les évaluations ont très souvent un caractère rétrospectif (évaluations ex post).

Il s'agit là d'enquêtes empiriques, c'est-à-dire d'enquêtes qui essaient de saisir les états de faits dans leurs manifestations concrètes. Les évaluations se distinguent ainsi des appréciations de mesures étatiques définies sur la base de simples considérations théoriques, p.ex. par des économistes qui se fondent sur des théories économiques. En règle générale, les évaluations ne sont pas orientées sur la construction de théories ou l'élaboration d'un savoir empirique largement applicable (p.ex. sur les instruments de l'action étatique), mais elles peuvent contribuer à l'une et à l'autre. Alors que, pendant longtemps, les évaluations ont été comprises d'abord comme un instrument permettant de constater les effets, on insiste à

nouveau plus fortement, depuis la fin des années 80, sur la référence théorique de la recherche évaluative.

Le caractère rétrospectif des évaluations n'est pas un obstacle à leur utilisation dans le changement et/ou la conception nouvelle d'une mesure, bien au contraire. La connaissance de l'effectivité, de l'efficacité et de l'efficience de la mesure en vigueur permet de définir des points de repère touchant la nécessité et l'orientation d'un changement ou d'un remaniement. Ces connaissances facilitent en outre l'appréciation des effets présumés d'une nouvelle mesure envisagée. Beaucoup d'évaluations ont, certes, un caractère rétrospectif, mais servent de base à l'action future.

Les évaluations prospectives (ou *ex ante*) visent à évaluer les effets ou l'efficience de mesures possibles ou prévues. Il ne s'agit plus d'enquêtes empiriques, mais de projections de connaissances existantes dans l'avenir. Ces évaluations peuvent avoir un caractère plutôt impressionniste. Elles peuvent se servir de certaines techniques comme par exemple celle des scénarios. Mais elles peuvent aussi reposer sur des modèles d'effets perfectionnés. Dans les domaines où existent des théories éprouvées sur la structure des effets (avant tout en politique économique, agricole, énergétique et de l'environnement), on peut utiliser les modèles correspondants tant pour la représentation rétrospective que (prenant pour hypothèse certaines conditions cadres) pour la simulation des effets de mesures changées. De manière tout à fait générale, on vise à lier plus étroitement les évaluations rétrospectives et prospectives. Les évaluations de l'efficience sont très souvent utilisées de manière prospective, afin d'estimer l'opportunité de grands projets d'investissement (p.ex. construction de routes, de voies de chemin de fer).

- Vous trouverez de plus amples détails sur l'engagement des évaluations en vue de changer ou de remanier des mesures au chapitre 1.5.

2.5 Que coûtent les évaluations et combien de temps durent-elles?

En règle générale, les évaluations sont coûteuses, en temps et en moyens financiers. Pour une étude approfondie (évaluation des effets ou de l'efficience), il faut compter avec une durée de six mois à deux ans et un coût de 50'000 à 300'000 francs. Pour un objet de recherche clairement limité et qui n'est pas trop complexe, on peut obtenir des résultats tout à fait utilisables dans un délai d'environ six mois et avec un budget inférieur à 100'000 francs. Des questions encore plus limitées, p.ex. des éclaircissements tout à fait simples concernant l'exécution, coûtent parfois moins de 50'000 francs. Les frais se répartissent en trois parties à peu près égales couvrant respectivement la préparation et la planification détaillée du projet, l'exécution proprement dite et la mise en valeur des résultats (présentations écrites, orales).

La recherche en sciences sociales n'est pas bon marché, car elle exige généralement un investissement en personnel très élevé. Pour ce qui est des techniques de recherche (cf. chap. 3.6), les interviews individuelles (p.ex. interviews en profondeur), les sondages portant sur un large échantillon et les exploitations statistiques avec important dépouillement de données ou élaboration de nouveaux modèles économétriques coûtent relativement cher. Inversement, les analyses de documents, les recherches documentaires, les sondages par téléphone ou par écrit (p.ex. interrogation de groupes cibles) et les exploitations statistiques plus simples sont comparativement bon marché. Il ne faut pas sous-estimer le coût des mesures de mise en valeur des résultats (p.ex. sous forme de manifestations de perfectionnement ou de brochures présentées de manière attrayante). Une stratégie plus chère peut parfaitement se justifier si elle promet des réponses meilleures et plus fiables à vos questions.

Les équipes d'évaluation sont parfaitement en mesure de respecter les calendriers et les budgets avec une marge d'erreur de 10%. Ce qu'on inclut trop peu dans la planification, ce sont les coûts de prise en charge des groupes d'accompagnement, de traduction et, de manière plus générale, de mise en valeur des résultats.

Documentation complémentaire

La meilleure introduction à l'objet et à la méthodologie des évaluations se trouve dans Rossi/Freeman 1993; une édition précédente de cet ouvrage, parfois entachée encore de certains défauts, a été traduite en allemand (Rossi/Freeman/Hofmann 1988). Le rapport du groupe de travail "Méthodes d'évaluation des politiques publiques" (1986) et Monnier (1987) peuvent être recommandés. Dans la perspective de la gestion de l'administration, il vaut toujours la peine de lire Wholey 1983. AGREB 1990, aux pages 12-14, examine les avantages et les inconvénients de la recherche interne et externe.

3. De quelles stratégies et techniques de recherche les évaluations se servent-elles?

3.1 Possibilités de classification des stratégies d'évaluation

Dans les limites de ce guide, il n'est pas possible de présenter dans le détail la méthodologie des évaluations. Un manuel spécial paraîtra à ce sujet en 1995/96 dans le cadre du PNR 27; il existe en outre différents ouvrages d'enseignement (cf. indications bibliographiques données à propos des chapitres ci-après).

Toutefois, une bonne partie des documents sur les évaluations sont formulés de manière plus détaillée que ne le souhaiteraient les prescripteurs d'évaluations. De plus, rares sont les ouvrages faits pour des pays tels que la Suisse. C'est pourquoi nous allons essayer de donner, dans les limites de cet ouvrage, une explication de base des principales stratégies d'évaluation et méthodes de recherche.

Il y a différentes possibilités de représenter les choses. Chacune part de critères de valeur implicites sur les méthodes appropriées et moins appropriées. Il serait faux de voir dans les considérations qui vont suivre un "livre de cuisine" sur les stratégies d'évaluation et de l'appliquer aveuglément. Chaque mode de représentation - celui-ci comme les autres - a ses avantages et ses inconvénients, ses possibilités et ses limites. Nous voyons là un point de départ, mais non un point final, à une compréhension approfondie de la méthodologie des évaluations.

Il y a un grand nombre de tentatives de classification en matière de recherche évaluative. Aux Etats-Unis, c'est notamment la distinction entre méthodes expérimentales, quasi expérimentales et "autres" qui domine (p.ex. Rossi/Freeman 1993). Toutefois, une transposition de ce système aux conditions européennes est difficile du fait que les expériences proprement dites sont plus rares en Europe qu'aux Etats-Unis et que, par conséquent, une grande partie de la méthodologie d'évaluation américaine ne peut être appliquée pratiquement que rarement. Aux Etats-Unis, les critères établis par la Société de recherche évaluative (ERS Standards Committee 1982) sont également significatifs; ces critères distinguent six types d'évaluation différents (évaluations prospectives, études d'évaluabilité, études d'exécution, études des effets, monitoring, méta-évaluations). Nous nous appuyons ici sur cette systématique, complétée par des éléments développés par le US General Accounting Office (1991a); elle se caractérise par des expériences pratiques en matière d'évaluations et apparaît facilement transférable à la situation de pays tels que la Suisse.

3.2 Notions de stratégie d'évaluation et de technique de recherche

Nous distinguons entre stratégie d'évaluation, technique de recherche (ou méthode de recherche) et design d'évaluation.

- La stratégie d'évaluation est la base de départ de l'enquête visant à répondre aux questions posées dans la perspective de l'exécution ou des effets de la mesure étatique. Elle est la réponse conceptuelle à la question de recherche.
- La technique de recherche est la méthode de sciences sociales (ou naturelles) choisie pour collecter et exploiter des données en vue de répondre à la question de recherche.
- Comme on le verra au chapitre suivant, lors de questions normatives et descriptives, la stratégie d'évaluation peut consister dans une large mesure en l'utilisation de certaines techniques de recherche. Face à des questions à orientation causale, il faut en règle générale une base de comparaison implicite ou explicite sur laquelle se fonde la stratégie d'évaluation.
- Le design d'évaluation est la structure d'enquête concrète mise en place pour répondre aux questions posées. Il contient des données sur la stratégie d'évaluation et les techniques de recherche choisies.

3.3 Aperçu des stratégies d'évaluation

3.31 Questions descriptives, normatives et à orientation causale

Les évaluations de l'exécution ou des effets peuvent contenir des questions descriptives, normatives ou à orientation causale.

Les questions descriptives concernent, comme leur nom l'indique, la description d'activités, d'objets ou d'états de faits. Par exemple, les rapports sur l'exécution fournis par l'administration ont en règle générale un caractère descriptif.

Dans le cas des questions normatives, il s'agit de comparer entre ce qui est et ce qui doit être. Les examens entrepris par les organes de contrôle partent généralement de questions normatives.

Lors de questions à orientation causale, il s'agit d'établir des rapports (de cause à effet) entre des activités et des situations. Les recherches scientifiques sont souvent fondées sur de telles questions.

Partant du domaine de la politique de protection des eaux, voici des exemples de ces trois genres de questions:

Figure 4

Questions descriptives, normatives et à orientation causale

| Question | Exécution | Objectifs ou effets |
|------------------------------------|--|--|
| descriptive | Combien y a-t-il de stations d'épuration en fonction en Suisse? | Comment la qualité des eaux s'est-elle développée au cours des dix dernières années? ⁶ |
| normative | Les prescriptions de protection des eaux sont-elles respectées dans le canton X dans le cadre de la procédure des autorisations de construction? | Dans combien d'eaux la qualité d'eau prescrite est-elle atteinte? ⁷ |
| à orientation causale | Dans quelle mesure la suppression des subventions à la construction de stations d'épuration a-t-elle exercé des effets? | Dans quelle mesure la politique de protection des eaux a-t-elle contribué à l'amélioration de la qualité de celles-ci? Quels effets secondaires se sont exercés sur le développement économique? |
| normative et à orientation causale | A quelles causes faut-il imputer les déficits (ou succès) d'exécution dans la protection des eaux? | Pourquoi n'a-t-on pas encore atteint le niveau prescrit de qualité des eaux dans les zones à forte densité démographique? |

3.32 Genres de questions et stratégies d'évaluation

La manière de formuler les questions - descriptive, normative ou à orientation causale - a une influence déterminante sur la stratégie d'évaluation. Les corrélations sont représentées à la figure 5 et seront commentées dans les chapitres suivants.

⁶ Il s'agit de la description de l'évolution au moyen d'un indicateur important. La question de savoir dans quelle mesure la qualité des eaux est imputable à la politique de protection des eaux ou à d'autres facteurs (p.ex. situation conjoncturelle) demeure ouverte.

⁷ Là aussi, on ne met en lumière que l'objectif quantitatif et on laisse ouverte la question de savoir si et dans quelle mesure la politique de protection des eaux a contribué au résultat.

Figure 5
Genres de questions et stratégies d'évaluation

| Genre de question | Stratégie d'évaluation | Combinaisons |
|---------------------------------|---|---|
| Questions descriptives | pas de stratégie d'évaluation spécifique (voir sous techniques de recherche) | combinaison avec questions à orientation causale et/ou normatives |
| Questions normatives | comparaison situation existante/idéale (mesure des écarts) | combinaison avec questions à orientation causale et/ou descriptives |
| Questions à orientation causale | stratégies explicitement comparatives | |
| | groupes de contrôle | combinaison avec comparaison longitudinale |
| | expériences groupes de contrôle non équivalents analyses de discontinuité-régression | |
| | comparaison transversale | combinaison avec comparaison longitudinale |
| | comparaison longitudinale (comparaison avant-après) | combinaison avec comparaison transversale |
| | stratégies implicitement comparatives, p.ex. - appréciation par les destinataires - interrogation de personnes clés - reconstruction complète du déroulement - recours aux prix fictifs | |

3.4 Stratégies d'évaluation pour les questions descriptives et normatives

Bien que, pour les évaluations, les questions normatives et à orientation causale constituent le principal défi, il y a dans chaque évaluation le plus souvent aussi des questions descriptives auxquelles il s'agit de répondre. Elles peuvent prendre les formes les plus diverses: Combien de temps faut-il en moyenne à un employé pour arriver au bout d'un dossier? Comment se déroule un entretien d'orientation pour les chômeurs? Comment se présente la situation des revenus des bénéficiaires de paiements directs dans l'agriculture? Quelles personnes fréquentent les centres de loisirs municipaux?

Les réponses aux questions descriptives peuvent être une description des états de faits et activités en cause. Une stratégie d'évaluation particulière n'est pas nécessaire. On fait intervenir tout l'éventail des techniques de recherche des sciences sociales (cf. chap. 3.6). Les instituts universitaires et les bureaux privés ont une expérience suffisante de l'approche des questions descriptives. Ils traitent ces questions avec compétence, même si parfois des questions difficiles se posent, telles par exemple celles-ci: Les données fiscales cantonales donnent-elles des renseignements fiables sur la situation des revenus des rentiers AVS? Quel est le meilleur moyen de déterminer la situation économique et sociale des familles à un parent? Les connaissances acquises dans le canton X à propos de la procédure d'étude d'impact sur l'environnement peuvent-elles être étendues à toute la Suisse? Combien de personnes étrangères travaillent-elles illégalement en Suisse? Ainsi, les questions à notre avis les plus délicates concernent la représentativité (et par conséquent aussi la transmissibilité) des données et des connaissances. En outre, la recherche en sciences sociales se heurte à un problème difficile avec les parties de la population "cachées" (p.ex. les immigrants illégaux), ainsi qu'avec les comportements interdits ou socialement proscrits (p.ex. mauvais traitements envers les enfants, emploi illégal, fraude fiscale, toxicomanie).

La réponse aux questions normatives fait partie intégrante de toute évaluation en tant qu'activité d'appréciation. Dans une telle question, il s'agit de comparer entre valeurs idéales et valeurs existantes et c'est là en fait que se trouve déjà la stratégie d'évaluation appropriée. Les sciences sociales disposent d'une expérience suffisante dans le domaine de la vérification des théories et, de manière générale, des techniques de mesure pour répondre avec compétence aux questions normatives.

Dans l'idéal, la réponse aux questions normatives présuppose les étapes suivantes (qui, dans le cas des évaluations concrètes, ne sont pas obligatoirement utilisées dans chaque cas ni dans cet ordre):

- Détermination des impératifs ou des objectifs: En théorie, les objectifs peuvent être déterminés par le recours aux actes de droit et aux documents qui s'y rattachent (messages, débats parlementaires). Mais ces textes sont rédigés généralement de manière assez vague, si bien qu'il faut se tourner vers d'autres matériels (déclarations politiques, critères de valeur des intéressés, etc.). Les règles d'interprétation juridique, notamment, peuvent aider à cerner les objectifs. Le cas échéant, l'équipe d'évaluation peut faire intervenir des impératifs/objectifs hypothétiques à titre de contrôle. Il est concevable aussi d'examiner des états de faits au moyen de plusieurs objectifs.
- Opérationnalisation des impératifs/objectifs: Ceux-ci sont traduits en objectifs de performance déterminés, afin de pouvoir être comparés avec des états de faits concrets. Par exemple, l'objectif de la stabilité du niveau des prix est fixé sur la base d'un taux de renchérissement maximum de 3 pour cent, ou l'objectif de la couverture du minimum vital est fixé à un revenu annuel de 30'000 francs pour une famille de quatre personnes habitant à la campagne.
- Mesure: Les objectifs de performance en tant que valeurs idéales sont comparés avec des états de faits concrets (valeurs existantes).
- Bilan global: Les résultats de plusieurs mesures en rapport avec les différents objectifs de performance sont récapitulés dans une appréciation globale.

Une question normative typique est celle de l'effectivité (conformité de l'exécution): Les autorités et les destinataires des mesures se sont-ils conformés aux dispositions juridiques et aux autres impératifs?

La question de l'efficacité est aussi une question normative, qui présente toutefois aussi un aspect d'orientation causale (voir plus loin). L'aspect d'orientation causale des questions normatives en référence à l'efficacité et à l'efficience est souvent négligé lors de la fixation de la stratégie d'évaluation, comme le montre l'exemple hypothétique ci-après.

Exemple: Pour réduire la consommation d'énergie, on projette d'introduire sur le marché des oeufriers spéciaux qui seront vendus à prix réduit dans le cadre d'une campagne. L'ob-

jectif de la campagne est d'écouler un million d'oeufriers (critère de l'effectivité) et de réduire ainsi la consommation d'énergie de 0,1 pour cent (critère de l'efficacité). Au bout d'un an, il apparaît que les deux critères ont été atteints. Mais une évaluation approfondie (à orientation causale) fait apparaître ensuite que seulement un oeufrier sur dix vendus est finalement utilisé en raison des problèmes posés par la détermination de la durée de cuisson nécessaire, et qu'en conséquence l'abaissement de la consommation d'énergie n'a été que de 0,01 pour cent. La réduction effective de la consommation constatée était due à d'autres facteurs (hiver doux).

Les stratégies d'évaluation visant à répondre à des questions à orientation causale (voir plus loin) touchent aussi implicitement des questions normatives, sans que toutes les équipes d'évaluation en soient pleinement conscientes. Les comparaisons implicites et explicites visant à déterminer les effets fixent la barre à des hauteurs diverses selon l'objet de la comparaison (groupe de contrôle, domaine à comparer, moment de la comparaison).

Exemple: L'évaluation, reposant sur une expérience, d'une nouvelle forme de thérapie pour les prisonniers peut donner des résultats très différents selon la situation de départ des détenus. Si d'autres thérapies ont déjà été entreprises auprès d'eux, il est possible que le succès soit maigre. Si, en revanche, il n'existe pas d'autres possibilités de thérapie, les chances sont plus grandes que l'introduction de la thérapie soit un succès.

Inversement, il peut arriver que des questions délicates se posent, dans le cas d'enquêtes d'abord descriptives ou à orientation causale, du point de vue des aspects normatifs. Dans deux évaluations menées dans le cadre du PNR 27, il est apparu clairement que, lors de l'exécution, des prescriptions juridiques n'avaient pas été respectées par les autorités et/ou les destinataires de la norme. Ces constatations étaient fondées en partie sur les informations confidentielles. Dans les deux projets, ces faits n'ont été mentionnés dans le rapport final que jusqu'à un certain point, car il ne s'agissait pas d'abord de signaler des erreurs individuelles, mais de représenter les conditions de succès de la mesure en cause et en partie aussi les processus d'apprentissage. Notamment face à des questions strictement réglées du point de vue juridique (p.ex. procédure d'autorisations de construction), les équipes d'évaluation doivent avoir conscience que les impératifs juridiques et l'exécution effective peuvent diverger et que ce fait risque de rendre plus difficile l'accès aux informations.

3.5 Stratégies d'évaluation pour les questions à orientation causale

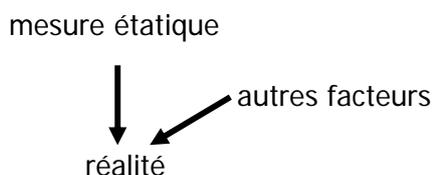
3.51 Stratégies d'évaluation explicitement et implicitement comparatives

Les questions à orientation causale concernent les rapports de cause à effet. Quels ont été les effets de l'obligation d'étude d'impact sur l'environnement sur la participation des associations écologiques à l'application du droit sur la protection de l'environnement? Quels effets les cautionnements et subventions accordés aux hôteliers ont-ils exercé sur le changement des structures dans l'hôtellerie? Quels ont été les effets des valeurs limites sévères imposées aux chauffages d'immeubles sur la situation concurrentielle de l'industrie suisse du chauffage? Ce sont là des questions typiques qui visent à éclaircir les relations de cause à effet.

Dans les évaluations des effets et en partie aussi dans celles de l'exécution, il s'agit de questions à orientation causale. En règle générale, il est très difficile d'y répondre parce que, entre la décision et l'entrée en vigueur d'une mesure, son exécution et ses effets dans la société, il existe la plupart du temps un tissu d'effets très enchevêtré, rarement accessible à l'observation directe. Dans l'idéal, il s'agit d'établir une relation fiable entre la mesure étatique et les changements (de la société, de l'économie, de l'environnement, etc.) qui lui sont imputables et d'isoler ainsi d'autres facteurs d'influence. La question est celle-ci: Qu'est-ce qui serait autrement si la mesure étatique n'était pas exécutée (comparaison "policy on" - "policy off")?

Figure 6

Situation des mesures étatiques par rapport à la réalité sociale



Comment, dans une réalité enchevêtrée et complexe, peut-on saisir des changements qui interviennent souvent lentement, souvent avec retard et parfois de manière furtive? Comment des causes déterminées peuvent-elles être isolées d'autres facteurs d'influence?

La réponse de la recherche évaluative se situe avant tout dans le développement de stratégies d'évaluation comparatives. On essaie de tirer des conclusions sur les relations de cause à effet par des variations en référence aux unités d'enquête (comparaison avec groupe de contrôle), aux unités spatiales (comparaison transversale) et à la perspective temporelle (comparaison longitudinale). Ces stratégies d'évaluation reposant sur une base comparative explicite seront exposées en détail dans les trois chapitres ci-après.

Naturellement, de telles stratégies ne peuvent être appliquées à toutes les mesures. Quand des mesures sont exécutées depuis longtemps sans changement et appliquées de la même manière dans tous les cantons et communes, et sont valables en outre pour tous les destinataires pris en considération, les comparaisons avec groupe de contrôle, les comparaisons transversales (à l'exception des comparaisons internationales) et les comparaisons longitudinales ne sont pas opportunes ou pas possibles.

Dans la mesure où, dans de tels cas, on n'est pas contraint de renoncer complètement aux indications sur les effets des décisions, il s'agit de tirer le meilleur parti de la situation. En tant que solution de rechange, on peut mettre en œuvre des stratégies d'évaluation implicitement comparatives. Ces stratégies ne se servent pas de variations dans le temps existant empiriquement (comparaisons longitudinales) et autres (comparaisons avec groupe de contrôle et transversales), mais recourent à des comparaisons de remplacement sous forme de constructions intellectuelles hypothétiques, d'examen inductifs de modèles d'effets, et de prix fictifs. Dans certaines circonstances, elles peuvent apporter des résultats tout à fait défendables. Les stratégies suivantes se présentent pour déterminer les effets:

- *Appréciation par les destinataires de la mesure*

Les destinataires (ainsi que, le cas échéant, d'autres intéressés) de la mesure peuvent (par des sondages écrits, des interviews en profondeur, des hearings, etc.) être interrogés sur leur appréciation des effets. Ils fonderont leur appréciation sur la base d'expériences précédentes (comparaison longitudinale implicite), de comparaisons transversales (avant tout avec des pays voisins) ou d'une construction intellectuelle hypothétique ("Quelle serait la situation si la mesure n'existait pas?").

Il est clair que les réponses peuvent être influencées par un comportement tactique (réponses à l'avantage de la personne questionnée) ainsi que par des erreurs d'appréciation. Ces dernières peuvent être dues au fait que, par exemple, on n'a pas pris en considération les déroulements d'effets contre-intuitifs (p.ex. la possibilité que l'aide sociale diminue l'initiative propre et le respect de soi des bénéficiaires) et les effets secondaires. Les équipes d'évaluation chevronnées sont capables de contrôler dans une certaine mesure ces sources d'erreur en faisant intervenir dans le processus de dialogue les arguments et questions correspondants. Mais on ne peut entièrement exclure que l'ignorance, les préjugés et l'état d'esprit du moment n'influencent les réponses des destinataires interrogés.

- *Consultation de personnes clés*

Des personnes clés telles qu'experts du domaine concerné, spécialistes de l'administration et représentants de groupes d'intérêts peuvent constituer une autre source d'informations sur les relations de cause à effet. Leur connaissance du processus doit contribuer à reconstruire les effets. En particulier, les personnes clés dans l'exécution des mesures peuvent avoir des connaissances précieuses sur des aspects centraux de la piste d'effets. Là aussi, la détermination des effets intervient en partie sur la base de constructions intellectuelles hypothétiques.

Mais, comme les destinataires (voir plus haut), les experts et les personnes clés de l'exécution peuvent commettre des erreurs d'appréciation. Leur prise de position peut être influencée par l'état d'esprit du moment (p.ex. euphorie ou scepticisme concernant certains développements). Précisément dans des réseaux communiquant étroitement entre eux, certains protagonistes, par leur force institutionnelle ou leur personnalité, peuvent influencer les opinions de telle sorte que même des interrogations menées sur une large base peuvent dégager une opinion très uniforme.

Les équipes d'évaluation chevronnées sont conscientes de ces dangers. Elles recherchent systématiquement d'autres déclarations auxquelles elles confrontent les personnes interrogées. Dans un processus itératif, les résultats dégagés dans une première étape sont encore une fois soumis aux personnes clés afin de les inciter, le cas échéant, à repenser leur opinion. Au besoin, le processus d'évaluation peut aussi être conçu comme un processus d'apprentissage commun dans lequel on veille à ce que chaque opinion ait la même valeur, quelle que soit la position hiérarchique de la personne qui l'exprime. Naturellement, même dans le cadre de tels processus d'évaluation conçus avec soin, il sera très difficile de reconstruire un résultat global fiable à partir de perspectives divergentes et nuancées concernant la structure des effets.

- *Reconstruction complète du déroulement/étude(s) de cas approfondie(s)*

Déjà l'interrogation de personnes clés peut être marquée par l'effort en vue de reconstruire de manière aussi précise que possible la structure des effets. Il s'agit toutefois d'une option autonome qui se sert de différentes techniques de collecte des données (analyses de documents, sondages, interviews en profondeur, analyses quantitatives de données, etc.). L'idée est de procéder à une reconstitution sans lacune d'un cas (cf. chap. 3.62). Ces connaissances étendues permettent d'exclure certains effets par voie inductive, dans la mesure où on ne trouve pas les indices correspondants qui sont la condition préalable à l'apparition des effets. Naturellement, une large marge de possibilités subsiste et il est très difficile de juger comment les différents éléments de la chaîne d'effets s'assemblent en un tout, l'effet.

- *Recours aux prix fictifs*

Un moyen de dépannage pour discerner les effets est constitué, dans les évaluations s'inspirant de la réflexion économique, par les prix du marché ou prix fictifs en tant qu'estimations de prestations qu'on peut théoriquement obtenir sur le marché. On peut par exemple déterminer approximativement les effets économiques de certaines constructions de routes en partant du renchérissement des terrains qu'elles mettent en valeur directement ou indirectement. On peut aussi examiner si et dans quelle mesure une réduction de la pollution sonore dans le voisinage d'un aérodrome se répercute sur les prix du terrain dans les communes concernées.

Naturellement, le marché n'est qu'un instrument très imparfait de détermination des effets. Les bulletins de vote économiques sont répartis de manière très inégale. Par exemple, on peut estimer que le démantèlement de la scène de la drogue en ville exerce des effets économiquement positifs sur les chiffres d'affaires des commerces du voisinage, alors qu'on ne prend pas en compte les frais encourus par les toxicomanes ou leurs familles en raison du déplacement de la scène de la drogue. L'analyse peut en soi être affinée

peu à peu pour prendre en compte aussi de telles objections. Néanmoins, des éléments impondérables demeurent, qui tiennent au fait que les "prix du marché" de certaines données institutionnelles (p.ex. prix du marché mondial, prix de cartels, prix de monopoles) sont dépendants et que ces contingences influencent le résultat de l'évaluation. En outre, il faut constater qu'une quantification monétaire de certains états de faits est la plupart du temps très difficile. Ce sont avant tout les évaluations des effets (présumés) de la déréglementation économique qui se servent de tels concepts de prix.

Ainsi, les stratégies d'évaluation implicitement comparatives présentent des risques et des chances quand il s'agit de répondre à des questions à orientation causale. Vous trouverez au chapitre 3.6 un certain nombre d'autres indications sur l'aptitude des méthodes de recherche à donner des éclaircissements sur les relations de cause à effet.

3.52 Comparaison avec groupe de contrôle

La stratégie d'évaluation de la comparaison avec groupe de contrôle (ou comparaison avec ou sans mesure) sert à répondre aux questions à orientation causale. La question typique pour une évaluation des effets est généralement celle-ci: Qu'est-ce qui serait différent dans la réalité si la mesure n'était pas mise en oeuvre? C'est précisément à cette question que veut répondre la comparaison avec groupe de contrôle. Dans le domaine de la recherche évaluative, elle est traitée à la rubrique des "méthodes expérimentales et quasi expérimentales". Nous mentionnerons ici les possibilités suivantes de comparaison avec groupe de contrôle:

Expérience (avec choix au hasard): A partir d'un ensemble, on soumet un échantillon de personnes choisies au hasard à une mesure déterminée (dans le jargon de l'évaluation: "traitement"). La situation constatée dans le groupe cible est comparée à celle d'un groupe de contrôle (lui aussi choisi au hasard) qui n'est pas soumis à la mesure. On déduit de la comparaison entre les situations des deux groupes les effets de la mesure.

Exemple: L'objectif d'une expérience est d'essayer une nouvelle forme de thérapie pour les prisonniers. On effectue une enquête auprès des détenus afin de savoir qui s'intéresse au traitement. Parmi ceux qui sont intéressés, on choisit au hasard un échantillon de personnes qui vont bénéficier de la thérapie; les autres inscrits servent de groupe de contrôle. La situation des deux groupes est saisie au moyen de questionnaires, d'observations effectuées par le personnel pénitentiaire et de données sur les taux de récidive à moyen terme. Les résultats sont confrontés les uns aux autres, ce qui permet de tirer des conclusions sur le succès ou l'échec de la nouvelle thérapie.

Groupes de contrôle non équivalents: Dans cette procédure, les groupes de contrôle ne sont pas créés par tirage au sort; au moyen de méthodes statistiques, à partir de différences de temps, de lieux et de personnes qui se produisent "naturellement" dans l'application de la mesure, on tire des conclusions sur les effets. Pour comparer aussi bien que possible les unités soumises à la mesure à celles qui ne le sont pas, on se sert de moyens statistiques tels que la constitution de "couples de jumeaux" ("matching") qui se ressemblent sous autant d'aspects importants que possible.

Exemple: Dans une évaluation sommaire cofinancée par le PNR 27, on a cherché à distinguer les effets du soutien accordé aux constructions agricoles par les crédits d'investissement et les subventions dans le canton de Zurich. Pour cela, on a confronté un groupe d'exploitations agricoles fortement subventionnées à des entreprises moins subventionnées, et on a comparé leurs revenus agricoles et leur endettement. Des différences sensibles sont apparues entre les deux groupes. L'enquête a bénéficié de la bonne situation des données dans le secteur agricole (étude Schnyder 1993).

Souvent, la comparaison avec groupe de contrôle est complétée par une comparaison longitudinale (pour le groupe cible et/ou le groupe de contrôle). Notamment dans le cas des

groupes de contrôle non équivalents, susceptibles de fournir des résultats contestables sur le plan de la méthode, on peut alors tirer des conclusions plus fiables sur les effets.

Exemple: dans l'enquête susmentionnée, on a également effectué une comparaison longitudinale pour le groupe d'entreprises qui avaient investi dans la construction d'étables. Il est apparu qu'environ 35 pour cent des entreprises ayant effectué des investissements avaient amélioré leur situation économique; toutefois, un nombre semblable d'entreprises ayant une situation économique de départ analogue ont été lourdement chargées par l'investissement, si bien que leurs revenus ne se sont pratiquement pas améliorés, ou se sont même détériorés.

Analyses de discontinuité-régression: Il s'agit d'une nouvelle méthode d'identification des effets d'une mesure, qui n'est valable que pour des groupes déterminés de personnes qui se distinguent des autres par des valeurs seuils déterminées. La méthode présuppose que l'équipe d'évaluation possède des connaissances statistiques approfondies.

Exemple: Dans le domaine des études d'impact sur l'environnement, les projets de construction prévoyant au moins 200 places de stationnement ont l'obligation d'établir un rapport d'impact sur l'environnement. Pour déterminer les effets de l'étude d'impact sur la durée de la construction, on peut partir de l'idée que l'accroissement du nombre de places de parking augmente aussi la durée de la procédure de construction. Pour savoir si l'obligation d'établir un rapport d'impact sur l'environnement a une influence significative sur la durée de la procédure, on examine si, dans la ligne de régression du rapport taille du parking/durée de la construction, il y a ou non une rupture à partir de 200 places de stationnement (cf. figure 1, chap. 1.6).

Bien que les comparaisons avec groupe de contrôle représentent une stratégie d'évaluation qui convient très bien à l'évaluation des effets, leur applicabilité est limitée à différents égards. Comme on l'a vu au chap. 1.6, certains principes éthiques et juridiques (égalité devant la loi) s'opposent à l'exécution d'expériences. Quand des mesures sont valables pour toute la population et appliquées uniformément dans tout le pays, il est souvent difficile de trouver des groupes de contrôle permettant de tirer des conclusions sur les effets de la mesure. Néanmoins, en liaison avec d'autres stratégies d'évaluation (avant tout la comparaison longitudinale), la comparaison avec groupe de contrôle (non équivalent) peut donner des résultats plus significatifs.

Documentation complémentaire

Rossi/Freeman 1993 et Mohr 1993 traitent de manière très approfondie la comparaison avec groupe de contrôle.

3.53 Comparaisons transversales

Par la comparaison transversale, une évaluation de l'exécution et des effets d'une mesure étatique est étendue au delà des unités spatiales (communes, cantons, pays). Dans la recherche évaluative, on trouve très peu de données sur les comparaisons transversales: en règle générale, elles sont traitées dans le cadre de la comparaison avec groupe de contrôle. Toutefois, dans la comparaison avec groupe de contrôle, on recherche d'abord les différences entre unités d'enquête (p.ex. destinataires des mesures tels qu'individus ou entreprises) au sein d'une zone homogène, alors que dans la comparaison transversale on compare les caractéristiques de l'exécution et des effets des mesures entre des zones qui ne sont pas obligatoirement homogènes (pays, cantons, communes).

La prise en compte réaliste des conditions, différentes d'une zone à l'autre, de l'exécution d'une mesure est le principal défi auquel répond la comparaison transversale. Il s'agit, à partir de la forme différente selon les zones d'une mesure déterminée, de tirer des conclusions sur ses mécanismes d'effets, en incluant dans l'appréciation l'influence des conditions de dé-

part et cadres (p.ex. force économique, facteurs démographiques, structures familiales, structure fédéraliste ou centraliste) sur le fonctionnement de la mesure. Il y a différentes manières de procéder pour prendre en compte l'influence des conditions cadres sur le mode d'action des mesures:

- On représente en détail les mesures (bases légales, programmes, activités et structures d'exécution, tendances de développement, etc.) prises dans des zones différentes (pays, cantons, communes), dans un secteur politique déterminé. Il s'agit en fait d'études de cas par zone (pour les études de cas particuliers et multiples, cf. chap. 3.62). La présentation des cas particuliers conserve un caractère isolé, toutefois des relations transversales se dégagent souvent (avant tout dans la perspective des aspects d'exécution). En partie, on essaie aussi d'exploiter les expériences ainsi faites pour en tirer des conclusions à propos des mécanismes d'effets fondamentaux. On renonce toutefois à une interprétation quantitative du rapport entre la forme de la mesure et les résultats. Cette procédure a été choisie en Suisse, en particulier, dans le cadre des études d'exécution (p.ex. Delley et al. 1982, Lunder et al. 1990).

Exemple: Balthasar/Knöpfel ont identifié trois modèles fondamentaux de contrôle des chauffages dans les cantons. Ils prennent pour exemple du premier modèle (contrôle des chauffages par des fonctionnaires) le canton de Lucerne, du second (contrôle des chauffages par des ramoneurs) le canton de Vaud, et du troisième (contrôle des chauffages par des entreprises de révision) le canton de Bâle-Ville. Ils ne cherchent pas à apprécier si un de ces modèles est dans l'ensemble supérieur aux autres (étude Balthasar/Knöpfel 1994).

- Si les zones incluses dans l'enquête sont nombreuses, on peut également procéder à des analyses quantitatives du rapport entre la forme de la mesure et des grandeurs d'effets importantes. A cet égard, il importe de s'assurer que l'analyse n'est pas systématiquement faussée par d'autres facteurs d'influence (p.ex. différences entre ville et campagne, entre régions économiquement fortes et faibles).

Exemple 1: Hausheer a enquêté, dans le cadre d'une recherche transversale, dans plus de 100 communes pour savoir si et dans quelle mesure le système de la taxe sur les sacs exerce des effets, en comparaison avec les solutions habituelles, sur le volume des déchets mixtes et les résultats des collectes sélectives (papier, métal). L'auteur est arrivé à la conclusion que la taxe sur les sacs conduit à une réduction sensible du volume des déchets mixtes et à un accroissement significatif de celui des collectes sélectives. Il a même été en mesure de prouver que les grandeurs mentionnées étaient également influencées par le montant de la taxe sur les sacs (Hausheer 1991).

Exemple 2: Balthasar/Knöpfel ont effectué une enquête approfondie dans 419 communes des cantons de Lucerne, Berne et Bâle-Campagne sur les facteurs qui influencent la diffusion de nouvelles installations de chauffage. Il est apparu que, notamment dans le canton de Lucerne, la part des chauffages modernes est très élevée et qu'en outre le degré de qualification des contrôleurs de chauffage dans les différentes communes exerce une influence significative sur le renouvellement des chauffages à mazout (Balthasar/Knöpfel 1994).

- Notamment dans les expériences pilotes menées de manière centralisée, on cherche à comparer des zones dans lesquelles une mesure déterminée est introduite avec des zones autant que possible semblables dans lesquelles ce n'est pas le cas. La structure d'enquête est ici analogue à celle des expériences (cf. chap. 3.52).

Exemple: Au début des années 80, la vitesse était limitée en Suisse à 60 km/h dans les localités. Pour éclaircir les effets de l'introduction du 50 km/h, on introduisit cette limitation dans quatre zones tests. On prit comme groupe de contrôle quatre zones aussi semblables que possible, où la vitesse continua à être limitée à 60 km/h. Les effets des deux limitations (sur la fréquence des accidents, leur gravité, la fluidité de la circulation, etc.) furent soigneusement observés et systématiquement comparés (Arbeitsgruppe Verkehrssicherheit 1983).

Les comparaisons transversales peuvent remplir toute une série de fonctions utiles. Déjà la représentation descriptive de diverses solutions peut aiguïser la conscience de nouvelles possibilités. Ainsi par exemple, dans la politique en matière de drogue, on cherche très sérieusement des moyens de diminuer la criminalité motivée par l'acquisition. De même, à partir de la représentation de cas particuliers (pays, cantons, communes) caractérisés par des succès ou des échecs marqués, on peut tirer des enseignements sur les conditions de succès d'une mesure. Dans de tels cas, les comparaisons transversales peuvent aussi se fonder sur d'autres formes de comparaison (comparaison avec groupe de contrôle, comparaison longitudinale) qui ont été utilisées dans différentes zones. L'instrument de la comparaison transversale peut être utilisé d'une manière particulièrement efficace pour répondre à des questions à orientation causale quand un grand nombre de zones sont étudiées sous l'angle de relations de cause à effet déterminées. Une telle structure d'enquête est alors identique à celle de la comparaison avec groupe de contrôle. Naturellement, il ne sera pas toujours possible de transférer les expériences ainsi faites à une autre zone, étant donné que, selon les circonstances, il peut exister des conditions cadres qui ne concordent pas avec celles des unités examinées. Il faut donc se garder de transférer mécaniquement les expériences étrangères.

Documentation complémentaire

Rose 1993 donne une vision excellente de l'aptitude des comparaisons transversales au transfert d'expériences d'une zone à l'autre. Dans la plupart des projets du PNR 27, on a également procédé à des comparaisons transversales entre cantons ou entre communes, comme par exemple dans Sangra 1992, Bierter/Binder 1993, Gerheuser/Schmid 1993, Balthasar/Knöpfel 1994, Hanser/Kuster/Cavelti 1994 et Hirsch et al. 1995.

3.54 Comparaison longitudinale

Avec la comparaison longitudinale, on essaie de distinguer les effets des mesures dans le déroulement temporel. La comparaison présuppose qu'une mesure est mise nouvellement en vigueur (p.ex. interdiction de ventes d'immeubles en cascade), changée (p.ex. changement dans le droit matrimonial), modifiée dans son dosage (p.ex. politique de la Banque nationale concernant la quantité de monnaie en circulation, niveau des subventions forestières) ou supprimée (p.ex. clause du besoin dans l'hôtellerie). On cherche à savoir quels changements interviennent dans les grandeurs cibles à la suite de la modification de la mesure. La question de savoir si ces changements sont imputables à la modification doit en règle générale faire l'objet d'un examen attentif (possibilité d'un parallélisme dû au hasard; possibilité que la modification de la mesure et des variables cibles puisse être imputée à une troisième grandeur, etc.). Quand d'autres développements se superposent à la modification de la mesure (évolution conjoncturelle, mutation des valeurs, mesures d'économies des pouvoirs publics, etc.), il est généralement très difficile de déterminer l'effet net de la mesure. C'est pourquoi la comparaison longitudinale a été utilisée dans la recherche évaluative en partie seulement comme deuxième ou troisième stratégie. Il y a néanmoins toute une série de conditions dans lesquelles cette comparaison est très utile. Nous mentionnerons à titre d'exemples les cas suivants:

- La mesure est une disposition nouvelle introduite dans plusieurs communes, cantons ou pays. Si on observe partout la même évolution des grandeurs cibles, le rapport entre la mesure et les effets peut être tenu pour bien assuré.
Exemple: Dans une étude du groupe de travail "évaluation législative", on a cherché à connaître les effets de l'introduction de la taxe sur les sacs à ordures sur le volume des déchets mixtes. Dans 11 des 21 communes observées, le volume des déchets mixtes a fortement diminué, dans 4 autres il a diminué faiblement; les communes restantes ne disposaient pas de chiffres détaillés. Le fait qu'on ait observé des résultats analogues dans presque toutes les communes a accru la fiabilité des conclusions (étude INFRAS 1990).

- L'ampleur des changements après l'entrée en vigueur d'une mesure est très grande, et il n'y a pas d'autres raisons qui expliqueraient de manière plausible ces changements.
Exemple: L'évaluation de la campagne sur le SIDA se fonde sur une comparaison longitudinale. Face à l'accroissement marqué de l'usage de préservatifs dans le déroulement temporel, il ne semble pas nécessaire de filtrer l'effet net de la campagne sur le comportement par rapport à d'autres facteurs de sensibilisation (maladie de personnes célèbres, rapports médicaux, conversations avec des connaissances).
- Il existe des bases théoriques sur le rapport entre la mesure et les variables cibles ainsi que d'éventuels autres facteurs d'influence. Les hypothèses théoriques sont empiriquement bien assurées. Dans ce cas, les changements identifiés empiriquement dans le déroulement temporel sont considérés comme une image fiable de la réalité s'ils concordent avec la théorie.
Exemple: En matière de protection des eaux, les corrélations entre l'épuration des eaux et la qualité de l'eau sont suffisamment connues pour que le rapport entre le développement du nombre et de la capacité des installations d'épuration, d'une part, et l'amélioration de la qualité de l'eau dans le déroulement temporel, de l'autre, puisse être reconstruit sans autre sur la base des substances repérées, et l'effet net de la politique de protection des eaux identifié.
- Il est possible de saisir statistiquement l'influence du changement de mesure sur certaines unités (parties de la population, entreprises, etc.) et de l'isoler ainsi d'éventuelles superpositions d'autres développements. La comparaison longitudinale, combinée à la comparaison avec groupe de contrôle non équivalent (voir plus haut), peut donner des résultats très significatifs.
Exemple: Dans une enquête sur les effets des cautionnements et subventions accordés conformément à la loi sur l'encouragement du crédit à l'hôtellerie et aux stations de villégiature, on a analysé dans le déroulement temporel les données d'exploitation de 48 hôtels ayant bénéficié d'un soutien. Ces valeurs ont été pondérées par rapport à la moyenne de la branche et ainsi implicitement comparées à l'ensemble des hôtels, afin de filtrer l'influence de la conjoncture, des conditions de neige, etc. (étude Hanser/Kuster/Cavelti 1994).

Une variante de la comparaison longitudinale est la comparaison entre situation existante et tendance. Dans la mesure où certaines grandeurs cibles suivent une tendance (p.ex. augmentation de la pollution atmosphérique du fait de l'accroissement du trafic), une simple comparaison longitudinale donne une image fautive. Mieux vaut comparer l'évolution qui se serait produite sans le changement de mesure à celle qui a suivi la mesure. Pour cela, il faut établir une prévision sur l'évolution sans la mesure (prévision fondée sur le statu quo), qui peut reposer sur une extrapolation de tendance ou une autre méthode d'estimation de l'avenir.

Documentation complémentaire

Schmitz 1989 et Widmer 1991 donnent une idée des enquêtes à base longitudinale. Parmi les études du PNR 27, nous mentionnerons Hanser/Kuster/Cavelti 1994 ainsi que Balthasar/Knöpfel 1994.

3.6 Techniques de recherche

3.6.1 Aperçu/présentation de techniques de recherche choisies

Divers auteurs utilisent la technique de recherche comme caractéristique centrale de la classification des stratégies d'évaluation. De notre point de vue, toutefois, lors de questions à orientation causale, la décision concernant le genre de comparaison adéquat (comparaison avec groupe de contrôle, transversale, longitudinale, stratégies d'évaluation implicitement comparatives) est d'une importance centrale et la technique de recherche doit être orientée là-dessus de manière optimale.

Il existe un nombre toujours plus grand de techniques de recherche. On peut les diviser en gros en méthodes de collecte et méthodes d'exploitation des données.

Nous énumérerons brièvement les méthodes suivantes de collecte des données:

- analyse de documents
- interview ouverte/semi-ouverte
- sondage écrit/téléphonique/personnel
- collecte de données statistiques (dépouillement de données d'exécution de l'administration, dépouillement de statistiques propres)
- enregistrement audio-visuel
- discussion en groupe (méthode interactive)
- observation participative

Parmi les méthodes d'exploitation des données, on peut mentionner celles-ci:

- étude de documentation/analyse de documents (appréciation qualitative)
- procédure d'interprétation herméneutique (p.ex. d'interviews orales)
- analyse du contenu (quantitative)
- exploitation statistique plus simple (moyenne, médiane, écart type, etc.)
- régression, régression multiple
- analyse de facteurs, analyse de clusters, analyse multidimensionnelle
- programmation linéaire
- modélisation causale avec modèles mathématiques, modèles économétriques, simulations.

Les techniques de recherche se subdivisent en méthodes qualitatives et quantitatives. Dans le cadre des méthodes qualitatives⁸, les données sont représentées sous forme de mots. Lors de l'utilisation de méthodes quantitatives, les données sont représentées sous forme de chiffres; il est nécessaire d'effectuer préalablement une "mesure", c'est-à-dire d'attribuer des chiffres aux événements ou états de faits.

Nous aborderons ci-après plus en détail trois techniques de recherche qui sont souvent utilisées dans le cadre des évaluations.

⁸ Dans le monde anglophone, les informations chiffrées dans lesquelles les chiffres ne peuvent pas être mis en rapport les uns avec les autres (au sens de mesure ordinale ou cardinale) comptent également parmi les méthodes qualitatives. Un exemple en est l'analyse du contenu des textes assistée par ordinateur.

Figure 7

Principales techniques de recherche, combinaisons

| Technique de recherche | Sous groupes | Formes spéciales, combinaisons avec d'autres techniques de recherche |
|--------------------------------|--|--|
| Etudes de cas | cas particulier cas choisis études de cas multiples | enquêtes interactives centrées sur les protagonistes études de cas intercommunales, intercantonales, internationales |
| Sondages | Sondages sur panels | |
| Analyses de données existantes | Analyses statistiques secondaires synthèses d'évaluations | études économétriques |

Documentation complémentaire

Atteslander 1985 et Lamnek 1988 donnent une introduction intéressante aux méthodes de la recherche sociale empirique et des méthodes qualitatives.

3.62 Etudes de cas

Nous entendons par étude de cas la description analytique d'un événement, d'une institution, d'un processus (d'exécution) ou d'une mesure. Les objets de recherche d'une étude de cas peuvent être par exemple:

- des procédures d'autorisation de construction choisies dans le canton X,
- les expériences de 5 chômeurs face au service de placement de l'Etat,
- les stratégies administratives en rapport avec la création d'un dépôt de déchets spéciaux,
- l'exécution de la loi sur l'assurance invalidité dans différents cantons.

La raison des études de cas tient en règle générale à la complexité de l'objet à étudier, qui exige une analyse plus en profondeur qu'en largeur. En conséquence, l'information recueillie sera abondante et complexe.

Dans le traitement des études de cas, on se sert d'une ou de plusieurs méthodes de collecte et d'exploitation des données (analyses de documents, interviews, observation participative, procédure d'interprétation herméneutique, exploitations statistiques plus simples, etc.). Ainsi, les études de cas sont moins une technique de recherche qu'un accès à celle-ci.

Nous distinguons entre les études de cas particuliers, de cas choisis et de cas multiples. Les études de cas particuliers visent habituellement à décrire un événement, un processus ou une mesure unique. Elles peuvent servir à l'illustration (p.ex. représentation de plusieurs procédures d'autorisation en cours parallèlement pour un grand projet de construction), et présenter alors des limites mouvantes avec l'"investigative journalism". Elles peuvent aussi servir à explorer des questions nouvelles qui n'ont pas encore été traitées. Troisièmement, le choix peut se porter sur un cas critique (p.ex. comportement des services compétents dans un cas de catastrophe), afin de tirer les leçons de l'événement et de pouvoir procéder à des améliorations. Dans le dernier exemple, il s'agit de l'étude de cas choisis selon des critères déterminés. On pose des questions normatives telles que celles-ci: Les objectifs de performance visés par la mesure (p.ex. taux d'occupation des participants aux programmes de perfectionnement pour chômeurs) sont-ils atteints? L'exécution s'est-elle déroulée comme prévu? De telles questions normatives - questions de contrôle - peuvent recevoir des réponses satisfai-

santes au moyen de cas choisis. En comparaison avec une analyse complète, une telle étude de cas ne pourra pas répondre à la question de savoir quelle est la fréquence des écarts entre impératifs idéaux et valeurs existantes. Mais, par un choix intelligent (p.ex. de "bons" et de "mauvais" cas), on pourra distinguer si des écarts se produisent, et pourquoi. Dans les études de cas multiples, on étudie plusieurs cas, soit pour rendre les déclarations plus crédibles, soit pour mieux tester certaines hypothèses, soit encore pour mettre en lumière les différences (entre cantons, moments, etc.). Des limites mouvantes existent entre les études de cas multiples qui concernent différentes zones, et les comparaisons transversales.

Les stratégies d'évaluation centrées sur les protagonistes, interactives, sont une combinaison entre études de cas et sondages (voir chapitre suivant). Elles incluent dans l'étude (souvent par des interviews ouvertes ou semi-ouvertes) les protagonistes les plus importants pour une mesure (instances d'exécution, destinataires, autres intéressés). Ce faisant, on se réfère souvent à leurs appréciations de l'efficacité. L'objectif n'est pas, comme dans beaucoup de sondages, de distinguer une opinion représentative, mais plutôt de saisir de manière aussi complète que possible les diverses formes d'argumentation possibles; dans ces conditions, le processus de recherche est généralement terminé quand, lors d'interviews supplémentaires, on n'enregistre plus de nouveaux arguments et on constate une "saturation" des formes d'argumentation. A partir du matériel trouvé, l'équipe d'évaluation peut établir certaines typologies. En règle générale, elle n'essaiera pas de reconstruire la réalité à partir des opinions existantes. Dans la mesure où les arguments n'ont pas déjà été trouvés de manière interactive (p.ex. dans des hearings, des groupes de discussion), l'équipe d'évaluation transmettra les résultats et les interprétations à tous les protagonistes, ou au moins aux plus importants, afin que les déclarations et propositions puissent être discutées en commun. Dans le pilotage de l'évaluation, l'équipe d'évaluation peut s'orienter sur les désirs et les questions d'intéressés importants, p.ex. dans le cadre d'un groupe d'accompagnement. En outre, le processus d'évaluation est en règle générale axé sur une amélioration des mesures et non pas simplement sur un bilan. L'approche centrée sur les protagonistes et interactive permet de poser les questions et les problèmes sur la table, de chercher en commun des solutions et de les discuter, l'équipe d'évaluation assumant alors la fonction d'animateur et d'expert. Cette stratégie d'évaluation convient avant tout pour les mesures dont l'exécution réussie et l'efficacité dépendent de la bonne collaboration entre autorités et intéressés.

Les études de cas conviennent généralement au traitement de questions descriptives. Si on choisit correctement les cas (étude de cas choisis selon des critères déterminés), elles peuvent également servir à répondre à des questions normatives (p.ex.: Y a-t-il des déficits d'exécution et, si oui, pourquoi?). Les études de cas peuvent exceptionnellement aussi servir à répondre à des questions à orientation causale, par exemple lorsqu'on examine, à propos d'un cas particulier, si l'enchaînement d'effets pris en compte est également observé dans la réalité. La question de savoir si, à partir d'études de cas, on peut obtenir des éclaircissements sur les effets pour un ensemble assez grand est contestée.

Les études de cas peuvent être combinées avec d'autres techniques de recherche, p.ex. avec des sondages, afin de recueillir des données tant en profondeur qu'en largeur.

Documentation complémentaire

Yin 1989 et General Accounting Office 1990 donnent un bon aperçu des études de cas. Rey et al. 1993 et Bierter/Binder 1993, tous deux des rapports finaux sur des projets mis en oeuvre dans le cadre du PNR 27 à orientation centrée sur les protagonistes, interactive, donnent des indications pratiques sur cette procédure. Un autre exemple est Gerheuser 1991.

3.63 Sondages

Lors d'un sondage, on saisit en règle générale, à partir de l'échantillon d'un ensemble, des données qui doivent permettre d'identifier l'apparition, le modèle de répartition et la relation réciproque entre certains événements et situations qui se manifestent naturellement⁹. L'objectif est de recueillir des informations à partir desquelles on pourra tirer des conclusions concernant des parties de la population qui n'étaient pas contenues dans l'échantillon. On utilise des échantillons aléatoires dans lesquels chaque segment de la population à examiner a une certaine probabilité d'être pris en considération dans l'échantillon. Dans les limites de la fiabilité statistique, les résultats peuvent être transférés à l'ensemble qui fait l'objet de la recherche. Dans la perspective de l'agrégation des données, on attache une grande importance à l'uniformité des données à recueillir auprès des unités observées dans l'échantillon. En règle générale, les informations sont recueillies à partir de questions fermées, c'est-à-dire que les personnes interrogées choisissent parmi les réponses possibles celle qui leur convient¹⁰. Les questions ouvertes sont possibles, mais elles exigent ensuite un codage des réponses lors de l'exploitation.

Les sondages peuvent être menés par écrit/par courrier, par téléphone ou par interrogation individuelle.

La définition de l'ensemble dont on tire l'échantillon est d'une importance capitale. Lors d'une vaste étude sur les effets de l'AVS, cela sera par exemple la population âgée de plus de 62 ou 65 ans. Dans une étude sur le groupe cible d'une mesure déterminée (p.ex. PTT, encouragement à la construction de logements), on éclaircira à chaque fois avec soin quelles couches socio-économiques et démographiques en font partie. On peut concevoir aussi de concentrer un sondage sur des couches de population dans lesquelles on pressent certains problèmes (p.ex. familles à un parent, chômeurs de longue durée, entreprises en faillite, etc).

Exemple: Dans une étude sur l'encouragement de l'accession à la propriété par la Confédération, on interrogea des propriétaires de maison aidés par des cautionnements de la Confédération, des personnes dont la demande d'aide fédérale avait été refusée, des propriétaires de maison ne bénéficiant pas de l'aide fédérale ainsi que des locataires sur leur situation et leur appréciation de l'aide fédérale. Il apparut que les propriétaires de maison bénéficiant de cautionnements de la Confédération avaient un profil sensiblement différent (revenus plus faibles, maisons plus petites, etc.) de ceux qui ne bénéficiaient pas de l'aide. L'interrogation des locataires donna également des informations révélatrices, p.ex. le fait que 62 pour cent des locataires interrogés n'avaient encore jamais songé à acquérir un logement (étude Schulz et al. 1993).

Les sondages sur panels constituent une combinaison des sondages avec les comparaisons longitudinales. Dans ce cas, on interroge plusieurs fois le même échantillon de personnes, sur une certaine durée. On peut déterminer ainsi des changements dans les états de faits et les opinions. Pour les décisions sur des mesures étatiques, les informations sur les tendances sont souvent plus utiles que les informations purement statiques. De telles informations peuvent aussi donner des indications sur les relations d'effets, dans la mesure où on peut distinguer quel état de fait s'est manifesté avant un autre.

L'exécution de sondages présuppose en règle générale des connaissances approfondies des états de faits. Celles-ci peuvent être acquises de diverses manières (étude de documentation, interviews en profondeur, etc.). Les sondages impliquent une étude préalable attentive. Parfois, des éclaircissements fondés sont nécessaires pour savoir de quel ensemble il faut logiquement tirer l'échantillon et quelle doit être la grandeur de l'échantillon pour chacun des divers segments. Ensuite viennent les prétests (examens expérimentaux) du question-

⁹ Nous n'examinerons pas spécialement ici les enquêtes exhaustives telles que, par exemple, les recensements populaires ou les enquêtes effectuées auprès des recrues.

¹⁰ Dans la mesure où on ne peut pas estimer d'avance toutes les réponses possibles, on crée dans le questionnaire une catégorie restante ("autres facteurs:.....").

naire, destinés à vérifier que les questions sont formulées de manière claire et neutre. Après l'exécution du sondage, une reprise souvent multiple des données est nécessaire afin d'obtenir les réponses manquantes ou de remplacer les personnes qui n'ont pas répondu par d'autres personnes convenant à l'enquête. Il s'agit de garantir que les résultats du sondage ne seront pas altérés par des réponses manquantes.

Les sondages permettent en règle générale de répondre à des questions à orientation descriptive. On peut saisir des états de faits (p.ex. état sanitaire, comportement en matière de dépenses des personnes âgées ou invalides) et des opinions (p.ex. degré d'indigence). En un temps relativement court et moyennant des frais supportables, on peut dégager des sondages l'appréciation de prestations étatiques (p.ex. remise du courrier, disposition des formulaires fiscaux, renseignements sur les prestations de l'AVS) par les utilisatrices et utilisateurs. En revanche, les sondages ne conviennent que dans certaines limites pour répondre aux questions à orientation causale. Parfois, on essaie d'obtenir des éclaircissements sur les connexions d'effets à partir des corrélations, c'est-à-dire des relations positives entre différentes grandeurs. Les sondages peuvent en particulier éclairer les relations de cause à effet quand ils fournissent des informations sur les motifs des personnes dans la perspective de certains comportements (observables) (p.ex. les raisons d'un déclin de l'utilisation des transports publics, qui peuvent résider dans l'augmentation des tarifs, les problèmes de sécurité, la longue durée des trajets ou le caractère peu pratique des correspondances).

Documentation complémentaire

Atteslander 1984 donne une introduction aux techniques de sondage; General Accounting Office 1991b traite de cet instrument dans le contexte des évaluations, et Wholey 1983 donne des indications sur la mesure dans laquelle les interviews par téléphone permettent de juger de la qualité des prestations de services étatiques. Parmi les projets du PNR 27, ce sont avant tout Schulz et al. 1993 et Sangra 1992 qui ont utilisé cette technique de recherche.

3.64 Analyse de données existantes

Les deux techniques de recherche qui viennent d'être mentionnées exigent en règle générale la collecte de nouvelles informations, que ce soit par des sondages, des interviews ou d'autres techniques de recherche (pour des études de cas). Etant donné que la collecte de données est en général coûteuse, il vaut toujours la peine d'examiner la qualification des données existantes; cette démarche facilite également la préparation de la recherche de nouvelles données.

Analyses statistiques secondaires

Beaucoup de données (p.ex. données d'exécution, recensements démographiques, recensements partiels, sondages) sont recueillies à des fins déterminées. Dans le cadre des analyses statistiques secondaires, on les utilise pour répondre à des questions qui n'avaient pas été prévues à l'origine. Dans la plupart des cas, les analyses statistiques secondaires peuvent servir à des fins descriptives.

Exemple: La proportion de propriétaires de maison et de locataires déterminée dans le recensement démographique et le nombre de nouvelles constructions relevé dans la statistique des constructions peuvent être retracés dans le temps, répartis par cantons et mis en rapport avec l'encouragement de l'accession à la propriété du logement par la Confédération ou les cantons (cf. étude Schulz et al. 1993).

Selon les circonstances, une combinaison de données recueillies à des fins descriptives peut fournir des informations sur certaines relations de cause à effet dans la perspective de mesures étatiques déterminées.

Exemple: Dans les années 60 et 70, différentes enquêtes furent menées sur l'origine sociale des étudiants. Elles firent apparaître que les étudiants des classes supérieures et moyennes étaient surreprésentés. Une comparaison plus récente des diverses statistiques

ainsi que des sondages effectués au cours des 30 dernières années sur l'origine sociale des étudiants a montré que, malgré le développement des systèmes de bourses dans les cantons, la surreprésentation des étudiants des classes supérieures et moyennes n'a guère changé (travail de séminaire non publié, Université de Berne).

La pertinence des analyses statistiques secondaires dépend de la situation des données en cause. Dans les domaines de l'économie et de l'agriculture, par exemple, les conditions préalables sont très bonnes. Néanmoins, les déclarations générales plus approfondies sont toujours difficiles et la situation des données doit être examinée attentivement dans chaque cas concret.

Synthèses d'évaluation

A propos de certaines questions ou mesures étatiques, il existe déjà des enquêtes. Dans le cadre de synthèses d'évaluation, les résultats de diverses études sont agrégés (réunis) dans le but de tirer des conclusions plus fiables que celles des différentes études.

Exemple: Il y a en Suisse toute une série de campagnes d'information et, dans certains cas (p.ex. campagne sur le SIDA, informations sur la protection civile), on effectue des contrôles des résultats. Dans le cadre d'une synthèse d'évaluation, on peut réunir les résultats de ces diverses enquêtes et cerner de plus près les conditions de succès de l'instrument "campagnes d'information" (description du projet Balthasar à l'attention du Fonds national).

Les synthèses d'évaluation peuvent avoir un caractère qualitatif, mais dépasser les études documentaires typiques. Les résultats sont pondérés et combinés qualitativement, sans qu'on vise à une agrégation statistique des différentes études. En outre, il y a la possibilité de combiner des études individuelles. Aux Etats-Unis, on a développé à cet effet la méthode de la "méta-analyse", qui consiste à établir une moyenne entre les volumes d'effets¹¹ déterminés quantitativement dans diverses études.

Documentation complémentaire

General Accounting Office 1992a et 1992b donnent des éclaircissements sur les analyses statistiques secondaires et les synthèses d'évaluation. Parmi les projets du PNR 27, Rieder (Les évaluations dans le domaine de la politique agricole I et II) ainsi que Stalder et al. 1994 ont effectué des analyses statistiques secondaires en se servant de la programmation linéaire et de l'économétrie.

3.7 Digression: Les modèles en tant que point de départ et résultat des évaluations

Les modèles ont une grande importance pour les évaluations, à la fois comme point de départ et comme résultat des évaluations ainsi qu'en rapport avec les techniques de recherche statistiques.

Les modèles sont des tentatives de représenter la réalité. Ils peuvent prendre une forme réelle (p.ex. la montre en tant que modèle de la rotation de la terre sous forme analogique ou digitale), discursive (p.ex. théories sur le changement des valeurs) ou mathématique/numérique (p.ex. modèle conjoncturel économique). Dans les évaluations, ce sont les deux dernières formes qui dominent.

Toute évaluation à orientation causale présuppose un modèle d'effets. Un modèle permet de centrer l'évaluation, et en même temps de spécifier des hypothèses importantes qui devront être examinées dans l'enquête empirique.

Exemple: La loi sur le travail à domicile prévoit l'obligation de rémunérer de manière égale les travailleurs/-euses à domicile et dans l'entreprise. Lors d'interviews d'employeurs et de travailleuses à domicile, on a cherché à savoir si la prescription mentionnée était respectée dans le cas des travailleuses à domicile soumises à la loi; à cette occasion, on a aussi

¹¹ Mesurés en considérant la différence de situation entre groupe cible et groupe de contrôle.

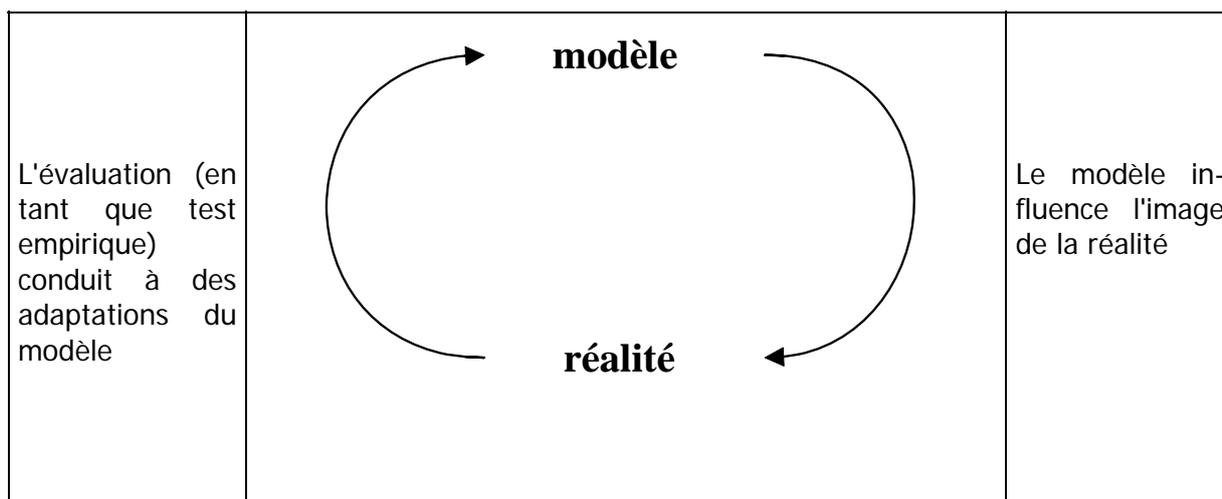
examiné de plus près les conditions préalables de succès auprès des instances d'exécution cantonales. L'enquête reposait sur le modèle d'effets selon lequel la réalisation de l'impératif d'égalité dépend des activités des instances d'exécution et des intérêts spécifiques des employeurs et des travailleuses à domicile - un modèle parfaitement plausible (étude Gerheuser/Schmid 1993). Mais, dans le cadre d'une enquête différemment centrée, on aurait pu aussi examiner d'autres aspects des effets, tels que, par exemple, la question de savoir si l'impératif d'égalité a également touché les travailleuses non soumises à la loi sur le travail à domicile ou si cet impératif a eu une influence sur le déplacement des places de travail vers des pays à bas salaires.

Les modèles sont indispensables à la recherche scientifique et aux évaluations à orientation causale, dans la mesure où ils permettent de limiter et de préciser la question de recherche; mais ils recèlent en même temps le risque de donner une direction unilatérale, voire fautive, à l'évaluation par l'image de la réalité qu'ils présupposent. Cela peut arriver par exemple quand, dans un modèle, seuls les effets souhaités d'une mesure sont inclus et que des effets secondaires importants sont occultés.

Les modèles constituent en même temps aussi le point final d'une évaluation au sens où tout résultat d'évaluation est une tentative de représenter des aspects de la réalité.

Figure 8

Les modèles en tant que point de départ et résultat des évaluations



Les modèles peuvent avoir des dimensions très diverses.

- Ils peuvent avoir un caractère statique ("Une augmentation du prix de l'essence de 5 pour cent conduit à une réduction de la consommation de 0,5 pour cent") ou dynamique ("Une augmentation du prix de l'essence conduit d'abord à une diminution des trajets en voiture, puis au développement de voitures consommant moins, et enfin à un effet d'accoutumance").
- Ils peuvent présenter un caractère partiel (p.ex. influence des prix des transports publics sur leur utilisation) ou global (utilisation des transports publics et privés dépendant du changement des valeurs, de l'évolution de la structure de l'habitat, de l'évolution des prix, de l'offre en prestations de transports publics, etc.).
- Ils peuvent prétendre à une validité limitée dans l'espace et dans le temps ("L'augmentation des amendes de parcage en 1993 a conduit à une diminution du stationnement sur les trottoirs de 21 pour cent) ou globale ("L'influence des peines sur le comportement délictueux dépend de la gravité de la peine et des chances d'être pris").

Les modèles qualitatifs sont présentés en mots, les modèles quantitatifs en chiffres. Ces derniers sont des systèmes d'équations mathématiques dans lesquels l'ampleur de l'influence entre les grandeurs (variables) est exprimée en valeurs numériques (dites paramètres). En

principe, les deux formes de modèle sont traduisibles dans l'autre "langue". Notamment dans le cas d'un grand nombre de variables, une représentation mathématique/numérique offre l'avantage d'une meilleure vision d'ensemble.

Les modèles quantitatifs, notamment, jouent un rôle considérable en ce qui concerne les attentes de capacité des évaluations. Selon les circonstances, ces modèles permettent d'évaluer les effets futurs de mesures existantes ou prévues, dans l'hypothèse de conditions cadres déterminées, et de simuler les effets possibles d'autres mesures.

Exemple: Une enquête économétrique relative aux effets de la politique des étrangers sur le changement des structures en Suisse entreprit de considérer tous les marchés importants (marché des biens et marché du travail). Le modèle développé réussit à retracer de manière satisfaisante l'évolution passée. Dans une simulation, on estima les effets possibles d'une libéralisation du marché du travail (immigration plus forte de travailleurs qualifiés, plus faible de travailleurs peu qualifiés). Il en résulta, selon le modèle, des effets marqués sur les secteurs économiques importants mais, fait étonnant, des effets insignifiants seulement sur l'évolution de la productivité (étude Stalder et al. 1994).

Les exigences posées aux modèles quantitatifs sont considérables. Il faut que des données suffisantes existent et que les relations entre les variables importantes demeurent stables sur d'assez longues périodes. Il est important en outre de pouvoir compter sur une pénétration scientifique à long terme du domaine en cause. De bonnes conditions préalables aux modèles quantitatifs existent par exemple dans les questions touchant l'économie, l'agriculture, l'énergie et en partie l'environnement.

Les modèles quantitatifs suscitent souvent chez les profanes, du fait de leur "langue", des attentes de précision (au sens de "concordance" avec la réalité et de capacité à prévoir les développements futurs). Même si les modèles quantitatifs facilitent un examen empirique, ils n'offrent nullement la garantie de pouvoir estimer de manière (plus) fiable les développements futurs.

Une question centrale des modèles quantitatifs est la relation réciproque entre les variables, notamment la mesure dans laquelle les changements des grandeurs (influencables par les mesures étatiques, dites exogènes) entraînent des changements dans les grandeurs qui en dépendent (dites endogènes). Ces coefficients de comportement, ou paramètres, peuvent être déterminés de trois manières; en conséquence, on distingue aussi trois genres de modèles:

Paramètres techniques, ou modèles préétablis du point de vue technique/des sciences naturelles: Dans ces modèles, les paramètres sont définis par les données techniques.

Exemple: Dans la politique de l'environnement, les modèles quantitatifs reposant sur des paramètres techniques jouent un rôle important. Sur la base des prévisions de la circulation et de la connaissance des effets de purification des catalyseurs (=paramètres techniques), ainsi que d'hypothèses concernant le renouvellement du parc automobile, on a pu estimer au milieu des années 80 les effets prévisibles de l'introduction du catalyseur sur la charge de pollution atmosphérique.

Paramètres déterminés (empiriquement) statistiquement, ou modèles statistiques positifs: Les coefficients de comportement sont (souvent à l'aide de connaissances théoriques) déterminés ou estimés par des méthodes statistiques (telles que calculs de régression, modèles de piste, analyses de facteurs, de variance et de clusters); souvent, on contrôle la plausibilité des résultats au moyen des connaissances théoriques, et au besoin on les adapte.

Exemple: L'analyse économétrique mentionnée dans l'avant-dernier exemple (étude Stalder et al. 1994) s'est servie de procédures statistiques, notamment du calcul de régression, pour estimer la relation entre les grandeurs importantes qui sont significatives dans l'intersection entre la politique des étrangers et la politique structurelle.

Paramètres fixés de manière normative, ou modèles normatifs: Sur la base de théories normatives, on fixe des paramètres (spécifiés mathématiquement et quantifiés par des données normées). Les paramètres ne sont pas vérifiés empiriquement, mais fixés comme hypothèses plausibles. On ne vise pas à un contrôle statistique de la validité (c'est-à-dire de

l'"exactitude" des rapports de cause à effet) du modèle. Dans la mesure où le modèle répond à des objectifs de prévision, on peut toutefois déterminer par la suite si les développements illustrés sur la base du modèle se produisent effectivement dans la réalité. Si c'est le cas, c'est un indice de la crédibilité du modèle également à des fins de prévision; mais le risque existe que, dans l'intervalle, les rapports entre des grandeurs importantes aient changé.

Exemple: Dans le domaine de l'agriculture, un modèle reposant sur la programmation linéaire devait enregistrer notamment les rapports entre l'élevage du bétail et la culture des céréales. Il était fondé sur l'hypothèse que les paysans ont un comportement économique rationnel. Le modèle put donner des indications sur le futur changement des structures paysannes et sur ses effets régionaux et sectoriels. Dans des simulations, on put notamment concrétiser le scénario d'une libéralisation selon le GATT (étude Felder/Rieder dans: Les évaluations dans le domaine de la politique agricole I 1994).

Souvent aussi, les modèles sont des combinaisons de ces trois types. Un modèle économique normatif reposant sur la programmation linéaire peut par exemple s'appuyer sur des corrélations de nature technique et scientifique (p.ex. grandeur des entreprises et produit des récoltes).

Documentation complémentaire

Rieder/Meyer 1994 donnent une brève introduction aux modèles quantitatifs en tant qu'instrument d'évaluation. Dans le PNR 27, les modèles quantitatifs ont été utilisés avant tout dans deux études (Les évaluations dans le domaine de la politique agricole I/II 1994, ainsi que Stalder et al. 1994).

3.8 Mélange de méthodes et triangulation

Pour formuler des déclarations sensées sur l'exécution et les effets de l'action étatique, il est opportun d'utiliser, dans le cadre des évaluations, des approches et des techniques de recherche diverses. Cela permet généralement de saisir d'une manière relativement satisfaisante les facettes de la réalité. Pour obtenir des résultats significatifs, un mélange de méthodes est souvent aussi plus avantageux que l'orientation sur une seule technique de recherche. Les équipes d'évaluation se servent par exemple d'analyses de documents, d'études documentaires, d'analyses statistiques secondaires rudimentaires et d'interviews pour se faire une idée générale du domaine de recherche, avant de faire intervenir d'autres techniques de collecte (p.ex. sondages) et d'exploitation (p.ex. régression multiple) des données.

Exemple: Balthasar/Knöpfel 1994 ont saisi les effets des prescriptions de protection de l'environnement sur les découvertes dans le domaine des installations de chauffage domestique, sur l'application technique de ces découvertes à l'amélioration des chauffages domestiques et sur leur diffusion. Ils ont enquêté, sous forme d'une étude de cas, sur le rapport entre les valeurs limites de protection de l'environnement et les découvertes; dans le cadre d'une étude de marché, ils ont examiné l'influence de l'homologation sur les innovations en matière de chauffage domestique; sur la base de sondages et d'analyses statistiques, ils ont cherché à connaître l'influence du contrôle des chauffages sur la modernisation des installations de chauffage. Par des analyses de séries temporelles, enfin, ils se sont efforcés de saisir les effets globaux des mesures administratives.

La triangulation se sert d'approches qui se recoupent, de manière à améliorer la fiabilité des résultats des évaluations. Elle peut comporter une ou plusieurs des formes suivantes:

1. Combinaison de stratégies d'évaluation: Lors de questions à orientation causale, une combinaison de comparaisons (comparaison avec groupe de contrôle, transversale, longitudinale) ou la combinaison de celles-ci avec une stratégie d'évaluation implicitement comparative peut améliorer la fiabilité de résultats faibles ou confirmer encore celle de résultats déjà bons.

Exemple: Une comparaison intéressante de stratégies d'évaluation et de techniques de recherche diverses a été rendue possible par deux des études données en mandat par le

groupe de travail "évaluation législative" (AGEVAL) concernant la prévoyance professionnelle. Dans le cadre d'une enquête (économétrique) statistique secondaire en série temporelle, on a déterminé les effets exercés sur le marché du travail par la loi fédérale sur la prévoyance professionnelle (étude Schätti 1990). Gerheuser 1991 a enquêté sur la même question en procédant à une interrogation (non représentative) de chefs du personnel d'entreprises.

2. Recoupement de techniques de recherche: On essaie par là d'assurer la solidité d'affirmations sur des états de faits. Il peut s'agir de vérifier les déclarations de certaines personnes de manières multiples (p.ex. par des interviews et par l'observation participative), d'intégrer un groupe de personnes de diverses manières lors de la détermination d'états de faits (p.ex. par des interviews en profondeur complétées par un sondage représentatif) ou d'exploiter plusieurs sources de données concernant des états de faits analogues (p.ex. analyses statistiques secondaires et sondages représentatifs).
Cf. étude Balthasar/Knöpfel plus haut.
3. Etablissement de diverses hypothèses d'effets (concurrentes): On sait que les hypothèses d'effets peuvent inciter à choisir les stratégies d'évaluation et les techniques de recherche de telle sorte que les hypothèses d'effets soient confirmées (p.ex. inefficacité d'un instrument d'action déterminé). L'établissement d'hypothèses d'effets concurrentes peut contribuer à diminuer ce risque.
4. Traitement de la même question par plusieurs équipes d'évaluation: Cette procédure est liée dans la plupart des cas à une des trois possibilités de triangulation susmentionnées. Pour un exemple, voir chiffre 1.
5. Attribution de tâches et de rôles divers au sein d'une équipe d'évaluation: Cette procédure est, elle aussi, généralement liée à l'une des trois possibilités de triangulation ci-dessus.

L'évaluation de la promotion économique valaisanne s'est déroulée de manière interactive, c'est-à-dire qu'elle a inclus les intéressés. Tout d'abord, on a recensé leurs estimations au moyen d'interviews. Ensuite, les résultats intermédiaires ont été approfondis dans le cadre d'un travail en groupe. A cette occasion, les fonctions d'animation, d'expertise et de méthodologie ont été réparties autant que possible entre diverses personnes au sein de l'équipe d'évaluation (étude Rey et al. 1993, p. 31).

Documentation complémentaire

Patton 1990 (pp. 186-198, 464-470) se fait le champion du mélange des méthodes et de la triangulation. Parmi les projets du PNR 27, avant tout Balthasar/Knöpfel 1994 se sont servis d'un mélange de méthodes et de la triangulation, et donnent des indications à ce propos.

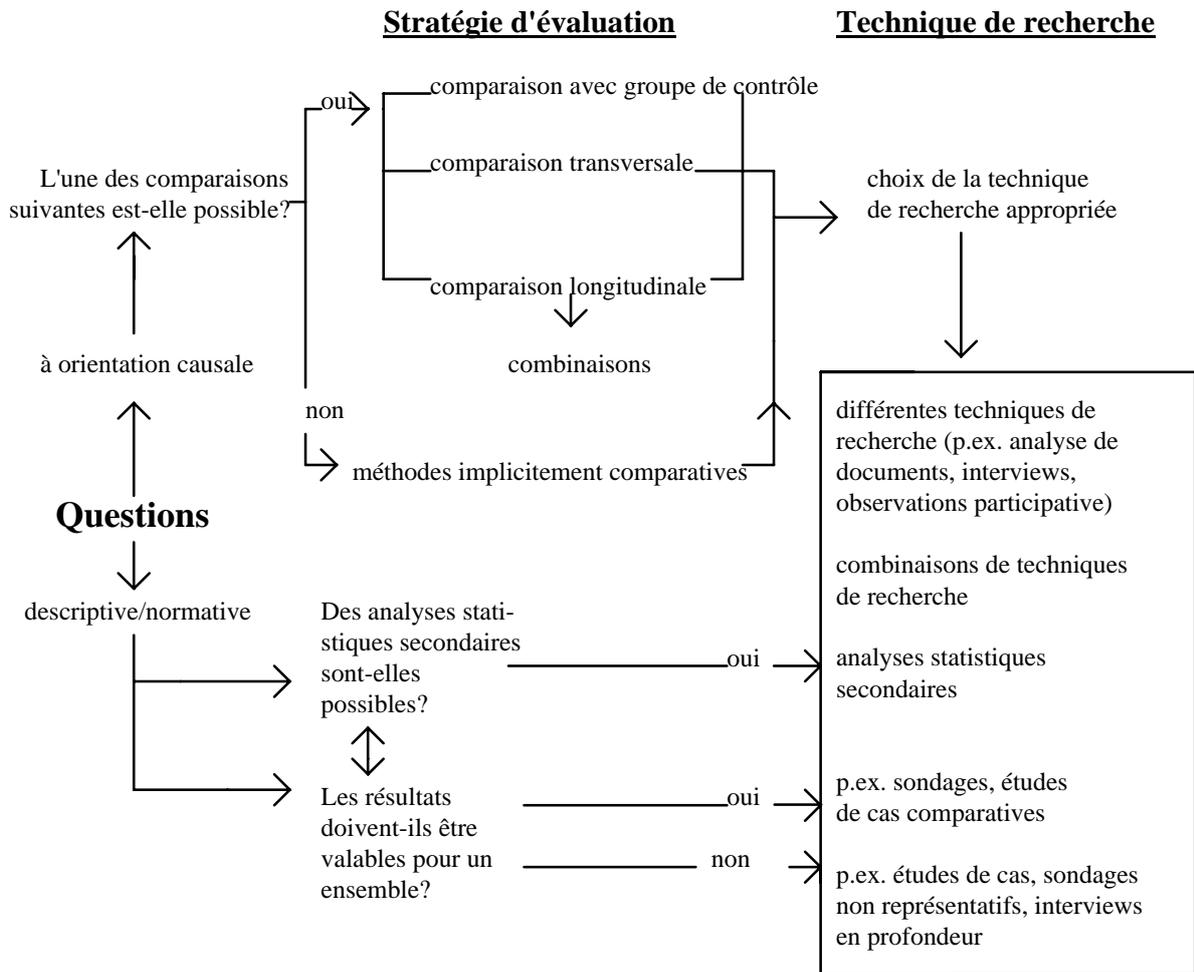
3.9 Indications sur le choix de la stratégie d'évaluation et de la technique de recherche

Le choix de la stratégie d'évaluation et de la technique de recherche est une tâche commune du prescripteur de l'évaluation et de l'équipe qui en est chargée. En règle générale, une équipe d'évaluation va aussi proposer des stratégies d'évaluation et des techniques de recherche déterminées parce qu'elle connaît bien les instruments d'analyse correspondants. La demande de plusieurs offres et leur appréciation critique peut aider à trouver la procédure optimale pour la question en cause.

Les évaluations présupposent de très bonnes connaissances des instruments des sciences sociales (et au besoin aussi des sciences naturelles). Même si, à la figure 9, on essaie de donner certaines indications sur la procédure, il faut souligner que la détermination de la stratégie d'évaluation et de la technique de recherche est un art, et par conséquent aussi un processus créatif, pour lequel il n'existe pas de recettes toutes prêtes.

Figure 9

Choix des stratégies d'évaluation et des techniques de recherche



Le choix de la stratégie d'évaluation et des techniques de recherche est un processus d'optimisation dans lequel il s'agit de situer dans une relation aussi harmonieuse que possible la question, les stratégies d'évaluation, les techniques de recherche et les impératifs financiers et temporels. C'est là une tâche extrêmement difficile, qui a néanmoins été maîtrisée par plusieurs équipes d'évaluation dans le cadre du PNR 27. Le meilleur moyen de cerner les questions et aspects qu'il s'agit de considérer consiste peut-être à partir de certaines faiblesses des évaluations que nous avons rencontrées au cours de ces dernières années dans des demandes et des projets formulés au sein et en dehors du PNR 27.

- Utilisation d'approches descriptives pour répondre à des questions à orientation causale. C'est là l'erreur de loin la plus fréquemment commise, selon notre expérience, lors du choix du design d'évaluation. La réponse aux questions à orientation causale doit être recherchée de préférence dans le cadre de stratégies d'évaluation explicitement comparatives; en tant que deuxième solution et pour le cas où des comparaisons ne sont pas possibles, on peut aussi utiliser des stratégies d'évaluation implicitement comparatives. Dans ce cas, il importe que l'équipe d'évaluation soit consciente du caractère significatif limité de ces stratégies, et n'en tire pas de conclusions irrecevables.
- Design d'évaluation trop complexe. L'élaboration d'un modèle d'effets différencié est une étape importante de toute évaluation. Avant de l'entreprendre (c'est-à-dire lors de la présentation de l'offre), l'équipe d'évaluation doit être consciente que, face à un modèle d'effets dotés de nombreuses variables, des réponses à orientation causale ne sont possibles que quand un très grand nombre de données d'observation peuvent être exploitées dans le déroulement temporel ou transversalement (comparaisons avec groupe de contrôle et

transversales). Sans cela, il n'est pas possible de déterminer l'influence des différentes variables sur le champ d'effets. Certaines équipes d'évaluation n'ont pas eu conscience de ces conditions.

- Echantillons trop petits pour les exploitations intercantionales de sondages. L'objectif de certaines évaluations est de formuler, à partir de sondages, des affirmations sur l'exécution cantonale. Mais, pour que cela soit possible, il faut avoir pour chaque canton un échantillon d'au moins 100 personnes par segment d'enquête. Ainsi, au niveau de toute la Suisse, un échantillon très nombreux est nécessaire¹².
- Transfert des résultats d'un segment à un ensemble sans garantie statistique préalable. Souvent, les équipes d'évaluation ont trop peu conscience qu'elles ont recueilli certaines données sur un segment partiel seulement (p.ex. dans certains cantons, auprès de certaines entreprises), et elles en tirent des conclusions pour un ensemble (p.ex. la Suisse entière, toutes les entreprises) plus vaste que ne le justifient les raisons statistiques.
- Techniques d'exploitation des données très coûteuses pour une qualité de données insuffisante. Pour certaines exploitations, on utilise parfois des procédures (économétriques) très coûteuses. Mais les dépenses encourues ne se justifient que si les données correspondent aux variables théoriques en cause et que leur qualité est suffisante.
- "Conversion" d'une évaluation de contrôle en évaluation d'amélioration. Nous aimerions mentionner ce point, même s'il ne touche qu'indirectement la méthodologie des évaluations. Dans plusieurs évaluations données en mandat par des instances de contrôle, l'équipe d'évaluation a essayé par la suite de transmettre des impulsions d'apprentissage aux intéressés. En règle générale, à ce moment, il est trop tard. Dans la mesure où une évaluation doit déclencher des impulsions d'apprentissage, les "mesures d'édification de la confiance" doivent intervenir au début et non à la fin d'une évaluation.

¹² Par des échantillons disposés par couches cantonales, on pourrait réduire un peu la taille des échantillons; toutefois, pour des segments sociaux partiels assez petits, la force de signification pour l'ensemble de la Suisse pourrait en être réduite.

Annexe 1 Explication des principales notions

Analyse coûts-avantages: procédure visant à déterminer l'efficacité d'une mesure en quantifiant en termes monétaires les coûts et les effets (=avantages).

Analyse coût-efficacité: détermination des coûts de différentes mesures (p.ex. réduction de la vitesse, développement du réseau routier, obligation de porter la ceinture de sécurité sur le siège arrière) pour obtenir un certain effet (p.ex. réduction du nombre des morts sur la route d'un nombre déterminé de personnes). Une telle analyse constitue une solution de rechange à l'analyse coûts-avantages quand les avantages ne sont pas quantifiables en termes monétaires.

Analyse des effets/contrôle des effets: enquête sur les effets des mesures étatiques. On examine ici aussi bien les effets conformes aux objectifs que les effets contraires. On accorde souvent une attention particulière aux effets d'aubaine, aux effets d'éviction et autres effets secondaires.

Biais: proportion dans laquelle une mesure, un échantillon ou une méthode analytique sous-estime ou surestime une valeur déterminée, ou part de conditions spécifiques concernant la théorie ou le contenu.

Comparaison avec groupe de contrôle: identification des effets d'une mesure étatique par comparaison entre la situation qui se présente dans le groupe cible et celle concernant un groupe de contrôle (non soumis à la mesure).

Comparaison longitudinale: enquête visant à tirer des conclusions sur les effets des mesures en observant des unités déterminées sur une certaine durée.

Comparaison transversale: comparaison de zones (pays, cantons, communes) dans la perspective de l'exécution et des effets de mesures étatiques déterminées.

Controlling: concept inspiré de l'économie d'entreprise et désignant un contrôle permanent des résultats. Alors que, dans l'économie privée, le controlling s'est développé en tant que complément et prolongement de la comptabilité d'entreprise, la notion est saisie de manière souvent beaucoup plus large dans le secteur public, pour inclure aussi le contrôle des effets. C'est sous cette forme qu'on utilise la notion de controlling dans le monde germanophone; les anglo-saxons utilisent avant tout les termes d'"évaluation", de "performance review", etc.

Degré de réalisation de l'objectif: rapport exprimé en termes quantitatifs entre les objectifs prescrits et les effets réellement intervenus.

Design d'évaluation: structure d'enquête concrète visant à répondre à des questions de recherche déterminées. Le design d'évaluation contient des données sur la stratégie d'évaluation et les techniques de recherche choisies.

Diachrone: désigne un concept d'évaluation qui étudie l'évolution de grandeurs déterminées dans le déroulement temporel (cf. comparaisons longitudinales).

Données qualitatives: données exprimées sous formes de mots.

Données quantitatives: information exprimée sous forme de données (exige une "mesure" en tant que procédure d'attribution de chiffres aux états de faits ou événements observés).

Echantillon représentatif: échantillon présentant la même répartition de caractéristiques que la population dont il est tiré.

Effectivité: ce critère se réfère à l'exécution de la mesure. Il désigne le degré de concordance entre le comportement légalement exigé ou prévu des responsables de l'exécution et/ou des destinataires et leur comportement observé. Selon le type de mesure, la question de l'effectivité se pose de manière particulière: degré de respect pour les injonctions et interdictions, degré d'application pour les normes de droit public qui nécessitent une mise en oeuvre par les instances administratives (p.ex. octroi d'autorisations), degré d'observation pour les normes de droit privé à caractère facultatif, mais aussi, par exemple, pour les aides financières, degré d'observation pour les informations et recommandations, etc. L'effectivité est une condition nécessaire mais non suffisante de l'efficacité.

Effets: changements dans l'économie et la société imputables à une mesure déterminée.

Efficacité: rapport entre les effets visés et les effets réellement intervenus (voir aussi → degré de réalisation de l'objectif).

Efficience: rapport entre l'output prévu ou effectif et les ressources prévues ou effectif et les ressources prévues ou effectivement investies (=efficacité en rapport avec l'output), ou rapport entre les effets prévisibles ou réels et les ressources prévues ou effectivement investies (=efficacité en rapport avec les effets). On atteint une meilleure efficacité soit en engageant moins de ressources pour obtenir un certain output ou certains effets en rapport avec le but, soit en obtenant avec les mêmes ressources plus d'output ou plus d'effets en rapport avec le but. Par efficacité, on entend souvent aussi, dans la recherche évaluative, le rapport, exprimé en grandeurs monétaires, entre les avantages/effets et les ressources/moyens investis (→ analyses coûts-avantages).

Enquêtes sur séries temporelles: enquêtes qui essaient de tirer des conclusions sur les effets des mesures par l'observation des changements d'unités déterminées sur une certaine durée.

Etude d'évaluabilité: une étude d'évaluabilité (ou pré-évaluation) sert à préparer et planifier une évaluation. Elle doit donner des éclaircissements, notamment, sur les objectifs de la mesure étatique, le modèle d'effets, la stratégie d'évaluation, les techniques de recherche et la situation des données. Une étude d'évaluabilité peut aussi avoir un sens sans être suivie d'une évaluation, et donner des indications pour une meilleure auto-évaluation des mesures.

Evaluation: le terme d'évaluation signifie l'appréciation d'un objet ou d'un état de fait. Il peut être utilisé dans un sens très large, qui ne limite ni l'état de fait à évaluer (produit, institution, mesure, domaine politique relativement vaste), ni l'objet de l'évaluation (décision d'achat, réorganisation, amélioration d'une mesure), ni la stratégie d'évaluation (méthodes de recherche des sciences sociales, appréciation impressionniste), ni l'instance d'évaluation (autorité responsable des mesures, autorité de surveillance, recherche scientifique libre, etc.). Dans le guide, nous nous limitons toutefois en premier lieu au secteur public (même si les stratégies d'évaluation peuvent parfaitement être appliquées à des questions intéressantes des entreprises, p.ex. programmes de formation interne, contributions à la promotion culturelle, etc.). Deuxièmement, nous nous concentrons sur les évaluations présentant un caractère méthodique, c'est-à-dire des évaluations qui (a) se fondent sur des réflexions et des affirmations reproductibles, (b) reposent sur une procédure systématique qui vise à saisir de façon complète les effets les plus significatifs, (c) visent à acquérir des indications aussi objectives que possible sur l'exécution et les effets, et (d) sont présentées sous une forme qui

atteste leur caractère méthodique. Troisièmement, nous nous concentrons sur l'évaluation de mesures. Il s'agit d'activités de l'Etat isolables analytiquement dans la perspective de buts déterminés, souvent formulés de manière vague, il est vrai, qui se réfèrent à des changements dans l'économie et la société. Elles comportent en règle générale des ressources en personnel et en moyens financiers.

Evaluation "boîte noire": étude et appréciation des effets sans qu'on cherche à savoir comment ils se produisent. On renonce notamment à une étude de l'exécution.

Evaluation ex ante: (→ évaluation prospective)

Evaluation ex post: (→ évaluation rétrospective)

Evaluation prospective: estimation des effets présumés de mesures prévues.

Evaluation rétrospective: saisie empirique des effets de mesures qui ont été mises en vigueur et exécutées.

Exécution: combinaison des activités de l'administration, d'une part, et des destinataires visés par les mesures étatiques, de l'autre.

Généralisabilité des résultats de l'évaluation: (→ transmissibilité)

Indicateurs de performance: statistiques qui permettent (ou doivent permettre) des déductions sur l'output des administrations et les effets qui en découlent, et contiennent des aspects tant monétaires que non monétaires.

Mesures étatiques: activités de l'Etat isolables analytiquement dans la perspective de buts déterminés, souvent formulés de manière vague, il est vrai, qui se réfèrent à des changements dans l'économie et la société. Une mesure comporte en règle générale des ressources en personnel et en moyens financiers.

Modèle d'effets: hypothèses (qui peuvent éventuellement reposer sur des théories scientifiques) sur la manière dont une mesure va exercer ses effets. Un modèle d'effets sert à concevoir ou à modifier une mesure; il constitue la base de l'évaluation des effets.

Monitoring (ou suivi): observation permanente, mise à disposition constante de données sur les aspects de l'exécution et des effets.

Output: biens fabriqués ou services fournis par l'administration (imprimés, ordonnances, autorisations, amendes, subventions accordées, etc.).

Piste d'effets: reconstruction des enchaînements d'effets à partir d'une mesure étatique donnée, éventuellement en référence à des aspects d'effets déterminés, spécifiques.

Recherche évaluative: expression utilisée dans une large mesure comme synonyme d'"évaluation", mettant toutefois au premier plan l'utilisation de méthodes des sciences sociales appliquées.

Sondage Delphi: interrogation d'experts à propos de questions touchant l'avenir. L'enquête se développe en plusieurs étapes. Premièrement, il y a en général une interrogation classique (écrite, en grande partie sur la base de questions standardisées). Dans une seconde étape,

les experts sont confrontés aux réponses de tous les autres, dans le but de les inciter à affiner leurs estimations. De nouveaux arguments peuvent venir compléter l'interrogation.

Stratégie d'évaluation: principale base de départ de l'enquête visant à répondre aux questions posées dans la perspective de l'exécution, des effets ou de l'efficacité de la mesure étatique. Réponse conceptuelle à la question de recherche.

Synchrone: désigne un concept d'évaluation qui ne fait pas varier le moment (→ comparaison longitudinale), mais étudie différents groupes, zones, etc. à un même moment (cf. comparaison avec groupe de contrôle, comparaison transversale).

Technique de recherche: méthode de sciences sociales (éventuellement aussi de sciences naturelles) choisie pour collecter et exploiter les données en vue de répondre à la question de recherche.

Transmissibilité des résultats d'évaluation: mesure dans laquelle les résultats d'une analyse d'effets peuvent être transférés (ou généralisés), au delà des unités examinées, à d'autres situations (autres groupes, autres unités spatiales, autres moments).

Validité externe: (→ transmissibilité)

Validité interne: mesure dans laquelle on peut identifier, dans une enquête/évaluation, les causes ou les effets de certains phénomènes.

Annexe 2 Glossaire des principales notions en quatre langues

| Français | Allemand | Italien | Anglais |
|-----------------------------------|------------------------------------|-------------------------|------------------------------|
| analyse coût-avantage | Kosten-Nutzen-Analyse | analisi costi-benefici | cost-benefit analysis |
| analyse coût-efficacité | Kosten-Wirksamkeits-Analyse | analisi costi-efficacia | cost-effectiveness-analysis |
| analyse des résultats | Erfolgskontrolle | analisi dei risultati | results analysis |
| effectivité | Effektivität | effettività | implementation, compliance |
| effet d'aubaine | Mitnahmeeffekt | impatto lordi | deadweights |
| effet, impact | Wirkung, Effekt | impatto | outcome, impact, effect |
| efficacité | Wirksamkeit, | efficacia | effectiveness |
| efficience | Wirtschaftlichkeit, Effizienz | efficienza | efficiency |
| emploi ménager des fonds | Sparsamkeit | economicità | economy |
| étude d'évaluabilité | Evaluierbarkeitsprüfung, Vorstudie | pre-valutazione | evaluability assessment |
| évaluation | Evaluation | valutazione | evaluation |
| évaluation endoformative | Verbesserungs-Evaluation | valutazione formativa | formative evaluation |
| évaluation récapitulative | Bilanz-Evaluation | valutazione di sintesi | summative evaluation |
| groupe cible, public cible | Zielgruppe | gruppo obiettivo | target group |
| groupe de contrôle, groupe témoin | Kontrollgruppe | gruppo di controllo | control group |
| mesure | Massnahme | metodo | measure, program |
| mise en oeuvre, exécution | Vollzug | esecuzione | implementation |
| objectif, but, finalité | Ziel | obiettivo | goal, objective, target, aim |
| partie prenante, protagoniste | Beteiligte | detentori di interessi | stakeholder |
| pilotage, suivi | Vollzugskontrolle, Monitoring | monitoraggio | monitoring |
| politique publique | Politik, öffentliche | politica | policy |
| résultats, réalisations, extrants | Ergebnis | risultati, rendimenti | outputs, results, products |

Annexe 3 Bibliographie

(+ = Résumés des recherches menées dans le cadre du PNR 27; * = Rapports finaux de projets du PNR 27; ** évaluations sommaires cofinancées par le PNR 27)

- AGREB 1990, Guide pour le Management des projets de recherche de l'administration fédérale, Berne, Office fédérale de l'éducation et de la science
- Arbeitsgruppe Verkehrssicherheit 1983, Versuch Tempo 50, Schlussbericht, OCFIM (A 22315), Berne
- Atteslander Peter 1985, Methoden der empirischen Sozialforschung, Berlin/New York
- Balthasar Andreas (Hrsg.) 1993, Einführung in die Evaluationsforschung für Politikwissenschaftlerinnen und Politikwissenschaftler, Forschungszentrum für Schweizerische Politik der Universität Bern
- Balthasar Andreas/Knöpfel Carlo 1994*, Umweltpolitik und technische Entwicklung, Eine politikwissenschaftliche Evaluation am Beispiel der Heizungen, Bâle, Helbing & Lichtenhahn
- Balthasar Andreas/Knöpfel Carlo 1994+, Umweltvorschriften und technische Entwicklung, PNR 27, Berne
- Bierter Willy/Binder Hans-Martin 1993*, Unternehmerische Innovationsprozesse, Die Wirksamkeit staatlicher Förderungsmassnahmen - eine Evaluation, Coire/Zurich, Rüegger
- Bierter Willy/Binder Hans-Martin 1994+, Unternehmerische Innovationsprozesse, Die Wirksamkeit staatlicher Förderungsmassnahmen - eine Evaluation, Berne
- Binder Hans-Martin 1994, Lernprozess als Ziel einer interaktiven Evaluationsstrategie, in: Bussmann Werner (Hrsg.), Lernen in Verwaltungen und Policy-Netzwerken, Coire/Zurich, 75-84
- Böhret Carl/Hugger Werner 1980, Test und Prüfung von Gesetzesentwürfen, Cologne
- Bussmann Werner 1992, Rechtsetzung unter Unsicherheit, in: Législation aujourd'hui, 2, 39-52
- Bussmann Werner 1994, Lernen in Verwaltungen und Policy-Netzwerken, Coire/Zurich, Rüegger Verlag
- Chelimsky Eleanor 1987, Linking Program Evaluation to User Needs, in: Palumbo Dennis J., The Politics of Program Evaluation, Newbury Park, Sage, 72-99
- Delley Jean-Daniel et al. 1982, Le droit en action, St. Saphorin
- Delley Jean-Daniel 1994* (avec la collaboration de Torricelli Fischer Claire et Rincon José-M.), Impact et limites des mesures de protection de l'air, le cas des installations industrielles et artisanales, travaux CETEL n° 42
- ECOPLAN 1993a**, Gestaltungsmöglichkeiten und Auswirkungen verschiedener Submissionsordnungen, Evaluation im Auftrag der Bau-, Verkehrs- und Energiedirektion des Kantons Bern
- ECOPLAN 1993b**, Schneeräumung von Urner Passstrassen, Evaluation im Auftrag der Baudirektion des Kantons Uri
- ERS Standards Committee 1982, Evaluation Research Society standards for program evaluation, in: Rossi Peter H. (Ed.), Standards of evaluation practice, New Directions in Program Evaluation, Vol. 15, San Francisco, Jossey-Bass
- Evaluationen in der Agrarpolitik I 1994*, Nationales Forschungsprogramm 27, Reihe Schlussberichte, Bern, mit folgendem Beitrag: Felder Stefan/Rieder Peter, Analyse des Angebotsverhaltens und der Produktionsstruktur im Agrarsektor mit Hilfe eines LP Strukturmodells
- Evaluationen in der Agrarpolitik II 1994*, Nationales Forschungsprogramm 27, Reihe Schlussberichte, Bern, mit folgenden Beiträgen: Baur Priska/Meyer Leo/Rieder Peter, Politökonomische Analyse der Getreidepolitik; Felder Stefan/Meyer Leo/Rieder Peter, Vorzeitige Stilllegung von Höchsttierbeständen in der Fleisch- und Eierproduktion 1980-

- 1984; Cavegn Gion/Meyer Leo/Rieder Peter, Analyse des Angebotsverhaltens der Getreideproduzenten
- Forster Sophie/Jeanrenaud Claude 1993**, La promotion économique neuchâteloise, bilan 1993, Neuchâtel
- Freiburghaus Dieter 1987, Präventivmassnahmen gegen die Arbeitslosigkeit in der Schweiz, Berne, Haupt
- General Accounting Office 1990a, Case Study Evaluations
- General Accounting Office 1990b, Prospective Evaluation Methods: The Prospective Evaluation Synthesis
- General Accounting Office 1991a, Designing Evaluations
- General Accounting Office 1991b, Using Structured Interviewing Techniques
- General Accounting Office 1992a, Quantitative Data Analysis: An Introduction
- General Accounting Office 1992b, The Evaluation Synthesis
- Gerheuser Frohmüt W. 1991, Die Wirkungen der beruflichen Vorsorge auf den Arbeitsmarkt, Coire/Zurich, Rüegger Verlag
- Gerheuser Frohmüt W./Schmid André 1993*, Bedingt wirksam: Das Lohngleichheitspostulat des Heimarbeitsgesetzes, Berne
- Gerheuser Frohmüt W./Schmid André 1993+, Bedingt wirksam: Das Lohngleichheitspostulat des Heimarbeitsgesetzes, Berne
- Hanser Christian/Kuster Jürg/Cavelti Guido 1994*, Auswirkungen der staatlichen Hotellerieförderung, Berne, OFCIM
- Hanser Christian/Kuster Jürg/Cavelti Guido 1994+, Auswirkungen der staatlichen Hotellerieförderung, Berne
- Hausheer Othmar 1991, Die Kehrichtsackgebühr – ein wirkungsvolles umweltpolitisches Instrument? Eine empirische Untersuchung, Coire/Zurich, Rüegger Verlag
- Hirsch Madeleine et al. 1995, Zukunft Skitourismus
- INFRAS 1990, Evaluation de la taxe sur les sacs à déchets urbains, cahier de l'environnement, no 136, OFEFP, Berne
- INFRAS 1990, Evaluation Kehrichtsackgebühr, Schriftenreihe Umwelt Nr. 136, BUWAL, Bern
- Lamnek S. 1988, Qualitative Sozialforschung, Band I und II, München
- Linder Wolf u.a. 1990, Verbrauchsabhängige Heizkostenabrechnung, Förderung erneuerbarer Energien, Forschungszentrum für schweizerische Politik der Universität Bern
- Littmann-Wernli Sabina 1991, Auswirkungen eines Diskriminierungsverbotes und eines verstärkten Kündigungsschutzes, überarbeitete Fassung eines Gutachtens im Auftrag des Bundesamtes für Justiz, in: Législation d'aujourd'hui 2, 31-82
- Lüscher Claude/Wyss Christian 1994*, Geringere Umweltbelastung dank Umweltverträglichkeitsprüfung? Berne
- Lüscher Claude/Wyss Christian 1994+ Geringere Umweltbelastung dank Umweltverträglichkeitsprüfung? Berne
- Mader Luzius 1988, Experimentelle Gesetzgebung, in: Grimm Dieter/Maihofer Werner, Gesetzgebungstheorie und Rechtspolitik, Opladen, Westdeutscher Verlag, 211 - 221
- Mader Luzius 1993, La législation expérimentale en Suisse, in: Morand Charles-Albert, Evaluation législative et lois expérimentales, Presses universitaires d'Aix-Marseille, 221-242
- Mastronardi Philippe 1991, Experimentelle Rechtsetzung im Bund, in: Revue de droit suisse, vol. 110, 449-469
- Meyer Leo/Rieder Peter 1994+, Evaluationen in der Agrarpolitik – Fallbeispiele im Getreidebau und der Tierhaltung, Berne
- Mohr Lawrence B. 1993, Impact Analysis for Program Evaluation, Newbury Park, Sage
- Monnier Eric 1987, Evaluations de l'action des pouvoirs publics, Paris
- Morand Charles-Albert (ed.) 1993, Evaluation législative et lois expérimentales, Presses universitaires d'Aix-Marseille
- Office fédéral du personnel 1993, Guide régissant l'engagement des conseillères et conseillers externes, Division développement du personnel et de l'organisation, Berne

- Patton Michael Quinn 1990, Qualitative evaluation and research methods, 2nd edition, Newbury Park, Sage,
- Price Waterhouse 1993**, Evaluation der Zürcher Gemeinschaftszentren (ZGZ), Evaluation im Auftrag des Sozialamtes der Stadt Zürich, non publié
- Rapport du groupe d'experts sur les scénarios énergétiques 1988, Les choix énergétiques, Publié par le Département fédéral des transports, des communications et de l'énergie (OFCIM)
- Rapport du groupe de travail "Méthodes d'évaluation des politiques publiques" 1986, Commissariat General du Plan, Evaluer les politiques publiques: méthode, déontologie, organisation, Paris
- Rey Michel et al. 1993*, La politique valaisanne de promotion économique: une évaluation privilégiant le point de vue des acteurs, Berne
- Rey Michel/Vodoz Luc 1993+, La promotion économique valaisanne, Berne
- Rieder Peter et al. 1990, Direktzahlungen an die Berglandschaft; Eine Gesetzesevaluation, Institut für Agrarwirtschaft, Zurich, VdF
- Rose Richard 1993, Lesson-drawing in Public Policy, A guide to Learning across Time and Space, Chatham House Publishers Inc.
- Rossi Peter H./Freeman Howard E., 1993, Evaluation: a Systematic Approach. Fifth Edition, Newbury Park, Sage
- Rossi Peter H./Freeman Howard/Hofmann Gerhard 1988, Programm-Evaluation: Einführung in die Methoden angewandter Sozialforschung, Stuttgart, Ferdinand Enke Verlag
- Sangra Emmanuel 1992*, Evaluation des actions d'information de la protection civile sur le comportement de la population en cas d'alarme, Editions J-M Blanc SA, Le Mont-sur-Lausanne
- Schaetti Stephan 1990, Berufliche Vorsorge und Arbeitsmarkt, WWZ-Studien, Nr. 18, Bâle
- Schmitz B. 1989, Einführung in die Zeitreihenanalyse: Modelle, Softwarebeschreibung, Anwendungen, Berne
- Schnyder Stefan 1993**, Über die Verteilung und Wirkung der Investitionshilfen im landwirtschaftlichen Hochbau des Kantons Zürich, Studie im Auftrag des Meliorations- und Vermessungsamtes der Stadt Zürich, non publié
- Schulz Hans-Rudolf/Muggli Christoph/Hübschle Jörg 1993*, Wohneigentumsförderung durch den Bund, Bern, Schriftenreihe Wohnungswesen, Band 55, OCFIM, Berne
- Schulz Hans-Rudolf/Muggli Christoph/Hübschle Jörg 1993+, Die Wirksamkeit der staatlichen Wohneigentumsförderung, Berne
- Stalder Peter/Frick Andres/Gaillard Serge/Salzgeber Renate 1994*, Die Auswirkungen der Ausländerpolitik auf den Strukturwandel der Schweiz, Berne
- Stalder Peter/Frick Andres/Gaillard Serge/Salzgeber Renate+, Die Auswirkungen der Ausländerpolitik auf den Strukturwandel der Schweiz, Berne
- Wholey Joseph S. 1983, Evaluation and Effective Public Management, Boston, Little Brown
- Widmer Thomas 1991, Evaluation von Massnahmen zur Luftreinhaltepolitik in der Schweiz, Eine quasi-experimentelle Interventionsanalyse nach dem Ansatz Box/Tiao, Zurich, Rüegger Verlag
- Yin Robert K. 1985, Case study research, Newbury Park, Sage