



Fonds structurels communautaires

*Évaluer les programmes socio-économiques*

**Principales techniques et outils d'évaluation**

**Collection**  
**MEANS**

**Volume 3**

Commission européenne



*Évaluer les programmes socio-économiques*

Principales techniques et  
outils d'évaluation

*Volume 3*

*MEANS est un programme de la Commission européenne.  
Il vise à améliorer les Méthodes d'Evaluation des Actions de Nature Structurelle.  
Le programme et la Collection MEANS ont été lancés et supervisés  
par MM Philippe Goybet et Miguel Angel Benito Alonso  
DG XVII/G2, Unité de Coordination de l'Evaluation  
Coordination-evaluation@fmb.dg16.cec.be*

*Sa réalisation a été confiée au Centre Européen d'Expertise en Evaluation (C3E)  
sous la direction d'Eric Monnier et de Jacques Toulemonde.  
Les opinions exprimées dans cet ouvrage n'engagent que les auteurs du texte :  
C3E - 13bis Place Jules Ferry - F-69006 Lyon - France - c3e@c3e.fr*

De nombreuses autres informations sur l'Union européenne sont disponibles sur Internet via le serveur Europa (<http://europa.eu.int>), et sur le site Inforegio (<http://inforegio.cec.eu.int>).

Une fiche bibliographique figure à la fin de l'ouvrage.

Luxembourg:

Office des publications officielles des Communautés européennes, 1999

ISBN 92-828-6627-0

© Communautés européennes, 1999

Reproduction autorisée, sauf à des fins commerciales, moyennant mention de la source.

IMPRIMÉ SUR PAPIER BLANCHI SANS CHLORE

*Printed in Italy*

## Préface



Développer des pratiques appropriées en matière d'évaluation est devenue l'une des actions prioritaires de la Commission afin de contribuer à la bonne utilisation des financements communautaires.

En effet, le respect des principes de bonne gestion financière qu'elle cherche à promouvoir dans son programme SEM2000 (Sound and Efficient Management) repose notamment sur la nécessité de procéder à l'évaluation régulière des actions communautaires.

Les Directions Générales en charge des politiques structurelles possèdent dans ce domaine une certaine expérience, car elles ont dû intégrer la démarche d'évaluation dès 1988, tout au long du cycle des interventions structurelles communautaires : lors de leur conception sous la forme d'une évaluation prospective, lors de leur mise en œuvre en tant que renforcement du système de suivi et, enfin, à l'issue de celles-ci afin de tirer les enseignements des impacts obtenus.

Cependant, la capacité de l'évaluation à contribuer à l'efficacité et à la transparence de ces actions doit s'appuyer sur la crédibilité des méthodes qu'elle emploie afin qu'elles puissent garantir la validité des résultats.

C'est pourquoi, il y a quatre ans, à l'initiative de la Direction Générale de la Politique Régionale et de la Cohésion (DG XVI), le programme MEANS a été conçu comme une réponse à ce souci de développer des méthodologies et des pratiques d'évaluation conciliant la rigueur scientifique et les spécificités propres aux interventions communautaires, tout en demeurant volontairement accessibles au plus grand nombre.

L'un des résultats les plus tangibles du programme MEANS est la publication de cette collection de six ouvrages rédigés par un groupe d'experts indépendants qui traitent des grands aspects de l'évaluation des programmes socio-économiques dont un bref aperçu est donné au début de ce volume. Cette collection capitalise des expériences, des échanges et des innovations méthodologiques dans le contexte des Fonds Structurels. Elle permettra aux responsables de l'évaluation de développer leurs

pratiques de manière autonome et aux commanditaires de mieux maîtriser cette démarche.

Témoign direct lors des grandes rencontres européennes entre spécialistes et utilisateurs de l'évaluation (Bruxelles 1995, Berlin 1996, Séville 1998) de l'engouement que suscitent ces nouvelles techniques, il m'appartient d'insister sur le fait que le développement de l'évaluation doit aller de pair avec la recherche d'une véritable insertion de ses résultats dans les processus décisionnels afin de savoir si les politiques menées sont équitables, efficaces et adaptées aux besoins.

En répondant à ces exigences, l'évaluation apparaît comme un moyen privilégié offrant tant aux autorités nationales que communautaires, les garanties nécessaires pour que les fonds communautaires soient utilisés au mieux des intérêts des citoyens de l'Union. Les enjeux de la Cohésion Economique et Sociale ainsi que les défis que représente le futur élargissement de l'Union européenne sont à ce prix.

Je me dois enfin de remercier le Comité des Personnalités Indépendantes constitué dès le lancement du programme qui par leur compétence et leur renom ont grandement contribué à asseoir la qualité et la pertinence méthodologique du programme MEANS. Il s'agit de : E. Chelimsky (USA), ex-Présidente de l'«American Evaluation Association» ; K. Kennedy (IRL), Directeur de l'«Economic and Social Research Institute» (ESRI) ; J. R. Cuadrado Roura (E), Président de l'«European Regional Science Association» ; C. Seibel (F), Directeur des Statistiques démographiques et sociales de l'Institut National de la Statistique et des Etudes Economiques (INSEE) ; L. Senn (I), Directeur du Département d'Economie Régionale de l'Université Bocconi ; E. Stern (UK), Directeur de l'unité d'évaluation du Tavistock Institute ; A. Sorber (†) (NL), Chef du «Policy Analysis Department of the Ministry of Finance» ; L. I. Strom (SV), Statens Institut för Regionalforskning (SIR) ; L. Tsoukalis (GR), professeur à l'Université d'Athènes ; H. Wollmann (D), professeur à l'Université de Humboldt à Berlin et membre fondateur de la «European Evaluation Society».

**E. Landaburu**





## La Collection MEANS :

### « Evaluer les programmes socio-économiques »

La démarche d'évaluation des interventions structurelles de l'Union européenne est entrée dans une phase de maturité ainsi qu'en atteste l'extension prise par les travaux réalisés tant par les Etats membres que par la Commission. Les obligations réglementaires issues des décisions du Conseil ont largement contribué à répandre la pratique de l'évaluation et ont également stimulé le développement méthodologique.

Pendant ces dix dernières années, la Commission s'est employée à promouvoir et améliorer les méthodes d'évaluation des interventions des Fonds structurels. Les principaux enseignements méthodologiques ont été débattus et diffusés dans le cadre du programme MEANS initié par la Commission dès 1991. Les acquis sont maintenant suffisamment substantiels pour nourrir une publication complète et accessible pour un public plus large : une Collection de six ouvrages traitant tous les aspects de l'évaluation des programmes socio-économiques. Ces ouvrages sont rédigés dans une perspective plus large que celle du strict cadre administratif actuel. En effet, celui-ci est amené à évoluer périodiquement alors que les standards professionnels ont une plus grande stabilité.

De manière générale, la Collection s'adresse à des lecteurs non-spécialistes désireux de disposer des informations indispensables à l'évaluation des programmes de développement socio-économiques. Ces programmes ont une multiplicité d'objectifs, une dimension de partenariat et un contenu multisectoriel, plusieurs caractéristiques qui appellent des méthodes d'évaluations adaptées et des solutions que n'offre pas encore la littérature technique internationale. La Collection complète comprend 6 Volumes conçus pour

être complémentaires et fournir les éléments indispensables au professionnel qui souhaite développer de façon autonome ses pratiques d'évaluation. Chaque Volume a été construit de telle manière qu'il puisse être lu et compris indépendamment des autres.

Le premier Volume de la Collection MEANS s'intitule « *Conception et conduite d'une évaluation* ». Il a une fonction de porte d'entrée dans la Collection, ainsi que dans le domaine de l'évaluation des programmes socio-économiques. Il permet au lecteur de comprendre les enjeux de l'évaluation et de savoir ce qu'il peut ou ne peut pas attendre d'une évaluation à un moment et dans un contexte donnés. Il apporte les éléments nécessaires pour piloter une évaluation de A à Z : répartir les responsabilités de l'évaluation, définir sa finalité et les grandes lignes de sa méthode, fixer son calendrier et son budget, apprécier la qualité des travaux et enfin gérer l'utilisation de ses conclusions.

Le second Volume porte sur « *le choix et l'utilisation des indicateurs pour le suivi et l'évaluation* ». Il traite des indicateurs utilisés pour le suivi et l'évaluation. Il montre comment les indicateurs peuvent être utilisés pour appuyer des décisions politiques, pour renforcer la gestion et pour mesurer la performance globale des programmes. L'ouvrage propose un cadre méthodologique et un guide pratique d'utilisation des indicateurs dans ces différentes situations.

Le troisième Volume présente une sélection des « *Principales techniques et outils d'évaluation* » applicables à l'évaluation des programmes socio-économiques. Ces outils sont issus des disciplines scientifiques pertinentes : économie, sociologie, management, géographie, etc. Chacun des 23 outils présentés est décrit sous forme d'une fiche comprenant sa description générale, les principes de sa mise en œuvre, une liste de ses points forts et faibles pour l'évaluation des Fonds structurels, un exemple

d'application et une bibliographie des ouvrages de référence.

Plusieurs méthodes adaptées à l'évaluation des programmes socio-économiques font l'objet d'une description plus complète dans le quatrième Volume intitulé « *Solutions techniques pour évaluer dans un cadre de partenariat* ». Plusieurs évaluations pilotes du programme MEANS ont permis de concevoir et de tester des techniques nouvelles dont le mode d'emploi est décrit en détail : cartographie conceptuelle d'impact, matrice d'impacts croisés, échelles de notation et analyse multicritère.

Le cinquième Volume traite de « *l'évaluation transversale des impacts sur l'environnement, l'emploi et les autres priorités d'intervention* » de l'Union Européenne et des Etats membres. Il aborde entre autres la question des impacts en termes d'égalité des chances hommes-femmes ou de compétitivité des PME. Le Volume offre les éléments nécessaires à l'analyse de ces impacts qui ne figurent pas nécessairement parmi les objectifs explicites de toutes les actions évaluées.

Enfin, le dernier Volume est constitué d'un « *glossaire des concepts et termes techniques* ». La mise au point de la Collection a en effet imposé un travail rigoureux d'homogénéisation et de définition des termes employés. Chacun des cinq premiers volumes est suivi d'un index dont les termes sont repris parmi les centaines de définitions rassemblées dans le glossaire.



# Table des matières

<b>Introduction</b>	<b>9</b>
<b>I Présentation des outils d'évaluation</b>	<b>15</b>
■ Outil, fonction, méthode	17
■ Propriétés des outils d'évaluation	19
■ Avantages et risques de la technicité	22
■ Etapes et fonctions d'une évaluation	24
■ Les situations d'évaluation	29
<b>II Outils pour l'évaluation d'ensemble d'un programme</b>	<b>35</b>
Exemples d'évaluations de programmes multisectoriels	37
Outils pour structurer l'évaluation	41
■ Analyse SWOT (Forces-Faiblesses-Opportunités-Menaces)	43
■ Cartographie conceptuelle d'impacts	49
■ Vote coloré	57
Outils pour observer les changements sur le terrain	61
■ Entretien individuel	63
■ Entretien de groupe ( <i>Focus Group</i> )	71
■ Etude de cas	77
Outils pour analyser les données	85
■ Système d'Information Géographique	87
■ Analyse structurelle - résiduelle ( <i>Shift-Share</i> )	93
■ Modèle entrées-sorties	99
■ Modèle macro économique	107
Outils pour porter un jugement	121
■ Panel d'experts	123
■ Analyse multicritère	129
<b>III Outils pour approfondir une question évaluative</b>	<b>135</b>
Exemples d'évaluations traitant une question de façon approfondie	137
Outils pour structurer l'évaluation	141
■ Cadre logique	143
■ Metaplan®	151

Outils pour observer les changements sur le terrain	157
■ Enquête par questionnaire	159
■ L'observation ethnographique	169
Outils pour analyser les données	175
■ Analyse factorielle	177
■ Enquête Delphi	183
■ Groupes de comparaison	187
■ Analyse de régression	195
Outils pour porter un jugement	201
■ Analyse Coût-Efficacité	203
■ Etalonnage (Benchmarking)	209
■ Analyse Coût-avantages	215
<b>IV Recommandations pour sélectionner et assembler les outils</b>	<b>227</b>
Guide pour repérer les outils adaptés	229
Recommandations pour assembler plusieurs outils	235
<b>Conclusion</b>	<b>241</b>
<b>Index</b>	<b>243</b>

## Introduction

Un professeur qui veut connaître le rôle et le rôle de l'évaluation, doit connaître les méthodes rigoureuses et rigoureuses de la recherche en éducation. Les sciences éducatives et sociales à l'ère technologique ont des documents de référence qui présentent un aperçu des outils utilisés pour évaluer. Parmi ces ouvrages on cite: *Principes de l'évaluation* (1998) ainsi que le *Collectif* (1997). Ces deux ouvrages contiennent des exemples et des modèles de programmes d'évaluation et de plans d'évaluation. Les données technologiques de l'évaluation sont présentées dans les pages de ce livre. Les données quantitatives ou qualitatives de l'évaluation, etc. Une typologie de l'évaluation est présentée en annexe par C. Robert (1998).

Il y a le plan pour un nouvel ouvrage sur les outils d'évaluation et il est, quelle est la valeur ajoutée ?

Cet ouvrage est d'abord nécessaire pour mettre les connaissances existantes à la portée des lecteurs qui travaillent dans le champ des programmes socio-économiques. En effet, les ouvrages disponibles sur ce sujet sont le plus souvent écrits par des experts de programmes sociaux ou de plans de développement. Le présent ouvrage est un petit guide de l'évaluation méthodologique générale et le contenu des livres suivants suivants.

En second lieu, il est nécessaire d'adapter les connaissances disponibles aux difficultés spécifiques de l'évaluation des programmes socio-économiques. En effet, les connaissances en matière d'évaluation sont souvent limitées à des programmes sociaux, à des plans de développement, à des programmes de développement socio-économique, etc. Le présent ouvrage est un petit guide de l'évaluation méthodologique générale et le contenu des livres suivants suivants.

© 2000 P.O. Box 1000, Vancouver, B.C. V6C 1A8  
Revised Edition, 2000  
Revised Edition, 2000  
Revised Edition, 2000



Un professionnel qui veut connaître la boîte à outils de l'évaluation doit consulter les nombreux textes et manuels consacrés à telle ou telle technique des sciences économiques et sociales. Il existe heureusement quelques documents de référence qui présentent un panorama des outils pertinents pour l'évaluation. Parmi les classiques, on citera : Rossi & Freeman (1993) ainsi que la Collection Program Evaluation Kit <sup>1</sup>. Tous ces ouvrages comprennent des « fiches techniques » illustrées par des exemples et organisées selon une typologie propre aux auteurs. Les distinctions typologiques les plus fréquemment rencontrées séparent les outils selon leur caractère quantitatif ou qualitatif, déductif ou inductif, etc. Une typologie particulièrement intéressante est présentée par C. Pollitt (Encadré 1).

Y a-t-il place pour un nouvel ouvrage sur les outils d'évaluation, et si oui, quelle est sa valeur ajoutée ?

Cet ouvrage est d'abord nécessaire pour mettre les connaissances existantes à la portée des lecteurs qui travaillent dans le champ des programmes socio-économiques. En effet, les ouvrages consacrés aux outils d'évaluation sont le plus souvent illustrés par des exemples de programmes sociaux ou de projets de développement. Le présent ouvrage crée un pont entre les ouvrages méthodologiques généraux et le contexte des Fonds structurels européens.

En second lieu, il est nécessaire d'adapter les connaissances disponibles aux difficultés spécifiques de l'évaluation des programmes socio-économiques. En effet, les textes relatifs aux outils d'évaluation font le plus souvent référence à des programmes simples, visant un public bien défini et cherchant à produire un changement limité à ce public. A l'inverse, un programme de développement socio-économique est typiquement multisectoriel et multiobjectifs. Il cherche à créer des transformations économiques ou sociales à l'échelle de tout un territoire. Pour atteindre son but, il adopte une stratégie intégrée, dirigée vers de nombreux publics (entreprises, salariés, chômeurs, collectivités, etc.). Il doit produire son impact global grâce à la bonne combinaison de multiples impacts spécifiques sur ses différentes cibles. Dans ce contexte, l'évaluation rencontre des problèmes techniques ardues auxquels la littérature existante apporte peu de solutions. Cet ouvrage est le premier guide méthodologique qui recense et qui présente ces solutions.

1 Rossi P.H, Freeman H.E. (1993) Evaluation: a Systematic Approach (5th edition) Newbury Park CA: Sage.  
Program Evaluation Kit, Center for the Study of Evaluation, University of California. Newbury Park CA: Sage

### **Encadré 1 - Quatre approches de l'évaluation**

Les approches expérimentales traitent les phénomènes de causalité sans les analyser directement. Les effets des programmes sont démontrés par comparaison entre un groupe expérimental et un groupe témoin. Les évaluations conduites dans ce cadre sont représentatives des pratiques américaines des années 60/70. Il s'agit d'évaluations de longue durée, basées sur des méthodes lourdes. Ces évaluations ne sont pas toujours concluantes. Elles ont donné lieu à des critiques nombreuses et parfois sévères de la part des commanditaires, à la recherche de conclusions opérationnelles, rapidement utilisables pour l'amélioration des programmes.

Les approches économiques sont fondées sur une vision individualiste de la société, inspirée de l'économie du bien-être. Dans cette optique, la valeur d'une action publique n'est rien d'autre que la somme des avantages qu'elle procure aux individus. Tous les individus sont supposés avoir un système de références et on suppose que ces systèmes de référence peuvent être agrégés. Ces hypothèses permettent de donner un caractère objectif aux évaluations réalisées par les techniques économiques. Malgré leur élégance, les techniques économiques ont été fortement critiquées, en termes pratiques aussi bien que théoriques.

Les approches naturalistes ou pluralistes partent du point de vue que le monde politique et social est une construction collective. Cette construction résulte de l'interaction de différents groupes sociaux qui ont des interprétations différentes des mêmes phénomènes et enjeux. C'est pourquoi les programmes publics sont perçus comme des compromis transitoires entre groupes d'acteurs. A l'aide de techniques appropriées, reposant beaucoup sur le travail en groupe, l'évaluation joue un rôle de médiation entre les différents points de vue; elle est un outil de conciliation des protagonistes, ses conclusions sont d'autant plus robustes qu'elles ont été produites de façon consensuelle.

Les approches pragmatiques adoptent des visions simplifiées des processus qu'elles ont pour objectif de décrire. Les références théoriques sont moins pures, plus éclectiques. Les objectifs en matière d'évaluation sont plus de nature gestionnaire que scientifique. A l'aide de techniques inspirées du management, les évaluations conduites dans cet esprit sont plus orientées vers la mise en œuvre et l'amélioration de l'efficacité. Elles n'ont pas autant de dimension cognitive que les autres formes d'évaluation.

Certaines propositions techniques présentées dans cet ouvrage n'existeraient pas sans les nombreuses recherches et expérimentations qui ont été conduites par la Commission européenne. Pour adapter les connaissances disponibles au domaine difficile des programmes socio-économiques, il a fallu faire des investissements méthodologiques importants : définir précisément les concepts de base (ex : outil, méthode), définir les fonctions susceptibles d'être remplies par les outils, construire de nouveaux outils (ex : modèles macroéconomiques spécialisés, cartographie conceptuelle des impacts) ou proposer de nouvelles façons

d'employer des outils classiques. Ces investissements ont un intérêt pour l'ensemble de la communauté professionnelle de l'évaluation. C'est la troisième et dernière raison qui justifie l'existence de cet ouvrage.

L'ouvrage contient une présentation détaillée de vingt-trois outils d'évaluation sous forme de fiches techniques. Le choix des outils fait un compromis entre les outils les plus fréquemment rencontrés et certains outils plus rares mais qui représentent un potentiel important pour l'évaluation des programmes socio-économiques. Tout compromis contient une part d'arbitraire et c'est pourquoi certains outils fréquemment employés sont simplement évoqués sans faire l'objet de fiches techniques (ex : analyse de variance).

Chaque fiche donne des informations sur les fonctions de l'outil, sur les conditions de son emploi, et sur les principales étapes de sa mise en œuvre. Elle renseigne sur les possibilités d'application dans le cadre des Fonds structurels, les points forts et les limites d'utilisation. Elle propose en outre un exemple issu de l'évaluation des Fonds structurels si cet outil a été utilisé dans ce cadre. Chaque fiche contient également une bibliographie commentée de quelques ouvrages qui font référence pour cet outil. Les fiches sont accompagnées de pictogrammes pour aider le lecteur à effectuer une présélection rapide des outils applicables dans une situation donnée. L'encadré 2 donne la signification de ces pictogrammes.

Cet ouvrage s'organise selon la structure suivante :

- Présentation des outils d'évaluation
- Outils pour l'évaluation d'ensemble d'un programme (12 outils)
- Outils pour l'approfondissement d'une question évaluative (11 outils)
- Recommandations pour sélectionner et assembler les outils

**Encadré 2 - Une lecture rapide de l'ouvrage pour présélectionner les outils**



Technicité forte



Technicité moyenne



Technicité faible



Usage rétrospectif



Usage prospectif



Coût faible



Coût moyen



Coût élevé



Usage stabilisé



Usage innovant



Evaluation d'ensemble



Approfondissement

# I

## Présentation des outils d'évaluation

*Cet ouvrage considère les termes outil et technique comme des synonymes, applicables à des procédures standardisées que tout évaluateur peut apprendre et appliquer dans le cadre d'une de ses missions.*

*La « boîte à outil » de l'évaluation contient plus que les 23 techniques qui sont présentées dans ce Volume. L'art de l'évaluateur implique donc de choisir tel ou tel outil selon la fonction qu'il est capable de remplir dans une situation donnée, et également de rejeter l'emploi d'outils inadaptés.*



## ■ Outil, fonction, méthode

Les praticiens de l'évaluation emploient toutes sortes de mots pour décrire leurs instruments ou leurs procédures de travail : technique, outil, approche, démarche, méthode, *design*, etc. Il est indispensable de s'accorder sur quelques définitions avant de commencer la lecture de cet ouvrage. Cela va être fait avec l'aide de l'exemple ci-dessous.

### Encadré 3 - Un exemple d'évaluation outillée

En 1989, le National Audit Office britannique a réalisé l'évaluation de huit programmes de développement régional dans le Pays de Galles. L'évaluation a comporté un calcul du coût par emploi net créé, c'est-à-dire qu'elle a été fondée sur l'analyse coût-efficacité, un outil classique de l'évaluation. Pour évaluer les huit programmes, le NAO a adopté une démarche en plusieurs étapes :

- 1 - Sélection d'un critère d'évaluation unique : les emplois nets créés.
- 2 - Construction d'un indicateur de mesure d'emploi net, selon la définition suivante : emplois nets = emplois bruts - effets d'aubaine - effets de déplacement + effets sur les fournisseurs.
- 3 - Sélection d'un échantillon d'environ 300 entreprises aidées.
- 4 - Enquête par questionnaire, passée par téléphone ; estimation de coefficients d'effets d'aubaine et de déplacement.
- 5 - Estimation de l'emploi total créé par extrapolation des coefficients obtenus à l'ensemble des emplois déclarés par toutes les entreprises aidées.
- 6 - Estimation de l'effet sur les fournisseurs par extrapolation à partir d'études portant sur des programmes semblables.
- 7 - Récapitulation des fonds publics mobilisés pour les programmes et analyse coût-efficacité en termes de coût par emploi créé.
- 8 - Comparaison avec le coût par emploi créé dans le cadre de programmes similaires, en l'occurrence un programme de développement régional en Ecosse.

Pour en savoir plus sur cette évaluation : S. Roberts and Ch. Pollitt, (1994) *Audit or Evaluation ? a National Audit Office VFM Study*, Public Administration.

Dans cet exemple, on reconnaît au moins deux outils couramment employés en évaluation :

- l'enquête par questionnaire (4<sup>e</sup> étape)
- l'analyse coût-efficacité (7<sup>e</sup> étape)

On remarque également que le travail d'évaluation a dépassé de beaucoup la simple application de deux outils. A chacune des huit étapes parcourues, une décision a été prise sur la façon de conduire le travail. Chaque décision était

dépendante du contexte précis de cette évaluation. Le N.A.O. a donc construit une méthode d'évaluation ad hoc pour réaliser la mission qui était la sienne.

Cet exemple va servir à définir trois des termes de base de cet ouvrage : « outil », « méthode », et « fonction ».

Un **outil** réalise un traitement standard et peut être employé dans n'importe quelle évaluation qui requiert ce traitement. Par exemple :

- une enquête par questionnaire permet de recueillir l'opinion d'un groupe sur un ou plusieurs points définis à l'avance ;
- une analyse coût-efficacité calcule l'argent public qui a été dépensé pour obtenir un effet donné ; le résultat se présente sous une forme qui permet des comparaisons entre programmes ou projets.

Une technique est plus ou moins la même chose qu'un outil, à tel point que les deux termes sont souvent employés comme synonymes dans ce Volume. On peut cependant considérer l'outil comme un objet (ex: la pelle) et la technique comme la façon d'employer cet objet (ex: le maniement de la pelle).

Une **méthode** est une procédure ad hoc, construite spécialement pour réaliser une évaluation donnée. Elle peut comprendre un ou plusieurs outils. Il arrive qu'elle ne soit absolument pas outillée. Dans son évaluation des programmes de développement régional du Pays de Galles, le NAO a construit une méthode adaptée aux questions qui étaient posées (Encadré 3). A l'inverse, pour réaliser l'appréciation ex ante des programmes proposés par les Etats membres, la Commission commande des expertises réalisées en quelques semaines. Les méthodes appliquées pour ces expertises ne comportent généralement aucun outil.

Ainsi défini, le terme « méthode » est proche du terme « *design* ». Dans la littérature américaine, « *evaluation design* » s'applique à la façon de conduire une évaluation donnée. Cependant, le terme s'applique le plus souvent aux fonctions d'analyse de causalité, c'est-à-dire à une partie seulement des fonctions que doit remplir une méthode d'évaluation. Les termes « approche » et « démarche » sont également voisins. Ils s'appliquent à la fois à une méthode donnée et à l'esprit dans lequel cette méthode a été construite.

La méthode construite pour réaliser une évaluation donnée peut toujours se décomposer en plusieurs étapes. Chacune de ces étapes réalise une tâche et remplit une **fonction** élémentaire. On retrouve les mêmes fonctions élémentaires de base dans la plupart des évaluations. Une fonction peut être remplie avec ou sans l'aide d'un outil. Par exemple :

- dans l'exemple du NAO (Encadré 3), la fonction de collecte des données est réalisée à l'aide d'une enquête par questionnaire.
- dans certaines évaluations de la Commission, la fonction d'estimation des effets est réalisée à l'aide d'un modèle macroéconomique.

Dans ce Volume, la méthodologie est définie comme la « science » de la construction des méthodes. Cet ouvrage a pour ambition de contribuer à faire émerger cette science. Il faut cependant admettre que la pratique professionnelle est encore plus proche de l'expertise que de la science. Le mot méthodologie est donc un peu ambitieux pour le moment.

## ■ Propriétés des outils d'évaluation

### Encadré 4 - La pelle de terrassier et l'analyse coût-efficacité

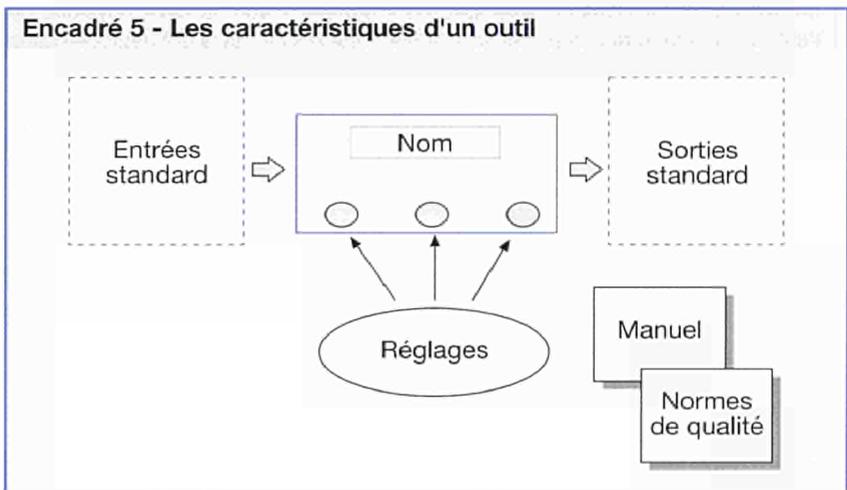
La pelle de terrassier permet de remplir une fonction standardisée, consistant à déplacer un matériau meuble d'un endroit à un autre, par exemple remplir une brouette de terre. Dans cet exemple, les « entrées » sont la force physique de l'utilisateur et la terre à déplacer, la « sortie » est la terre qui a été déplacée dans la brouette. Il n'existe a priori pas beaucoup de formes de pelles ni de façons différentes de s'en servir. On peut dire que cet outil réalise une transformation stable. Les principales variantes sont la taille de la pelle, la distance du déplacement et la nature plus ou moins meuble du matériau à déplacer. Le cas de la pelle peut être transposé terme à terme dans le domaine de l'évaluation, par exemple à l'analyse coût-efficacité.

Les entrées de l'analyse coût-efficacité sont doubles. Il faut d'abord disposer d'une mesure de l'impact obtenu (mesure à une seule dimension, comme, par exemple, les emplois nets créés) puis d'une mesure des sommes dépensées. La sortie est simple, c'est le ratio de coût par unité d'impact. La transformation est stable, il s'agit d'une simple division. Les degrés de liberté concernent surtout la manière de prendre en compte le coût : coût budgétaire pour tous les partenaires publics, coût budgétaire pour le seul acteur public qui commande l'évaluation, coût d'opportunité des fonds publics, coût limité à la durée du programme ou coût étendu à toutes les aides publiques, etc.

Un outil sert à traiter une matière première (des entrées) pour en faire un produit (des sorties). Les entrées et les sorties sont des informations, des données, des faits, des opinions, des jugements, etc. Les entrées et les sorties peuvent être de nature différente. Par exemple, une enquête par questionnaire collecte des opinions et produit des données quantitatives à propos de ces opinions. Un panel d'expert est alimenté par des

informations de toutes sortes et il produit un jugement, etc. L'outil se définit par le fait que son mode de traitement est stable, c'est-à-dire que les entrées ont une forme standard, que les sorties sont toujours de même nature et que les transformations sont réalisées selon les mêmes principes. Cependant un outil présente toujours une certaine souplesse qui lui permet de s'adapter à ses situations d'emploi. Ces « réglages » sont toutefois assez réduits pour que l'outil reste identifiable.

Un outil peut être librement utilisé par quiconque en a l'usage. La plupart des outils employés en évaluation sont à la fois libres et gratuits, mais certains outils doivent être achetés. Par exemple, l'outil d'animation de réunion Metaplan® qui est présenté dans cet ouvrage est une marque déposée. Un outil possède un mode d'emploi. La façon de s'en servir est décrite dans un ou plusieurs manuels. Il existe des cours de formation qui permettent un apprentissage et une appropriation par des débutants. Des normes de qualité permettent de vérifier si l'outil a été employé de façon professionnelle.

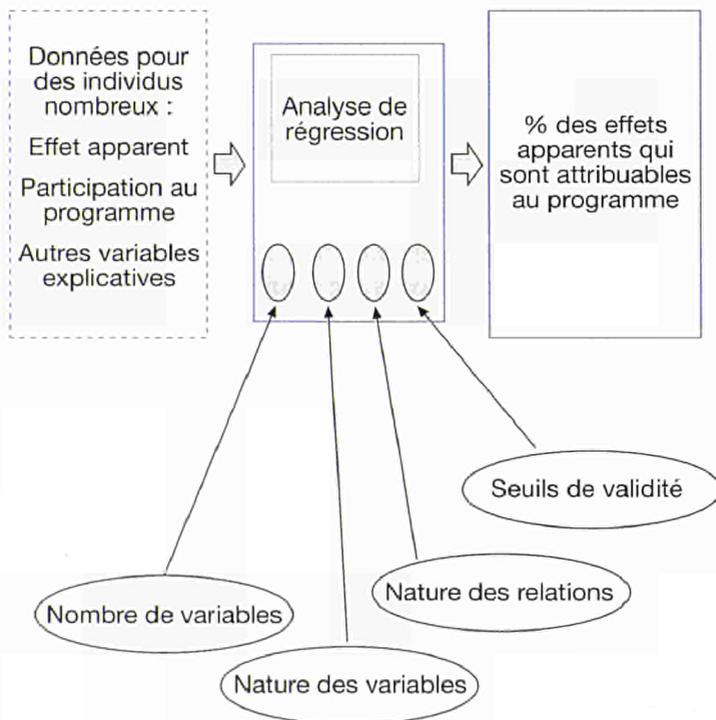


Pour résumer, on reconnaît un outil aux caractéristiques qui figurent dans l'encadré 5 : il est identifiable par un nom, il réalise un traitement dont les entrées et les sorties sont standard, il dispose de réglages mais en nombre limité, il dispose également d'un manuel et de normes de qualité. Toutes ces propriétés rendent l'outil transférable d'une évaluation à une autre et d'un professionnel à un autre. En particulier, son usage est public et n'est pas réservé à son inventeur.

Ce schéma s'adapte parfaitement à certains des outils qui sont présentés dans ce Volume comme le montre l'exemple de l'analyse de régression développé ci-après. Dans cet exemple, on reconnaît les caractéristiques de base d'un outil, à savoir :

- des entrées standard (données quantitatives sur l'effet apparent, sur la participation au programme et sur les autres variables explicatives, ces données étant disponibles pour un nombre important d'individus),
- quelques réglages (nombre et nature des variables, nature des relations, seuils de validité),
- des sorties standard (répartition des effets apparents ou bruts entre les effets nets attribuables au programme et les effets des autres variables explicatives).

#### Encadré 6 - L'analyse de régression comme outil d'évaluation



Par ailleurs, l'outil est bien identifié par un nom (analyse de régression), quelles que soient les variantes correspondant aux différents réglages (régression linéaire, régression logistique, etc.). Enfin, l'outil est décrit dans de nombreux manuels et dans toutes les langues européennes.

Parmi les 23 fiches qui sont présentées dans ce Volume, plusieurs présentent une partie seulement des caractéristiques d'un véritable « outil ». Par exemple, certaines sont relativement peu standardisées (ex : analyse SWOT), d'autres ne disposent pas de véritable manuel (ex : vote coloré, Metaplan®), d'autres encore n'ont pas de normes de qualité stabilisées (ex : analyse multicritère).

A part quelques outils qui sont nés avec l'évaluation (analyse coût-avantages, analyse multicritère), les outils sont généralement empruntés aux grandes disciplines scientifiques. Ainsi, l'enquête par questionnaire est empruntée à la sociologie, l'analyse de régression a été créée par des statisticiens, le modèle d'entrées-sorties (*input-output*) vient de l'économie, le panel d'expert emprunte beaucoup aux pratiques juridiques, le système d'information géographique et l'observation ethnographique font référence aux sciences du même nom.

Lorsqu'une technique est « empruntée » à une discipline pour être rangée parmi les outils de l'évaluation, cela s'accompagne souvent d'adaptations et de simplifications. D'un point de vue scientifique pur, on regrettera toujours de ne pas retrouver la forme « canonique » de la technique originelle, mais l'évaluation est avant tout une activité pragmatique. Les fiches techniques présentées dans cet ouvrage ne doivent donc pas être jugées avec la rigueur académique de leurs disciplines d'origine, mais bien en fonction de leur utilité pour les praticiens de l'évaluation.

### ■ Avantages et risques de la technicité

Le recours à un outil a de nombreux avantages en termes d'assurance sur le résultat obtenu, de sécurité sur le contenu de l'évaluation, de maîtrise des coûts, de comparabilité des résultats et de contrôle de qualité

Le recours à un outil fournit des garanties quant à la forme des résultats, à condition que la mise en œuvre de l'outil soit effectuée de manière conforme aux règles de l'art. Ainsi, une enquête par questionnaire bien conduite (par exemple avec passation confidentielle par téléphone) permet de produire une estimation correcte des effets nets.

Le choix de certains outils oblige à structurer toute l'évaluation. Par exemple, le fait de choisir une analyse coût-efficacité implique que l'évaluation va privilégier un effet principal (ex : emploi). Dans ce cas, l'outil apporte une sécurité au commanditaire qui obtient la certitude que cet effet principal sera privilégié, si telle est sa volonté.

L'utilisation d'un outil connu facilite l'estimation du temps et du budget nécessaires à l'évaluation car les praticiens disposent de références pour fonder leurs estimations. Par ailleurs, le fait qu'une évaluation repose sur des outils standardisés facilite le contrôle des conditions de son exécution.

Plusieurs évaluations basées sur les mêmes outils produisent des résultats plus faciles à comparer. Dans l'exemple de l'Encadré 3, il a été possible de comparer les coûts par emploi créé entre plusieurs régions parce que toutes les évaluations avaient utilisé l'analyse coût-efficacité.

En résumé, une évaluation plus technique est souvent plus simple à commander, plus facile à contrôler, plus crédible pour les lecteurs du rapport et plus prévisible pour le commanditaire. On comprend que tous ces avantages rendent les outils d'évaluation très attractifs. Mais ces avantages sont inséparables d'un certain nombre de risques.

Les risques de l'évaluation « outillée » peuvent être caricaturés par les deux phrases qui suivent :

- « Avec un marteau en main, je vois des clous partout » - risques d'erreur dans le choix d'un outil.
- « Puisque j'ai un marteau, je peux construire toute la maison » - risques de survalorisation d'un outil.

Un outil doit toujours être adapté à son contexte d'utilisation et aux fonctions qu'on veut lui faire remplir. Le risque, pour celui qui a un marteau en main, est de s'en servir pour enfoncer une vis. Dans un exemple étudié dans le cadre du programme MEANS, le commanditaire et l'équipe d'évaluation se sont promis d'utiliser l'analyse de régression. Au bout de quelques semaines, ils se sont rendu compte que les données disponibles étaient insuffisantes, ils ont été dans l'incapacité de mettre en œuvre l'outil annoncé et ont dû réorganiser le travail d'évaluation dans l'urgence. Pour savoir si un outil d'évaluation est adapté à son contexte, il faut examiner l'objet à évaluer, les données disponibles, le budget de l'évaluation, le calendrier de remise des résultats. Choisir un outil sans tenir compte de toutes ces dimensions risque d'entraîner des conséquences graves :

absence de conclusions, dépassement de délai, conclusions pas crédibles, conclusions non-pertinentes, etc.

Il faut également reconnaître et accepter les faiblesses des outils employés. Il vaut mieux réaliser une évaluation moins technique en précisant les limites du travail, plutôt que fonder des conclusions sur l'emploi d'outils sophistiqués mais mal maîtrisés.

Puisque la technicité est attirante, on a souvent tendance à survaloriser les outils employés en évaluation. Ainsi, l'évaluation du N.A.O. (Encadré 3) a été présentée comme un exercice de "*value for money*", c'est-à-dire d'analyse coût-efficacité. L'ensemble de l'évaluation a pris le nom d'un des outils mis en œuvre. Cet outil est devenu emblématique. Ce phénomène est largement répandu et s'applique à de nombreuses évaluations qui s'appelleront selon les cas : évaluation coût-avantages, évaluation économétrique, évaluation qualité/prix, etc.

Assimiler une évaluation à son outil emblématique est toujours une erreur. Un outil d'évaluation permet de remplir une fonction et la plupart du temps, une seule. Privilégier une fonction revient à négliger toutes les autres fonctions qui sont nécessaires à l'évaluation et à affaiblir la qualité de l'ensemble du travail.

## ■ Etapes et fonctions d'une évaluation

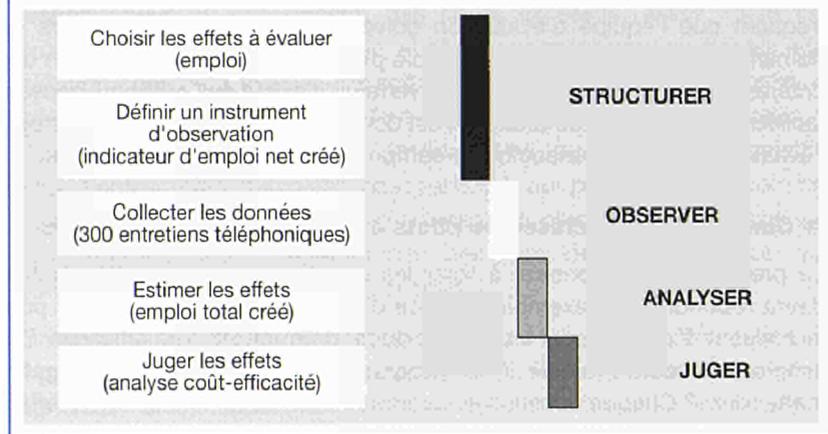
En général, un outil remplit bien la fonction pour laquelle il a été conçu et remplit mal les autres. Un tel outil est considéré comme spécialisé ou spécifique. Un exemple d'outil spécifique est celui du modèle macroéconomique qui ne peut pas remplir d'autre fonction que celle de l'estimation d'impacts à l'échelle d'un pays ou d'une région.

Cependant, il existe des outils qui sont capables de remplir plusieurs fonctions d'évaluation, ce sont des outils génériques. On verra dans cet ouvrage que la technique de l'étude de cas peut servir à la collecte de données mais aussi à d'autres fonctions. De même, on peut mettre en place un panel d'experts pour lui faire remplir une fonction de jugement, mais aussi d'estimation d'impact.

Toute évaluation comporte un certain nombre d'étapes auxquelles sont associées une ou plusieurs fonctions. En reprenant l'exemple de l'évaluation réalisée par le NAO, on voit que l'évaluation a comporté au moins cinq fonctions.

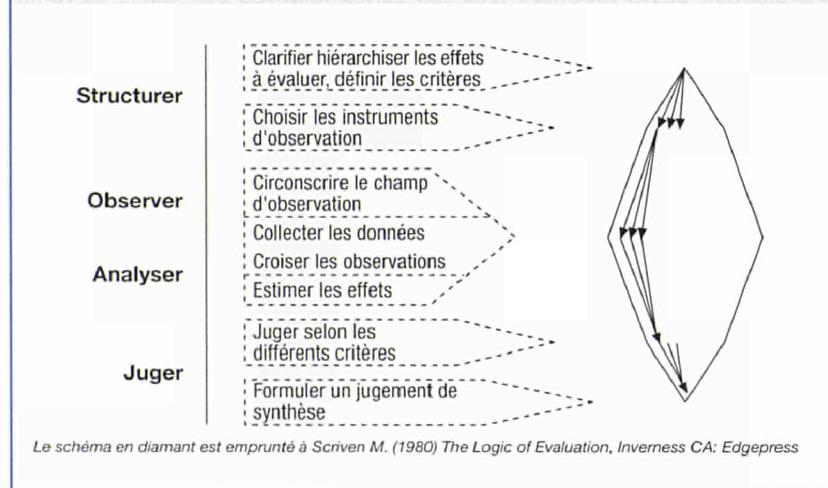
### Encadré 7 - Quelques étapes au sein d'une évaluation

(Evaluation des programmes de développement économique au Pays de Galles)



En analysant de la même façon un grand nombre d'évaluations, on peut identifier les fonctions qui se rencontrent le plus fréquemment. Elles sont présentées ci-dessous en huit étapes de base et regroupées en quatre grandes étapes : structurer l'évaluation, observer les changements sur le terrain, analyser les données, juger le programme.

### Encadré 8 - Les étapes standard d'une évaluation



Le schéma en diamant est emprunté à Scriven M. (1980) *The Logic of Evaluation*, Inverness CA: Edgepress

Les étapes de structuration interviennent normalement avant le travail d'évaluation proprement dit (avant la rédaction du cahier des charges). On peut donc s'interroger sur leur place dans le présent ouvrage. En fait, il est fréquent que l'équipe d'évaluation doive revenir sur ces étapes lors du démarrage de son travail, par exemple pour approfondir l'explicitation des critères de jugement. Par ailleurs, on verra qu'il existe des outils qui peuvent faciliter le travail de structuration. Il est donc logique que cet ouvrage intègre l'évaluation dans son ensemble, y compris les toutes premières étapes.

### ● Clarifier et hiérarchiser les effets à évaluer, définir les critères

La première étape consiste à lister les questions auxquelles l'évaluation devra répondre. Par exemple : est-ce que les contribuables en ont pour leur argent ? quelles sont les parties du programme les plus efficaces ? le programme est-il justifié ? le programme a-t-il des effets négatifs inattendus ? Chaque question se subdivise éventuellement selon les effets à évaluer. Par exemple : la création d'emploi, l'investissement induit, les gains de parts de marché, l'augmentation d'employabilité d'une catégorie de population ou encore l'accroissement d'attractivité d'une région. Chaque question se réfère également à un critère de jugement. Par exemple, l'effet recherché et/ou obtenu est-il pertinent, conforme aux objectifs, efficient, durable, équitable ?

Les outils applicables à cette étape permettent de lister les questions, de les préciser, de les hiérarchiser et de sélectionner celles auxquelles l'évaluation devra répondre (voir les fiches relatives aux outils suivants : Cadre logique, Cartographie conceptuelle des impacts, Analyse SWOT, Metaplan®, Vote coloré).

### ● Définir les instruments d'observation

Cette seconde étape consiste en une traduction des questions d'évaluation en indicateurs ou en descripteurs observables sur le terrain. Dans l'exemple des programmes de développement régional du Pays de Galles (Encadré 3), la question était la suivante : « L'effet des programmes sur l'emploi a-t-il été suffisant en regard de l'argent dépensé ? ». Elle a été traduite sous la forme d'un indicateur d'efficacité mesuré comme suit : « coût d'un emploi net créé pour le budget national ».

Cet ouvrage ne propose pas d'outils pour cette étape. On pourra se reporter au Volume 2 consacré aux indicateurs et au Volume 4 pour mieux connaître la construction et l'utilisation des descripteurs.

### ● Circonscrire le champ de l'observation

Cette troisième étape consiste à délimiter un champ d'observation assez restreint pour ne pas générer des coûts excessifs, assez vaste pour produire une information crédible, et assez contrasté pour révéler les effets que l'on cherche à évaluer. Il s'agit par conséquent d'un compromis, qui intègre des choix tels que : faut-il observer toutes les catégories de destinataires ? veut-on aussi observer une ou plusieurs populations équivalentes mais qui n'ont pas participé au programme ? comment sélectionner des cas pour obtenir l'image la plus complète possible du programme ? faut-il sélectionner des cas de *best practice* pour comprendre les facteurs clés de succès ?

On parle ici du champ d'observation, c'est-à-dire de la délimitation des « groupes » à observer (entreprises, stagiaires, habitants, exploitants agricoles, etc.), lequel ne doit pas être confondu avec le champ de l'évaluation (délimitation de l'intervention soumise à évaluation).

### ● Collecter les données

La quatrième étape consiste à collecter les données primaires sur le terrain. De nombreux outils d'enquête sont en concurrence pour remplir cette fonction. Ils sont empruntés aux diverses disciplines des sciences sociales qui réalisent des collectes de données : sociologie, politologie, ethnographie, etc.

Cette étape est indispensable si l'évaluation ne peut pas entièrement s'appuyer sur les données secondaires telles que les évaluations et recherches antérieures, les données statistiques et les données du suivi.

Les outils applicables à cette étape permettent de collecter des faits et des opinions et de les restituer sous forme qualitative ou quantitative (voir les fiches relatives aux outils suivants : Entretien individuel, Entretien de groupe, Etude de cas, Enquête par questionnaire, Observation ethnographique).

### ● Croiser les observations

La cinquième étape permet de présenter les données primaires et/ou secondaires de manière à suggérer des conclusions et des rapprochements de façon intuitive ou inductive.

Si l'on prend l'exemple de l'évaluation d'un stage de formation pour chômeurs de longue durée, le croisement des observations consistera par exemple à décomposer le taux de retour à l'emploi selon les branches professionnelles pour voir s'il y a des différences significatives et pour mieux comprendre les facteurs de succès.

Les outils de base pour cette fonction sont les graphiques et les tableaux croisés. Cet ouvrage propose aussi plusieurs outils plus sophistiqués (voir les fiches relatives aux outils suivants : Système d'information géographique, Analyse factorielle).

### ● **Estimer les effets**

La sixième étape est celle qui permet d'estimer les effets nets du programme. Cette étape requiert une analyse de causalité qui suit presque toujours une démarche déductive, c'est-à-dire une démarche fondée sur la vérification d'hypothèses sur les causes et les effets. La littérature technique de l'évaluation privilégie souvent les outils d'analyse quantitative qui permettent de remplir cette fonction (ex : analyse comparative, analyse de régression).

Sur ce point, cet ouvrage se démarque de la littérature existante. D'une part, il met à égalité les outils d'analyse qualitative ou quantitative en exposant les forces et les faiblesses de chacun. D'autre part, il ne survalorise pas l'analyse de causalité que l'on a trop tendance à considérer comme le rôle unique de l'évaluation. Le parti pris est que les fonctions d'observation et de jugement sont au moins aussi importantes que la fonction d'estimation des effets nets.

Les outils applicables à cette étape permettent d'estimer les effets nets du programme, au passé ainsi qu'au futur, et au niveau micro ainsi qu'au niveau macroéconomique (voir les fiches relatives aux outils suivants : Analyse structurelle-résiduelle, Modèle d'entrées-sorties, Modèle macroéconomique, Enquête Delphi, Groupes de comparaison, Analyse de régression).

### ● **Juger les effets selon chaque critère**

L'évaluation doit pouvoir dire si les effets que le programme a produits ou va produire sont suffisants ou insuffisants. Dans l'exemple des programmes de développement régional du Pays de Galles (Encadré 3), l'équipe

d'évaluation avait à juger la création d'emplois en regard du critère d'efficacité. L'outil employé a été l'analyse coût-efficacité. Cet outil permet d'affirmer qu'un programme est d'autant plus satisfaisant que son coût par emploi créé est plus faible.

Pour que l'évaluation soit complète, il fallait aussi pouvoir dire si le coût par emploi créé était suffisamment faible. La norme choisie a été le coût moyen des emplois nets créés par des programmes comparables en Ecosse. La comparaison entre le Pays de Galles et l'Ecosse a permis de dire que les programmes évalués étaient suffisamment efficaces.

Ce Volume présente plusieurs outils permettant de traiter l'étape du jugement (voir les fiches relatives aux outils suivants : Analyse coût-efficacité, Etalonnage).

#### ● **Formuler un jugement de synthèse**

La huitième et dernière étape de l'évaluation consiste à formuler les conclusions en faisant la synthèse finale des jugements partiels portés au nom des différents critères. Cette fonction est nécessaire quand plusieurs critères ont été retenus. Pour porter un jugement synthétique sur le programme, il est nécessaire de combiner, pondérer ou hiérarchiser les différents effets que l'évaluation a pu estimer. Dans le cadre des programmes socio-économiques, cette fonction est rendue d'autant plus difficile qu'il y a toujours plusieurs partenaires et que ceux-ci n'accordent pas la même importance aux mêmes critères.

Il existe plusieurs outils qui remplissent totalement ou partiellement cette fonction (voir les fiches relatives aux outils suivants : Panel d'experts, Analyse multicritère, Analyse coût-avantages).

#### ■ **Les situations d'évaluation**

Un outil est fait pour remplir une fonction donnée. Il y a donc une relation forte entre outil et fonction. Cependant, il est fréquent que plusieurs outils puissent remplir une même fonction (voir l'encadré suivant). Lorsque plusieurs outils d'évaluation sont disponibles, se pose la question du choix de l'outil le plus adapté à la situation.

Par exemple, dans une situation d'intervention publique massive à l'échelle d'un vaste territoire, la fonction d'estimation des effets sera remplie par un

modèle macroéconomique. Ainsi, le modèle HERMIN a été construit et utilisé avec succès pour estimer les impacts macroéconomiques des Fonds structurels en Irlande, au Portugal et en Espagne. Dans une situation différente, par exemple si les sommes engagées sont faibles en comparaison avec l'économie du territoire considéré, ou si le territoire est petit, le modèle macroéconomique n'est pas adapté à l'estimation des effets.

#### **Encadré 9 - L'importance de la situation pour le choix d'un outil**

La pelle est un outil dont la fonction consiste à déplacer un matériau meuble d'un endroit à un autre. Pour remplir une brouette de 50 litres, on utilise une pelle de terrassier. S'il faut remplir un camion de 25 tonnes, il vaudra certainement mieux employer une pelle mécanique.

Dans ces deux exemples, la fonction de l'outil est la même, mais la situation est très différente. L'outil doit être adapté à la fois à sa fonction et à sa situation d'emploi.

Comme le suggère l'exemple du modèle HERMIN, une première différence est celle de l'échelle **des effets** à évaluer qui peut être micro (un petit groupe de destinataires potentiels) ou macro (une grande région ou un pays). L'évaluation d'un impact au niveau d'une économie nationale nécessite l'emploi d'outils particuliers, notamment les modèles macroéconomiques, et disqualifie certains autres, par exemple l'étude de cas.

Une autre différence importante est celle du **partenariat**. En effet, dans le cadre d'une évaluation qui implique plusieurs décideurs, le choix des effets à évaluer et des critères d'évaluation revêt une grande importance et doit respecter les divers points de vue en présence. L'évaluation multidécideurs nécessite l'emploi d'outils ouverts, comme par exemple les techniques de travail en groupe ou l'analyse multicritère. A l'inverse, elle s'accommode très mal d'une technique fermée comme l'analyse coût-efficacité.

La troisième différence est celle de l'utilisation de l'évaluation : formative ou récapitulative. L'étude de cas, qui aide à comprendre et à tirer rapidement les leçons de l'expérience, est un outil adapté à l'évaluation formative. Par contre, cet outil produit des conclusions difficiles à synthétiser pour les décideurs et le public. Il est donc moins adapté à l'évaluation récapitulative.

### Encadré 10 - Les trois temps de l'évaluation

Dans le cadre des programmes des Fonds structurels européens pour la période 1994-99, l'évaluation intervient à trois moments.

L'évaluation ex ante porte sur le projet de programme. Elle concerne la cohérence et la pertinence du projet ainsi que le réalisme des effets escomptés.

Les évaluations à mi-parcours portent sur les effets des toutes premières réalisations du programme. Elles sont utilisées en commun par les différents niveaux de décision. Elles complètent le suivi et visent à réorienter le programme en cours de route.

Les évaluations ex post sont principalement destinées à rendre compte des effets des programmes.

La quatrième différence est celle de la **temporalité** (voir Encadré 10). L'évaluation ne peut pas faire appel aux mêmes outils pour évaluer des effets qui sont déjà obtenus ou qui sont encore attendus.

La cinquième différence est une des plus importantes pour le choix des outils. Il s'agit de la **complexité** du programme. La majorité des programmes socio-économiques sont multisectoriels, c'est-à-dire qu'ils interviennent sous plusieurs formes, visent plusieurs publics différents et produisent de multiples sortes d'impacts. Or, la plupart des outils d'évaluation ont traditionnellement été employés dans des situations simples (un seul secteur d'intervention, un seul public, un impact principal attendu).

Cette dernière différence est très importante pour le choix des outils. Elle est également propre aux programmes socio-économiques. C'est la raison pour laquelle elle a été privilégiée dans la rédaction de cet ouvrage. La présentation des outils est donc structurée en deux situations principales :

- Evaluer globalement l'ensemble d'un programme. Cette situation est celle d'un programme régional multisectoriel que l'on souhaite évaluer globalement dans toute sa complexité. Cette situation fait notamment place à l'évaluation des impacts de niveau macro.
- Approfondir une question évaluative. Cette situation rejoint le cas de l'évaluation d'un programme simple, qui est décrit dans la plupart des manuels d'évaluation.

De manière générale, plus la complexité est grande, plus les fonctions de clarification et de structuration doivent être traitées avec attention, plus l'observation empirique est coûteuse et plus l'analyse de causalité est difficile. La pratique normale est une approche par découpage du

programme en catégories d'interventions homogènes. Chaque intervention est évaluée sous l'angle de sa pertinence, de son efficacité ou de son efficience. L'évaluation fait ensuite une synthèse au niveau du programme. La difficulté de cette démarche est que les différentes composantes du programme (mesures, actions, etc.) doivent être évaluées selon des critères communs à tout le programme. En effet, si tous les critères sont spécifiques, la synthèse est impossible. Une variante de cette situation est celle où l'on doit estimer les effets macroéconomiques à l'échelle d'une région ou d'un pays. Cela nécessite en premier lieu d'estimer les effets micro de chaque composante du programme selon une grille commune, puis d'intégrer ces effets dans l'ensemble des mécanismes macroéconomiques pour en déduire un effet global qui n'est pas la somme des effets microéconomiques.

Différents types de questions peuvent nécessiter des approfondissements :

- Evaluation centrée sur une composante élémentaire du programme, par exemple un axe prioritaire, une mesure, une action. Dans ce cas, les réalisations sont homogènes et les destinataires peuvent être facilement identifiés et interrogés. La collecte et l'analyse de données primaires sont faciles, de même que le jugement par comparaison avec des interventions similaires dans d'autres régions.
- Evaluation centrée sur un public, les PME ou les chômeurs de longue durée par exemple. Ce cas ressemble beaucoup au précédent. Il s'en distingue par le fait que plusieurs composantes élémentaires du programme peuvent toucher le même public.
- Evaluation centrée sur un thème prioritaire. Dans ce cas, on évalue la pertinence ou l'impact de toutes les composantes du programme sur un thème particulier, comme l'emploi, l'environnement ou l'égalité des chances. La difficulté de ce type de situation est que les interventions sont multisectorielles, ainsi que les résultats et les mécanismes d'impact. Cette situation n'est donc pas loin de présenter autant de difficultés que l'évaluation d'ensemble d'un programme complexe. Elle est traitée dans le Volume 5 de cette Collection.
- Evaluation centrée sur un territoire particulier. Dans ce cas, on évalue la pertinence ou l'efficacité de toutes les mesures du programme sur une partie restreinte de la zone couverte par le programme (ex : quartiers en difficultés, zones les plus défavorisées, etc.). Cette situation présente exactement les mêmes difficultés que l'évaluation d'ensemble d'un programme complexe.

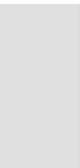
Dans la suite de cet ouvrage, le terme « évaluation approfondie » désignera seulement les deux premiers découpages qui produisent des situations simples, dans lesquelles les outils traditionnels d'évaluation sont bien adaptés.



## II

# Outils pour l'évaluation d'ensemble d'un programme

*Les textes relatifs aux outils d'évaluation font le plus souvent référence à des programmes simples, visant un public bien défini et cherchant à produire un changement limité à ce public. A l'inverse, un programme de développement socio-économique est typiquement multisectoriel et multiobjectifs. Il cherche à créer des transformations économiques ou sociales à l'échelle de tout un territoire. Pour atteindre son but, il adopte une stratégie intégrée, dirigée vers de nombreux publics (entreprises, salariés, chômeurs, collectivités, etc.). Il doit produire son impact global grâce à la bonne combinaison de multiples impacts spécifiques sur ses différentes cibles. Dans ce contexte, l'évaluation rencontre des problèmes techniques auxquels la littérature existante apporte peu de solutions. Cette section recense et présente douze outils adaptés à cette situation. Beaucoup d'entre eux sont innovants ou sont employés de manière innovante.*



## Exemples d'évaluations de programmes multisectoriels

Les trois exemples qui suivent décrivent des situations où le programme à évaluer était clairement multisectoriel. Ces évaluations ont employé un ou plusieurs outils de façon performante.

Dans le premier exemple, le programme apporte un financement important à une région en retard de développement. Il finance des infrastructures de base aussi bien que des aides aux entreprises. Ses objectifs sont nombreux et pas toujours précis. Une des questions auxquelles a répondu l'équipe chargée de l'évaluation intermédiaire est la suivante : « Est-ce que le contexte économique régional évolue dans le sens de l'atteinte des objectifs ? » L'évaluation a permis de savoir si les problèmes économiques régionaux étaient en voie de résorption ou d'aggravation.

### Encadré 11 - L'évaluation intermédiaire d'un programme régional en Espagne

En 1996, la Generalitat de Valencia a lancé l'évaluation intermédiaire d'un programme de développement socio-économique régional (Programme Opérationnel d'Objectif 1) cofinancé par les Fonds structurels européens. Le programme comportait 25 mesures destinées à agir, ensemble ou séparément, sur plusieurs catégories de populations cibles. Les objectifs du programme étaient nombreux, peu hiérarchisés et rarement quantifiés.

L'évaluation a débuté par une série de réunions au cours desquelles les partenaires ont établi une liste des dix principaux impacts attendus. Cette liste a été obtenue avec l'aide de la technique de la **cartographie conceptuelle**. Chaque impact attendu a reçu un nom (ex : « diversification et efficience des ressources énergétiques »).

A chacun des dix impacts on a associé un ou plusieurs indicateurs de contexte (ex : « nombre d'entreprises régionales consommant du gaz naturel »). L'équipe d'évaluation a ensuite rassemblé des données statistiques pour connaître l'évolution globale de ces indicateurs au niveau régional. On a analysé l'évolution tendancielle de chaque indicateur par la technique de **comparaison avant-après**.

Dans cet exemple, la technique de la **cartographie conceptuelle** a permis de « clarifier et de hiérarchiser les effets à évaluer » (première fonction d'évaluation), tandis que la technique de **comparaison avant-après** a permis de croiser des données pour montrer des évolutions intéressantes (cinquième fonction).

Dans le second exemple, la situation est celle d'un programme qui apporte un financement pour le développement d'une région française. Ici aussi, les interventions sont multisectorielles et multiobjectifs. L'équipe chargée de l'évaluation intermédiaire souhaitait montrer la dimension spatiale des réalisations et des résultats.

#### **Encadré 12 - L'évaluation intermédiaire d'un programme régional en France**

Les Programmes Intégrés Méditerranéens (PIM) ont été décidés dans les années 80 pour aider les régions fortement concurrencées par l'entrée de l'Espagne et du Portugal dans la Communauté Européenne. Le PIM du Languedoc-Roussillon était composé de plusieurs sous-programmes, orientés vers des secteurs différents : l'agriculture et la pêche, l'industrie, l'artisanat, le tertiaire, la forêt et l'environnement, ainsi que le tourisme. Plus de la moitié des fonds mobilisés ont été orientés vers l'agriculture et la pêche, ces secteurs étant les plus concernés par l'intégration de l'Espagne et du Portugal.

Le PIM de la région Languedoc-Roussillon a fait l'objet d'une évaluation intermédiaire. Les PIM ayant été conçus pour être complémentaires des politiques régionales, l'évaluation devait notamment analyser la synergie entre ces deux niveaux de politique. Cette évaluation s'est particulièrement intéressée à la localisation et à l'intégration spatiale des différentes mesures du programme. L'équipe d'évaluation a rassemblé toutes les données du suivi ainsi que certains indicateurs statistiques dans une base de données en notant leur localisation. Un système d'information géographique a été utilisé à cette occasion pour la première fois à des fins d'évaluation. L'équipe d'évaluation a présenté ses résultats sous forme de cartes. La lecture de certaines cartes a révélé un certain nombre de décalages entre la localisation des financements et celle des zones rurales les plus menacées.

Dans cet exemple, la technique du système d'information géographique a permis d'effectuer un croisement des données qui a produit des conclusions intéressantes (cinquième fonction).

Le troisième exemple concerne l'évaluation globale des aides des Fonds structurels dans la République d'Irlande. La situation est celle d'un ensemble de programmes couvrant de nombreux secteurs et apportant un financement massif à l'échelle d'un pays. Les objectifs globaux du programme sont de nature macroéconomique. Ils sont peu nombreux et précisément définis. L'équipe chargée de l'évaluation ex ante devait estimer les impacts macroéconomiques probables.

### **Encadré 13 - L'évaluation ex ante d'un ensemble de programmes en République d'Irlande**

Pour la période 1989-93, toutes les aides européennes au développement de l'Irlande ont été coordonnées au sein d'un « Cadre Communautaire d'Appui ». La Commission européenne et le gouvernement national ont commandé conjointement une évaluation ex ante de cet ensemble très complexe. L'évaluation devait estimer les impacts potentiels du CCA sur l'économie irlandaise. Une question importante était de savoir si une intervention massive (voisine de 5% du PIB irlandais) n'aboutirait pas de manière indirecte à une dégradation des finances publiques à cause des besoins de cofinancement.

L'évaluation a été réalisée par l'Economic and Social Research Institute (ESRI) de Dublin. L'équipe d'évaluation a estimé les impacts microéconomiques des aides en s'appuyant sur des travaux de recherches et sur des évaluations d'aides similaires. Ces impacts primaires ont été introduits dans un modèle macroéconomique permettant de simuler le fonctionnement de l'économie nationale. L'ESRI avait précédemment créé ce modèle dans le cadre du programme européen HERMES. Un travail préalable a dû être effectué afin d'adapter le modèle employé à la situation d'évaluation, notamment en désagrégeant sa structure sectorielle initiale pour mieux tenir compte des effets des aides européennes. Le modèle a permis d'estimer les impacts à l'échelle de l'économie nationale, en prenant en compte de nombreux mécanismes macroéconomiques. Le risque d'une augmentation excessive de la dette publique a été confirmé.

Dans cet exemple, l'équipe d'évaluation a appliqué la technique du modèle macroéconomique pour estimer les effets (sixième fonction). Il faut cependant noter que l'outil employé incorpore une série de choix quant aux variables à modéliser ou à négliger. L'outil a donc implicitement servi à clarifier et hiérarchiser les effets à évaluer (première fonction).

Ces trois exemples ont permis d'illustrer l'emploi de plusieurs outils dans des situations où l'évaluation porte sur l'ensemble d'un programme multisectoriel. La section qui suit va montrer que la boîte à outils de l'évaluation est suffisamment fournie pour que la plupart des fonctions des évaluations de ce type puissent être outillées.



## Outils pour structurer l'évaluation

Lorsqu'un programme socio-économique est complexe, la phase de structuration de l'évaluation est particulièrement importante car elle conditionne le caractère complet et compréhensible des résultats. En outre, si le contexte est partenarial, cette phase est indispensable pour clarifier les critères d'évaluation qui conviennent aux différents partenaires et détermine fortement l'acceptabilité des conclusions.

L'analyse SWOT est généralement utilisée ex ante pour définir la stratégie d'un programme. Cet outil aide à valider les objectifs et à définir plus précisément les effets à rechercher ainsi que les interventions souhaitables. Il peut également aider à évaluer une stratégie déjà établie ou même déjà appliquée.

La cartographie conceptuelle d'impacts permet de définir les effets à évaluer et les indicateurs associés, dans le cas où les objectifs sont nombreux et pas encore stabilisés. L'outil peut également être très utile si les objectifs ont été fixés de manière peu précise. Cet outil est adapté au travail en partenariat car il fonctionne à partir de l'agrégation des points de vue individuels en dégagant un consensus entre les partenaires.

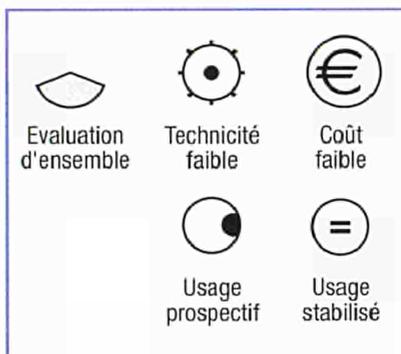
Le vote coloré est particulièrement adapté pour clarifier et hiérarchiser les questions évaluatives dans une situation d'évaluation associant plusieurs décideurs. Il fournit également un moyen rapide et robuste pour choisir les critères d'évaluation ainsi que pour valider les conclusions de l'évaluation et pour formuler des recommandations. Cet outil est utile lorsque les protagonistes sont nombreux et qu'il faut prendre en compte et respecter leurs différences de point de vue.

La matrice d'impacts est un outil qui donne une vision simplifiée de l'ensemble d'un programme complexe. Son principe consiste à regrouper les différentes composantes du programme en dix à vingt sous-ensembles (les lignes de la matrice), à clarifier les principaux impacts attendus (les colonnes de la matrice). Chaque cellule de la matrice reçoit une note représentant l'importance probable d'un impact donné pour une composante donnée. L'outil permet de s'assurer que tous les effets essentiels du programme ont été envisagés avant de hiérarchiser ceux qui devront faire l'objet d'un effort particulier d'évaluation. Cet outil n'est pas décrit dans les pages qui suivent. On pourra se reporter au Volume 4 de la Collection qui est consacré aux techniques innovantes.

Certains outils classiques de structuration ne sont pas adaptés à la situation d'évaluation d'ensemble d'un programme complexe. Par exemple, la technique de l'arbre d'objectifs ou celle du cadre logique ne permettent pas de prendre correctement en compte la multiplicité des actions et de leurs impacts escomptés.

## ■ Analyse SWOT (Forces-Faiblesses- Opportunités-Menaces)

L'analyse SWOT est un outil développé depuis les années 50 pour aider les entreprises à définir leur stratégie dans un environnement concurrentiel fluctuant et contraignant. Cet outil d'aide à la décision doit son nom au fait qu'il examine les forces et faiblesses (*Strengths, Weaknesses*) internes à l'entreprise ainsi que les opportunités et menaces du marché (*Opportunities, Threats*). Il fait partie des outils classiques de l'analyse stratégique, de même que, par exemple, la matrice BCG (*Boston Consulting Group*) qui présente de nombreuses similarités.



### Exemple : L'appréciation ex ante du plan Objectif 6 suédois

L'analyse SWOT a été utilisée lors de l'élaboration du plan de développement régional pour l'Objectif 6 en Suède. Les résultats produits à travers cette méthode ont permis de mettre en évidence les grandes tendances environnementales, structurelles et démographiques caractéristiques du contexte de la région retenue. L'identification des stratégies et des priorités, ainsi que la conception du plan d'action, ont été réalisées par la suite sur la base des facteurs majeurs identifiés à travers l'analyse SWOT, dont l'aperçu suivant n'est qu'un résumé.

**Forces :** Grandes espaces naturels préservés, ressources naturelles importantes.

**Faiblesses :** Région isolée, faible densité de population, vulnérabilité aux fluctuations économiques, à celles des budgets publics, faiblesses structurelles des entreprises.

**Opportunités :** Tourisme, qualité des ressources forestières, production piscicole en eau douce, prospection minière.

**Menaces :** Déclin démographique, insuffisance du nombre d'emplois.

Les recommandations issues de cette analyse basée sur des données socio-économiques régionales ont été jugées par les responsables comme pertinentes pour alimenter leur réflexion prospective. Leur utilité pour justifier les choix faits dans le cadre du Plan de Développement Régional semble avérée. Depuis cette première tentative d'utilisation concrète de l'analyse SWOT dans le cadre des programmes financés par la Commission européenne, cet outil a été utilisé dans de nombreuses régions.

Source : Prior appraisal of the regional development plan for the objective 6 region of Sweden, Nordland Research Institute, Bodo, juin 95.

Dans le secteur public, les grandes villes ont été les premières à employer cet outil dans les années 1980 pour réfléchir à différents scénarios de développement. L'utilisation de l'analyse SWOT s'est plus récemment étendue aux diagnostics des territoires et à l'évaluation ex ante de programmes régionaux.

### ● A quoi sert l'outil ?

L'analyse SWOT vise à intégrer dans la réflexion sur un programme les caractéristiques intrinsèques du territoire concerné, mais aussi les facteurs déterminants de l'environnement dans lequel celui-ci prend place. L'outil est utilisé pour diminuer les zones d'incertitude liées à la mise en œuvre d'un projet ou d'un programme applicable au territoire concerné. Il permet ainsi de définir une stratégie plus pertinente par rapport au contexte dans lequel l'action va se déployer. La finalité de l'outil est donc :

- la mise en évidence des facteurs dominants et déterminants, internes et externes à la collectivité, et propres à influencer le succès de son projet ;
- la production de lignes stratégiques pertinentes par la mise en relation du projet avec son environnement.

L'analyse SWOT peut être prolongée en s'inspirant des outils dits « de gestion de portefeuilles », comme la matrice B.C.G., pour examiner la validité d'une stratégie proposée ou en cours d'application et pour recommander, le cas échéant, des modifications. Le classement des différentes possibilités d'action tient compte de leur faisabilité (atouts disponibles en interne = forces et faiblesses), ainsi que de leur potentiel (attraits par rapport à l'environnement extérieur = opportunités et menaces).

L'utilité de tels outils dans les processus évaluatifs réside dans leurs capacités à dégager une image systémique des relations que le programme évalué entretient avec son environnement direct.

### ● Dans quels cas l'utiliser ?

Développés il y a une vingtaine d'années par les spécialistes du management privé, les outils d'aide à la décision stratégique trouvent aujourd'hui un nouvel essor dans la réflexion stratégique publique. Construits à l'origine en fonction des notions de produits, de clients, de marché et d'avantages compétitifs, leur utilisation s'étend actuellement aux collectivités publiques dans la mesure où leurs politiques visent à créer des

avantages compétitifs à l'échelle du territoire dont elles sont responsables. Les notions de forces, de faiblesses, d'opportunités et de menaces peuvent en effet s'appliquer à une économie régionale dans le cadre d'une compétition européenne, aussi bien qu'à une entreprise en compétition sur ses marchés.

L'analyse SWOT permet de dégager les lignes stratégiques les plus pertinentes par rapport à un objectif global de développement économique. L'utilisation de l'outil est donc particulièrement indiquée lors de la conception d'un programme et lors de son évaluation ex ante car elle permet une amélioration de la qualité de l'insertion de celui-ci dans son contexte.

### ● **Comment le mettre en œuvre ?**

La mise en œuvre d'une démarche stratégique de type SWOT comprend six étapes distinctes.

**Etape 1. Radiographie de l'environnement du programme :** Cette opération va permettre de déceler les grandes tendances et problèmes susceptibles d'affecter l'avenir du territoire considéré, à l'aide d'indicateurs de type socio-démographiques, économiques, politiques ou physiques. Les indicateurs de disparités régionales sont particulièrement utiles pour révéler les opportunités et les menaces. Cette phase ne saurait être exhaustive, son objectif étant limité à l'obtention d'images globales, capables d'illustrer les grandes tendances et problèmes auxquels va devoir se confronter la communauté considérée.

**Etape 2. L'inventaire des actions possibles :** Les actions possibles sont formulées en termes généraux, et construites en fonction des grands problèmes identifiés.

**Etape 3. L'analyse externe des opportunités et menaces.** Elle liste les paramètres de l'environnement, qui ne sont pas sous le contrôle direct des collectivités, et dont on suppose qu'ils peuvent influencer fortement le développement socio-économique.

**Etape 4. L'analyse interne des forces et faiblesses :** Elle complète l'analyse externe en procédant à l'inventaire des facteurs qui sont au moins en partie sous le contrôle de la collectivité, et qui peuvent constituer des moteurs ou des freins pour le développement.

**Etape 5. Le classement des actions possibles :** Il vise à mettre en évidence les actions les mieux à même de réduire les problèmes de développement (lignes stratégiques) en s'appuyant sur les forces et en réduisant, voire en éliminant, les faiblesses, ceci dans une perspective de maximisation des opportunités et minimisation des menaces.

**Etape 6. L'évaluation d'une stratégie :** Cette étape facultative peut être réalisée s'il convient de juger la pertinence d'une stratégie déjà appliquée ou en projet. Elle est conçue sur la base d'une analyse de « portefeuille d'activité ». A l'instar d'une entreprise avec ses produits et ses marchés, un programme socio-économique contient en effet un ensemble d'interventions, dont certaines combinent des forces et des opportunités tandis que d'autres cherchent à compenser des faiblesses ou à prévenir des menaces. L'équipe d'évaluation place les interventions sur un plan à deux axes : (1) faisabilité interne, forces, faiblesses et (2) environnement externe, opportunités, menaces. La discussion de la carte ainsi produite permet de juger la pertinence de la stratégie évaluée.

#### ● **Application aux Fonds structurels**

L'analyse stratégique est adaptée à la clarification des enjeux de développement socio-économique des régions auxquelles s'appliquent les Fonds structurels. Elle est indiquée pour élaborer des stratégies de développement régional, qui doivent être centrées sur les forces et les opportunités, en essayant de pallier les faiblesses et les menaces. Un outil comme l'analyse SWOT propose un point de départ pour définir les lignes stratégiques d'application en fonction d'un environnement défini.

SWOT peut également servir aux gestionnaires comme outil de jugement de la pertinence d'une stratégie en projet ou en cours d'application. Dans le cadre d'une évaluation en cours de programme, cet outil permet de vérifier si la stratégie mise en œuvre reste pertinente par rapport aux besoins de l'économie régionale.

#### ● **Points forts et limites d'utilisation**

- L'analyse SWOT appliquée à l'action publique est orientée vers la recherche d'une stratégie efficace.
- L'outil matriciel est issu de théories du management déjà anciennes et souvent critiquées pour leur caractère simplificateur. Il faut être conscient de cette limite lors de leur transposition au management public, afin de ne pas se laisser enfermer dans un cadre réducteur.

- En tant qu'outil d'évaluation, il permet de réaliser un classement simple des activités en fonction de leur pertinence. Sa principale faiblesse provient de la procédure, souvent subjective par laquelle l'équipe d'évaluation classe les activités. L'association des partenaires à ce classement est une façon de renforcer la crédibilité et l'utilité de l'analyse.
- L'analyse SWOT nécessite une volonté délibérée de la part des différents acteurs participant à son application de déboucher sur un consensus stable. Le processus d'élaboration de lignes stratégiques n'acquiert de valeur que sous cette condition. Dans le cas contraire, ce modèle peut tendre à produire un diagnostic erroné et/ou inapplicable.

● **Pour en savoir plus...**

Bryson J.M. et Roering W.D. (1987) « Applying Private Sector Strategic Planning in the Public Sector », *Journal of the American Planning Association*, n°53, pp. 9-22.

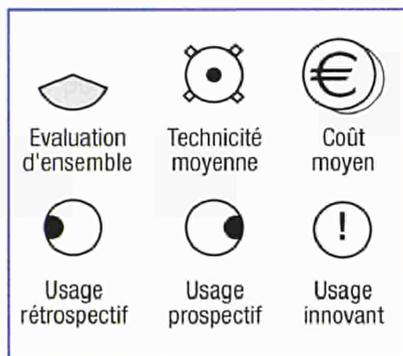
Examen des conditions d'applicabilité des méthodes de planification stratégique d'entreprises face aux contingences du secteur public

Quinn, J.B. & al (1988) *The Strategy Process*, New Jersey : Prentice hall.  
Etude critique sur la planification stratégique



## ■ Cartographie conceptuelle d'impacts

La cartographie conceptuelle d'impacts est une adaptation de la technique dite de *Concept mapping* utilisée depuis une dizaine d'années pour l'évaluation des programmes. Celle-ci a d'abord été utilisée comme aide à la définition de stratégies ou de réorganisation des entreprises avant d'être appliquée à la conception de politiques publiques, notamment de programmes sociaux.



### Exemple : La clarification des impacts escomptés du Programme Opérationnel FEDER 1994-99 de la Generalitat de Valencia.

Cet exemple correspond à celui de l'Encadré 11.

La cartographie conceptuelle d'impacts utilisée dans le cadre de l'évaluation intermédiaire du P.O. FEDER 1994-99 de la Generalitat de Valencia poursuivait deux objectifs :

- la clarification des impacts spécifiques et globaux du programme afin de déterminer des critères de jugement sur l'efficacité des interventions pour l'évaluation intermédiaire.
- la proposition d'indicateurs d'impacts intermédiaires et globaux afin de compléter la batterie d'indicateurs du programme.

L'évaluation s'est déroulée en deux phases centrées sur les impacts intermédiaires, puis sur les impacts globaux, en s'appuyant respectivement sur les gestionnaires des mesures et les responsables de l'ensemble du programme.

Les fonctionnaires de la Generalitat en charge de la mise en œuvre des mesures au sein du programme ont été répartis en trois groupes en fonction des domaines d'intervention (aides directes aux entreprises, aides indirectes aux entreprises, infrastructures). A l'issue d'un cycle de trois réunions, chacun des groupes a déterminé les critères d'évaluation spécifiques des actions du domaine d'intervention concerné. Les mêmes groupes ont ensuite déterminé quels pourraient être les indicateurs d'impact intermédiaire à même de fournir l'information nécessaire pour porter un jugement en fonction des critères spécifiques d'évaluation retenus.

Dans une seconde phase, les responsables du programme dans son ensemble (représentants de la Commission européenne, de l'administration centrale espagnole

et des Autorités Régionales) ont, selon le même processus, construit des critères globaux d'évaluation du programme en s'appuyant sur les résultats des travaux des groupes. Le processus s'est achevé par la sélection d'indicateurs d'impacts globaux correspondants.

Source : Programme MEANS 1996 - Instituto Valenciano de Investigaciones Economicas, « Projet pilote de cartographie conceptuelle des impacts »

Les logiciels de traitement des données sont utilisés pour obtenir une image graphique des représentations que les acteurs ont de la réalité socio-économique en général et des impacts escomptés en particulier. Le programme MEANS a développé une application spécialement conçue pour la clarification des impacts d'un programme de développement régional et la recherche d'indicateurs associés, dans le cadre du partenariat (voir le Volume 4 de cette Collection consacré aux techniques innovantes).

#### ● A quoi sert l'outil ?

La cartographie conceptuelle d'impacts est un outil d'aide à la clarification des impacts escomptés d'une intervention publique. La clarification/hierarchisation est obtenue en tenant compte des représentations de l'ensemble des protagonistes. Le travail du groupe est facilité par le caractère systématique de la démarche qui évite des digressions infructueuses et par la visualisation de la proximité ou de la distance entre les impacts. Ainsi on peut animer la réflexion de groupes hétérogènes de taille importante (jusqu'à 200 personnes) et obtenir une structuration et hierarchisation des principaux impacts attendus d'une intervention. La technique de la carte conceptuelle d'impacts permet ainsi dans un contexte ex ante de déterminer les objectifs visés et les priorités. Dans un contexte ex post, elle permet d'élaborer le référentiel de l'évaluation (critères d'efficacité et indicateurs d'impact).

#### ● Dans quels cas l'utiliser ?

Cet outil est précieux pour structurer le contenu d'un programme complexe. Il a été par exemple utilisé dans un quartier d'Irlande du Nord pour définir un programme social en association avec les communautés protestantes et catholiques concernées. Il apporte une aide en matière d'évaluation chaque fois qu'il est nécessaire de renforcer l'évaluabilité d'un programme, c'est-à-dire chaque fois qu'il est nécessaire de rendre les objectifs d'un programme plus explicites.

Cette technique est particulièrement pertinente lorsque le renforcement de l'évaluabilité d'un programme doit se faire dans le cadre du partenariat entre plusieurs décideurs/financeurs. Elle permet ainsi de parvenir à un référentiel commun d'évaluation malgré les divergences d'objectifs qui peuvent exister entre ces partenaires.

Deux caractéristiques fréquentes des programmes peuvent ainsi motiver le recours à la cartographie conceptuelle d'impacts : le manque de clarté et de précision des objectifs d'un programme et des impacts escomptés d'une part, et l'insuffisance des indicateurs d'impact prévus par le programme, d'autre part. Sa mise en œuvre étant relativement lourde, il convient d'en réserver l'usage aux programmes complexes où les impacts escomptés sont nombreux.

### ● Comment le mettre en œuvre ?

**Etape 1. Constitution d'un groupe des parties prenantes du programme.** Cette étape est essentielle car elle va déterminer le cadre dans lequel vont se dérouler les étapes postérieures du processus. Le groupe constitué doit être équilibré (répartition équitable des différents acteurs du programme), riche (connaissance approfondie de l'intervention par les membres du groupe) et stable (participation continue à l'ensemble du processus).

**Etape 2. Recensement des impacts escomptés du programme sous forme de descripteurs.** Il s'agit, lors de cette étape, de produire une liste exhaustive de l'ensemble des impacts escomptés du programme sous forme de courts énoncés qualitatifs appelés « descripteurs d'impact » (Exemple : « les acteurs du tourisme se sont professionnalisés »).

Cette liste doit inclure aussi les éventuels impacts défavorables du programme qui ne sont pas escomptés mais qui sont pourtant prévisibles.

Deux sources alimentent cette liste :

- le repérage par les chargés de l'évaluation de l'ensemble des descripteurs d'impact figurant dans les documents de programmation ;
- les descripteurs d'impact formulés par les membres du groupe en fonction de la connaissance qu'ils ont des interventions. Il s'agit des impacts des mesures ou actions des programmes que chacun des membres du groupe s'attend à constater à court, moyen ou long terme. La formulation peut se faire dans le cadre d'interviews ou dans le cadre d'une première réunion du groupe.

Une réunion du groupe est indispensable pour valider les descripteurs proposés et arrêter en commun la liste définitive des descripteurs.

**Etape 3. Pondération des descripteurs et constitution de piles.** Dans le cadre d'une deuxième réunion du groupe, il est remis à chacun des membres la liste récapitulative des descripteurs, ainsi qu'un jeu de bostons (chacun des bostons comportant un descripteur). Il est demandé à chacun des membres du groupe individuellement :

- de pondérer les différents descripteurs, c'est-à-dire de noter les descripteurs en fonction de leur importance stratégique pour le programme. La note sera d'autant plus élevée que l'impact décrit occupe une place importante dans la réussite du programme.
- de constituer d'autre part des piles de descripteurs en fonction de leur proximité conceptuelle; chacun des membres du groupe va ainsi procéder à des regroupements de descripteurs conceptuellement proches, autrement dit qui relèvent d'une même « famille » d'impacts.

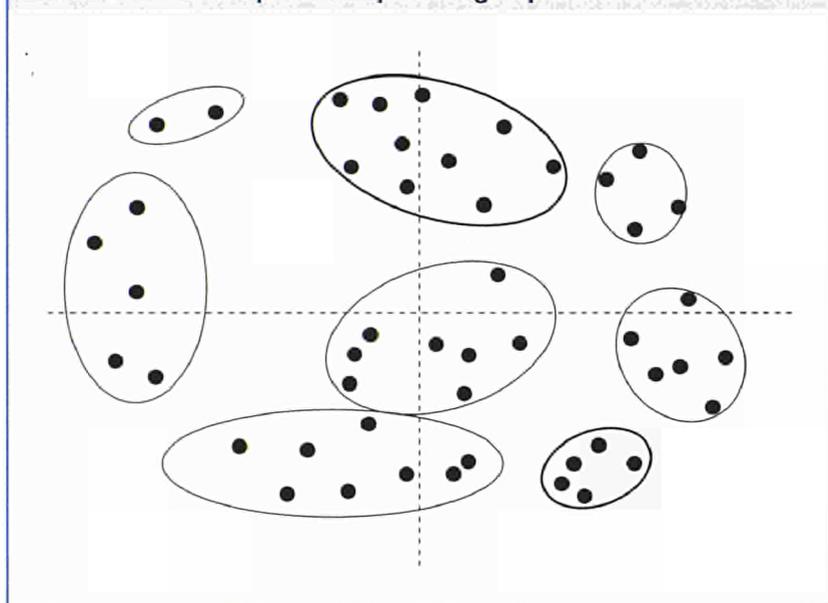
**Etape 4. Calcul de la carte conceptuelle des impacts.** Les données produites par chacun des membres du groupe (pondérations et piles) vont faire l'objet d'une analyse statistique conduite grâce à un logiciel.

Au terme d'une analyse multidimensionnelle de positionnement non métrique, les descripteurs vont être représentés sous la forme d'un nuage de points, dans un espace à deux dimensions. Seront proches, dans ce nuage de points, les descripteurs qui ont été souvent placés dans une même pile par les participants. Seront les plus éloignés, les descripteurs qui n'ont jamais été mis dans une même pile par les participants.

Au terme d'une seconde analyse qui consiste en une classification hiérarchique en grappes, le logiciel va identifier des groupes de descripteurs. Seront placés dans un même groupe les descripteurs qui ont été le plus souvent associés par les différents membres du groupe. Ces groupes sont ainsi constitués de descripteurs proches d'un point de vue conceptuel.

Enfin, la pondération moyenne des descripteurs inclus dans chacun des groupes est calculée. Le résultat final est ainsi une carte comprenant des groupes pondérés de descripteurs (le poids est représenté dans le diagramme ci-dessous par l'épaisseur du trait).

#### Encadré 14 - Descripteur d'impacts regroupés en familles



**Etape 5. Définition des critères d'efficacité.** Dans le cadre d'une troisième réunion, la carte conceptuelle d'impacts obtenue est examinée par le groupe. La composition des différents groupes de descripteurs est ainsi discutée, certains descripteurs pouvant être déplacés pour rendre la composition plus cohérente. La carte calculée par le logiciel ne constitue ainsi pas un résultat final mais plutôt une base de discussion.

Une fois la composition des familles validée, il est procédé à leur dénomination. L'objectif est que les membres du groupe s'accordent sur une dénomination pour chacune des familles qui synthétise l'ensemble des descripteurs composant la famille. Cette dénomination correspond à un critère d'évaluation. Il y aura ainsi autant de critères d'évaluation que de familles de descripteurs d'impact. Les familles étant pondérées, les critères obtenus seront également pondérés, ce qui permet d'utiliser directement ces critères pour un jugement multicritère.

La réflexion du groupe peut être poursuivie afin de préciser la temporalité des familles d'impact (impacts qui sont attendus à court, moyen ou long terme), et les relations de causalité supposées entre ces familles d'impact

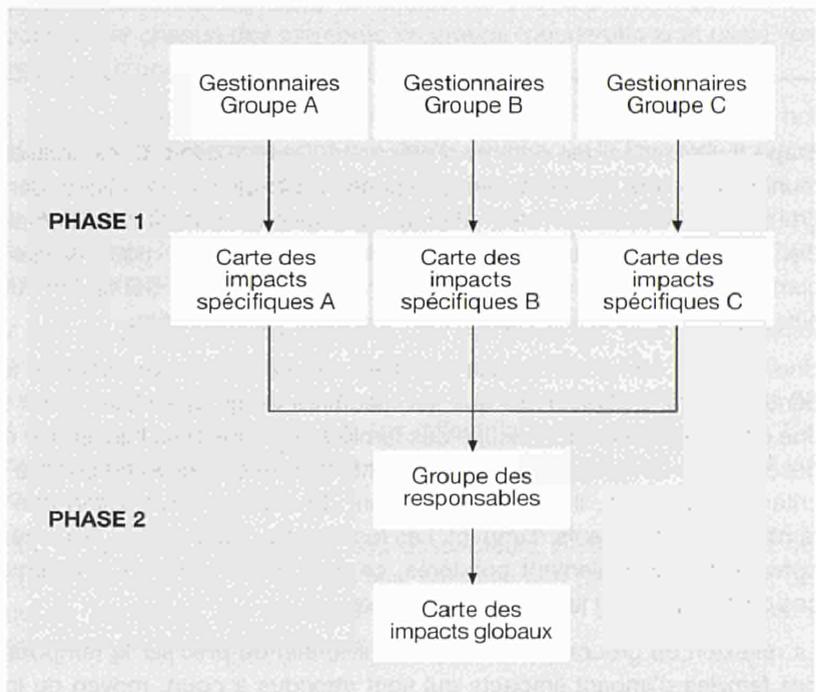
(ex : telle famille d'impact constitue un impact intermédiaire qui contribuera à l'atteinte d'un objectif plus global).

La réflexion du groupe peut enfin être approfondie afin de proposer des indicateurs d'impact directement liés aux différents critères d'évaluation déterminés. Les descripteurs inclus dans chacune des familles constituent une aide pour définir ces indicateurs. Certains de ces descripteurs, et notamment ceux auxquels ont été attribuées les plus importantes pondérations, peuvent ainsi inspirer directement la définition d'un indicateur.

La discussion par le groupe de chacun des indicateurs proposés va porter sur la pertinence de l'indicateur du point de vue de son lien avec le critère d'évaluation et de la disponibilité de sources d'information.

### ● Application aux Fonds structurels

#### Encadré 15 - Elaboration partenariale des cartes d'impacts



Cet outil se révèle particulièrement bien adapté au contexte des programmes structurels. Ce type de programme poursuit les objectifs définis par l'Union européenne et adaptés en fonction des caractéristiques de l'Etat membre ou de la région bénéficiaire. Les Documents Uniques de Programmation étant le produit de négociations, il est fréquent que les objectifs des programmes ne soient pas suffisamment précis pour pouvoir construire une évaluation fiable. Le recours à cet outil est un appui pour la réalisation de la première étape de clarification. Il permet la clarification des questions d'évaluation, la hiérarchisation des effets et la définition des critères d'évaluation dans un contexte de partenariat qui facilite l'appropriation des résultats de l'évaluation.

La carte conceptuelle d'impacts est précieuse pour préparer les évaluations intermédiaires des programmes. En effet, elle permet d'établir de façon systématique des familles d'impacts transversales à plusieurs mesures et ainsi de construire des critères pertinents pour l'évaluation du programme dans son ensemble. On peut en particulier adopter la démarche en deux phases, comme celle mise en œuvre à Valencia et que l'on peut résumer sous la forme de l'Encadré 15.

#### ● Points forts et limites d'utilisation

- La médiation technique (logiciel informatique) se révèle précieuse pour parvenir à un accord malgré l'hétérogénéité des points de vue initiaux.
- La dynamique de groupe générée par la technique se révèle utile et positive pour accroître la connaissance mutuelle des actions des différents gestionnaires et peut ainsi favoriser des ajustements de la mise en œuvre. Mobilisant les parties prenantes dès la phase préalable des travaux d'évaluation, elle facilite grandement les étapes postérieures de l'évaluation et l'acceptation des résultats.
- L'outil permet de renforcer l'évaluabilité des programmes. Il aide à parvenir à un référentiel commun d'évaluation, tout en prenant en compte les points de vue des différents partenaires. Les discussions de groupe successives qui sont une partie de la mise en œuvre de cet outil assurent la légitimité du référentiel obtenu.
- Les résultats peuvent être produits en quelques jours. La technique est ainsi compatible avec le cycle de l'évaluation.
- Les ressources nécessaires (même si elles restent modestes) et la mobilisation de nombreuses parties prenantes lors de réunions

successives ne se justifient que dans la mesure où les enjeux de l'évaluation du programme sont suffisamment importants.

- La constitution du groupe de partenaires est une phase sensible au cours de laquelle il faut s'assurer du bon équilibre entre les parties prenantes. Des efforts importants peuvent être nécessaires en cas de relations conflictuelles.
- Si la technique d'élaboration de la carte conceptuelle d'impacts limite les risques liés au leadership de quelques membres du groupe, ce risque demeure lors des phases de production des descripteurs et d'exploitation de la carte conceptuelle.
- Cet outil permet le traitement de toutes les fonctions de l'étape de structuration.

#### ● Pour en savoir plus...

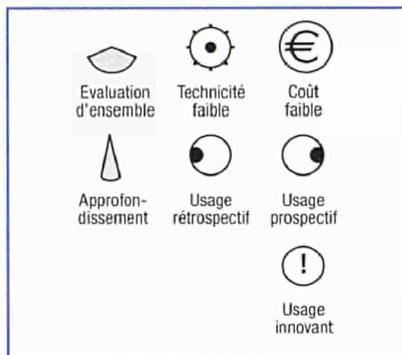
Trochim, W.M.K. (1989) 'An introduction to concept mapping for planning and evaluation', in : *Evaluation and Program Planning*, 12 (1) : 1-16.  
Présentation de l'outil et de son intérêt pour l'évaluation

Knox, C. (1995) 'Concept Mapping in Policy Evaluation : A Research Review of Community Relations in Northern Ireland', in : *Evaluation*, 1(1)  
Un exemple intéressant d'utilisation de l'outil dans un contexte européen

Dans la Collection MEANS : le Volume 4 consacré aux techniques innovantes

## ■ Vote coloré

Dans le domaine de l'évaluation, le travail en groupe est souvent indispensable, particulièrement lorsque plusieurs partenaires sont associés à la commande et au pilotage de l'évaluation. Les différentes techniques de vote permettent de structurer la réflexion collective, que ce soit dans le but d'en améliorer l'efficacité, d'en stimuler la créativité ou de dégager une position commune. Comme ces techniques sont nombreuses, nous parlerons plus particulièrement ci-dessous de « l'Abaque de Régnier » ou vote coloré.



### Exemple : L'évaluation d'un programme de développement économique en Suisse

Dans le cadre d'une recherche dédiée à la mesure des effets des actions publiques en Suisse, la CEAT (Communauté d'études pour l'aménagement du territoire) a réalisé l'évaluation d'une politique cantonale de promotion économique. Deux groupes de travail ont été constitués avec des acteurs et des opérateurs de la politique, le premier consacré aux « objectifs et stratégies » de cette politique et le second à sa « mise en œuvre et coordination ».

L'équipe d'évaluation a utilisé l'Abaque de Régnier comme technique d'animation des groupes de travail. Les groupes ont travaillé à partir d'affirmations, souvent provocatrices, rédigées par l'équipe d'évaluation en vue de pousser les participants à se positionner. En faisant voter les participants sur ces affirmations, la technique a créé un point de départ pour faire exprimer les positions et prendre connaissance des autres avis présents dans le groupe. Le débat a permis de définir une position commune sur les objectifs de la politique de promotion économique.

Source : Alberton S., Rey M. et L. Vodoz (1993) La politique valaisanne de promotion économique : une évaluation privilégiant le point de vue des acteurs. Berne : Programme National de Recherche 27.

## ● A quoi sert l'outil ?

Les techniques de vote facilitent l'expression de tous les participants sur un sujet donné. C'est la manière d'obtenir ces avis qui varie selon les techniques. Le vote coloré mise sur la prise de connaissance visuelle de

tous les avis présents au sein du groupe grâce à une présentation en couleur des résultats du vote individuel. On évite ainsi que la forte personnalité de certains dissimule l'avis de participants plus discrets. Le vote coloré permet aussi de dégager rapidement les questions consensuelles de celles qui sont conflictuelles. On gagne ainsi du temps en concentrant la réflexion du groupe sur les aspects problématiques.

Les techniques de vote permettent d'éviter certains inconvénients du travail en groupe : pressions à la conformité, convergence des opinions autour des valeurs centrales, prédominance de certains acteurs, etc.

### ● Dans quels cas l'utiliser ?

Un groupe de travail peut utiliser ces techniques pour n'importe quel type d'évaluation : ex ante pour délimiter les enjeux ou les besoins à satisfaire, in itinere pour clarifier les objectifs comme dans l'exemple ci-dessus et, le cas échéant, ex post pour valider les conclusions de l'évaluation.

Il arrive fréquemment que les objectifs et les critères de jugement d'un programme ne puissent être légitimement clarifiés par une seule personne. Dans ce cas, les techniques de vote facilitent le travail d'un groupe qui réunit des compétences et des niveaux de responsabilité très variés. Autrement dit, les réunions de travail qui encadrent la réalisation d'une évaluation sont un moment idéal pour appliquer les techniques de vote. Dans cette optique, le vote coloré sera particulièrement utile à la hiérarchisation des objectifs du programme, la définition des critères d'évaluation et la sélection des critères les plus importants.

### ● Comment le mettre en œuvre ?

**Etape 1 : Formulation des énoncés.** Les énoncés sont des affirmations exprimées de manière concise, généralement en une seule phrase. La formulation des énoncés est faite préalablement par l'équipe d'évaluation, sur la base d'entretiens personnels ou sur la base d'une première discussion de groupe.

Dans l'exemple ci-dessus les énoncés ont été formulés comme suit :

- « Pour les entreprises industrielles, les aides financières sont dérisoires »
- « Pour l'industrie, la promotion économique doit être destinée aux entreprises "high-tech" »

- « La promotion économique doit bénéficier à toutes les zones du canton de manière égalitaire »
- etc.

A ce stade, il est utile d'effectuer une vérification collective de la compréhension des énoncés.

**Etape 2 : Vote individuel.** Les participants attribuent une couleur à chaque énoncé. La signification des couleurs est simple :

rouge	rose	jaune	vert clair	vert foncé	blanc	noir
désaccord total	plutôt en désaccord	pas d'opinion tranchée sur le sujet	plutôt d'accord	accord total	sans opinion sur le sujet	ne veut pas participer au vote

**Etape 3 : Retranscription des votes sur un tableau aimanté ou sur écran d'ordinateur.** Le résultat est ainsi immédiatement présenté aux participants sous forme d'un tableau rempli de pastilles de couleur. Les tendances sont bien visibles et cela permet à chacun de savoir s'il existe un consensus dans le groupe.

**Etape 4 : Discussion du groupe.** L'animateur procède à la sélection des énoncés soumis à discussion. Le plus souvent on retiendra en priorité ceux ayant fait l'objet de dissensus.

**Etape 5 : Reformulation des énoncés avec le groupe.** Cette étape finale permet de dégager une position collective sur le sujet.

### ● Application aux Fonds structurels

La gestion des Fonds structurels est partenariale, ce qui veut dire que de nombreuses réunions à vocation créatrice peuvent s'appuyer sur des techniques de vote pour structurer et dynamiser leur travail (réunions consacrées à la sélection de projets, à l'évaluation des programmes à leur réaménagement ou à leur renouvellement, etc.)

Le vote coloré est, on l'a vu, particulièrement bien adapté aux situations où une position commune ou un consensus sont requis. Il permet aussi de hiérarchiser les propositions émanant d'un groupe, ce qui peut faciliter les conclusions d'une réunion.

### ● Points forts et limites d'utilisation

- Le vote coloré est apprécié par les groupes lorsqu'ils recherchent un consensus dans un domaine relativement conflictuel. Ainsi, la transparence et la simplicité de la technique assurent une bonne confiance dans le processus.
- Les résultats de ce vote sont immédiatement visibles et les tendances sont, grâce aux couleurs, faciles à percevoir. La possibilité de voter « blanc » ou « noir » permet de mettre en avant le manque d'information ou les points de désaccord sur l'objet de discussion.
- Une forte responsabilité repose sur l'animateur, ce qui requiert une grande expérience de sa part.
- La faible technicité renforce la confiance des participants (NB. : la dernière version de l'Abaque de Régnier est informatisée). L'aspect ludique de cette méthode a également l'avantage de détendre l'atmosphère des réunions. Il existe cependant un risque de rejet si les participants trouvent la technique trop scolaire.
- L'éventuelle prédominance d'un acteur dans les débats est atténuée par la nature non verbale de l'outil.

### ● Pour en savoir plus...

Anzieu, D. et Martin, J.-Y. (1994), *La dynamique des groupes restreints*. Paris : PUF. 395 p.

Ouvrage de référence pour le fonctionnement des groupes restreints. Contient en annexes diverses méthodes de « réunions-discussions ».

Mucchielli, R. (1992), *La conduite des réunions*. Paris : Librairies techniques. 120 p.

Ouvrage de vulgarisation contenant des exercices pratiques.

Viveret, P. (1992) 'Les outils d'intelligence collective et l'évaluation', in : *L'évaluation en développement*. Paris : La Documentation Française. 179-183.

Article résumant des techniques de groupe.

## Outils pour observer les changements sur le terrain

Lorsqu'un programme est complexe et qu'on veut l'évaluer dans son ensemble, le champ d'observation à balayer est très large et rend l'observation exhaustive très onéreuse. Il existe cependant des outils qui permettent de collecter les données nécessaires à l'évaluation.

Il est possible de réaliser une observation de terrain en organisant une série d'**entretiens individuels** avec des opérateurs (observation indirecte) ou avec des destinataires du programme (observation directe). Cette technique est particulièrement adaptée lorsque les informations que l'on cherche à collecter sont relativement factuelles.

Les **entretiens de groupes** (*focus group*) réalisés également avec des opérateurs ou des destinataires permettent de collecter des opinions et des données. En particulier, un entretien de groupe bien conduit permet de collecter des opinions avec une bonne fiabilité.

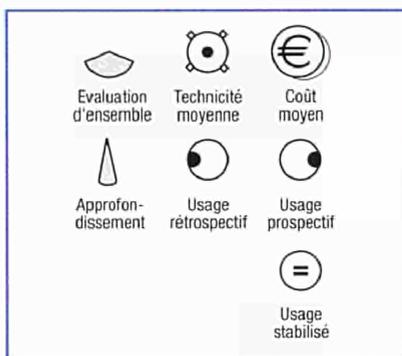
Il est possible de réaliser une observation de terrain systématique grâce à des **études de cas** soigneusement sélectionnées. Dans une situation d'évaluation d'un programme complexe, l'étude de cas semble être le seul outil qui permette à la fois d'observer les effets chez les destinataires directs ou indirects et de respecter la complexité des différentes composantes du programme.

Certaines techniques d'observation classiques sont difficilement applicables dans la situation d'évaluation d'ensemble d'un programme complexe. Par exemple, l'**enquête par questionnaire** auprès des destinataires directs du programme est trop onéreuse si, comme c'est souvent le cas, il existe une dizaine de catégories différentes de destinataires.



## ■ Entretien individuel

La technique de l'entretien individuel permet de recueillir des informations qualitatives et les opinions des personnes concernées par un programme ou un projet particulier, son contexte, sa mise en œuvre, ses résultats ou son impact. Plusieurs formes d'entretien méritent d'être distinguées, chacune répondant à un besoin différent : l'entretien par conversation informelle; l'entretien semi-directif avec guide ; l'entretien directif, approche la plus rigide.



La présentation qui suit est centrée sur l'entretien semi-directif. C'est la technique la plus utilisée dans le cadre de l'évaluation des Fonds structurels.

### Exemple : L'évaluation du programme de North Jutland Technology (NordTek)

Le programme NordTek a pour objet le développement du Nord Jutland par l'introduction de technologies de l'information dans le tissu industriel régional. Le programme NordTek a deux finalités qui sont la création d'emplois et la compensation d'une image négative de région périphérique. A cette fin, la stratégie de développement mise en œuvre par ce programme repose sur quatre objectifs qui seront atteints par trois types d'actions.

Les objectifs sont : renforcer les réseaux locaux d'acteurs, améliorer la qualification des travailleurs en regard des nouvelles technologies, orienter la production régionale vers la demande locale et développer l'attraction des investissements étrangers. Les actions mises en œuvre pour cela sont des mesures en faveur de la connaissance et de l'utilisation des technologies nouvelles, des mesures pour la formation à l'utilisation des technologies nouvelles et des mesures d'aide à l'investissement des PME.

Le programme a été évalué avec des entretiens semi-directifs auprès de 70 entreprises aidées. Le but de ces enquêtes est de comprendre la façon dont les entreprises utilisent les services reçus, et ce que ces derniers leur apportent. Les entreprises qui ont fait l'objet d'entretiens ont été sélectionnées parmi 250 entreprises aidées de 5 des projets les plus importants du programme de NordTek (les différents projets qui composent le programme concernent des thèmes précis, comme le conseil, la formation, la mise en réseau...). Les entreprises sélectionnées devaient avoir reçu une aide relativement importante pendant au moins six mois et être représentatives des différentes formes d'aide qui ont été octroyées.

L'approche choisie est une méthode utilisant un guide d'entretien, afin de rassembler la totalité des informations concernant les entreprises. Le guide d'entretien présente également une plus grande rigueur qui facilite le recouplement des résultats des entretiens réalisés par plusieurs enquêteurs différents. Après l'entretien, l'enquête se poursuivait par une conversation plus informelle au cours d'une visite de l'entreprise, la durée totale étant d'environ deux heures. La personne interrogée recevait ensuite le résumé de l'entretien sous forme d'un document de 4-8 pages, pour validation.

Les thèmes abordés au cours de l'entretien étaient les suivants : les motifs de l'adhésion au projet, le bénéfice attendu de l'aide et le bénéfice reçu, les possibilités d'obtenir une aide provenant d'une source différente, ...

Les résultats des entretiens ont permis d'identifier les raisons qui ont conduit les entreprises à participer aux projets de NordTek. Globalement, il ressort de ces entretiens que l'assistance apportée par NordTek est dans certains cas unique et ne peut être trouvée ailleurs tandis que pour certains aspects, elle concurrence des sociétés locales, voire d'autres projets NordTek. Les services de mise en réseau et de conseil de NordTek se développent sur une période de plusieurs mois, et même plusieurs années ; la connaissance personnelle, la proximité géographique sont dès lors perçues comme des facteurs importants pour les entreprises. Il reste cependant difficile de quantifier les bénéfices que les projets Nortek ont pu apporter aux entreprises, à cause des nombreux facteurs externes pouvant influencer le processus dans lequel les services évoluent.

Source : Olsen, Leif et Olaf Rieper. NordTek-evalueringen, AKF Forlaget, 1991.

### ● A quoi sert l'outil ?

Ce type d'enquête est un moyen de bien connaître et d'examiner l'avis des acteurs (destinataires et autres parties prenantes) sur un projet ou un programme : son adéquation avec les besoins, les résultats par rapport aux attentes... L'enquête par entretien est également utilisée en sciences sociales comme outil d'investigation des logiques de l'utilisateur.

L'entretien individuel est une technique exploratoire permettant d'obtenir des informations pertinentes sur les raisonnements, les conceptions et représentations que se font les personnes interrogées sur un projet, un programme, ... Au-delà des représentations subjectives, il sert aussi à recueillir des informations sur des pratiques individuelles.

C'est un outil particulièrement précieux pour déterminer de quelle façon l'intervention a été mise en œuvre et pour identifier les bonnes pratiques. Il est donc souvent utilisé dans le cadre d'évaluations endoformatives

destinées à tester une théorie d'action et/ou à aider à affiner une politique ou un programme. Il peut également être utile pour fournir une évaluation récapitulative d'un programme destiné à apporter des changements dans l'attitude ou la perception des acteurs (tels que des transferts de technologie ou des programmes de formation) par opposition à des résultats plus tangibles (tels que la création de nouveaux emplois ou la construction de nouveaux équipements).

### ● Dans quels cas l'utiliser ?

L'entretien est utilisé dans un cadre exploratoire, autrement dit lorsque l'on ne dispose pas d'hypothèses a priori ou de connaissances suffisantes sur un projet ou un public cible qui permettraient d'avoir recours à une enquête par questionnaire.

L'entretien est une technique également pertinente lorsque le public concerné par le programme ou projet évalué est trop réduit pour faire l'objet d'un sondage statistiquement représentatif.

L'entretien semi-directif avec guide sera utilisé lorsque l'on recherche des informations relativement standardisées sur le programme étudié. Cette technique nécessite une certaine connaissance préalable du sujet.

La technique de l'entretien individuel peut s'utiliser à tout stade du processus d'évaluation : pour établir une théorie sur le programme évalué, identifier les problèmes et besoins d'une région et les améliorations nécessaires pendant la mise en œuvre du programme. Il peut se révéler particulièrement efficace quand il est utilisé pour recueillir les opinions d'acteurs clés tels que les gestionnaires de programmes ou les principaux décideurs.

### ● Comment le mettre en œuvre ?

Afin de bien conduire un entretien, il est nécessaire de respecter les étapes suivantes :

**Etape 1. La sélection des personnes à interroger :** l'échantillon nécessaire à la conduite d'un entretien est de taille plus réduite que pour une enquête menée par questionnaire ; les informations obtenues sont en effet validées par le contexte et non pas par la probabilité d'occurrence propre aux questionnaires. Le nombre d'entretiens dépend du thème de l'étude, de la variété des comportements face au thème étudié, et des moyens

disponibles (en général de 20 à 60 entretiens). L'échantillon retenu repose le plus souvent sur la sélection des composantes dites caractéristiques de la population (échantillon diversifié). L'échantillon peut être sélectionné soit par accès direct (via des listes administratives par exemple) ou en faisant intervenir des tiers, ce qui permet une sélection plus ciblée mais présente des risques de distorsion.

**Etape 2. La conception du plan d'entretien :** le plan d'entretien comprend l'élaboration d'un guide d'entretien. Il s'agit de préciser les thèmes que l'enquêteur souhaite aborder. Il n'est pas indispensable de suivre les indications du guide d'entretien dans un ordre précis ; ce guide constitue plutôt une sorte de « check-list », qui permet à l'enquêteur de vérifier qu'il traite bien les questions essentielles. L'enquêteur peut moduler son intervention en fonction de la personne interrogée, formuler de nouvelles questions. Les premiers entretiens pourront donner lieu à des ajustements ou à des modifications dans la définition des questions si celles-ci ont posé problème aux personnes interrogées. Très souvent, il est également utile, afin de ne pas perdre de temps pendant l'entretien, d'avoir recueilli au préalable certaines informations de base auprès des personnes interrogées. De même, on pourra fournir préalablement à la rencontre le plan des questions qui seront abordées afin qu'elles puissent réunir toutes les données objectives. Il est généralement préférable d'établir un contrat spécifiant le but de l'entretien, sa durée, le niveau de confidentialité, l'usage qui sera fait des conclusions, etc.

**Etape 3. La sélection et la formation des enquêteurs :** la bonne conduite d'un entretien nécessite un grand professionnalisme de l'enquêteur. Celui-ci doit justifier de compétences pour communiquer, pour écouter et pour la prise de notes. Pour faciliter le déroulement de l'entretien, et faire en sorte que la personne interrogée se sente à l'aise, il peut être utile de s'assurer de la proximité sociale de l'enquêteur et de l'interrogé. Les enquêtes les moins structurées nécessitent en outre de la part de l'enquêteur une connaissance importante du terrain.

**Etape 4. Le déroulement de l'entretien :** un enquêteur doit se montrer « respectueux » de la personne interrogée et des informations collectées, et doit savoir exposer clairement le sujet et le contexte. Le contact initial est très important. Il établit une base de communication. L'enquêteur doit veiller à ne pas influencer la personne interrogée, en l'approuvant ou en orientant ses réponses. Il est possible d'enregistrer l'entretien, afin d'être fidèle aux assertions propres à la personne interrogée, et de s'assurer au

moment du compte-rendu que les remarques les plus importantes n'ont pas été omises. Toutefois, l'utilisation d'un magnétophone peut parfois « intimider » la personne interrogée. Il faut également prévoir, dans l'estimation du budget et du calendrier, le temps et le coût nécessaire à la transcription des enregistrements. Le compte-rendu (ou le résumé) devra être approuvé par la personne interrogée qui pourra alors éventuellement compléter toute réponse restée inachevée. Il est recommandé d'adresser une lettre de remerciement aux personnes interrogées (notamment afin de s'assurer de leur bonne volonté au cas où leur coopération serait à nouveau nécessaire dans le cadre d'études futures).

**Etape 5. L'analyse des résultats** : cette dernière phase consiste à analyser les discours, interpréter, comparer les informations communiquées par les personnes interrogées, rechercher les avis communs et divergents, afin de dresser le bilan de l'évaluation. Il est possible de faire un résumé entretien par entretien, centré sur la cohérence propre à chaque entretien, ou une analyse thématique transversale, plus appropriée à la recherche de modèles capables d'expliquer les pratiques individuelles.

Pour produire les résultats et les discuter ensuite, l'enquêteur établit une grille d'analyse, à partir de la lecture des entretiens et d'hypothèses descriptives. Cette grille constitue un outil explicatif, appliqué à chaque compte-rendu d'entretien.

### ● Application aux Fonds structurels

Les techniques d'entretien sont largement utilisées dans le cadre de l'évaluation des interventions structurelles et en particulier pour l'évaluation des programmes. L'entretien avec les gestionnaires et les destinataires reste l'un des outils les plus communément utilisés dans l'évaluation intermédiaire des programmes structurels. Le principe consiste à sélectionner plusieurs destinataires ou gestionnaires, en fonction des caractéristiques de l'action mise en œuvre et du public concerné.

Utilisé comme outil d'analyse, l'entretien est l'un des seuls outils qui permette l'accomplissement d'un tour d'horizon rapide des programmes à évaluer. Les entretiens les moins structurés - conversation informelle avec les gestionnaires et autres personnes concernées par le programme - pourront être utilisés pour identifier les parties du programme qui devront être approfondies. Les entretiens avec les décideurs peuvent représenter une étape importante dans la mise au point de l'évaluation et la détermination des résultats clés attendus par ces décideurs.

Les entretiens peuvent également s'avérer pertinents et fournir des informations utiles dans le cadre de programmes à caractère social pour lesquels les destinataires sont souvent peu motivés pour remplir des questionnaires, comme les mesures d'insertion professionnelle d'anciens chômeurs ayant suivi un stage.

#### ● Points forts et limites d'utilisation

- L'avantage de ce type d'enquêtes est qu'il fournit des informations approfondies sur les valeurs, les faits et comportements des personnes interrogées ; il permet d'identifier, de mettre en relation un ensemble d'éléments, donnant lieu à une étude relativement exhaustive sur un thème donné.
- Un entretien individuel prend en compte les facteurs situationnels et individuels ; pour cette raison, il reste difficile de tirer des conclusions générales. Les entretiens individuels peuvent permettre une identification exhaustive des effets et des causes possibles ; ils ne sont par contre pas utilisables pour mesurer des impacts ou hiérarchiser les causes.
- Un entretien bien mené peut permettre de mieux comprendre les mécanismes de mise en œuvre, les relations causales propres à un programme et d'identifier des réussites exemplaires ou des défaillances manifestes. Dans ce cas, la technique produit une information facile à communiquer sous forme d'encadrés illustratifs.
- Pour être conduite efficacement, cette pratique demande beaucoup de temps et nécessite la contribution de professionnels. Des compétences particulières sont requises pour concevoir, conduire et interpréter un entretien ; une maîtrise insuffisante de ces compétences produirait une information sans valeur. Or, les entretiens ne sont pas toujours pratiqués avec le professionnalisme requis pour produire des résultats efficaces.
- Le bénéfice de cet outil dépend des connaissances des personnes interrogées et de leur volonté à bien vouloir répondre. Les informations peuvent être également faussées du fait du choix des informateurs.

#### ● Pour en savoir plus...

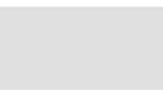
Blanchet, A., Gotman, A. (1992) *L'enquête et ses méthodes : l'entretien*. Nathan Université  
Ouvrage entièrement consacré à l'enquête par entretien, vue sous un angle pédagogique

Casley, D., Kumar, K. -(1991) *Collecte, analyse et emploi des données de suivi et d'évaluation*. Banque Mondiale.

Présentation claire et accessible des techniques d'entretien -Ch. 2 et 3, 12 à 47.

Patton, M.Q., (1987) *How to use qualitative methods in Evaluation*, Sage Publications.

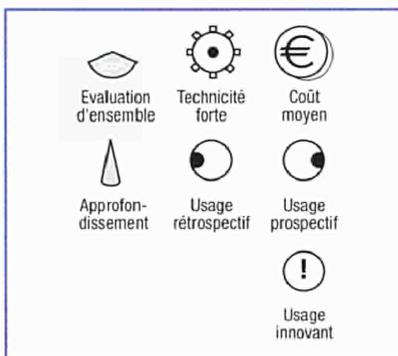
Consacré à l'ensemble des techniques qualitatives, il donne un aperçu de leur application dans le domaine de l'évaluation.



## ■ Entretien de groupe (Focus Group)

Un entretien de groupe (*Focus Group*) fait intervenir plusieurs personnes ayant des caractéristiques homogènes et fournit des informations qualitatives au cours d'une discussion ciblée.

La technique de l'entretien de groupe a été imaginée par les spécialistes du marketing pour analyser l'impact potentiel de stratégies publicitaires et messages commerciaux. La technique fait appel à l'interaction et à la créativité des participants pour enrichir et consolider l'information collectée. En particulier, elle peut aider les participants à se forger une opinion sur un sujet qu'ils connaissent peu. Elle est spécialement intéressante pour analyser des thèmes ou domaines qui suscitent des opinions divergentes que l'on doit tenter de réconcilier ou qui ont trait à des questions complexes qui doivent être explorées en profondeur.



### Exemple : L'évaluation thématique du DOCUP Objectif 3, 1994-99 Royaume Uni

Dans le cadre d'une évaluation du DOCUP objectif 3 Royaume-Uni, la technique de l'entretien de groupe (*focus group*) a été utilisée en complément d'autres techniques. Les objectifs généraux de cette évaluation étaient de vérifier, d'une part si les actions mises en œuvre au sein du DOCUP atteignaient effectivement les destinataires souhaitables et, d'autre part, si la structure du DOCUP favorisait l'émergence de projets efficaces en termes d'intégration des chômeurs aidés sur le marché du travail.

La méthode construite pour cette évaluation comportait un volet quantitatif basé sur l'exploitation des données de suivi et un volet qualitatif basé notamment sur quatre entretiens de groupe dans quatre régions représentatives de la diversité des situations au Royaume-Uni : l'Ecosse, les "West Midlands", Londres et le Sud-ouest. Ces groupes étaient constitués de représentants de porteurs de projets (notamment centres de formation, "regional councils", instituts d'enseignement supérieurs, organisations non gouvernementales, "Training and Enterprises Council"). Les thèmes et questions abordés étaient structurés comme suit :

- examen des sources d'information disponibles pour l'élaboration de projets correspondant aux besoins des publics cibles ainsi que pour choisir les moyens d'atteindre ces publics ;

- mise en évidence des points forts et faibles des modes de sélection, d'approbation et de gestion, notamment financière, des projets et leur incidence sur l'efficacité en termes de ciblage ;
- discussion sur les moyens de mesure et d'appréciation des résultats de ces actions en fonction de la spécificité des publics cibles et des demandes des employeurs ayant fait l'objet d'un débat ;
- et enfin, débat sur l'efficacité du DOCUP comme instrument pour susciter des projets adaptés aux besoins des destinataires.

Ces discussions ont permis de faire émerger des idées et des considérations nouvelles, couvrant souvent des thèmes connexes aux questions posées tels que l'existence et l'intérêt de "projets intégrés", les modes de fonctionnement en réseaux de porteurs de projet, "l'additionnalité" des projets,... Les résultats tirés de ces entretiens de groupe, bien que non représentatifs au sens statistique du terme, ont permis de dégager des "idées à creuser" dans des domaines divers tels que : l'accès à des publics difficiles, les conditions facilitant l'émergence de projets pertinents en termes de ciblage,...

Source : Targeting and the ESF Objective 3 1994-1996 Single Programming Document (SPD), final report to the Department for Education and Employment (DfEE) ECOTEC, juin 1996

### ● A quoi sert l'outil ?

Les entretiens de groupe commencent à être largement utilisés aujourd'hui, plus généralement dans le domaine des sciences sociales et notamment pour l'évaluation. En jouant sur l'interaction et la confrontation des points de vue, cette technique permet de révéler les perceptions des participants sur un thème donné, que ce soit la pertinence du programme évalué, sa mise en œuvre, ses résultats ou ses impacts.

La technique de l'entretien de groupe constitue une application particulière des techniques d'animation de réunions. L'outil est standardisé en vue de créer une interaction optimale pour la collecte d'informations. D'autres techniques de groupe sont utilisées en sciences sociales, comme la technique Delphi, ou celle du groupe nominal (recensement de tous les points de vue individuels, ensuite soumis au groupe et discutés), qui ne comportent pas de véritable discussion interactive.

L'entretien de groupe permet de réunir simultanément ou de manière séquentielle plusieurs protagonistes d'un programme (gestionnaires, destinataires et/ou opérateurs) et de collecter de nombreuses informations qualitatives dans un temps relativement court, grâce à une démarche

« d'intelligence commune ». La confrontation des réalités vécues et des représentations des participants au groupe permet de mieux comprendre le phénomène à observer.

En impliquant les acteurs ou les destinataires d'un programme dans les travaux d'évaluation, les conclusions de l'étude seront en outre plus facilement crédibles et acceptées. La technique de l'entretien de groupe peut également constituer un outil de validation d'une collecte de données, ou venir compléter des données quantitatives.

### ● Dans quels cas l'utiliser ?

Les entretiens de groupe sont adaptés aux cas où les thèmes à traiter suscitent des opinions divergentes mais où la discussion reste possible ; les sujets traités seront « d'intérêt public » de manière à pouvoir être débattus en groupe.

### ● Comment le mettre en œuvre ?

Dans la même logique que les entretiens individuels, plusieurs étapes sont nécessaires afin de bien préparer et conduire le entretiens de groupe. Une fois le thème de l'étude précisément défini, les étapes suivantes sont mises en œuvre :

**Etape 1 - la sélection des participants :** le choix des participants aux groupes dépend des caractéristiques de la population et du thème traité. Il est préférable de constituer plusieurs types de groupes (entre 3 et 5) au sein des catégories retenues. En effet, se limiter à travailler avec un seul groupe comporte des risques quant à la validité de l'étude : la dynamique de groupe peut par exemple être faussée à cause de la composition hétérogène de l'échantillon. Les participants seront sélectionnés de manière à assurer une certaine homogénéité au groupe. Le nombre de participants est généralement compris entre 6 et 10 par groupe, de manière à ce que chacun puisse prendre la parole et afin de s'assurer que des sous-groupes ne se constituent pas. En règle générale, les participants ne se connaissent pas.

Les participants sont parfois rémunérés ; cependant, plutôt qu'une rémunération incitative, il est préférable de proposer aux participants de rembourser leur frais de déplacement, s'il y a lieu, et de leur servir une collation en fin de séance.

**Etape 2 - la conception du guide d'entretien** : compte tenu des délais et du nombre de personnes, il est important de bien définir et limiter les sujets abordés. Les guides pour chaque groupe seront différents. Le guide des entretiens de groupe devra être court, car il faut plus de temps lorsque plusieurs personnes participent à la discussion et répondent aux mêmes questions. La plupart des questions posées sont des questions ouvertes ; elles seront également soigneusement établies et sérieuses, les plus générales venant en premier lieu.

**Etape 3 - le choix et la formation des animateurs** : l'animateur doit faire preuve de créativité, de motivation et assurer de bons rapports avec les participants. La formation des animateurs tiendra compte des attitudes à adopter vis-à-vis du groupe (respect des idées, ...), des connaissances qu'il doit intégrer quant aux objectifs de l'étude, aux techniques d'analyse des données. L'animateur doit également faire preuve de compétences pour communiquer, pour la prise de notes, etc. Il est utile de prévoir une deuxième personne par groupe, particulièrement si la séance n'est pas enregistrée sur magnétophone pour prendre des notes pendant l'entretien pendant que l'autre anime et facilite la discussion. Des « co-animateurs » peuvent observer le déroulement de la discussion, soit en étant physiquement présents, soit d'une autre pièce (à travers une glace sans tain, par exemple), et transmettre ainsi à l'animateur des recommandations sur la conduite de la réunion.

**Etape 4 - le déroulement de la discussion** : La discussion peut être lancée de manière assez ouverte, en introduisant le sujet de la session et en posant une question simple d'intérêt général, permettant à chaque participant d'émettre un premier avis ou une remarque sur le sujet. Le guide d'entretien sert à rappeler à l'animateur tous les points essentiels qu'il doit aborder. Afin de conserver tous les avantages de l'entretien, il faut en posséder le compte-rendu rédigé. Il est fortement recommandé d'enregistrer, voire filmer, l'entretien, bien que cette dernière formule soit souvent perçue comme une « intrusion » par les participants. Certaines salles sont conçues de manière à ce que les participants ne remarquent pas l'enregistrement.

**Etape 5 - l'analyse et le compte-rendu des résultats** : cette dernière phase consiste à interpréter et comparer les informations communiquées par les participants et à rechercher les avis communs et divergents au sein de chaque groupe. Les informations collectées sont codifiées de manière à organiser les résultats au vu des objectifs de l'enquête. L'interprétation des données doit tenir compte et distinguer deux aspects majeurs de la

discussion : ce que les participants considèrent intéressant et ce qu'ils jugent important. L'analyse dépendra du nombre de groupes interrogés, et de la nature des entretiens (par exemple, l'entretien de groupe est-il mené selon une approche structurée ou non ?). Les résultats en provenance des différents groupes sont comparés afin d'identifier d'éventuelles convergences. La rédaction du rapport pourra intégrer les citations les plus marquantes des participants, ainsi qu'un résumé des discussions.

### ● Application aux Fonds structurels

La technique de l'entretien de groupe peut s'utiliser pour tester une mesure innovatrice (évaluation ex ante), clarifier les objectifs d'un projet, établir une théorie d'action pour le programme en cours d'évaluation, identifier les problèmes et besoins d'une région et les améliorations nécessaires pendant la mise en œuvre du programme. Cette technique peut également être utilisée en fin de programme dans le cadre d'une évaluation ex post, afin de collecter des informations qui permettront d'identifier et/ou d'interpréter les effets et impacts du programme concerné, de fixer de nouvelles priorités et orientations.

### ● Points forts et limites d'utilisation

- Ce type d'enquêtes fournit des informations approfondies sur les valeurs, les opinions des personnes interrogées.
- Le fait de regrouper plusieurs personnes fournit un certain équilibre dans les réponses données et permet à l'équipe d'évaluation de cerner plus facilement l'opinion générale qui se dégage sur un programme donné. Du fait de la participation de plusieurs personnes, les entretiens de groupe assurent un certain « contrôle de qualité » sur la collecte des données, en jugeant les « pour et contre » des arguments avancés par chacun, évitant ainsi les opinions extrêmes.
- En peu de temps (environ une heure trente - deux heures), il est possible de collecter une quantité importante d'informations qualitatives.
- Des compétences particulières sont nécessaires pour gérer la dynamique de groupe et obtenir une discussion équilibrée évitant l'influence dominante des animateurs. Il existe également un risque de « polarisation », c'est-à-dire que les participants seront, après discussion du groupe, influencés par l'opinion qui était majoritaire au début de la discussion.

- Les discussions peuvent parfois être biaisées, du fait que les participants (destinataires) de politiques publiques sont sujets à un effet de « clientélisme », et produiront un jugement a priori positif.
- Le nombre de personnes limite cependant le nombre de questions en raison du temps nécessaire pour y répondre ; par ailleurs, des conflits peuvent émerger entre les différentes personnes, et diminuer la qualité et l'efficacité de l'entretien.
- Il est possible que la participation à un groupe d'entretien change la perception des personnes, soit en raison de « l'effet Hawthorne » (le fait que le comportement des personnes observées change parce qu'elles se savent observées), soit parce que l'interaction avec d'autres participants leur fournit de nouveaux raisonnements ou leur offre de nouvelles perspectives. Ainsi par exemple, participer à un groupe d'entretien peut réellement améliorer les performances des gestionnaires de programme. De cette façon, la méthodologie de l'entretien de groupe peut avoir un impact sur le programme en cours d'évaluation ou sur les programmes qui suivront.

#### ● Pour en savoir plus...

Anzieu, D, Martin, J.Y., (1994) *La dynamique des groupes restreints*, PUF Le psychologue

Présentation complète, par des spécialistes de la psychosociologie, du concept du groupe restreint, des principaux phénomènes de groupe, de certains domaines d'application.

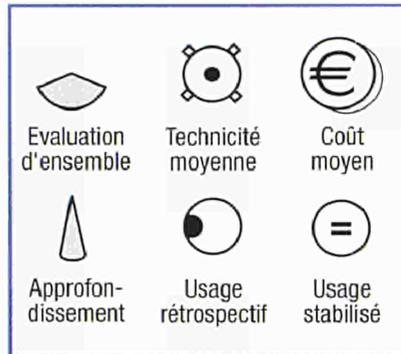
Aubel, J., (1992) *Guide pour des études utilisant les discussions de groupe*, BIT.

Guide pratique de la conduite d'entretien de groupe.

M. Moore, C., (1994) *Group techniques for idea building*, Sage publications  
Approche plus détaillée sur les techniques de groupe, qui présente des techniques spécifiques (Nominal Group Technique, Ideawriting, Interpretive Structural Modeling,...).

## ■ Etude de cas

Cet outil est basé sur l'étude approfondie de données rassemblées à propos de cas spécifiques. Il peut s'agir d'individus, de programmes, d'organisations, de projets, de groupes de personnes, ou encore de processus de prise de décision. Cette approche utilise une série de techniques de collecte de données qualitatives et quantitatives. L'étude de cas est indiquée pour les évaluations visant l'analyse détaillée de processus complexes de mise en œuvre de programmes et la représentation de leurs impacts. L'étude de cas s'applique également aux situations qui requièrent une étude de terrain exploratoire en raison du manque d'informations disponibles.



### Exemple : l'évaluation du programme de NordTek, Danemark

Ce programme, destiné aux petites et moyennes entreprises, comportait une centaine de projets offrant des services aux entreprises dans les domaines de la formation, du conseil, de l'accès à des installations informatiques modernes, ... Il s'agit donc d'un programme que l'on peut qualifier de « soft » par opposition aux programmes « hard », visant à l'amélioration des infrastructures. Les résultats de tels programmes sont difficiles à évaluer précisément car ils sont immatériels.

Six projets de taille importante ont été sélectionnés en tant que cas à étudier. Les projets devaient, selon les critères de sélection, être représentatifs de la diversité au niveau des stratégies d'intervention, mais aussi au niveau de leur localisation. De plus, les projets sélectionnés devaient avoir bénéficié de financements importants dans le cadre du Programme NordTek et être orientés exclusivement vers les entreprises (et non des instituts de recherche ou des administrations). Les six projets évalués représentaient une part tellement importante du programme total que l'analyse des impacts de ces interventions permettait de couvrir l'ensemble des phénomènes importants. Il a été reconnu sans ambiguïté que les projets choisis représentaient les « meilleurs cas » du programme.

La collecte de données a été faite à deux niveaux opérationnels : au niveau du projet proprement dit et au niveau des entreprises. En ce qui concerne le projet, chacun des cas a été étudié à 4 ou 5 reprises sur le terrain par les évaluateurs en l'espace de deux ans. A chaque fois, les gestionnaires étaient interrogés, et chaque projet devait garder les traces des services proposés et des entreprises concernées dans un « cahier d'enregistrement » conçu spécialement pour l'évaluation. L'analyse des

documents ainsi obtenus était retransmise aux chefs de projet et discutée lors des visites. Au niveau des entreprises, des données ont été collectées au cours d'entrevues auprès d'un échantillon de 20 PME qui avaient utilisé les services offerts. Enfin, 40 PME qui n'ont pas utilisé les services proposés ont été interrogées dans une dernière étape de l'évaluation.

Source : Olsen, Leif et Olaf Rieper. Nordtek-evalueringen, AKF Forlaget, 1991.

### ● A quoi sert l'outil ?

L'étude de cas peut répondre aux finalités suivantes :

- illustration : l'étude de cas est un outil qui permet d'ajouter du réalisme à une évaluation, si sa présentation est réalisée sous une forme narrative. Il faut cependant choisir le cas avec soin, car il doit être représentatif de l'ensemble d'un programme ou illustrer un point spécifique (par exemple, une action particulièrement efficace ou une approche ayant révélé de sérieuses faiblesses qu'il conviendrait donc d'éviter à l'avenir) ;
- exploration : émettre des hypothèses pour de futures investigations, identifier les différents points de vue des parties prenantes ;
- analyse critique : vérifier et valider une affirmation relative à un programme, un projet ou une stratégie ;
- analyse de mise en œuvre : examiner la diffusion des services et ses mécanismes, souvent dans des lieux différents ;
- analyse des impacts des programmes : comprendre la nature des processus qui engendrent les impacts.

Les résultats d'une étude de cas sont toujours présentés sous une forme narrative, comme une histoire, et donnent ainsi au lecteur une « vision de l'intérieur » du cas étudié, lui procurant une impression d'authenticité. Ainsi, l'étude de cas possède une finalité analytique aussi bien que communicative. Les lecteurs sont plus susceptibles de se sentir concernés par des cas où les programmes et les intervenants sont identifiés. Cependant, et particulièrement si l'étude de cas est utilisée pour illustrer des échecs et en tirer des leçons, il pourra être nécessaire de conserver un caractère anonyme à tout ou partie des informations.

Les études de cas peuvent souvent être conçues de façon cumulative pour faciliter la réponse aux questions d'évaluation. Un même cas peut être étudié à plusieurs reprises pour fournir une analyse plusieurs fois mise à jour. Les cas peuvent être descriptifs, normatifs ou destinés à mettre en

évidence les liens de causalité. Ils peuvent être particulièrement utiles comme outil pédagogique (par exemple, ils peuvent aider les responsables officiels à repérer des méthodologies d'évaluation différentes).

### ● Dans quels cas l'utiliser ?

L'étude de cas a été largement utilisée dans le domaine de l'évaluation au cours de la décennie passée. Aujourd'hui, on reconnaît que cette approche fournit des informations valables aussi bien pour l'évaluation des programmes que pour la diffusion de nouvelles connaissances. Les études de cas utilisant des procédures de sélection sophistiquées (par exemple, « étude de cas multiples avec protocole itératif ») ont tendance à remplacer les enquêtes quantitatives à grande échelle réalisées dans des contextes culturels diversifiés.

L'étude de cas est une méthode d'analyse holistique qui s'applique aux situations complexes. Cela signifie que son utilisation est adaptée à la compréhension approfondie des comportements et des phénomènes sociaux, en prenant comme cadre de référence les personnes et les organisations analysées. Les études de cas sont précieuses pour identifier les effets des programmes de manière inductive, en permettant de développer des hypothèses à propos des phénomènes qui relient les causes aux effets. Ces hypothèses doivent ensuite être étayées par les informations issues des différentes études de cas et testées par la recherche d'explications alternatives.

Cela peut s'avérer utile pour observer les résultats et les impacts escomptés, mais aussi pour mettre à jour ceux qui n'étaient pas prévus. Cette approche est moins pertinente pour mettre en évidence les relations causales, bien qu'elle puisse être utilisée pour démontrer la probabilité de leur existence.

L'étude de cas vise l'illustration la plus complète possible d'une situation donnée, afin de donner une image précise du déroulement des phénomènes en cours et d'en appréhender les fondements. Ceci est obtenu par la description puis l'analyse d'exemples restitués dans leur contexte. Cela signifie que ce type d'analyse doit être fondé sur des sources de données multiples, comme les entretiens, les observations au fil du temps, les statistiques, les informations physiques, etc. Un recoupement des données doit également être effectué pour permettre d'en vérifier la cohérence. La notion de « contexte » englobe tous les facteurs qui pourraient affecter le

cas étudié. Ainsi, par exemple, les impacts d'un projet spécifique sur les destinataires sont influencés par un grand nombre de facteurs externes.

### ● Comment le mettre en œuvre ?

La quantité de travail nécessaire pour une étude de cas peut être extrêmement variable. Il faut garder à l'esprit que l'étude de cas doit être suffisamment riche pour donner au lecteur un compte rendu fidèle des faits. Cependant, l'étude de cas fait partie des méthodes les moins standardisées et peut inclure une série d'approches différentes dans des situations différentes. La conduite d'une étude de cas suppose la mise en œuvre des quatre étapes suivantes :

**Etape 1. La sélection des cas à étudier :** Il existe au moins trois critères de sélection de cas : la commodité/l'accès, l'objectif à atteindre et la possibilité d'élargir un cas particulier. La sélection des cas est une étape critique pour généraliser et répondre aux questions d'évaluation. Il est difficile de justifier une sélection fondée uniquement sur la commodité (facilité de l'accès aux données), le sondage probabiliste est parfois difficilement réalisable. Aussi, le choix sur la base de la finalité est approprié dans la plupart des cas, comme le montre le tableau suivant :

A quelles questions peut-elle répondre ?	Base de sélection
Que se passe-t-il aux extrêmes ? Qu'est-ce qui explique de telles différences ?	cas contrastés
Qu'est-ce qui explique l'efficacité d'un projet ?	les meilleurs cas
Pourquoi un projet ne fonctionne-t-il pas ?	les cas les moins bons
Comment les différents types de projets peuvent-ils être comparés entre eux ?	par sous-ensembles
Parmi les exemples choisis pour représenter des variations importantes, que se passe-t-il et pourquoi ?	cas représentatifs
Dans un site typique, que se passe-t-il et pourquoi ?	cas typiques
En ces circonstances particulières, que se passe-t-il et pourquoi ?	cas particuliers

(adapté de GAO, 1990, p.23)

**Etape 2. Collecte de données et processus :** La collecte de données regroupe en principe toutes les données disponibles sur le cas, y compris

celles provenant des documents sur le projet, les comptes-rendus des réunions de projet, et également les données collectées aux différents niveaux opérationnels : les entretiens avec les chefs de projet et le personnel, l'observation du site du projet, les enquêtes auprès des destinataires des services obtenus grâce au projet. Ces données doivent être collectées, enregistrées (élaboration d'un « registre ») et rassemblées/réorganisées de façon à pouvoir être intégrées dans le rapport final.

**Etape 3. Les minutes du cas :** La rédaction des « minutes » du cas reprend et organise l'ensemble des données brutes relatives au cas en un ensemble d'informations exploitables. Les informations sont éditées, les redondances sont éliminées, les différentes parties sont reliées entre elles. Le compte-rendu est organisé pour être facilement consultable, soit de manière chronologique, soit de manière thématique. Le rapport doit inclure toutes les informations nécessaires pour l'analyse ultérieure, c'est-à-dire pour construire le récit de l'étude de cas.

**Etape 4. Récit :** La monographie du cas devrait donner au lecteur un accès immédiat aux informations, à la situation particulière du cas - la situation d'un projet - et lui faire comprendre le projet dans son ensemble. Chaque récit de cas, dans un rapport d'évaluation, devrait être isolé (la taille pouvant varier entre une et cinq pages). Cependant, dans les dernières étapes de l'analyse, il se peut que les cas soient utilisés de façon contrastée et comparés en fonction des objectifs de l'évaluation.

### ● Application aux Fonds structurels

Le dispositif d'étude de cas multiples est particulièrement adapté aux analyses des diverses régions et pays membres, mais aussi aux évaluations thématiques : la flexibilité de chaque étude de cas permet de dresser un portrait adéquat des particularismes d'un lieu donné ou d'un projet. La formulation d'un ensemble commun de questions relatives à l'évaluation facilite l'analyse des résultats obtenus à partir de multiples études de cas. En effet, les résultats s'avèrent plus solides s'ils sont produits à propos d'une variété de lieux (par la réutilisation d'études de cas). De même, la spécificité des cas d'échecs ou de succès apparaît alors de manière plus évidente.

Il convient de rappeler que l'analyse transversale de cas consiste en des examens qualitatifs croisés et une description basée sur le cadre de référence établi par les questions d'évaluation. Un processus cumulatif

peut être recherché lorsque l'évaluation porte par exemple sur les programmes opérationnels des régions d'Objectif 1 et que les conclusions doivent être synthétisées à l'échelle de plusieurs Etats membres.

La présentation des résultats de plusieurs récits de cas pourrait être une barrière à l'utilisation généralisée. On peut résoudre cette difficulté au moyen d'un résumé graphique. Ce résumé fournit un compte-rendu sommaire du récit et une présentation graphique des résultats en lien avec chacune des questions. Sous une telle forme, la réponse à chacune des questions d'évaluation est disposée, pour chaque cas, sur une seule page : une présentation graphique en haut, un résumé du récit court mais riche, avec les principaux résultats, et une conclusion concise. Le résumé du cas transversal pourrait être ébauché de la même manière, suivi des conclusions et recommandations.

Un autre type d'étude de cas pouvant s'appliquer aux Fonds structurels est « l'approche intégrée ». Cette approche tient compte par exemple de l'étude des résultats et des impacts dans le contexte d'un projet spécifique. Dans l'exemple du NordTek mentionné ci-dessus, et pour chacun des projets, un nombre de destinataires (PME) était sélectionné pour une mini étude de cas. De cette manière, le rapport fait par les PME sur l'utilisation des résultats obtenus a pu être utilisé dans le contexte du projet. Ainsi, la variation des résultats parmi les différentes PME peut s'expliquer par les différences de situation des PME concernées.

### ● Points forts et limites d'utilisation

- L'étude de cas est pertinente pour donner un aperçu des processus et des complexités qui sont impossibles à découvrir autrement. Elle peut même faire prendre conscience aux personnes extérieures, particulièrement aux gestionnaires européens peu impliqués dans ce domaine, de la réalité des actions quotidiennes et leur fournir une vision plus claire de la manière selon laquelle le programme est mis en pratique une fois la décision prise avec les autorités nationales.
- Selon la stratégie de sélection retenue, il peut s'avérer possible, si plusieurs études de cas sont effectuées, d'extrapoler au-delà des résultats de ces études. Pourtant, il est impossible de généraliser au sens statistique les résultats à une population globale, de passer de quelques projets à un programme, ou de quelques destinataires à tous les destinataires d'une certaine catégorie.

- Cette approche est moins appropriée pour mesurer l'amplitude des impacts ou pour en inférer la causalité.
- En raison des coûts de mise en place d'une bonne étude de cas (requérant des sources de données multiples, des évaluateurs compétents), il convient de se limiter à un nombre d'observations réduit. Mais par ailleurs, on peut la réutiliser et l'appliquer à d'autres contextes et ainsi faire des économies d'échelle.
- La crédibilité des résultats de l'étude de cas est susceptible d'être affaiblie si la méthode n'est pas mise en œuvre correctement : incomplétude, sélection arbitraire des informations, troncage des commentaires, distorsion des résultats... Pour accentuer la fiabilité d'une étude de cas, quelques précautions peuvent être recommandées. Par exemple : demander une relecture des études de cas par les personnes concernées, afin de vérifier la précision, la véracité des données et leur interprétation ; faire écrire des commentaires distincts par deux évaluateurs, en s'appuyant sur un même cas ; faire intervenir des professionnels extérieurs (comme des journalistes) dans la rédaction du commentaire.

● **Pour en savoir plus...**

Dufour, S., Fortin, D., Hamel, J. (1994), *L'enquête de terrain en sciences sociales : l'approche monographique et les méthodes qualitatives*, Montréal : Saint-Martin.

GAO (1990), *Case Study Evaluations*. Washington DC : General Accounting Office.

Bonne approche générale de la méthode.

Yin, R.K. (1994), *Case Study Research - Design and Methods*, 2nd ed. Newbury Park : Sage Publications.

Ouvrage complet sur la technique de l'étude de cas.



## Outils pour analyser les données

Une fonction importante de l'évaluation est celle qui consiste à croiser les données primaires et secondaires qui ont été collectées. Dans la situation concernée, ces données sont généralement nombreuses et peu homogènes, ce qui crée des difficultés pour l'emploi d'outils sophistiqués. La plupart des évaluations se contentent d'employer les outils classiques que sont les **tableaux** et les **graphiques** (voir par exemple : Henry G.T. ed., 1997, *Creating Effective Graphs : Solutions for a Variety of Evaluation Data*, *New Directions for Evaluation*, no. 77).

Une technique innovante et particulièrement intéressante dans cette situation est celle du **système d'information géographique** qui permet de croiser les données en tenant compte de leur localisation. Les cartes produites par cet outil servent à effectuer des recoupements visuels souvent très suggestifs.

En raison de sa souplesse d'utilisation, la technique de l'**étude de cas**, déjà citée plus haut, est une des rares qui permette l'analyse exhaustive, bien que qualitative et non hiérarchisée, de l'ensemble des effets d'un programme complexe.

La plupart des outils d'analyse quantitative sont inadaptés en raison de l'hétérogénéité des données ainsi que de la multiplicité des publics et des processus de production des impacts. Par exemple, il paraît très difficile d'employer dans cette situation des outils tels que l'analyse factorielle, l'analyse de régression ou les groupes de comparaison.

Par contre, les outils macroéconomiques sont bien adaptés dans le cas de vastes programmes couvrant l'ensemble d'un territoire, sous réserve que la situation se conforme aux critères qui suivent. Les dépenses engagées dans le cadre du programme doivent être significatives par rapport à l'économie du territoire considéré. Ce territoire doit être suffisamment vaste (un pays ou une grande région), ou suffisamment fermé (une île), pour que le fonctionnement global de son économie puisse être isolé. Lorsque ces conditions sont réunies, l'équipe d'évaluation a la possibilité d'employer plusieurs outils empruntés à l'économie régionale ou à la macroéconomie.

L'**analyse structurelle-résiduelle** (*Shift-Share*) consiste à projeter les tendances économiques nationales sur l'économie d'une région. Cet outil

permet d'estimer une situation contrefactuelle et fournit, par différence avec la situation observée, une évaluation de l'impact global du programme.

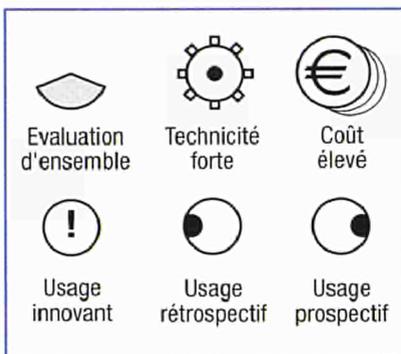
Le modèle d'entrées-sorties (*input-output*) et le modèle macro-économique permettent de simuler des scénarios d'évolution de l'économie d'une région. Ces outils sont généralement appliqués ex ante sur deux jeux d'hypothèses (avec / sans programme) et permettent de donner une estimation globale des impacts macroéconomiques probables.

## ■ Système d'Information Géographique

Les systèmes d'information géographique (SIG) sont des outils qui permettent de travailler avec les données qui peuvent se rapporter à une unité localisée (point, zone, limite, etc.). A l'origine, l'outil a été développé dans les années 1970 pour stocker et traiter l'information liée à

l'environnement physique (environnement naturel et construit). Plus récemment, l'intérêt des utilisateurs des SIG s'est étendu aux informations socio-économiques, le SIG devenant un point de rencontre entre toutes les informations qui peuvent être référées directement ou indirectement à un emplacement géographique. Les SIG permettent maintenant de collecter, réunir, cumuler, analyser, exploiter, exposer et mettre à jour toutes données et informations spatialement référées, dans le but de répondre aux problèmes de planification et de gestion de l'espace.

Les progrès dans le domaine informatique, notamment la cartographie assistée par ordinateur, ont permis le développement de systèmes d'information très performants. La gestion des interventions publiques peut ainsi prendre en compte un degré supplémentaire de complexité et intégrer un nombre considérable de facteurs interdépendants pour mieux saisir la réalité. Grâce aux SIG cette réalité peut être décrite et analysée en combinant des informations provenant des trois dimensions thématique, spatiale et temporelle.



### Exemple : L'évaluation du PIM de Languedoc-Roussillon en France

Cet exemple correspond à celui de l'Encadré 12.

La technique du Système d'Information Géographique a été utilisée dans le cadre d'une évaluation conduite sur le Programme Intégré Méditerranéen (PIM) Languedoc-Roussillon cherchant notamment à analyser les effets spatiaux du programme. Cet outil a été choisi car la demande d'évaluation était de rendre compte des synergies territoriales d'un programme composé d'actions multiples reposant sur une logique régionale (attribution de subventions dans des zones sélectionnées).

L'éventail des domaines touchés par ce programme est large. Il s'agit des problèmes de développement régional comme notamment le maintien, la diversification et

l'adaptation de l'agriculture, le développement de nouveaux emplois, du tourisme, la protection de l'environnement, cela dans les perspectives spatiale (sur des zones géographiques définies), thématique (par activité économique) et temporelle (référence à l'état initial).

Pour apprécier l'impact des mesures du PIM ainsi que les effets de synergie, il était nécessaire d'observer les transformations produites par le programme, puis de les expliquer. Dans ce but, un « état initial » a été constitué pour décrire la situation de référence. Cette étape a nécessité un important travail de recueil de toutes les données et indicateurs de la structure socio-économique sur le territoire considéré ; collecte néanmoins indispensable pour procéder à une comparaison avec la situation observée après l'attribution de subventions. La base de données ainsi constituée, les évaluateurs ont tenté de vérifier, à l'aide de représentations géographiques notamment, le respect des logiques de localisation, d'intégration spatiale et d'intégration spatio-temporelle des actions entreprises.

L'utilisation de représentations cartographiques associant les faits matériels et relationnels (faits sociaux, politiques, économiques) est pertinente lorsque la programmation repose elle-même sur l'identification de zones sensibles sur un territoire donné. Ainsi, par exemple, les actions de prévention des feux de forêt (reboisement, protection) ont été corrélées sur le plan cartographique avec les zones d'incendies (fréquence) afin d'apprécier la logique d'intégration spatiale des actions.

Si la restitution par une carte est pertinente dans ce cas, il convient de rappeler que très souvent le résultat demande à être interprété car la lecture d'une carte n'est pas immédiate. En effet, dans un premier temps, seules les actions du programme en question et de l'année en cours ont été enregistrées. Les actions de même type entreprises par d'autres organismes entre le moment de la « situation de référence » et de « la situation observée » n'ont pas été relevées ; ce qui a conduit à une interprétation biaisée de la situation réelle.

Source : Commission Européenne DG XVI (1993) L'évaluation du PIM Languedoc-Roussillon, MEANS pilote.

### ● A quoi sert l'outil ?

D'une manière générale, un système d'information géographique permet de rassembler toutes sortes d'informations à caractère spatial et de les présenter sous une forme optimisant leur lisibilité. Les SIG peuvent intégrer des données socio-économiques, des données en provenance des recensements, des résultats de sondages et enquêtes, ainsi que des données de suivi. La représentation spatiale peut être directe (coordonnées cartographiques, adresse postale précise) ou indirecte (code postal, aire de recensement).

En termes d'évaluation, la principale valeur ajoutée du SIG est de permettre des croisements de données que les autres techniques ne permettent pas. En particulier, l'outil peut rapprocher des informations dont les références géographiques sont de nature différente, par exemple : adresse des entreprises aidées et zones urbaines en difficulté, ou bien zones éligibles aux aides et zones sensibles du point de vue de l'environnement. Le SIG permet également d'estimer la valeur d'un indicateur pour une zone donnée, lorsqu'on connaît sa valeur à des échelles statistiques différentes (Ce point est détaillé dans le Volume 2 de cette Collection, consacré aux indicateurs.)

### ● Dans quels cas l'utiliser ?

Créer un SIG est un investissement lourd, surtout en termes de saisie des données. Il est donc difficile de faire un tel investissement dans le cadre d'une évaluation dont le budget et la durée sont limités. Au contraire, un tel investissement est justifié dans le cadre du suivi.

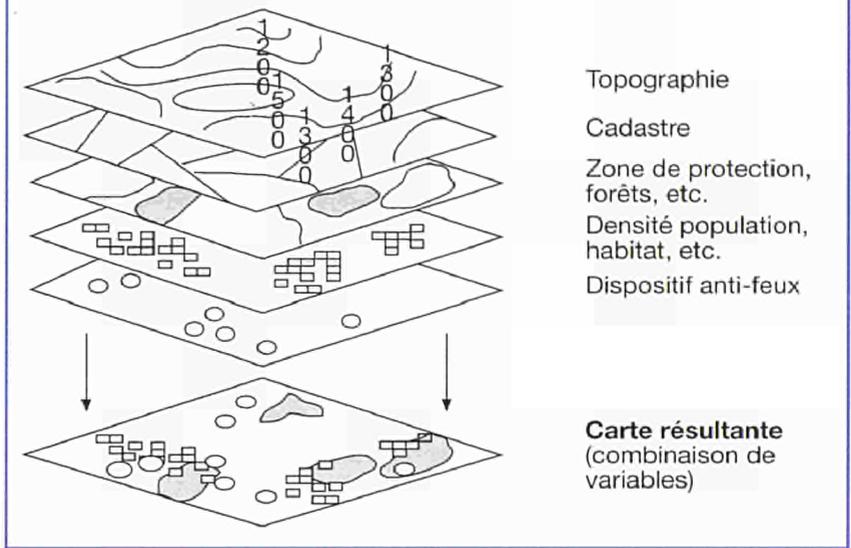
S'il existe un SIG, l'évaluation aura intérêt à l'utiliser. La collecte des données d'évaluation devra comporter une référence géographique systématique, ce qui permettra de réaliser des croisements avec les autres informations déjà contenues dans le système. Les conclusions de l'évaluation pourront également être présentées sous forme de cartes.

### ● Comment le mettre en œuvre ?

Les SIG utilisent une structure de base de données spéciales : toutes les informations sont reliées à un système de coordonnées géographiques caractérisées par des nœuds, des lignes et des aires (donc aussi des points, des arcs, des polygones, etc.).

Les différentes étapes d'un SIG sont les suivantes :

**Etape 1. La collecte des données de bases :** La première phase est la collecte de l'information. Les données de base proviennent de sources très variées : il peut s'agir de cartes géomorphologiques, de photographies aériennes, de données socio-démographiques existantes, etc. Dans certains cas, lorsque les données brutes ne sont pas directement stockables sur support informatique (par exemple une banque d'échantillons), elles sont enregistrées au moyen de fiches indicatrices. Cette étape comprend la digitalisation et la saisie des données géographiques et des données alphanumériques associées.

**Encadré 16 - Schéma d'un SIG**

**Etape 2. le stockage et la gestion de la base de données géographiques:** Cette fonction inclut la mise en mémoire des données sous forme alphanumérique, la programmation des mécanismes de rappel pour servir les besoins des étapes suivantes, telles que les requêtes des utilisateurs. Il s'agit d'organiser les données (création d'index) pour en faciliter l'accès à partir de critères géographiques, thématiques, topologiques, etc.

**Etape 3. l'exploitation et l'analyse des données :** Un large spectre de techniques est nécessaire pour la transformation et l'utilisation des données dans le cadre d'une problématique spécifique. L'utilisateur procède soit par « interrogation », c'est-à-dire qu'il pose une question directe au SIG, soit par « extraction » lorsqu'il désire constituer un fichier partiel pour un usage particulier. De plus, le SIG permet la combinaison de données, l'identification de corrélations entre certaines variables dans le but de mettre en exergue les caractéristiques d'une zone géographique.

**Etape 4. la restitution des résultats :** Le SIG comprend différents systèmes graphiques permettant l'organisation des données dans des modèles, tels que des cartes, des graphiques, des images, des projections de séquence, etc. Ces derniers facilitent la communication et permettent une utilisation conviviale des données.

## ● Application aux Fonds structurels

Cet outil a un grand potentiel d'intégration et de synthèse des informations. Sa capacité à intégrer la dimension territoriale est pertinente pour les Fonds structurels. Elle permet de centraliser les résultats en provenance des offices statistiques, des travaux de suivi et des enquêtes d'évaluation effectués sur le territoire éligible.

Un SIG permet aussi une visualisation aisée des informations relatives au programme. Cette particularité peut rendre des services importants lors de la présentation des résultats aux divers comités et groupes de travail concernés.

Ce potentiel de synthèse facilite l'illustration de la cohérence des mesures mises en œuvre dans le cadre des programmes, ce qui est un atout lors de la publication de documents de synthèse.

## ● Points forts et limites d'utilisation

- Les SIG ont surtout une faculté « didactique » : les résultats peuvent être restitués de manière conviviale, ce qui contribue à la sensibilisation des acteurs. Pourtant le langage cartographique n'est pas univoque ; les signes et les légendes sont des codes qui varient en fonction des concepteurs et qui nécessitent d'être explicités.
- Outre le fait que la création d'une base de données géographique complète s'avère souvent coûteuse en temps et en argent, l'utilisation des SIG peut se révéler complexe. En effet, les données existantes sont souvent collectées sur des entités spatiales différentes (découpages administratifs en communes, cantons, etc.) ce qui nécessite de fastidieuses transformations ou parfois même une nouvelle collecte. Il arrive également que les données existantes aient été stockées sur des systèmes d'information dont le manque de compatibilité en empêche l'utilisation.
- Avant d'être intégrées dans un SIG, les données issues du suivi doivent être traitées par le biais de techniques statistiques parfois lourdes.
- Dans la pratique, les SIG permettent le traitement d'un nombre considérable de données. Ce support informatique a également l'avantage de permettre la transmission des données d'un utilisateur à l'autre. Le caractère centralisateur de ce support facilite la mise à jour des données et permet la manipulation d'informations géographiques à des échelles variées (de la parcelle à la région, etc.).

- Les SIG sont également utiles pour d'autres développements, notamment lorsqu'ils sont combinés avec d'autres techniques d'analyse, et particulièrement l'analyse factorielle.
- L'expérience tend à montrer que l'outil peut susciter des conflits entre responsables territoriaux en raison de la lecture partisane qu'ils font des cartes produites.

● **Pour en savoir plus...**

Collet, C. (1992), *Système d'information géographique en mode image*, coll. « Gérer l'environnement », Lausanne : Presses polytechniques et universitaires romandes. 186 p.

Ouvrage simple mais très efficace pour l'initiation aux SIG.

Fisher, M. et Nijkamp, P. (1993), *Geographic Information Systems, Spatial Modelling and Policy Evaluation*, Berlin : Springer-Verlag. 280 p.

Ouvrage réunissant des articles de fond avec une orientation spécifique sur les applications à l'évaluation.

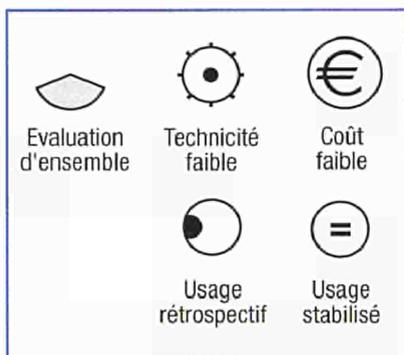
Martin, D. (1996), *Geographic Information Systems, socioeconomic applications*, London : Routledge.

Ouvrage d'introduction aux SIG avec une orientation spécifique sur les applications socio-économiques.

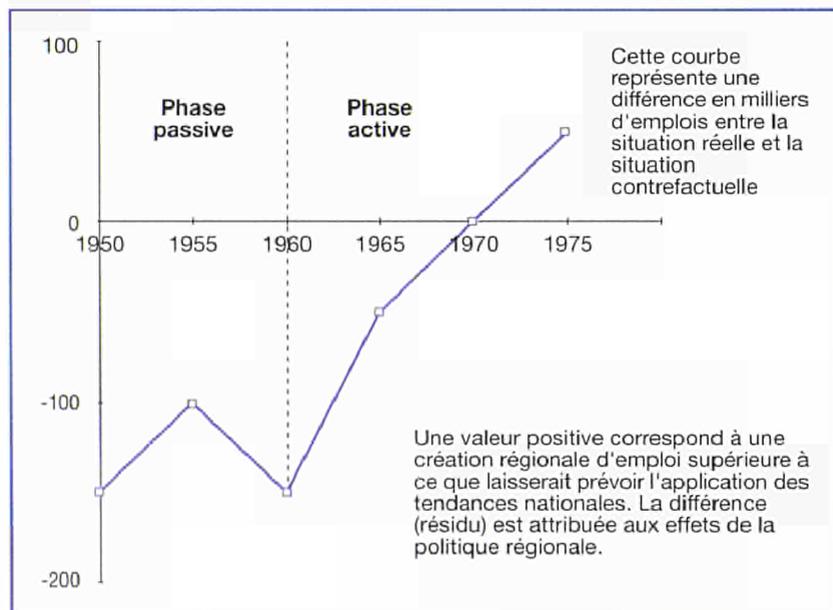
## ■ Analyse structurelle - résiduelle (*Shift-Share*)

L'analyse structurelle-résiduelle est utilisée depuis les années soixante pour l'évaluation des politiques régionales. Il s'agit d'une technique d'évaluation *top-down*, car elle projette les tendances observables au niveau d'un pays sur l'économie d'une région. Le point de départ de l'analyse structurelle-résiduelle

n'est pas constitué par l'étude des conséquences d'un programme, mais par les tendances observées aux niveaux national et régional. L'analyse structurelle-résiduelle construit une base de comparaison de ce qui se serait passé en l'absence du programme à l'aide d'un raisonnement qui ne porte que sur des grandeurs macroéconomiques.



**Exemple : L'impact sur l'emploi des politiques de développement régional en Ecosse.**



Le cas de l'Ecosse est un exemple classique d'analyse structurelle-résiduelle. Il est d'autant plus parlant qu'il présente deux phases bien distinctes au niveau de la politique régionale : une phase passive au cours des années cinquante, pendant laquelle il n'existait pas de politique régionale, suivie d'une phase plus active pendant les années soixante où le gouvernement britannique a abondamment aidé le développement de l'Ecosse. L'application de l'analyse structurelle résiduelle à l'évaluation de la politique régionale britannique en Ecosse a permis de dresser le diagramme ci-dessus.

On voit que, dans les années soixante-dix, l'emploi écossais tend à augmenter plus vite que ne le laisserait supposer l'application des tendances nationales. Ceci permet de conclure à un impact positif des politiques régionales sur l'emploi de cette région.

Source : Diamond D., Spence N. (1983), *Regional Policy Evaluation. A Methodological Review and the Scottish Example*, Gower House : Ashcroft.

### ● A quoi sert l'outil ?

L'analyse structurelle-résiduelle permet de donner une indication sur les effets d'une politique dans une région donnée. L'effet de la politique en question est déterminé en comparant la situation réelle de la région (*policy-situation*), avec une reconstitution de ce qu'aurait été la situation régionale si elle avait subi les tendances nationales, et seulement les tendances nationales (situation contrefactuelle ou *policy-off situation*).

### ● Dans quel cas l'utiliser ?

L'analyse structurelle-résiduelle est destinée à évaluer les effets d'une politique régionale au niveau macroéconomique. Utilisée dans le cadre d'une analyse rétrospective, cette technique s'appuie sur des indicateurs statistiques simples et ne nécessite aucune enquête auprès des groupes cibles du programme.

### ● Comment le mettre en œuvre ?

L'analyse structurelle-résiduelle se déroule en deux étapes : la construction de la situation contrefactuelle et la comparaison. Ces deux étapes sont illustrées avec un exemple d'analyse en termes d'emploi (c'est l'utilisation la plus fréquente de l'analyse structurelle-résiduelle).

**Etape 1. Construction de la situation contrefactuelle** : elle repose sur une modélisation simple. On applique les tendances de l'emploi national à la structure des emplois régionaux. La formulation de cette procédure est la suivante :

$$IM_1 = IM_0 + NS$$

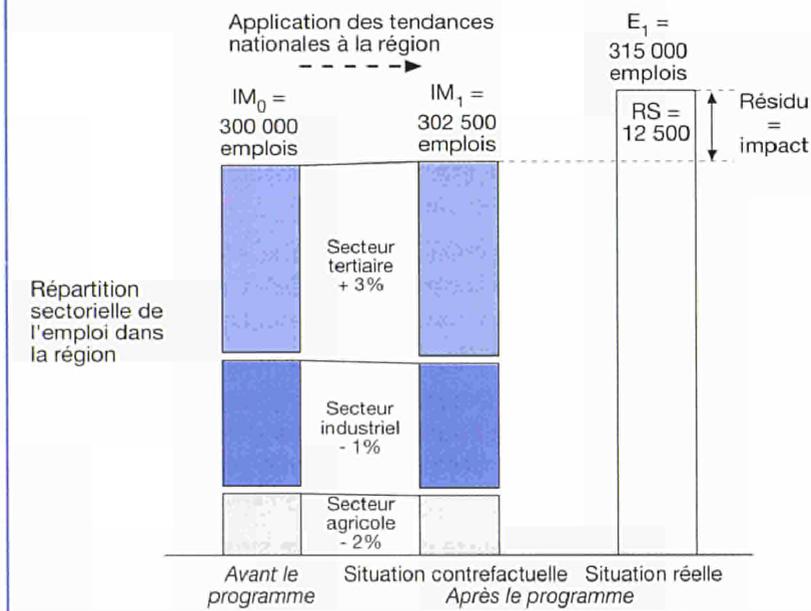
où : IM caractérise la distribution sectorielle de l'emploi de la région (*industry mix*) et NS est la variation de l'emploi qui se produirait si chaque secteur régional évoluait au rythme observé pour ce secteur au niveau national (*national share component*).

### Encadré 17 - Exemple d'une analyse résiduelle-structurale de l'emploi régional

Le nombre d'emploi dans une région dans laquelle les pouvoirs publics interviennent en faveur de l'emploi s'élève à 300 000 ( $IM_0$ ) dont 100 000 emplois industriels, 150 000 dans le tertiaire et 50 000 dans l'agriculture et l'agro-alimentaire en période  $t = 0$ .

Au niveau national l'emploi évolue dans les secteurs industriel, tertiaire et agricole aux rythmes -1%, 3% et -2% respectivement entre  $t = 0$  et  $t = 1$ . Dans l'hypothèse que l'emploi régional évolue au même rythme que l'emploi national, l'augmentation de l'emploi régional est de 2 500 (NS). Le niveau d'emploi dans la situation contrefactuelle s'élève donc à 302 500 ( $IM_1$ ).

Par contre, pendant la période  $t = 0$  à  $t = 1$ , l'emploi dans la région aidée a progressé de 5%, soit 15 000 emplois. Le nombre d'emplois se chiffre alors à 315 000 ( $E_1$ ). Selon cette analyse, le nombre d'emploi créé grâce à l'intervention est de 12 500 (RS).



Etape 2. Comparaison pour déterminer l'impact de l'action : C'est à ce niveau que peut avoir lieu la comparaison entre la situation contrefactuelle, obtenue au cours de l'étape précédente ( $IM_t$ ) et la situation réelle ( $E_t$ ). La différence entre ces deux situations (le résidu) permet la mise en évidence des impacts de l'action politique.

$$RS = E_t - IM_t$$

On admet ici que l'impact du programme équivaut au résidu (*regional shift component*).

### ● Application aux Fonds structurels

L'analyse structurelle-résiduelle ayant été conçue pour évaluer les effets des politiques régionales, elle est, en principe, directement applicable aux Fonds structurels. A ce titre elle peut servir dans les évaluations ex post des programmes de développement régional, à chaque fois que l'accent est mis sur l'étude d'un impact particulier tel que l'emploi ou la valeur ajoutée. L'analyse structurelle-résiduelle est susceptible d'être utilisée pour les programmes mis en œuvre dans le cadre du FEDER en raison de leur nature régionalisée. Etant donné que cette technique est peu exigeante, elle présente l'avantage d'être facile d'emploi. Les données statistiques qu'elle requiert sont disponibles pour la plupart des régions européennes, notamment par le biais de données d'emploi et de valeur ajoutée de niveau NUTS 1 et 2 mises à disposition par EUROSTAT.

### ● Points forts et limites d'utilisation

- L'analyse structurelle-résiduelle permet d'aboutir à des conclusions quantifiées en matière d'effets des politiques.
- La conduite d'une analyse structurelle-résiduelle n'est pas standardisée (dates de début et fin, décompositions sectorielles, référence nationale) ; il en résulte des difficultés à comparer les résultats d'études conduites sur des programmes différents.
- Il s'agit d'un outil peu coûteux, car il peut s'appuyer sur des données statistiques très répandues. Il suffit de disposer d'un indicateur d'impact (emploi ou valeur ajoutée), avec une ventilation sectorielle aux niveaux régional et national.
- La logique de la mise en œuvre d'une analyse structurelle-résiduelle est souvent simplifiée à l'excès. La plupart du temps, les variations observées au niveau de l'emploi régional peuvent être expliquées par d'autres

facteurs que la simple projection des tendances nationales. Parmi ces facteurs on peut citer le dynamisme de la région à l'exportation, la productivité, la qualification de la main d'œuvre, la pression démographique, les flux d'immigration, l'intensité en Recherche et Développement. Il n'est pas crédible d'attribuer l'effet de tous ces facteurs à la seule politique régionale, à moins que celle-ci ait une emprise vraiment importante sur tous les agents économiques régionaux.

- Dans l'analyse structurelle-résiduelle l'impact est constitué par le résidu, qui est donné par la différence entre la situation réelle et la situation contrefactuelle issue de l'application des tendances nationales. Or ce résidu peut être attribué à la fois au programme en cours d'évaluation et à toutes les autres actions politiques nationales et européennes ainsi qu'à la trajectoire passée de la politique régionale. En théorie, l'analyse structurelle-résiduelle s'applique mieux si on évalue une politique régionale de très grande ampleur à la suite d'une période marquée par l'absence de politique régionale.

#### ● Pour en savoir plus

Bartels, C.P.A. (1982), Measuring Effects of Regional Policy. An introduction, *Regional Science and Urban Economics* 12, 41p.

Une revue d'ensemble des techniques permettant d'analyser l'impact des effets sur l'emploi. L'analyse shift-share y est abordée de façon critique. Très accessible.

Steevens, B.H., Moore, C.L. (1980), 'A Critical Review of the Literature on Shift-Share as a Forecasting Technique', in : *Journal of Regional Science*, 20(4), 420-435.

Un article très formalisé qui présente les principaux modèles shift-share et leurs possibilités d'utilisation pour la prospective.

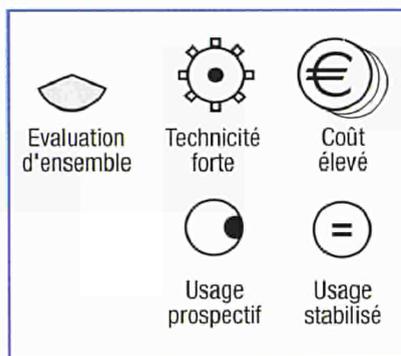
Tervo, H., Okko, P. (1993), 'A Note on Shift-Share Analysis as a method of estimating the employment effects of regional economic policy', in : *Journal of Regional Science*, 23(1), 115-121.

Un article critique vis-à-vis de l'utilisation de la technique shift-share pour évaluer les effets sur l'emploi régional. Basé sur des démonstrations formalisées.



## ■ Modèle entrées-sorties

À l'origine, les matrices entrées-sorties (*Input-Output*) ont été développées pour analyser les échanges entre les différents secteurs au sein d'économies relativement fermées (faiblement exportatrices et importatrices). Elles fournissent un éclairage sur la manière dont le système productif satisfait la demande finale, laquelle se décompose entre consommation, investissement et exportations. Une matrice entrées-sorties représente les liens entre toutes les ressources dont dispose une économie et toutes les consommations finales, c'est-à-dire l'ensemble des utilisations qu'elle en fait. La matrice peut revêtir des formes simplifiées (trois secteurs : industrie, services, agriculture) mais aussi plus complexes (plus de 500 branches). Il s'agit de l'une des seules techniques applicables à l'évaluation des impacts sectoriels des interventions structurelles, à cause de la finesse de découpage du tissu productif qui est permise. Une matrice d'entrées-sorties peut être assimilée à un modèle macroéconomique très simplifié quant aux mécanismes économiques représentés, mais très détaillé du point de vue sectoriel.



### Exemple : L'impact des Fonds structurels sur les pays et régions d'Objectif 1

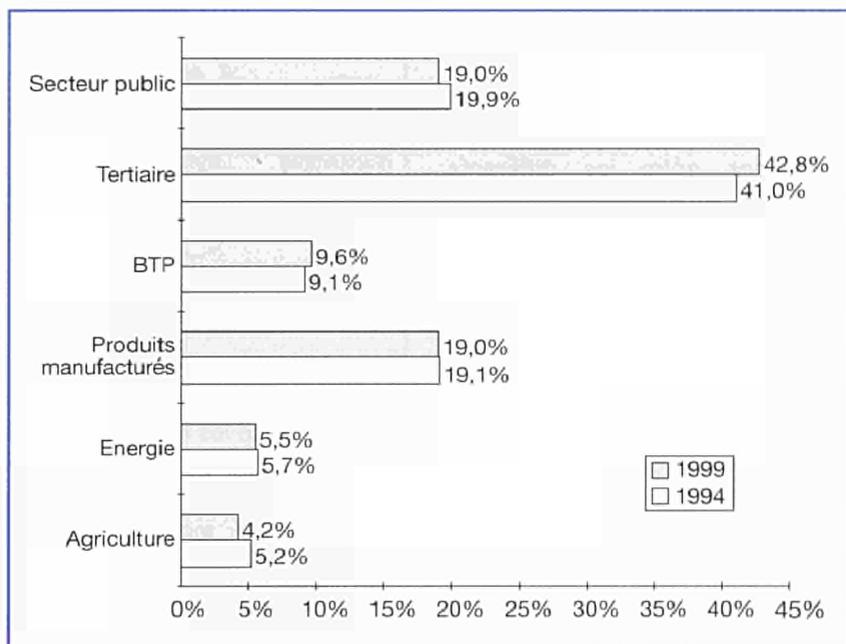
Une équipe d'universitaires allemands a adapté l'analyse entrées-sorties à l'étude de l'impact des Fonds structurels sur les pays et régions Objectif 1, à l'aide de matrices fournies par EUROSTAT pour l'année de référence 1985 et des projections pour les années 1991, 1994 et 1999. Dans cet exemple l'équipe a cherché à estimer l'impact sur la croissance économique globale et sur la répartition sectorielle. Voici les résultats les plus représentatifs de cette étude :

Sans les Fonds structurels, la croissance économique moyenne des zones d'Objectif 1 aurait été inférieure de 0,7 points absolus.

En 1999, les Fonds structurels devraient induire un accroissement de 1,9% de l'emploi dans les zones d'Objectif 1.

L'analyse a été différenciée selon les pays ; les effets sur la croissance et sur l'emploi seraient les plus forts en Grèce et au Portugal.

Entre 1994 et 1999, les modifications de la structure sectorielle de la production (en termes de PIB) de ces régions peuvent être représentées comme suit :



Source : J. Beutel, Université de Konstanz

### ● A quoi sert l'outil ?

Les matrices entrées-sorties sont utilisées principalement dans le domaine de la prévision économique, où elles servent à valider des scénarios de politiques en prenant appui sur la structure technique du pays et l'état de la demande finale. Dans le cadre d'une évaluation on peut les utiliser avec ou sans politique, de la même manière qu'on utilise un modèle macroéconomique (voir fiche correspondante). L'outil permet d'estimer des impacts différenciés selon les différentes branches d'une économie. Parmi les nombreuses applications des matrices entrées-sorties à l'évaluation des programmes de développement, on s'attardera sur l'exemple cité ci-dessus car il a l'avantage de présenter des conclusions harmonisées à l'échelle européenne.

Le modèle a été produit par l'Université de Konstanz et EUROSTAT. Il a été employé pour estimer les effets nets des principaux programmes financés par les Fonds structurels.

### ● Dans quel cas l'utiliser ?

De façon générale cet outil s'emploie à l'échelle nationale, car c'est à ce niveau que sont recueillies les données statistiques nécessaires à la construction des matrices. Des matrices régionales existent et se prêtent aussi à l'analyse entrées-sorties. Mais leur valeur statistique est moindre, car il est difficile de suivre des flux de biens qui ne sont pas soumis à déclaration douanière.

### ● Comment le mettre en œuvre ?

Les lignes qui suivent décrivent la mise en œuvre de l'outil depuis sa construction jusqu'aux simulations effectives dans le cadre de l'évaluation. En réalité on construit rarement une matrice d'entrées-sorties pour les besoins d'une évaluation. La mise en œuvre de l'outil commence donc en pratique à l'étape 4.

**Etape 1. La construction d'une matrice des transactions :** on décrit les liens entre les différentes branches d'une économie. En lignes de cette matrice, on trouve les sorties (*outputs*) qui concernent les ressources fournies par un secteur donné à chacun des secteurs d'activité et à la demande finale. En colonnes de cette matrice, on trouve les (*inputs*) en provenance des différents secteurs, c'est-à-dire ce que consomme un secteur pour assurer sa production.

Les valeurs prises dans cet exemple sont monétaires. Ainsi, par exemple, la branche Agriculture fournit une valeur de 150 à la branche Industrie pour assurer sa production tandis qu'une valeur de 100 est destinée à la demande finale. La branche Industrie, quant à elle, consomme pour une valeur de 150 en produits intermédiaires en provenance de la branche Agriculture pour assurer sa production, cette dernière conduisant à une création de valeur ajoutée de 320. En additionnant les consommations intermédiaires de produits agricoles effectuées par les différents secteurs et la demande finale pour la branche agricole, on obtient le total des utilisations de la production agricole. La demande finale peut elle-même être décomposée entre la consommation finale, l'investissement et les exportations. La sommation des entrées intermédiaires consommées par la branche agricole pour assurer sa production et de la valeur ajoutée agricole permet d'obtenir la production agricole. Les totaux en ligne et en colonnes sont donc égaux.

Sorties Entrées	Agri- culture	Industrie	Services	Consom- mations inter- médiaires	Demande finale	Total des utilisa- tions
Agriculture	50	150	50	250	100	350
Industrie	150	150	100	400	320	720
Services	100	100	200	400	600	1000
Produits inter- médiaires	300	400	350	1050	1020	2070
Valeur ajoutée	50	320	650	1020		
Production	350	720	1000	2070		

**Etape 2. La construction d'une matrice des coefficients techniques :** Elle est obtenue à partir de la matrice des transactions. Les coefficients techniques ne sont rien d'autre que des valeurs issues de la matrice des transactions divisées par le total des productions. On voit ainsi la part relative de valeur créée par chaque branche pour l'ensemble des autres branches. Les coefficients techniques servent à fournir des indications sur la structure technique de l'économie modélisée ; ils permettent de voir quels sont les secteurs à forte valeur ajoutée, à forte productivité, à fort taux d'exportation. Ces coefficients sont artificiellement figés pour les étapes suivantes.

**Etape 3. La construction d'une matrice inversée :** Elle est le résultat d'une transformation matricielle qui permet de calculer des coefficients multiplicateurs, ces coefficients résumant l'ensemble des effets indirects.

**Etape 4. Les hypothèses d'impact primaire :** L'impact primaire du programme fait l'objet d'une estimation quantitative, convertie sous la forme d'une augmentation de demande finale (augmentation de dépense publique, augmentation de consommation, augmentation d'investissement d'infrastructure, etc.). L'hypothèse d'augmentation de demande comporte une décomposition sectorielle de façon à pouvoir être introduite dans la matrice.

**Etape 5. L'estimation de l'impact total :** L'impact total est calculé à partir des hypothèses d'impact primaire et de la matrice inversée. A partir d'un

impact primaire sur un secteur, l'outil permet d'estimer l'ensemble des effets indirects sur tous les secteurs, notamment en termes de consommation, d'importations, de production et d'emploi.

### ● Application aux Fonds structurels

Parmi les différentes applications des modèles entrées-sorties, on privilégie la présentation du modèle cité plus haut, connu sous le nom de « modèle de Beutel ». Cet exemple est particulièrement intéressant car le modèle a été spécialement construit pour l'évaluation et il présente l'intérêt de fournir des estimations harmonisées à l'échelle européenne.

Le modèle de Beutel a principalement recours à une série de tableaux entrées-sorties harmonisés pour l'année 1985, qui a été mise en place par EUROSTAT pour les économies de l'Union européenne. Ces tableaux entrées-sorties sont complétés par des tableaux régionaux, notamment dans le cas de l'Espagne et de l'Italie. Les pays couverts sont le Portugal, la Grèce, l'Irlande, l'Espagne, l'Italie et le Royaume-Uni (Irlande du Nord). L'Allemagne est par ailleurs incluse dans la période de planification 1994-1999. Etant donné que la base de données 1985 est clairement dépassée pour toute application pratique utile, il faut avoir recours à une méthode d'actualisation. Beutel a développé à cet effet une technique de projection qui a été adoptée par EUROSTAT (voir approfondissement en fin de section).

L'impact primaire des financements communautaires est estimé en divisant les dépenses totales en trois catégories : investissements immobiliers, investissements en biens d'équipement et les salaires. On établit ensuite une comparaison entre le développement de l'économie avec et sans financements communautaires. En ce qui concerne l'évaluation des interventions dans le cadre de l'Objectif 1, l'approche Beutel fournit des estimations portant sur les variables macroéconomiques centrales suivantes :

- L'emploi : La comparaison entre les cas avec et sans intervention est établie non seulement en termes d'emploi total mais une distinction est également faite entre les salariés d'une part et les travailleurs indépendants de l'autre.
- Le PIB et son taux de croissance : Il est possible d'établir une comparaison entre les taux de croissance avec et sans intervention. Une autre comparaison intéressante est faite entre la croissance annuelle

avec intervention et les cas où les interventions sont interrompues pendant une année spécifique.

- **Formation du capital fixe brut** : Dans cette variable sont présentés les changements induits dans les investissements par les interventions et son taux de croissance annuel.
- **Valeur ajoutée sectorielle** : La répartition sectorielle et les évolutions induites dans 25 secteurs économiques par les interventions sont calculés chaque année. Il existe une présentation standard de l'évolution des 6 secteurs qui connaissent une transformation structurelle : agriculture, sylviculture et pêche, combustibles et énergie, BTP, services privés, services publics.
- **Exportations et importations** : L'impact des interventions sur les exportations et les importations est calculé chaque année et subdivisé en exportations vers et importations en provenance des pays membres de la Communauté d'une part, et en exportations vers et importations en provenance de pays tiers, d'autre part.

### ● **Points forts et limites d'utilisation**

- La contribution la plus intéressante des matrices entrées- sorties concerne l'impact sur la répartition sectorielle et sur le commerce extérieur. Les fuites dues aux importations du pays bénéficiaire en provenance des pays membres de l'Union européenne et des pays tiers peuvent revêtir une certaine importance pour la prise de décision politique de l'Union européenne. De même, la connaissance de l'impact de la demande sectorielle peut se révéler importante, notamment si l'on estime que certains secteurs doivent être stimulés pour accélérer le taux de croissance.
- La valeur de cette approche réside avant tout dans sa large couverture et dans son traitement cohérent des principales économies de l'Union européenne, bénéficiaires d'une aide dans le cadre de l'Objectif 1.
- L'intérêt de l'outil est limité à l'estimation des effets sur la demande. Ainsi, il ne tient pas compte de l'un des effets les plus importants visés par les interventions dans le cadre de l'Objectif 1, à savoir les effets durables sur le potentiel de production. La plupart des effets sur l'offre qui sont susceptibles d'accroître de manière durable le taux de croissance des régions assistées et de leur permettre de rattraper leur retard sont totalement ignorés (création de nouvelles capacités de production,

amélioration de la qualification de la main d'œuvre, création d'infrastructures, productivité dans l'ensemble de l'économie, diffusion du progrès technique, intensité en technologie accrue dans la production). Ce sont tous ces effets sur l'offre qui transforment de manière durable et irréversible la capacité productive.

- On observe une tendance à la surestimation de l'impact des Fonds structurels sur la demande car l'impact des interventions n'est pas modéré par les effets de rétroaction, notamment en ce qui concerne les changements de prix émanant des marchés des produits, des marchés financiers et celui du travail.

### ● Pour en savoir plus

Brauers, W. K. (1995), *Prévisions économiques à l'aide de la méthode Entrées - Sorties*, Paris : Economica, 111 p.

Un ouvrage d'introduction à l'analyse entrées-sorties, destiné à un large public. Met en relief les principales techniques, retrace les débats qui ont lieu à ce sujet, donne des exemples d'application. Va à l'essentiel.

Miller, R. E. et Blair, P. D. (1985), *Input-Output Analysis, Foundations and Extensions*, New-Jersey : Prentice-Hall, 464 p.

Destiné aux initiés, cet ouvrage décrit les fondements de l'analyse entrées-sorties et offre un panorama des différents types d'analyses qui peuvent être réalisées à partir de cette technique. Chacun pourra aisément sélectionner les articles qui satisferont leur curiosité.

Kurz, H. D., Dietzenbacher, E. et Lager, C. (1998), *Input-Output Analysis*, Cheltenham : Edward Elgar, 3 Volumes, 1504 p.

Cet ouvrage retrace l'évolution de l'analyse entrées-sorties en reprenant une collection d'articles publiés depuis les origines de ce type d'analyse.

### ● Approfondissement : La construction des matrices du modèle de Beutel

Dans le modèle Beutel, les matrices entrées-sorties d'EUROSTAT sont actualisées en s'appuyant sur les prévisions économiques de la Direction générale des affaires économiques et financières de la Commission européenne. Ces prévisions sont dérivées du modèle QUEST (voir la fiche sur les Modèles macroéconomiques). On a également recours aux

prévisions de 25 activités de production, préparées par ERECO (European Economic Research and Advisory Consortium, composé de l'Ifo-Institut/Munich, BIPE/Paris, Prometeia/Bologne, NEI/Rotterdam, WIFO/Vienne et Cambridge Econometrics/Cambridge).

Les données utilisées pour la mise à jour du modèle comprennent le tableau entrées-sorties de l'année de référence, y compris la matrice des importations, la projection des taux de croissance de la valeur ajoutée par secteur (agriculture, industrie, services) et la projection des composantes de la demande finale (consommation, investissements, exportations, importations). Le tableau entrées-sorties actualisé est estimé par une procédure d'itération qui fournit des estimations sur la consommation intermédiaire des biens et des services importés, sur la composition structurelle de la demande finale par produit, sur la production intérieure par produit et sur les importations par produit.

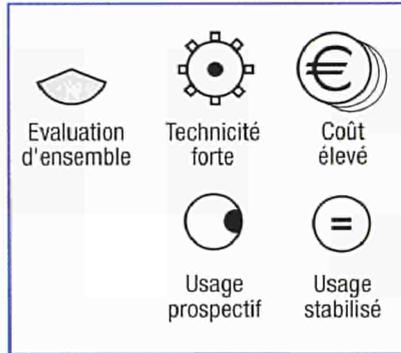
Cette technique tente d'éviter les changements arbitraires des coefficients d'entrée en traitant toutes les activités sur un pied d'égalité. Apparemment, en dehors de ses besoins limités en données et du fait qu'elle a recours à des sources officielles, cette technique présente également l'avantage de produire des mises à jour robustes à faible coût.

## ■ Modèle macro économique

Le but de la modélisation est de représenter le fonctionnement de l'ensemble d'une économie et des différents marchés qui la composent (ex : marché des produits, marché du travail et marché de la monnaie, etc.). Un modèle schématise le fonctionnement d'une économie en équilibre, mais il est aussi capable de simuler l'évolution de cette économie à la suite d'un choc exogène (ex : variation du prix du pétrole, variation des taux de change, prélèvement fiscal), le rajustement des variables économiques modélisées et l'atteinte d'un nouvel équilibre des marchés.

Les modèles macroéconomiques sont utilisés par les administrations pour la préparation de leurs budgets et pour l'étude des choix politiques lorsque les enjeux économiques sont de grande ampleur. Certains organismes de prévision, publics et privés, utilisent également des modèles pour des prévisions économiques à court et moyen terme.

L'utilisation des modèles à des fins d'évaluation est beaucoup plus récente et relève encore en partie du domaine de l'innovation méthodologique, ce qui n'empêche pas de rencontrer des réussites exemplaires telles que celle de l'exemple qui suit.



### Exemple : L'évaluation de l'impact probable de l'aide économique européenne à la République d'Irlande

L'évaluation de l'impact probable de l'aide économique européenne à la République d'Irlande (CCA 89-93) a été réalisée par l'« Economic and Social Research Institute » de Dublin (ESRI) à l'aide d'un modèle macroéconomique que cet institut avait construit dans le cadre du programme HERMES. L'évaluation devait permettre d'estimer l'impact probable de l'aide européenne sur l'économie irlandaise dans son ensemble en particulier sur des variables telles que le PIB par habitant ou l'endettement public.

Le modèle HERMES avait été construit pour simuler le fonctionnement de l'économie irlandaise dans un contexte de crise de l'énergie. Il a été légèrement modifié à l'occasion de l'évaluation pour mieux prendre en compte les mécanismes d'impact

des programmes européens. Le modèle utilisé pour l'évaluation contient plus de 600 équations qui représentent des hypothèses portant sur le comportement des agents économiques concernés directement ou indirectement par l'aide européenne. Les conclusions de l'évaluation ont été les suivantes :

En l'an 2000, l'impact de l'intervention communautaire est estimé à 2,7% du PNB irlandais.

Néanmoins, cet impact ne serait pas suffisant (à lui seul) pour permettre à l'économie irlandaise de rattraper la moyenne européenne.

Il est intéressant de souligner que cette évaluation a fait l'objet de débats politiques lors du référendum irlandais sur le traité de Maastricht et qu'elle a servi d'argument dans le cadre de la renégociation de l'aide européenne pour la période 1994-99. On voit donc que malgré la fragilité inhérente à toute estimation, l'évaluation macroéconomique a servi de base à des débats constructifs.

Source : Bradley J. et Fitzgerald J., *Medium-Term Review 1989-1994*, Economic and Social Research Institute, Dublin, 1989.

### ● A quoi sert l'outil ?

Un modèle macroéconomique représente de façon simplifiée les mécanismes d'un système économique. Employé en évaluation, l'outil permet de simuler un scénario avec intervention publique et un scénario sans intervention publique, et donc d'estimer les conséquences de l'intervention par différence entre les deux scénarios. Les modèles sont entièrement quantifiés et fournissent donc une estimation de l'ampleur des impacts sur toutes les variables macroéconomiques qui ont été modélisées.

Il va de soi qu'un modèle ne peut pas tout prévoir, ni même garantir une bonne fiabilité de ses simulations. L'outil reste cependant pertinent si on désire estimer des effets macroéconomiques car aucune technique ne peut prétendre à une meilleure fiabilité. Il fournit des valeurs de référence sur l'impact de mesures spécifiques difficiles à obtenir selon une méthode alternative en raison de la complexité des mécanismes en action (par exemple, impact sur l'emploi et l'investissement privé). Ces valeurs de référence permettent d'alimenter les discussions et de concrétiser l'ampleur réelle des enjeux. Finalement, l'apport de l'outil est plus souvent de nature pédagogique. L'emploi du modèle et l'examen des résultats obtenus obligent les parties prenantes de l'évaluation à s'intéresser aux mécanismes macroéconomiques et à mieux les comprendre.

### ● Dans quel cas l'utiliser ?

L'estimation des effets peut intervenir *ex ante*. Dans ce cas, on compare deux scénarios prévisionnels (avec et sans). C'est l'usage le plus fréquent en évaluation. Bien que cela n'ait presque jamais été fait, on peut aussi employer un modèle dans une situation *ex post*. Dans ce cas, la comparaison porte, d'une part, sur l'évolution économique passée telle que le modèle l'a recalculée de façon approchée et, d'autre part, sur un scénario sans intervention publique, reconstitué *ex post* par le modèle (simulation de la situation contrefactuelle). Dans le cadre de l'évaluation intermédiaire, l'outil peut être exploité pour guider les opérations de recalibrage éventuel du programme (par exemple, répartition des fonds entre aides à l'investissement, développement des infrastructures et soutien à la R&D).

### ● Comment le mettre en œuvre?

La mise en œuvre d'une évaluation utilisant un modèle macroéconomique ne doit être envisagée que si l'on dispose des connaissances, des compétences et d'un délai suffisant. Les étapes sont en effet nombreuses et très techniques.

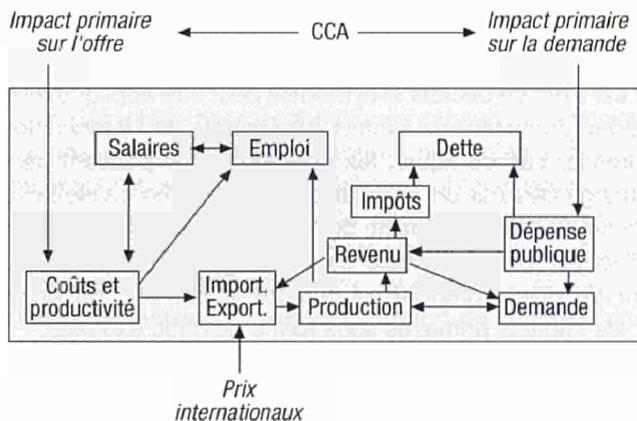
**Etape 1. Adaptation du modèle à l'évaluation :** La construction d'un modèle est un investissement lourd qui nécessite souvent plusieurs années. On ne construit donc jamais un modèle macroéconomique pour les besoins spécifiques d'une évaluation. Il peut arriver que plusieurs modèles soient disponibles et que la première étape consiste à en choisir un, mais ce cas de figure est exceptionnel. Généralement, le choix de l'équipe d'évaluation est dicté par l'existence et la disponibilité d'un seul modèle. Le problème est alors de décider si le modèle peut être appliqué tel quel ou s'il doit être adapté pour mieux simuler les impacts de l'intervention évaluée. Dans le premier cas de figure, les adaptations ne peuvent généralement être que marginales car on ne peut pas remettre en cause les conditions d'équilibre du modèle sous peine de devoir le reconstruire entièrement. En général, l'adaptation du modèle se limite à quelques modifications qui permettent de mieux prendre en compte le contenu du programme et d'intégrer les impacts primaires sous forme de choc exogène.

**Etape 2. Calage du modèle :** Avant toute utilisation, et particulièrement s'il a été modifié, le modèle doit être ajusté aux dernières évolutions constatées et aux dernières statistiques connues. Un modèle est au point

lorsqu'il est capable de rendre compte correctement de ce qui s'est passé dans les cinq à quinze années précédentes. En pratique, le modèle devra être réestimé régulièrement pour tenir compte des révisions intervenues dans les informations de base et permettre ainsi un « calage » sur les données les plus récentes.

**Etape 3. Estimation des impacts primaires :** L'exemple qui suit (Encadré 18) schématise les mécanismes macroéconomiques pris en compte dans un modèle. On voit par exemple que l'emploi dépend de la production des entreprises, de la productivité et du niveau des salaires. Ces trois variables sont endogènes, c'est à dire internes au modèle. Pour évaluer l'effet d'une intervention publique, par exemple son effet sur l'emploi, il faut considérer que cette intervention publique introduit des changements exogènes dans le modèle (on emploie le terme choc exogène). L'équipe d'évaluation doit donc sélectionner quelques variables ou coefficients qui font partie du modèle et qui vont être directement affectés par l'intervention publique. Ces grandeurs constituent des « ponts » entre l'intervention évaluée et le modèle. Dans l'exemple qui suit, ce sont la dépense publique et la productivité. Les impacts primaires des interventions des Fonds structurels peuvent également intervenir sous forme d'investissements dans les structures ou en capital humain. L'équipe d'évaluation doit ensuite estimer les impacts primaires de l'intervention évaluée (impact microéconomique) et les introduire dans le modèle.

### Encadré 18 - Impact primaire et mécanismes macroéconomiques



**Etape 4. Simulations :** Dans un contexte d'évaluation, on exploitera toujours le modèle au moins deux fois. La première fois pour caler la situation avec intervention (cas d'une évaluation ex post) ou pour simuler cette situation (cas d'une évaluation ex ante), et la seconde fois pour simuler la situation sans intervention. Les impacts de l'intervention sont estimés par différence entre les deux simulations.

**Etape 5. Présentation des résultats :** La dernière phase est celle de la mise en forme et de la présentation des résultats. Compte tenu de la complexité de l'outil et des très nombreuses variables qu'il est possible de quantifier, il est nécessaire de réaliser un travail spécifique pour présenter les résultats de façon simple et facilement accessible. Cette phase doit donc être considérée comme une des plus importantes de la mise en œuvre de l'outil.

#### ● Application aux Fonds structurels

L'évaluation macroéconomique d'un programme des Fonds structurels est conditionnée par l'existence préalable d'un modèle pour le territoire éligible. Les modèles macroéconomiques disponibles se limitent généralement aux modèles nationaux pour des raisons de disponibilité des données statistiques et de coût de modélisation. Les modèles construits à l'échelle régionale ou locale sont rares, notamment parce que les flux économiques sont très vite dilués hors du territoire concerné. La modélisation garde cependant du sens pour des petites économies à condition qu'elles soient relativement fermées. C'est pourquoi les économies insulaires sont souvent équipées de modèles macroéconomiques.

L'emploi de modèles est donc généralement possible pour les grands programmes dont la zone éligible correspond aux frontières nationales (Portugal, Irlande, Grèce) ou à de vastes ensembles régionaux (Sud de l'Espagne ou de l'Italie). Sauf exception, les modèles sont inadaptés à des programmes dont la zone éligible est une petite région ouverte.

Par ailleurs, l'emploi d'un modèle a surtout du sens si l'intervention du programme atteint une masse suffisante par rapport à l'économie modélisée. On considère que pour atteindre cette masse critique, l'importance économique du programme doit être au moins de l'ordre de 1% du PIB de l'économie du territoire.

Dans le cadre des interventions structurelles européennes, l'évaluation macroéconomique se limite à quelques expériences dont les deux plus connus font appel aux modèles QUEST et HERMIN qui vont être présentés ci-après.

Dans le cas du modèle QUEST II, les dépenses des Fonds structurels sont considérées à la fois comme une augmentation des dépenses publiques et comme un accroissement du stock de capital national. De ce point de vue, les dépenses engagées sont donc considérées comme une augmentation exogène de la consommation publique (effet de demande), et comme un accroissement de l'investissement public, conduisant à augmenter la capacité de production (effet d'offre).

Ce type d'approche produit une estimation assez grossière des effets d'offre dans la mesure où tout investissement est supposé produire le même impact sur la capacité de production, qu'il soit public ou privé, qu'il intervienne dans le domaine des infrastructures de base, de la R&D ou de la formation.

Les développements ultérieurs du modèle gagneraient à différencier les impacts selon qu'ils prennent leur origine dans telle ou telle intervention publique. Cela présenterait l'avantage de prendre en compte des effets d'offre et de demande des interventions des Fonds structurels d'une manière claire et assez directe. Les effets d'offre seraient simulés séparément et directement via l'impact sur la production potentielle. La combinaison des effets d'offre et de demande indiquerait dans quelle mesure la structure de l'économie est susceptible de tirer profit de l'augmentation de la production potentielle.

Le modèle HERMIN, quant à lui, a été utilisé dans le cadre d'évaluations en Espagne, en Irlande et au Portugal. Comme dans l'expérience réalisée avec le modèle QUEST, les effets de demande des dépenses réalisées dans le cadre des Fonds structurels ont été introduits sous la forme d'augmentations exogènes des dépenses publiques. Pourtant, les évaluations réalisées avec HERMIN ont introduit une innovation intéressante en simulant explicitement les effets d'offre des dépenses publiques, avec une tentative d'estimation séparée pour les effets de trois catégories d'intervention : infrastructures de base, capital humain et environnement productif. Les impacts primaires des Fonds structurels en termes d'offre sont introduits de plusieurs manières au sein du modèle. Ils prennent d'abord la forme d'une augmentation de la productivité qui affecte

un ou plusieurs paramètres de la fonction de production. Dans les versions espagnole et irlandaise, cette productivité est « incarnée » dans le travail (un paramètre). Dans le modèle portugais, c'est la productivité totale des facteurs qui est augmentée (tous les paramètres). L'augmentation de productivité est une fonction de l'investissement généré par l'aide européenne sous forme de capital privé, de capital humain et de capital d'infrastructures publiques. On fait l'hypothèse qu'il y a un rapport constant (appelé élasticité) entre l'augmentation d'investissement et l'augmentation de productivité. Une difficulté réelle de l'évaluation tient au choix de la valeur de ce rapport. Les observations empiriques qui permettraient de l'estimer sont très peu fiables et, pour éviter l'arbitraire total, il faut souvent avoir recours à des études similaires utilisant des données américaines ou internationales. On voit que cette forme d'estimation des effets d'offre repose sur une base dont la crédibilité est fragile.

Une seconde hypothèse sur les impacts primaires est que l'action des Fonds structurels contribue à attirer plus d'investissements étrangers et donc à accroître la compétitivité des exportations. Pour réaliser une approximation de cette chaîne d'effets, on estime leur équivalent en termes d'augmentation de la demande mondiale. Ici encore, on fait l'hypothèse qu'il y a un rapport constant entre l'augmentation des investissements générés par les Fonds européens et l'augmentation des exportations. Le choix d'une valeur pour ce rapport est ici plus problématique encore. D'abord, l'arbitraire est inévitable, étant donné l'absence totale, non seulement pour les pays concernés mais sur le plan international, de travaux empiriques applicables. Plus grave encore, le résultat final semble être très sensible à la valeur choisie pour ce rapport. C'est particulièrement vrai pour l'impact sur le PIB à long terme.

Pour apprécier le potentiel de l'outil HERMIN pour l'évaluation des effets macroéconomiques, il ne faut pas perdre de vue que ce travail est en cours et que nombre des difficultés actuelles sont susceptibles d'être résolues au fur et à mesure que le projet avancera. Le travail des différentes équipes sur l'estimation de l'impact du CCA devra permettre d'affiner les hypothèses d'impact, les paramètres des modèles, la qualité des données, de même que la comparabilité des résultats.

#### ● Points forts et limites d'utilisation

- L'objectif global des Fonds structurels est la cohésion économique et sociale. Cet objectif est essentiellement défini de façon macro-

économique. Le modèle macroéconomique est pratiquement le seul outil qui permette de répondre formellement à la question de savoir si la politique européenne atteint ou non son objectif global.

- Dans le cadre de l'évaluation des programmes socio-économiques, une des limites majeures des modèles est qu'ils ne rendent pas compte des « effets d'offre » de façon crédible. Par exemple, si on cherche à décrire l'effet des nouvelles infrastructures sur la productivité des entreprises, on pourra difficilement trouver une équation qui traduise cet effet et qui soit ajustée de façon convaincante aux statistiques passées.
- Une limite majeure de l'outil tient à la construction des modèles. Cette construction représente un investissement important et demande toujours plus d'un an. En outre, un modèle nécessite une mise à jour continuelle des paramètres sous la supervision de spécialistes ayant un savoir-faire important. L'expérience montre que les instituts de statistique et de prospective sont obligés de « recalculer » plusieurs fois par an leurs modèles afin de les réajuster en fonction des statistiques récentes.
- Il est exceptionnel que l'on construise un modèle pour les besoins spécifiques d'une évaluation. L'emploi de l'outil est donc conditionné par l'existence préalable d'un modèle. Quand cette opportunité se rencontre, le modèle existant doit souvent être adapté aux questions posées par l'évaluation.
- Chaque modèle macroéconomique est basé sur un jeu d'hypothèses qui lui sont propres. Lorsque deux modèles sont construits pour un même territoire, ils ne produisent jamais exactement les mêmes simulations.
- La faiblesse de beaucoup d'évaluations macroéconomiques tient à la façon dont les effets primaires microéconomiques sont évalués avant d'être introduits dans le modèle. En ce sens, l'évaluation macroéconomique rencontre la même difficulté que les autres évaluations de programmes multisectoriels : il faudrait pouvoir résumer les effets d'une multitude d'interventions hétérogènes sous la forme de quelques impacts microéconomiques standardisés et quantifiés. Ce travail, très difficile à réussir pour une évaluation normale, n'est que la première étape d'une évaluation macroéconomique. C'est l'étape qui a été définie ci-dessus comme l'estimation des impacts primaires. Quelle que soit la justesse de ses hypothèses internes, le modèle ne fait que transformer des estimations d'impacts primaires en estimations macroéconomiques. Il cumule donc les difficultés des deux niveaux d'estimation.

## ● Pour en savoir plus

Pour faciliter l'accès à cet outil, les paragraphes qui précèdent ont volontairement écarté toute la technicité de la construction des modèles. Pour le lecteur qui souhaite aller plus loin, il est possible de lire de nombreux ouvrages sur la macroéconomie et la modélisation. Une liste en est donnée ci-dessous.

L'emploi des modèles est encore expérimental pour l'évaluation des programmes socio-économiques, et on trouvera difficilement un manuel expliquant de façon détaillée les techniques d'estimation des impacts macroéconomiques. Cette fiche est donc prolongée pour traiter les modèles QUEST et HERMIN de façon plus approfondie.

Kebabdjian, G. (1994), *Les modèles théoriques de la macroéconomie*, Paris : Dunod, 194 p.

Un manuel de macroéconomie qui aborde toutes les tendances actuelles sans a priori ; c'est un bon outil de synthèse, exigeant.

Gray, A.W. (1995), *EU Structural Funds and other public sector investments. A guide to evaluation methods*, Gill & Macmillan, 106 p.

Aborde un ensemble de modèles adaptés à l'évaluation des Fonds structurels, et explique les problèmes dans ce domaine. Très accessible.

Daloz, J.P. et Goux, C. (1985), *Macroéconomie appliquée. Du simple au complexe*, Paris : Cujas, 156p.

Ouvrage de base, très pédagogique, introduisant de façon très simple l'économétrie.

Andersen, K.V. (1995) *Information Systems in the Political World : Implementation, Use and Implications of Economic Modelling*, Amsterdam : IOS Press, 283p.

Ce livre examine l'utilisation des modèles économétriques pour l'analyse de politique et l'évaluation en Europe et aux Etats-Unis, au niveau national et au niveau régional. Il s'intéresse aux aspects sociaux et politiques des processus de construction et d'emploi des modèles.

## ● **Approfondissement : les modèles QUEST et HERMIN**

### ◆ **QUEST : un modèle multinational**

Le modèle QUEST (Quarterly European Stimulation Tool) a été développé par la DG II de la Commission européenne. Celui-ci a pour principal but de réaliser des prévisions à moyen et long terme. Le modèle est employé pour quantifier des scénarios dans divers domaines : politique monétaire, politique budgétaire, étude de l'incidence économique des changements intervenant sur la scène internationale tels que les mouvements au niveau des taux de change, taux d'intérêt des devises, prix du pétrole et autres prix mondiaux. De fait, la manière dont est modélisée la dimension multinationale constitue une caractéristique essentielle du modèle.

Le modèle QUEST couvre tous les Etats membres de l'Union européenne (le Luxembourg étant associé à la Belgique), les Etats-Unis et le Japon par le biais de sous-modèles nationaux. Pour le reste du monde, seuls les échanges internationaux sont modélisés. Ils sont répartis en sept principaux pays de l'OCDE et six autres régions. Les équations des exportations bilatérales pour les échanges de marchandises sont calculées, sur la base des volumes d'importations et des prix à l'exportation, pour les 26 pays/régions. Sur cette base, on peut représenter le schéma des échanges internationaux de marchandises par le biais d'une série d'équations sur les liens commerciaux. Ce bloc des relations commerciales permet de suivre un choc économique dans les relations commerciales internationales. Il est séparé des sous-modèles nationaux. Il est ainsi possible d'étudier un impact sur un pays, soit pris isolément, soit imbriqué dans l'ensemble des échanges internationaux.

Les sous-modèles nationaux ont tous la même structure de base. Ils ont un nombre relativement faible d'équations de comportement dont les paramètres ont été ajustés pour chacun des pays. La structure de base du QUEST réunit trois types de relations : le bloc de l'offre, les composantes de la demande et le rapport salaires/prix.

#### *(a) L'offre*

Le bloc de l'offre est principalement chargé de déterminer la capacité de production. On part du principe que la capacité de production est limitée approximativement par le stock de capital national. Il faut signaler que QUEST est un modèle à secteur unique qui n'établit donc pas de différences entre les productions des divers secteurs. La capacité de production sert,

en conjonction avec la production réelle, à définir le degré d'utilisation des capacités. Cette dernière variable joue un rôle important dans le modèle et détermine en partie la demande en termes d'investissements et d'importations, et donc la demande globale ainsi que la formation des prix.

La capacité de production dépend du capital privé et public, de l'emploi et de l'énergie, ainsi que d'une tendance exogène à l'augmentation de la productivité de tous ces facteurs. Les effets d'offre sur la capacité de production se matérialisent complètement en six ans environ.

### *(b) La demande*

La consommation privée est déterminée par le revenu des ménages, en valeur actualisée et en tenant compte des anticipations, ainsi que par la valeur de leur patrimoine. Le revenu immédiatement disponible entre aussi en compte dans la mesure où une partie des ménages a des contraintes de liquidité.

L'investissement varie positivement en fonction de la valeur actualisée des profits attendus, et négativement en fonction du coût du capital.

Les importations dépendent de la demande finale intérieure ainsi que de leur prix relatif par rapport aux prix intérieurs.

### *(c) Les salaires et les prix*

Les entreprises sont supposées fixer leur prix en appliquant une marge commerciale à leur coût marginal. Les exportations dépendent à la fois des coûts de production intérieurs et du niveau des prix de la concurrence internationale. Il existe aussi une rigidité des prix à court terme, modélisée à l'aide d'une contrainte de coût d'ajustement.

Les salaires nominaux sont établis en fonction de quatre facteurs : (1) le niveau de désincitation dépendant des minima sociaux et de la valeur du temps libre, (2) la productivité du travail, (3) les tensions existantes sur le marché du travail et (4) le niveau général des prix. Les salaires ne s'ajustent pas immédiatement aux changements du contexte économique dans la mesure où les accords salariaux sont signés pour plusieurs années.

Les taux d'intérêt et les taux d'échange sont également traités comme des variables endogènes dans le modèle QUEST, à l'inverse de beaucoup d'autres modèles. Cela permet par exemple de simuler les effets d'éviction (le fait que les besoins de financement de la dépense publique entraîne une hausse des taux d'intérêt réels, et donc une baisse de l'investissement privé).

### ◆ **HERMIN : un modèle à quatre secteurs**

Les modèles HERMIN (HER repris du modèle macrosectoriel HERMES dont ils s'inspirent et MIN pour indiquer que ce modèle est de petite dimension) couvrent l'Irlande, où le prototype a été développé à l'origine, le Portugal, l'Espagne et la Grèce. Le prototype d'HERMIN provient de HERMES, un système de modèles complexes (onze secteurs, plus de 600 équations), né d'une collaboration trans-européenne conduite par la DG XII et couvrant, avec des degrés divers de performance, tous les Etats membres de l'Union européenne. Les modèles HERMES avaient été développés pour analyser les politiques énergétiques et se sont donc concentrés sur le rôle de l'énergie en tant que facteur de production. Ils comprenaient donc une simulation soignée des processus relatifs à la production.

Les modèles HERMIN sont d'échelle petite à moyenne (4 secteurs avec 100 à 150 équations chacun). Ils ont été élaborés par des équipes nationales dans chacun des quatre pays périphériques de l'Union européenne. Un des principaux objectifs de ces modèles est l'étude comparative de l'impact des Fonds structurels sur la périphérie de l'Union européenne. La structure de base du HERMIN réunit trois types de relations : les composantes de l'offre, les composantes de la demande et le rapport salaires/prix.

#### (a) *L'offre*

Le modèle HERMIN comprend quatre secteurs : le secteur des biens échangeables, le secteur des biens non échangeables, le secteur agricole et le secteur public. Ces secteurs présentent les caractéristiques distinctes suivantes :

Le secteur des biens échangeables est soumis à la concurrence internationale et sa production est principalement déterminée par la demande mondiale (exprimée en termes de production mondiale) et la compétitivité (mesurée par le rapport des coûts unitaires du travail sur les coûts unitaires mondiaux). Les prix sont déterminés de manière externe sur les marchés internationaux.

Le secteur des biens non échangeables couvre les services, le bâtiment, la construction et les équipements collectifs. Il est soumis à la demande intérieure et les prix sont déterminés par la marge bénéficiaire sur les coûts du travail.

La production des secteurs des biens échangeables et des biens non échangeables est modélisée sous la forme d'une fonction où la valeur

ajoutée dépend du capital et du travail. Les niveaux de l'emploi et des investissements dépendent alors du niveau de la production. La demande des deux facteurs de production tient également compte de leur coût respectif (sous l'hypothèse d'élasticité de substitution constante). Le progrès technique est assimilé à une tendance temporelle exogène, différente dans chaque secteur.

Le secteur agricole couvre l'agriculture, la sylviculture et la pêche. Il est considéré comme essentiellement exogène avec des prix et des quotas de production définis dans le cadre de la Politique Agricole Commune de l'Union européenne.

Dans le secteur public, qui comprend l'administration, l'éducation et la santé, le niveau d'activité dépend de décisions gouvernementales. Les dépenses sont contraintes par les règles d'équilibre budgétaire et de remboursement de la dette.

#### *(b) La demande*

La consommation est déterminée par le revenu personnel disponible réel et par la richesse financière réelle. Les investissements privés non résidentiels sont dérivés de la fonction de production en tant que demande de capital fixe. Les investissements résidentiels sont une fonction du revenu disponible par habitant. Enfin, les investissements publics sont une variable exogène. Le secteur monétaire n'est pas modélisé et le taux de change, de même que le taux d'intérêt intérieur, sont des variables exogènes.

Les exportations agricoles sont estimées de manière exogène. Les exportations du secteur des biens échangeables sont une fonction de la production de ce secteur ainsi que du taux de croissance du PNB (reflet du cycle économique) dans le cas de l'Espagne. Le tourisme et les exportations de services sont une fonction de la production mondiale. Les importations sont déterminées de manière résiduelle pour faire la différence entre la production totale des quatre secteurs et la demande finale (consommation privée, consommation publique, investissements et exportations).

#### *(c) Les salaires et les prix*

Les salaires sont fixés par une équation simulant le processus de négociations salariales entre les employeurs et les syndicats. Les négociations sur les taux salariaux nominaux sont influencées par la

situation du marché du travail, telle qu'elle est reflétée par le taux de chômage actuel ou par son taux de variation. Elles subissent également l'influence de la « ponction fiscale », c'est-à-dire la différence entre le salaire brut avant impôts exprimé en prix de production et le salaire net de toutes déductions exprimé en prix à la consommation. Finalement, le niveau des prix de production et celui de la productivité du travail affectent également le processus des négociations salariales. Dans le cas de l'Irlande, le marché du travail est supposé être homogène et les salaires du secteur des biens échangeables dominent tous les autres secteurs. Dans le cas de l'Espagne, les salaires du secteur des biens non échangeables sont évalués à part, mais les secteurs public et agricole sont également dominés par le secteur des biens échangeables. L'offre de main d'œuvre est déterminée différemment dans les différents pays. Les migrations jouent un rôle important dans le cas de l'Irlande, notamment en direction et en provenance du Royaume-Uni. Dans le cas de l'Espagne, une distinction est faite entre les taux d'activité des femmes et des hommes.

## Outils pour porter un jugement

La situation d'évaluation d'ensemble d'un programme complexe est caractérisée par le fait que les effets attendus sont nombreux et que tous les partenaires n'attachent pas la même importance aux mêmes effets. Certains outils de jugement sont particulièrement bien adaptés à cette situation.

Par exemple, le **panel d'experts** construit un jugement synthétique sur le programme à évaluer. Cet outil, lorsqu'il est mis en œuvre de façon optimale, accroît la crédibilité et l'acceptabilité des conclusions de l'évaluation en respectant les différences entre points de vue tout en construisant des conclusions consensuelles.

L'**analyse multicritère** présente des avantages importants en termes d'acceptation de la multiplicité des points de vue. L'outil intègre un mécanisme d'identification et d'explicitation des points consensuels.

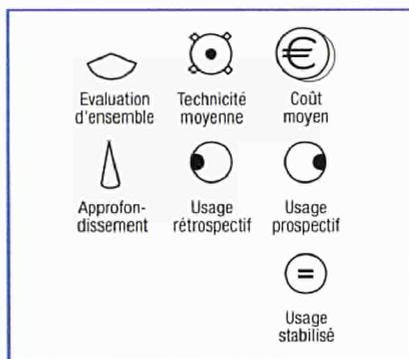
Les outils de jugement les plus connus sont l'analyse coût-avantages et l'analyse coût-efficacité. Pour des raisons diverses, ces outils n'ont pas la souplesse nécessaire pour prendre en compte la complexité du jugement à porter sur l'ensemble d'un programme socio-économique. L'**étalonnage** (*benchmarking*) est un outil de jugement plus souple qui serait peut-être applicable à l'évaluation d'ensemble d'un programme complexe. Cependant, cet outil a été employé jusqu'à présent dans des évaluations simples. C'est la raison pour laquelle il est présenté plus loin.

Certains outils présentés plus haut, comme des aides à la structuration de l'évaluation, peuvent également être employés pour formuler un jugement. Par exemple, l'**analyse SWOT** peut aider à juger la pertinence d'une stratégie, de même que le **vote coloré** peut aider à valider les points clés de la synthèse finale de l'évaluation dans un contexte de partenariat.



## ■ Panel d'experts

Le panel d'experts réunit plusieurs spécialistes indépendants et reconnus dans les domaines correspondant au programme évalué. Les experts produisent collectivement un jugement sur la valeur du programme et de ses effets. Cet outil permet de disposer rapidement et à faible coût d'un jugement global intégrant les principales informations connues sur le programme ainsi que de multiples expériences antérieures et extérieures.



Le panel peut être considéré comme un outil d'évaluation dans la mesure où il existe une procédure standard et répliquable pour le constituer, pour le réunir et pour l'amener à produire ses conclusions. L'outil est inspiré des jurys universitaires, ce qui explique qu'il soit apparu dès le début des années 70, dans le domaine de l'évaluation des programmes de Recherche-Développement.

Par « panel d'expert », nous entendons ici un groupe de travail spécialement constitué et réuni à l'occasion de l'évaluation. On trouvera plus loin la description d'un outil proche, l'enquête Delphi, qui s'appuie aussi sur des experts mais qui diffère sur de nombreux points.

### Exemple : Evaluation de la qualité et de la pertinence des programmes de développement en Suède

Le Conseil national suédois du développement technique finance depuis plusieurs années des programmes de développement des compétences et des connaissances fondamentales. Depuis 1981, la qualité de ces programmes est évaluée régulièrement.

La première étape consiste à sélectionner cinq experts internationalement reconnus et représentatifs des divers points de vue intéressant le programme. Ces experts s'engagent à participer à toutes les réunions du panel. L'un d'entre eux est désigné comme président.

Dans une seconde étape le panel examine la documentation du programme concernant les ressources mobilisées (personnel, équipements, financements), les thèmes de recherches, les résultats, les publications, les coopérations. Le panel auditionne également quelques destinataires.

L'étape suivante est constituée par des visites sur le terrain, réalisées individuellement par chaque expert ou par groupe de deux.

A l'étape finale, le panel élabore un rapport commun. Si les points de vue divergent, le panel discute jusqu'à obtention d'un consensus. L'expérience a montré que la recherche des compromis adoucit inévitablement les conclusions. Cependant, l'évaluation est renforcée par la légitimité des conclusions du panel. Les gestionnaires du programme sont invités à faire part de leurs remarques et commentaires. Le rapport final est publié.

Cet outil est considéré comme efficace et peu coûteux. Il est largement appliqué dans tous les pays nordiques, et plus généralement dans tous les pays qui évaluent leurs politiques scientifiques. L'OCDE applique également cet outil à de nombreux autres domaines politiques.

Source : Christensen, H.C (1987) "Evaluation of research Programmes", in : E. Ormala (éd.) Evaluation of Technical Research and Development, 88-108, Espoo, Nordforsk.

### ● A quoi sert l'outil ?

Le panel d'expert sert principalement à juger un programme, mais c'est un outil générique. Les termes de référence donnés au panel peuvent en effet comprendre un grand éventail de questions allant de la pertinence des objectifs du programme jusqu'à l'estimation de ses effets réels ou probables.

### ● Dans quels cas l'utiliser ?

On fait appel aux panels d'experts pour dégager un consensus sur des questions complexes et mal structurées, pour lesquelles les autres outils ne permettent pas de fournir des réponses univoques ou crédibles.

L'utilisation des groupes d'experts permet, dans un laps de temps de quelques mois, de rassembler l'essentiel des points de vue et des connaissances pertinentes pour l'évaluation. L'outil est donc recommandé lorsqu'il existe une expertise suffisante dans le domaine et que l'évaluation est complexe.

L'outil est également très bien adapté à des petits programmes simples pour l'évaluation desquels il n'est pas justifié de consacrer trop de ressources.

### ● Comment le mettre en œuvre ?

Etape 1 - Identification d'une liste d'experts potentiels : Les membres du panel doivent être des spécialistes reconnus dans au moins un des champs

concernés par le programme. Ils doivent avoir une large expérience du domaine et être indépendants des commanditaires. Ils doivent par ailleurs avoir la disponibilité et la volonté de s'impliquer dans l'évaluation.

**Etape 2 - Sélection et mandatement des experts :** Le panel est généralement composé de six à douze membres appartenant aux différents « champs d'expertise ». La tendance actuelle est d'élargir le spectre des intérêts et de rechercher la plus grande diversité possible de points de vue au sein du panel. Les experts sont nommés intuitu personae et ne représentent pas leur institution. Chaque expert signe un contrat qui, selon les cas, prévoit ou ne prévoit pas de rémunération. Le président du panel est choisi par le commanditaire ou élu par ses pairs. Il est impératif de confier le secrétariat du panel à une personne qui a la disponibilité suffisante, ce qui n'est généralement pas le cas des experts eux-mêmes.

**Etape 3 - Investigations :** Les experts se réunissent entre trois et six fois à intervalle d'environ un mois. Toutes les dates de réunion doivent être programmées dès le début du travail. Les débats internes au panel sont placés sous le sceau de la confidentialité. Les membres du panel consultent les documents du programme ou du projet de programme (rapports d'activités, études préalables, enquêtes), ils auditionnent les responsables du programme et quelques destinataires typiques. Ils peuvent aussi réaliser des visites sur le terrain, généralement en binômes pour limiter les risques de parti-pris.

**Etape 4 - Synthèse:** Le panel produit un rapport et formule des conclusions et des recommandations acceptées collectivement. En cas de désaccord, il peut être utile d'exprimer la conclusion majoritaire et d'annexer un commentaire de l'expert minoritaire. Le panel doit tenir une dernière réunion après réception des commentaires du commanditaire sur son rapport provisoire. Il est maintenant envisageable qu'une partie du travail du panel, notamment dans sa phase finale, fasse appel aux techniques de communication à distance (Internet ou visioconférence).

### ● Application aux Fonds structurels

L'outil peut être utilisé pour réaliser la totalité d'une évaluation ex ante, à condition que le délai disponible soit d'au moins trois mois. Par exemple, on pourra demander à un panel d'expert d'estimer l'impact probable du programme en termes d'emploi, ou de juger les mérites du programme en termes de potentiel de synergie.

L'outil peut également être intéressant chaque fois que l'on veut renforcer l'étape de structuration ou de jugement. Dans ce cas, le panel d'expert peut intervenir au début et à la fin de l'évaluation, en combinaison avec d'autres outils mis en œuvre pour la collecte ou l'analyse des données (voir un exemple dans l'Encadré 37).

Le panel d'experts permet de formuler un jugement indépendant et faisant autorité, ce qui est particulièrement intéressant dans un contexte de partenariat, notamment s'il existe des divergences de vues entre partenaires. Dans un tel contexte, il est possible de demander à chaque partenaire de choisir une partie des experts en fonction de leur proximité avec son point de vue. On peut aussi proposer une liste d'experts excédant les besoins du panel et donner à chaque partenaire le droit de rayer un ou deux noms sur la liste.

#### ● **Points forts et limites d'utilisation**

- Le panel d'experts est un outil relativement peu coûteux, rapide et flexible. Ses conclusions jouissent d'une bonne crédibilité lorsque l'on dispose d'experts reconnus.
- Le panel d'experts peut remplir de nombreuses fonctions, mais il est préférable de limiter son travail à une partie seulement de l'évaluation : structuration des objectifs, estimations des effets ou jugement. Plus le travail du panel est défini de façon précise, plus la portée de son travail sera reconnue. La fiabilité de l'outil peut être affaiblie si les questions posées aux experts sont trop vastes.
- Le risque de biais d'empathie est important dans la mesure où les panels sont trop souvent restreints aux spécialistes du domaine couvert par le programme (pairs) et sont donc peu enclins à critiquer la pertinence des objectifs ou à s'intéresser à d'éventuels effets pervers.
- La confrontation des avis mène souvent à la sous-évaluation des points de vue minoritaires. En effet, le mode de fonctionnement consensuel sur lequel repose la dynamique du panel produit une convergence d'opinions autour des valeurs majoritaires qui ne sont pas forcément les plus pertinentes.
- L'outil est cependant susceptible de produire des conclusions créatives et des recommandations riches s'il est associé à des techniques d'animation adaptées telles que celles qui sont présentées dans ce Volume (Metaplan®, vote coloré).

### ● Pour en savoir plus

Callon, M., Laredo, P. et Mustar, P. (1995), *La gestion stratégique de la recherche et de la technologie*. Paris : Economica. 31-88.

Présentation dans les deux premiers chapitres de cet ouvrage du modèle de panel d'experts utilisé par la Commission des communautés Européennes.

Cozzens, S.E. (1987), 'Expert Review in Evaluating Programs', in : *Science and Public Policy*, 14(2), 71-81.

Inventaire et analyse de l'expérience américaine d'utilisation des panels d'experts pour l'évaluation. L'intérêt des conclusions dépasse le cadre des politiques scientifiques.

Nadeau, M-A. (1988), *L'évaluation de programme*. Laval, Québec : Presses de l'université de Laval, 349-352

Ce manuel présente les méthodes et les outils de l'évaluation de programme et consacre quelques pages aux panels d'experts.

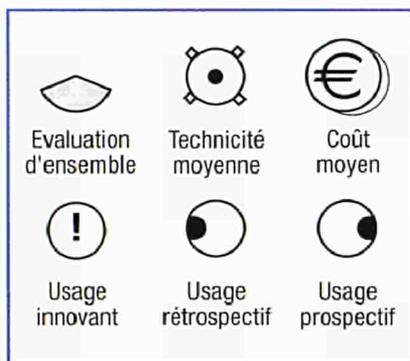
Witkin, B.R et Altschuld, J.W (1995). *Planning Conducting Needs Assessments*, Thousand Oaks : Sage publications. 193-203.

Cet ouvrage présente une série d'outils applicables à l'évaluation des besoins et consacre une fiche aux panels d'experts.



## ■ Analyse multicritère

L'analyse multicritère est utilisée pour porter un jugement comparatif sur des projets alternatifs ou des mesures hétérogènes. Elle permet de prendre en compte simultanément plusieurs critères intervenant dans des situations complexes, et ainsi d'aider les « décideurs » à intégrer dans un cadre prospectif ou rétrospectif les



différentes options reflétant les opinions des acteurs concernés par un processus de décision. Elle implique la participation des décideurs et aboutit à des conseils opérationnels ou recommandations.

### Exemple : L'évaluation d'un programme de développement régional dans le Hainaut

L'analyse multicritère a été utilisée lors de l'évaluation d'un programme de développement régional cofinancé par les trois Fonds structurels européens et par le gouvernement de la région wallonne. Le programme s'applique à la province du Hainaut, il couvre la période de 1994 à 1996 et son budget s'élève à plus d'un milliard d'EURO, financé à près de 50% par l'Europe. L'outil d'analyse multicritère a été utilisé pour sa spécificité « multijuge » qui permet d'utiliser des systèmes de pondération différents, propres à chaque partenaire. Huit critères d'évaluation ont été définis en accord avec les représentants des deux niveaux de gouvernement : (1) viabilité des entreprises aidées ; (2) employabilité des personnes formées ; (3) diversification sectorielle de l'économie ; (4) ouverture extérieure de l'économie ; (5) insertion environnementale des activités ; (6) répartition territoriale des activités ; (7) services aux entreprises ; (8) synergie entre les composants du programme.

L'efficacité des mesures au sein du programme a été appréciée par les gestionnaires des mesures qui ont attribué durant des entretiens des notes d'efficacité selon ces huit critères. Les résultats ont montré une grande dispersion des notes d'efficacité selon les mesures et selon les critères. Puis, au moyen d'une procédure d'entretien formalisée, on a conduit six « juges » (2 membres de la Commission européenne, 2 du Gouvernement Wallon, 2 de la Région du Hainaut) à établir une pondération des critères utilisés en fonction de l'importance qu'ils attribuent à chacun de ces critères. Les résultats ont été de nouveau très dispersés : un des juges a donné le poids de « 3 » au critère viabilité des entreprises alors qu'un autre a donné à ce même critère le poids de « 26 », ce qui témoigne de la variabilité des centres d'intérêt.

Après traitement informatique, les mesures ont finalement été classées par ordre d'efficacité en tenant compte du fait qu'il existe un classement pour chaque juge (puisque chacun a son système de pondération des critères). Malgré la dispersion des avis tant des gestionnaires sur les notes d'efficacité que des juges sur la pondération des critères, les divers classements ont convergé vers un résultat pratiquement identique sur la valeur relative des mesures les unes par rapport aux autres. Cela a apporté un éclairage indiscutablement utile aux décideurs, notamment dans un contexte où la perspective était de réallouer une partie du budget en cours de programme.

Source : Cahier MEANS n°4 (1995) Appliquer la méthode multicritère à l'évaluation des programmes structurels. Bruxelles : Union européenne.

### ● A quoi sert l'outil ?

L'outil vise à structurer et à combiner les appréciations qui interviennent dans les processus de prise de décision. Il convient d'entendre ici la décision comme un processus jalonné de multiples choix. Les décisions relatives à ces choix conditionnent en grande partie la décision finale.

L'analyse multicritère vise notamment à mettre en évidence les logiques et les convictions subjectives des acteurs par rapport à une question particulière. Elle permet également de synthétiser les opinions exprimées par ces mêmes acteurs afin de déterminer les critères prioritaires ; d'analyser des situations conflictuelles ou encore, d'aboutir à des recommandations ou à des conseils opérationnels.

### ● Dans quels cas l'utiliser ?

L'analyse multicritère est proche des techniques adoptées dans le domaine du développement organisationnel ou du management des systèmes d'information, telle que la méthode des Facteurs Critiques de Succès, développée dans les années 70. De même, elle s'apparente à l'analyse coût-avantages tout en évitant de ramener à une base unitaire commune (monétaire) des phénomènes très disparates. Cette amélioration est fondamentale car elle permet de prendre en compte des critères dont on ne connaît pas, a priori, le coût (paysage, bruit, satisfaction, etc.).

L'analyse multicritère est particulièrement utile lors de la formulation d'un jugement sur des problèmes complexes. C'est le cas lorsque les critères de jugement s'opposent (par exemple emploi contre environnement) ou si l'arbitrage entre les critères est difficile.

D'une manière générale, cette technique est utilisée lors d'évaluations ex-ante de projets publics et de leurs variantes (le choix d'un tracé autoroutier, la construction d'une nouvelle infrastructure, etc.). Plus rarement, l'analyse multicritère sera appliquée lors d'évaluations intermédiaires ou ex post de programmes.

### ● **Comment le mettre en œuvre ?**

L'analyse multicritère est un outil de comparaison qui consiste à prendre en compte plusieurs points de vue. Elle se décompose en plusieurs phases décrites ci-dessous d'une manière chronologique. Une répétition des phases est cependant possible pour permettre des corrections.

**Etape 1. Définition des projets ou actions à juger :** Inventaire des actions réalisées ou envisagées, ou des éléments sur lesquels va porter le jugement comparatif.

**Etape 2. Définition d'une batterie de critères de jugement :** Le chargé d'évaluation lui-même ou les commanditaires de l'évaluation vont préciser les critères à prendre en compte. Une attention particulière doit être portée à la définition des critères afin d'être le plus exhaustif possible et de bien cerner la question. Les critères doivent traduire les préférences des décideurs ou encore les différents « points de vue » afin de résumer et regrouper diverses dimensions caractéristiques permettant d'évaluer une action.

**Etape 3. Analyse des impacts des actions :** Lorsque les projets et les critères ont été définis, il convient de procéder à une estimation quantitative ou à une description qualitative de l'impact de chaque projet selon les critères. Pour cela, on peut utiliser de brefs énoncés décrivant les différents niveaux d'impacts : les « descripteurs d'impact ».

**Etape 4. Jugement des effets des actions selon chacun des critères retenus :** La notation des impacts peut être effectuée par le chargé d'évaluation à partir de données quantitatives, ou, plus subjectivement, par des experts ou par les destinataires de l'évaluation eux-mêmes.

**Etape 5. Agrégation des jugements entre eux :** Cette combinaison peut se faire selon trois principes :

- Jugements personnels : il n'est fait aucune synthèse entre les différents critères de jugement. Chacun des destinataires de l'évaluation construit son jugement personnel à partir de l'analyse et s'en sert pour argumenter son point de vue.

- **Aide à la coalition** : une technique informatisée classe les actions les unes par rapport aux autres. Une action en surclasse une autre dans la mesure où elle est mieux placée pour une majorité de critères (maximum d'alliés) et où elle a moins de « notes éliminatoires » du point de vue des autres critères (minimum d'opposants).
- **Aide au compromis** : une pondération des critères est proposée par le chargé d'évaluation ou négociée par les destinataires de l'évaluation. Il en découle un classement des actions selon leur note pondérée.

La principale option, quant à l'analyse multicritère, est celle de l'implication ou non du chargé d'évaluation dans la définition des critères, dans leur notation et dans leur pondération. Si l'équipe d'évaluation intervient activement dans l'analyse, la crédibilité des résultats en souffre. En revanche, lorsque les destinataires de l'évaluation participent à la définition des critères, on constate que chaque partenaire prolonge la discussion jusqu'à ce qu'on ait trouvé au moins un critère de jugement qui place son action favorite en première position.

### ● **Application aux Fonds structurels**

On l'a vu dans l'exemple ci-dessus, l'analyse multicritère est bien adaptée à une conduite partenariale de programmes structurels et à leur évaluation. Les avis de membres nationaux et supranationaux peuvent être exprimés conjointement sans pour autant perdre de leur spécificité et devoir faire des concessions trop grandes au regard de leurs échelles de valeur.

L'expérience du Hainaut a été effectuée dans le sens d'une utilisation possible de la technique par la Commission, dans le cadre de l'évaluation des Fonds structurels. Elle a produit des conclusions encourageantes quant à son éventuelle mise en œuvre et a permis notamment de mettre au point une variante de la méthode appelée « multicritère-multijuge » qui permet à chacun des partenaires de construire son jugement à partir des critères et des poids qu'il souhaite.

### ● **Points forts et limites d'utilisation**

- L'analyse multicritère offre un cadre pour que les acteurs prennent part à la prise de décision et à la résolution de problèmes. Son caractère négocié et son traitement explicite des critères de jugement contribuent à la mise en forme d'une réalité non structurée. Elle permet donc de prendre en considération les valeurs et les avis individuels de plusieurs

acteurs. Elle permet également de traiter de manière quantitative des relations fonctionnelles comprises dans un réseau complexe.

- L'intervention d'un expert, la marge de manœuvre laissée aux décideurs et l'analogie avec la procédure de vote font de cette technique un outil adapté à une démarche partenariale.
- La technique peut aider à définir une coalition ou un compromis, mais elle ne dicte pas le jugement individuel ou collectif des partenaires.
- Les analyses multicritères ont plus souvent la préférence des décideurs car ils y sont aisément associés à travers un cadre technique relativement simple, à la différence des analyses coût-avantages qui mesurent artificiellement tous les effets en termes monétaires.
- Des problèmes de mise en œuvre spécifiques peuvent limiter l'usage de l'analyse multicritère, ou nécessiter la présence d'experts.
- Cette technique n'est pas toujours employée de façon interactive, comme elle le devrait, et a tendance à figer des critères qui sont dans la réalité mouvants.

#### ● Pour en savoir plus...

Roy, B. et Buysou, D. (1993), *Aide Multicritère à la décision : Méthodes et Cas*, Paris : Economica, 695 p.

Ouvrage très complet, lecture à plusieurs niveaux possible grâce à un marquage des éléments essentiels.

Saaty, T.L. (1984), *Décider face à la complexité*, Paris : Entreprise Moderne d'Édition, 231 p.

Ouvrage décrivant la technique à partir d'exemples variés, réflexion sur le processus de décision.

Schärlig, A. (1990, 2<sup>ème</sup> éd.), *Décider sur plusieurs critères, panorama de l'aide à la décision multicritère*, Lausanne : Presses polytechniques et universitaires romandes, 303 p.

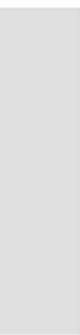
Manuel complet et très accessible.



### III

## Outils pour approfondir une question évaluative

*Il est possible d'approfondir une question évaluative lorsque cette question concerne une seule intervention ou un petit nombre d'interventions similaires, lorsque le public concerné est homogène et lorsque les effets à évaluer sont peu nombreux. Si une question évaluative est suffisamment simple, il est alors possible de la traiter de façon approfondie, c'est-à-dire de réaliser une collecte des données primaires auprès des destinataires directs, de réaliser une analyse approfondie des impacts et de rechercher des interventions comparables pour servir de référence de jugement. La quasi totalité des outils d'évaluation sont applicables, sans nécessité d'innovation, dans cette situation qui est décrite dans la plupart des manuels d'évaluation. Cette section recense et présente onze outils en insistant sur leurs conditions d'application à l'évaluation des programmes socio-économiques.*



## Exemples d'évaluations traitant une question de façon approfondie

Les exemples présentés ci-dessous décrivent des situations qui se présentent de façon simple, soit que l'évaluation porte sur une action ou un groupe d'actions homogène, soit qu'elle se concentre sur un thème unique tel que l'obtention d'un résultat ou d'un impact particulier (pour plus d'informations sur les évaluations thématiques, on pourra se référer au Volume 5 de la Collection qui y est consacré). Lorsque l'objet de l'évaluation est suffisamment simple, il est possible d'approfondir la collecte des données primaires, l'analyse des impacts et la formulation du jugement. La quasi totalité des outils d'évaluation sont applicables dans cette situation. Les trois exemples qui suivent en donnent des illustrations.

Dans le premier exemple les responsables d'un ministère français désiraient obtenir une estimation comparée des effets sur l'emploi de plusieurs interventions. La question d'évaluation était la suivante : « Certaines mesures d'aide à l'emploi sont-elles plus efficaces que d'autres ? »

### Encadré 19 - Evaluation de l'impact des aides à l'embauche en France

En France, le Ministère de l'Emploi réalise fréquemment des enquêtes par questionnaire passées par téléphone auprès d'échantillons d'entreprises destinataires des différentes mesures d'aides à l'emploi. Chacune de ces enquêtes permet d'estimer un coefficient d'effet d'aubaine et, le cas échéant, des coefficients d'effets de substitution ou d'attraction. Ces coefficients permettent d'extrapoler les données du suivi (effets bruts) pour passer à une estimation des effets nets. Un panel d'experts a été réuni à plusieurs reprises pour confronter l'ensemble des coefficients résultant des diverses enquêtes et pour proposer des règles générales d'estimation des effets d'aubaine en fonction de la nature des aides.

Dans cet exemple, la technique de l'enquête par questionnaire a été utilisée pour l'observation. La technique du panel d'experts a contribué à consolider l'estimation des effets.

L'exemple suivant est tiré de l'expérience conduite par le programme MEANS en Catalogne. Une évaluation des aides aux coopératives a été demandée à la suite d'un audit de la mise en œuvre qui a émis des doutes sur l'efficacité de ces aides en terme de création et de maintien d'emplois.

#### **Encadré 20 - Evaluation ex post des aides aux coopératives catalanes**

En 1992, les autorités catalanes avaient développé un programme d'aide aux coopératives en considérant que celles-ci pouvaient avoir un impact important en matière d'emploi, en raison d'impératifs de rentabilité moindres que ceux des entreprises privées. Cependant, un doute subsistait sur l'efficacité de ces aides et le Département du Travail a commandé une évaluation du programme.

En raison du nombre de coopératives, de leur diversité et de la faible connaissance des destinataires, il a été décidé de faire une **enquête par questionnaire** auprès de la totalité de la population-cible, soit environ 3000 coopératives. Les réponses obtenues ont été traitées par une **analyse factorielle** afin de constituer une typologie des coopératives. Grâce à cette analyse, on a pu mieux comprendre les caractéristiques des coopératives qui ont créé des emplois, mais aussi de celles qui n'ont pas créé d'emplois. L'analyse a montré que les coopératives les plus susceptibles de transformer l'aide reçue en création ou maintien d'emploi sont celles qui sont déjà anciennes, qui sont membres d'une fédération de coopératives, qui ne sont pas situées dans la zone côtière, qui ont une taille moyenne... En revanche il apparaît aussi que ce sont des coopératives qui investissent de manière autonome par rapport aux aides.

Dans cet exemple, l'outil d'observation a été l'enquête par questionnaire. L'analyse factorielle a permis de remplir de façon très performante la fonction de croisement des observations.

Le troisième exemple décrit un travail réalisé pour évaluer l'efficacité d'un programme de conseils aux PME au Royaume-Uni.

### Encadré 21 - Evaluation du conseil aux entreprises au Royaume-Uni

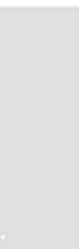
Un programme financé par le Ministère du Commerce et de l'Industrie britannique s'adressait aux PME auxquelles il offrait des prestations de conseil en gestion, de formation et d'information. Les impacts de ce programme ne pouvaient pas être exprimés en emplois créés, puisqu'il visait à augmenter la productivité, ce qui conduisait à court terme à détruire des emplois. Il a donc fallu évaluer le programme selon d'autres critères comme la productivité, la rentabilité ou l'investissement induit.

La méthode développée pour l'évaluation a été la suivante :

1. Un premier questionnaire a été envoyé aux 75 opérateurs du programme (prestataires de service). Ce questionnaire abordait les prestations de conseil apportées aux destinataires et les catégories d'entreprises ciblées.
2. Des entretiens individuels avec 20 opérateurs ont permis d'approfondir la description des stratégies développées par les opérateurs.
3. Des études de cas ont été réalisées pour six opérateurs. Dans le cadre de ces études de cas, les coordonnées des destinataires ont été relevées.
4. Un échantillon de 300 entreprises destinataires a été constitué. Un autre échantillon de 200 entreprises non-destinataires a également été constitué selon la technique des groupes de comparaison (binômes appariés).
5. Un second questionnaire a été adressé aux entreprises destinataires et à celles du groupe apparié. Une comparaison statistique entre les entreprises des deux groupes a permis d'obtenir des indications quantitatives sur les effets du programme en termes de chiffre d'affaires, de productivité et d'investissements induits. Afin d'éviter des déclarations « biaisées », les questionnaires comportaient des questions nombreuses et partiellement redondantes (questionnaire triangulé). L'enquête intégrait la recherche d'effets d'aubaine.

Cette évaluation a employé un grand nombre d'outils dans le cadre d'une démarche très complète d'observation de terrain. La collecte des données a fait appel à deux enquêtes par questionnaire, à une série d'entretiens individuels et à des études de cas. La fonction d'estimation des effets a été traitée grâce à la technique des groupes de comparaison (par binôme apparié).

La récapitulation de ces trois exemples montre que les outils d'évaluation ont surtout été employés pour les fonctions de collecte et d'analyse de données. Les pages qui suivent montreront que l'on peut également choisir d'autres outils pour les fonctions de structuration et de jugement.



## Outils pour structurer l'évaluation

Dans une situation d'évaluation approfondie, les outils qui suivent constituent des aides efficaces pour clarifier les effets à évaluer, sélectionner les critères et les indicateurs de l'évaluation :

Le **cadre logique** est un outil qui a été conçu pour l'évaluation ex ante des projets d'aide au développement. Il permet d'aboutir rapidement à une clarification de la logique d'action, à la définition des impacts escomptés et à la construction d'indicateurs d'impact.

**Metaplan®** est un outil de travail en groupe très répandu. Dans une situation d'évaluation, cet outil est particulièrement bien adapté pour clarifier et hiérarchiser les effets attendus.

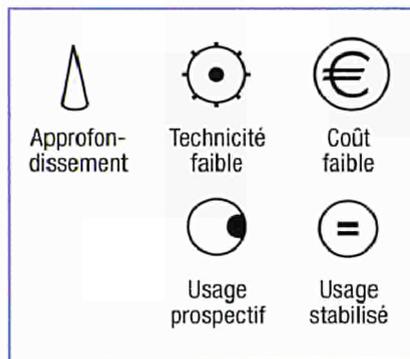
Le **vote coloré** (outil décrit précédemment) est également applicable, notamment pour le choix des questions évaluatives.

A l'inverse, des outils comme la matrice d'impacts ou la cartographie conceptuelle d'impact sont trop lourds pour une situation caractérisée par la simplicité de l'objet à évaluer.



## ■ Cadre logique

La technique du cadre logique est à la fois un exercice de structuration des éléments composant un projet (ou programme simple) et une technique d'analyse de la cohérence interne et externe de celui-ci. Le produit de cette technique, le cadre logique, est une présentation formelle et matricielle du fonctionnement interne du projet, des moyens de vérification de l'atteinte de l'objectif et des facteurs endogènes et exogènes conditionnant son succès.



### Exemple : Cadre logique d'un projet de pêche artisanale

Cet exemple concerne un projet financé par l'Union européenne dans le cadre de sa politique d'aide au développement.

Les objectifs globaux (ou finalités) du projet sont les suivants : amélioration du niveau de vie, sécurité alimentaire accrue, augmentation des recettes d'exportation, augmentation de l'emploi, utilisation des ressources de façon durable.

Les objectifs spécifiques du projet (ou buts) sont de procurer une augmentation de revenus aux acteurs économiques du secteur de la pêche artisanale. Pour évaluer l'atteinte de cet objectif, on pourra s'appuyer sur une série d'indicateurs objectivement vérifiables : évolution des niveaux de revenu de la pêche, répartition des revenus par acteur économique (pêcheurs, industriels, transporteurs, grossistes, artisans), par sexe et par zone de pêche.

Parmi les facteurs qui conditionnent l'atteinte des objectifs, on note particulièrement la nécessité de disposer de biens de consommation en quantité suffisante et à des prix abordables, ainsi que l'accessibilité suffisante des autres services (santé, éducation, conseils, etc.).

Les résultats attendus sont : une production optimale, un traitement optimum du poisson, une commercialisation satisfaisante de la production de pêche et des zones de pêche artisanale définies de manière adéquate. Pour évaluer l'obtention de ces résultats, on pourra s'appuyer sur une série d'indicateurs objectivement vérifiables : nombre de prises, nombre de produits vendus, distance entre la zone de pêche et les côtes, types de poissons pêchés, durée de la pêche, équilibre de la répartition des prises entre les pêcheurs, durée de la vente, temps écoulé entre la livraison et la vente, profondeur des ressources, accord des pêcheurs sur les zones de pêche.

Parmi les facteurs qui conditionnent l'obtention des résultats, on note particulièrement le fait que la demande doit être suffisante, ainsi que les niveaux de prix à la production.

Les activités qui font partie du projet sont les suivantes : création de coopératives de pêcheurs et de traitement du poisson, notamment par les femmes, mise à disposition de matériel de pêche, mise à disposition de séchoirs et d'entrepôts, organisation du transport, information sur la demande et les marchés, formation des acteurs, et notamment des femmes, aux nouvelles techniques, formation aux techniques de l'emballage, information suffisante sur les zones de pêche potentielles, négociation d'accords entre les pêcheurs, organisation d'un système de contrôle sur les pratiques de la pêche, etc.

Source : Commission européenne - DG VIII.

### ● A quoi sert l'outil ?

En tant que technique d'appui à la définition objective d'un projet, à sa formulation en termes opérationnels, à sa mise en œuvre, à son suivi et à son évaluation, le cadre logique est un outil qui permet :

- la description d'un projet de façon systématique, synthétique et compréhensible ;
- une articulation et une hiérarchisation claire des objectifs poursuivis ;
- une mise en évidence des conditions à réunir pour le succès du projet ;
- une clarification des hypothèses propres au projet ;
- une identification des liens de causalité entre ressources, résultats et objectifs ;
- et enfin, une identification, dès la conception du projet, des critères de succès et des moyens de vérification de l'atteinte des objectifs.

Cette technique se révèle d'autant plus efficace quand elle est préparée de façon participative et entreprise dès l'identification du problème que le projet est censé résoudre. Elle ne cessera d'évoluer et de s'élargir en fonction de l'environnement dans lequel le projet s'insère, et ainsi, par extension, permettra l'organisation complète de l'ensemble des phases du « cycle de vie du projet ».

Le cadre logique est un outil qui facilite l'élaboration des étapes clés : il permet d'établir un budget détaillé, de répartir les responsabilités, de planifier les activités, de concevoir le plan de suivi et les termes de référence des évaluations concomitantes et ex post.

### ● Dans quels cas l'utiliser ?

Conceptualisée il y a une trentaine d'années aux Etats-Unis, la technique du cadre logique est devenue progressivement l'instrument

méthodologique de prédilection des concepteurs de projets de développement. Cette approche est actuellement en vigueur dans la plupart des organismes nationaux et supranationaux de coopération, Union européenne comprise.

Son succès doit beaucoup à sa capacité de décrire de façon ordonnée, causale et logique le fonctionnement interne d'un projet dans un environnement donné. C'est cette faculté qui lui a permis de rapidement s'imposer dans le domaine de la coopération au développement. La technique du cadre logique s'avère idéale dans le cas de programmes simples, dont les objectifs et les principaux acteurs (des bailleurs de fonds aux exécutants) sont clairement identifiés.

### ● Comment le mettre en œuvre ?

L'application d'un cadre logique à un projet nécessite une analyse complète du contexte dans lequel il s'insère, des objectifs poursuivis et des stratégies à mettre en œuvre.

**Etape 1. Analyse du contexte :** le contexte est le cadre général dans lequel se situe le projet. En analysant ce contexte on identifie les problèmes que le projet doit contribuer à résoudre. Pour ce faire, il faut notamment :

- une définition de l'environnement politique, institutionnel, social et économique (politique du gouvernement, caractéristiques et perspectives du secteur, etc.) ;
- une analyse des motivations et des attentes des destinataires et des acteurs principaux ;
- ainsi qu'une identification et une hiérarchisation des problèmes à résoudre et de leurs causes probables.

**Etape 2. Analyse des objectifs :** les problèmes que le projet doit résoudre conduisent à définir les objectifs du projet. Pour formuler ou reformuler ceux-ci, on s'attache :

- aux perspectives et à la description de la situation future lorsque les problèmes auront été résolus ;
- à hiérarchiser les objectifs en termes de buts et de finalités. Cette distinction entre but et finalité n'est pas uniquement discursive. Elle répond à un besoin méthodologique en ce sens que la construction d'un cadre logique nécessite, pour des raisons de cohérence, de limiter le

nombre d'objectifs principaux (ou finalités) et d'objectifs spécifiques (ou buts). L'emploi d'outils spécifiques de hiérarchisation des objectifs, de type « arbre d'objectifs », s'avère ici nécessaire.

- à la visualisation de l'ensemble des relations entre les ressources mobilisées (intrants) et les finalités.

**Etape 3. Analyse de la logique du projet :** La logique d'un projet comprend généralement quatre 'niveaux de valeurs' différentes : les ressources ou intrants (ressources humaines, financières, matérielles, réglementaires, etc.), les produits (extrants), les buts principaux (objectif spécifique) et les finalités (objectifs globaux). Les niveaux sont reliés entre eux par des hypothèses de causalité (de type : si...alors...). Tandis que les trois premiers niveaux sont internes au projet, le quatrième, celui des finalités, relie le projet à un programme ou à une politique plus large, fréquemment défini hors de la sphère de gestion.

Ainsi, la logique verticale, ou logique d'intervention, s'attache à clarifier les intentions du projet, tout en précisant les incertitudes liées à leur matérialisation. Elle permet de décrire l'articulation logique des différents 'niveaux de valeurs' ainsi que les conditions critiques qui influencent leur réalisation. Ces conditions critiques, internes ou externes au projet, sont l'expression des contraintes nécessaires pour passer d'un 'niveau de valeur' à l'autre du cadre logique. Or ces passages dont les liens de causalité sont exprimés sous forme d'hypothèses se doivent d'être formulés.

La logique horizontale vise à identifier et décrire avec précision ce qui est produit à chacun des quatre 'niveaux de valeur', et ceci à l'aide de deux concepts : les Indicateurs Objectivement Vérifiables (IOV) et les Sources de Vérification (SdV) de ces indicateurs.

Les IOV correspondent à l'ensemble des critères susceptibles de démontrer que ce qui est attendu est bien produit. Ils sont en quelque sorte des « opérationnalisations » des déclarations d'intention de la logique d'intervention, constituant ainsi des preuves tangibles de réussite pour chacun des quatre niveaux : intrants, extrants, buts et finalités.

Les SdV de ces indicateurs sont décrites en précisant les moyens et les sources d'informations nécessaires à l'obtention des données recherchées.

L'imbrication des éléments du cadre logique peut être schématisée dans la forme suivante :

### Encadré 22 - Les éléments de base du cadre logique

Logique d'intervention	Indicateurs objectivement vérifiables	Sources de vérification	Conditions critiques
<p><b>FINALITE</b>  <b>Impact global</b>                      Raisons d'être du programme dans lequel le projet s'insère</p>	<p>Mesure de la diminution des besoins qu'il faut satisfaire</p>	<p>Sources d'information externes au projet</p>	<p>Hypothèse de causalité du programme                      « si...but, alors...finalité »</p>
<p><b>BUT</b>  <b>Impact spécifique</b>                      Raison d'être du projet</p>	<p>Éléments de réussite du projet du point de vue des destinataires directs</p>	<p>Sources d'information externes au projet</p>	<p>Hypothèse de causalité du projet                      « si...extrants, alors... but »</p>
<p><b>EXTRANTS</b>                      Ensemble des réalisations concrètes à produire pour atteindre le but</p>	<p>Indicateurs permettant de quantifier et de qualifier l'ordre de grandeur des extrants et la corrélation avec les dates de production prévue</p>	<p>Sources d'information internes au projet</p>	<p>Hypothèse de réalisation                      « si...intrants, alors...extrants »</p>
<p><b>INTRANTS</b>                      Ensemble des ressources nécessaires pour produire les extrants concrets nécessaires pour atteindre le but</p>	<p>Suivi des intrants</p>	<p>Sources d'information internes au projet</p>	<p>Conditions préalables au démarrage du projet</p>

Source : adapté de ACIDI (1981) Guide d'utilisation de la Méthode du Cadre Logique dans la gestion et l'évaluation des projets de coopération de l'ACIDI, Toronto : ACIDI.

### ● Application aux Fonds structurels

L'apport de cet outil pour l'évaluation des Fonds structurels est sa capacité de clarification et de structuration de l'action publique et de ses intentions. L'apport de l'outil est d'autant plus net que l'action publique peut se structurer autour d'un seul objectif principal, compris ici comme finalité. La technique du cadre logique intervient dans ce cas comme une démarche d'analyse de la cohérence de l'ensemble.

Dans le cadre d'un programme complexe, l'outil conduit à structurer les objectifs sous forme d'un arbre dont le tronc est la finalité et qui se ramifie en sous-programmes, axes, mesures et projets. Une telle interprétation complète par cascade a l'avantage de la simplicité, mais se révèle souvent inopérante pour appréhender une réalité complexe.

### ● Points forts et limites d'utilisation

- En tant que processus analytique de conception et de suivi de projet, la technique du cadre logique se révèle, à l'usage, un cadre structurant apprécié par les praticiens. C'est un outil synthétique qui clarifie les articulations entre moyens et fins, les contraintes exogènes et les solutions pour se prémunir contre leur influence. De plus, l'outil facilite la mise en place d'un système de contrôle et de suivi dès la conception du projet. Il permet ainsi de délimiter précisément les responsabilités de chaque intervenant dans l'exécution du projet.
- Cette technique fournit un cadre pour l'examen de la logique d'un projet, mais ce sont les responsables du projet qui doivent remplir le cadre. L'outil doit donc être mis en œuvre avec la participation des partenaires du projet.
- La technique du cadre logique est plus difficilement applicable dans le cas de projets complexes et de programmes à visées multiples.
- Bien que cet outil se prête davantage à une utilisation dans le cadre d'évaluation ex ante et de suivi, pour les évaluations ex post, il peut être utilisé en complément d'autres approches moins systématiques qui sont plus focalisées sur les principales questions d'évaluation.
- La présence de plusieurs décideurs (avec des intérêts parfois divergents) peut conduire à une conception de programme aux « contours flous » et imprécis dans ses mécanismes. Seules la présence de phases intermédiaires de réflexion collective et l'émergence régulière de

consensus forts peuvent garantir que l'on a bien tiré parti du potentiel de l'outil.

- Si l'application de la technique du cadre logique est perçue comme un simple exercice technocratique obligatoire au niveau de la conception du projet, alors sa capacité à refléter l'évolution de l'état de celui-ci au cours de son application en sera fortement amoindrie. Cela peut porter préjudice à l'ensemble des documents obtenus par extension, comme par exemple les termes de référence d'une évaluation.

#### ● Pour en savoir plus...

Agence canadienne de développement international (1981), *Guide d'utilisation de la Méthode du Cadre Logique dans la gestion et l'évaluation des projets de coopération de l'Agence canadienne de développement international*, Toronto : ACDI.

Manuel de référence de la division de l'évaluation de l'Agence canadienne de développement international

Commission des communautés européennes (1993), *Gestion du cycle de projet, approche intégrée et cadre logique*, Bruxelles : Commission des communautés européennes, série méthode et instrument pour la gestion du cycle de projet, n° 1.

Brochure de présentation de la gestion du cycle de projet

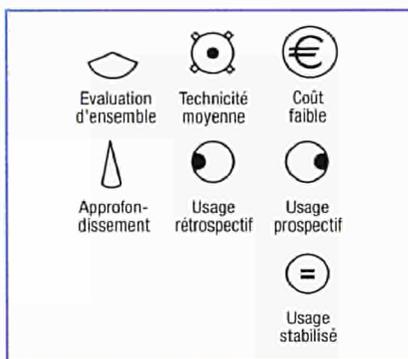
Swedish International Development Cooperation Agency, Department for Policy and Legal Services (1996), *Guidelines for the Application of LFA in Project Cycle Management*.

Guide d'utilisation de la technique du cadre logique.



## ■ Metaplan®

La technique Metaplan® (une marque déposée d'un cabinet allemand) permet de structurer des discussions de groupe. Conçue à l'origine pour faciliter la prise de décision collective dans les milieux de l'urbanisme, elle a donné lieu à de nombreuses variantes. Elles sont toutes basées sur les mêmes principes : solliciter les avis des participants à une réunion, structurer et synthétiser l'information recueillie et la regrouper sous une forme consensuelle grâce à une pondération par les participants des idées issues de la réunion.



### Exemple : Lancement d'une évaluation intermédiaire du programme Objectif 2 en Franche Comté

La première réunion du groupe de pilotage de l'évaluation avait comme principal objectif de produire la liste des principaux impacts attendus dudit programme. L'outil METAPLAN® a été utilisé à cette fin.

La préparation du plan de discussion et l'animation de la séance ont été pris en charge par des consultants extérieurs formés à l'outil. Les participants (13 personnes) étaient essentiellement des administratifs de collectivités locales dont un représentant du commanditaire de l'évaluation.

La séance qui a duré environ deux heures, avait donc pour objectif de faire ressortir les principaux impacts attendus du programme, sachant que ces impacts devraient servir de base à l'évaluation et en particulier à la formulation de questions évaluatives spécifiques qui devaient faire l'objet d'évaluations approfondies.

Les trois questions qui ont servi à bâtir le plan de discussion de la séance METAPLAN® ont été formulées de la sorte : « Que diriez-vous de l'utilité du programme ? », « Sur quel impact voulez-vous une information ? » et enfin « Quels sont les deux impacts à évaluer ? », cette dernière question permettant de hiérarchiser les priorités.

Grâce à cette structuration, le groupe de pilotage a rapidement établi de façon consensuelle une liste de cinq impacts qui sont apparus comme particulièrement importants à évaluer ; à savoir la « Requalification », « l'Adaptation des formations », « l'Attrait du territoire », la « Reconversion/Diversification et les « Emplois d'initiative locale ».

Source : Rapport C3E (1998) L'évaluation intermédiaire des interventions structurelles dans la zone d'objectif 2 de Franche-Comté.

### ● A quoi sert l'outil ?

Metaplan® permet de rassembler et de structurer de l'information provenant de points de vue multiples, et d'en produire une synthèse acceptable par la plupart des participants à la réunion. Un tel outil répond à plusieurs types de besoins : extraire des informations et les agréger grâce à des questions posées de manière adéquate, tout en respectant les points de vue de tous les membres du groupe ; augmenter l'acceptation des décisions prises au sein du groupe, du fait de la prise en compte de la totalité des points de vue, y compris ceux qui ne sont pas dominants au sein du groupe.

### ● Dans quels cas l'utiliser ?

Metaplan® est un outil de discussion et non de décision. Il permet d'identifier les freins et les résistances au bon accomplissement d'une action ou d'une stratégie et de mettre en avant des solutions susceptibles de résoudre ces problèmes. Il permet de structurer une action, composée de priorités multiples. Avec cet outil, il est possible de diminuer les jeux de pression dans un groupe, notamment grâce à la participation de toutes les parties prenantes et au recours de techniques de discussion et de pondération. Enfin, un tel outil peut participer à la construction d'une représentation commune des problèmes et des solutions à mettre en œuvre, ce qui permet de susciter une volonté commune d'agir.

### ● Comment le mettre en œuvre ?

Le déroulement des étapes doit obligatoirement être assuré par un modérateur externe, neutre et bien formé à l'outil. C'est à lui qu'incombe la préparation du plan de discussion, l'animation de la réunion et la présentation des résultats.

**Etape 1. Préparation du plan de discussion :** Le plan de discussion est composé de trois questions qui vont fournir le substrat de la session. La formulation et l'articulation de ces questions est importante en ce sens qu'elles doivent être suffisamment ouvertes pour laisser des latitudes de réponse, et tournées de telle manière que les participants puissent et aient envie d'y répondre. L'ordre dans lequel sont posées les trois questions est également important. La première question ou « question d'appel », sert à illustrer le fonctionnement de la méthode aux participants et à se doter d'un référentiel commun. La seconde question ou « question carte » est au

cœur de la démarche : elle est un approfondissement de la première question ; son traitement sera plus élaboré en ce sens que les réponses seront discutées plus en détail. La troisième question, dite « question de pondération », servira à choisir parmi les réponses données, afin d'établir des priorités d'action. Ces trois questions doivent être préparées en avance par le modérateur qui veille à leur bonne articulation et à la progressivité de la démarche.

**Etape 2. Déroulement de la séance de travail :** La séance proprement dite se fait en groupe de travail unique. En fonction de la taille du public, plusieurs sous-groupes peuvent être créés lors du travail sur la deuxième question (question carte). Il est par contre important que le travail organisé autour des première et troisième questions (question d'appel et question de pondération) se déroule en plénière.

Le modérateur pose la « question d'appel » et demande des réponses structurées en affirmations simples, qu'il note au fur et à mesure sur des post-it (30 au maximum) et qu'il fixe sur une feuille de paperboard en formant des « pavés » qui regroupent les réponses de nature analogue ou relevant d'une même logique. Cette première question permet aux participants de créer une représentation commune des problèmes et de mettre les mêmes mots sur les mêmes concepts et de comprendre la méthode de discussion.

La « question carte » est la plus importante car elle traite du véritable objectif de la réunion et servira de support à la discussion de groupe. Pour apporter une réponse à cette question, les participants seront divisés en groupes de 2 à 3 personnes, chargées d'apporter des réponses sous la forme de courtes phrases qu'une personne du groupe écrit sur des post-it (3 à 5 par groupe). Le modérateur collecte ces réponses puis, en les lisant une par une distinctement, les affiche sur le paperboard en essayant de les regrouper par « pavés ». Les réponses peuvent être contestées par les participants ou faire l'objet d'une demande de précision. Dans ce cas le modérateur appose ou dessine un « éclair » sur le post-it concerné (règle de l'éclair) et des éléments de précision à la question peuvent être apportés et faire l'objet d'un post-it de couleur différente apposé à coté de l'éclair. Au fur et à mesure et avec l'aide des participants, le modérateur termine de composer les « pavés » de réponses (6 post-it par « pavé » au maximum) qui doivent correspondre à des logiques analogues. Les participants doivent ensuite se concerter pour trouver des titres décrivant le contenu des différents « pavés ».

La « question de pondération » va permettre de montrer quelle importance les différents participants accordent aux différents « pavés ». La réponse à cette question s'effectue à l'aide de gommettes que les participants collent sur les pavés qui leur paraissent le mieux répondre à la question.

Voici un exemple pour ces trois types de questions :

<p><i>1. Question d'appel</i></p> <p>« Qu'attendez-vous d'une politique de formation ? »</p>	<p><i>2. Question carte</i></p> <p>« Sur quels éléments plus précis aimeriez-vous que l'évaluation d'une politique de formation porte ? »</p>	<p><i>3. Question de pondération</i></p> <p>« Quels sont les éléments qui devraient être traités en priorité par l'évaluation ? »</p>
--	---	---

La première question permet de lancer le débat sur les impacts d'une politique de formation et d'engager la réflexion sur les typologies d'effets possibles. Les réponses données permettront aussi de guider les réponses qui seront données à la question carte.

La seconde question permet d'approfondir les éléments abordés au cours de la première, mais aussi d'ajouter des éléments d'information supplémentaires.

La troisième question permet de sélectionner les éléments de réponse qui apparaissent comme étant prioritaires.

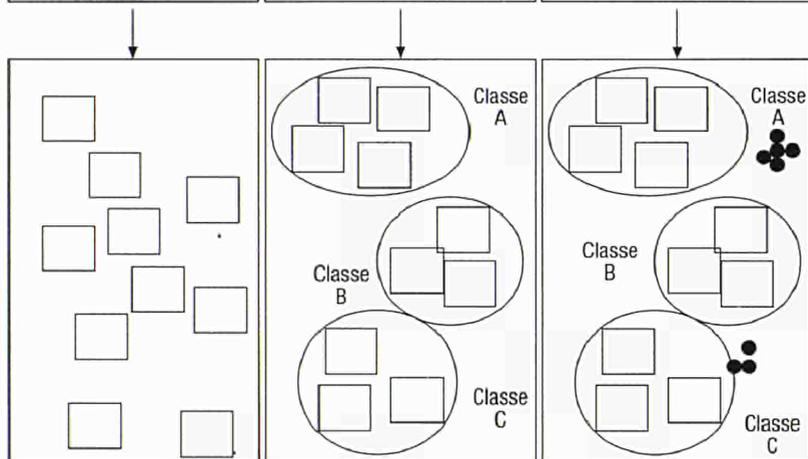
Voici une description stylisée du déroulement des deux dernières phases d'un Metaplan® :

### Encadré 23 - Structuration d'une évaluation avec la technique Metaplan®

Les réponses à la « question carte » sont notées sur des Post-it et positionnées sur le Paperboard

Les réponses sont groupées en « pavés » relevant d'une logique commune. Un titre est donné à chaque pavé

Les participants hiérarchisent les pavés en réalisant une pondération avec des gommettes



#### ● Application aux Fonds structurels

Metaplan® est un outil qui présente plusieurs possibilités d'utilisation dans le cadre de l'évaluation des Fonds structurels :

Il peut servir à la clarification de la logique des actions et à l'identification des mesures qui nécessitent un approfondissement particulier. C'est un outil de structuration qui peut être utilisé par les gestionnaires lors de réunions préparatoires au lancement d'une évaluation.

Il peut servir d'appui à la définition des plans d'évaluation et des cahiers des charges. Metaplan® facilite la définition des questions d'évaluation (par exemple à partir des questions proposées plus haut) et, à partir de ces questions, la sélection des critères d'évaluation.

L'avantage de cet outil est tout à fait approprié à une démarche pluraliste d'évaluation. Celle-ci permet d'associer les parties prenantes dans la sélection des critères d'évaluation et ainsi d'augmenter l'utilisation des résultats de l'évaluation, les critères étant acceptés par une majorité des parties-prenantes.

### ● Points forts et limites d'utilisation

- Metaplan® est un outil de structuration et de formalisation, qui permet une appropriation des conclusions par les personnes qui ont participé à la réunion ; ceci quel que soit le nombre de personnes. Il facilite donc la construction d'un consensus.
- Metaplan® nécessite la présence d'une personne formée pour donner de bons résultats. En effet, la formulation et l'enchaînement des trois types de questions sont importants et doivent être préparés quant au fond et à la forme. Il en va de même pour l'animation du groupe. Du fait de son apparente simplicité, il peut être tentant d'organiser des séances de travail avec cet outil, malgré un manque de formation. Les résultats du travail risquent alors d'être décevants et de présenter des répercussions sur l'ensemble du groupe.
- Cet outil permet l'identification des points de discordance et permet de surmonter les situations de blocage qui tendent à apparaître dans les réunions. Chaque idée proposée est en effet prise en compte et discutée, ce qui permet en règle générale d'éviter les situations où quelques individus dominent l'ensemble du groupe.
- Si Metaplan® est un outil qui permet la formalisation des idées, ce n'est pas un outil de création d'idées nouvelles. Cet outil n'exprime rien de plus que ce qui est dans l'esprit des participants à la séance. Si un accord implicite sur les solutions à mettre en œuvre existe (exemple : pensée unique), des solutions fausses mais faisant l'objet d'un consensus peuvent émerger. En d'autres termes, un tel outil peut contribuer à verrouiller des situations.

### ● Pour en savoir plus...

Schnelle, E. (1979), *The Metaplan®-Method : Communication Tools for Planning and Learning Groups*. Hamburg: Quickborn.

## Outils pour observer les changements sur le terrain

L'enquête par questionnaire est l'outil d'observation le plus simple. Administré par la poste, par téléphone, ou plus récemment par courrier électronique, il permet de recueillir des données nombreuses, homogènes et susceptibles de faire l'objet de traitements statistiques. Cet outil n'est pas adapté à l'évaluation ex ante.

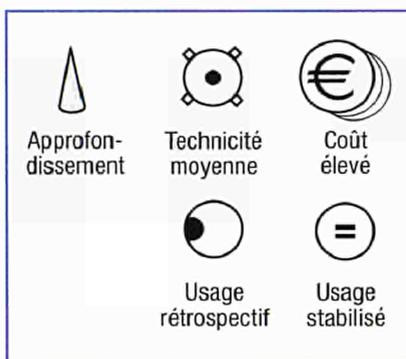
Plusieurs outils présentés précédemment sont également adaptés à l'observation de terrain dans le cas d'une évaluation approfondie : **entretien individuel**, **entretien de groupe** (*focus group*) et **étude de cas**.

Enfin, l'**observation ethnographique** permet de collecter une information complète et approfondie dans le cas où les populations sont difficiles d'accès ou réticentes à participer à une enquête.



## ■ Enquête par questionnaire

L'enquête par questionnaire consiste à poser une série de questions standards, selon un schéma préétabli, à un échantillon d'individus considéré comme représentatif de la population à observer ou quelquefois à l'ensemble de cette population de façon exhaustive.



L'enquête par questionnaire et l'échantillonnage se sont développés à partir des premières prévisions électorales précises obtenues entre les deux guerres par les instituts de sondage d'opinion. Considérée comme un des outils de base des sciences sociales, l'enquête par questionnaire est aujourd'hui utilisée par la majorité des instituts publics et privés d'étude, de recherche, de statistique et d'évaluation.

### Exemple : Une enquête réalisée auprès d'employeurs sur les effets des politiques en faveur de l'emploi

Cet exemple correspond à celui de l'Encadré 19.

Le Ministère français du Travail, de l'Emploi et de la Formation Professionnelle utilise fréquemment des enquêtes par questionnaire, en général téléphoniques, pour évaluer les interventions en faveur de l'emploi. Une enquête a été menée en 1990 pour évaluer une mesure d'aide à l'embauche de chômeurs longue durée sous forme d'exonération de charges sociales pour les employeurs. Un questionnaire a été administré, par téléphone, aux employeurs d'un échantillon représentatif d'environ 1000 chômeurs destinataires de la mesure. L'échantillon ne comportait que des employeurs et des chômeurs directement concernés par la mesure.

Dans chaque entreprise, la personne interrogée devait avoir une connaissance suffisante de la décision d'embauche et du rôle joué par l'exemption de charges sociales dans la prise de décision. Pour éviter les réponses biaisées les employeurs ont été avertis que le questionnaire serait traité de façon anonyme. Le questionnaire comprenait plusieurs questions croisées susceptibles de révéler ce qui se serait passé dans la situation contrefactuelle, c'est-à-dire en l'absence d'exonération. Parmi les réponses proposées : aucune embauche supplémentaire, embauche d'une personne présentant des caractéristiques différentes, embauche de la même personne mais dans des conditions différentes, etc. L'enquête a permis d'estimer, pour chaque entreprise, le niveau de l'effet d'aubaine (c'est-à-dire le nombre de chômeurs qui auraient été embauchés même sans l'existence de l'exonération), ainsi

que l'effet de substitution (c'est-à-dire dans quelle mesure l'exonération a incité l'employeur à recruter un chômeur éligible plutôt qu'une autre personne non éligible). L'observation de l'ensemble de l'échantillon a permis d'estimer des coefficients moyens d'effets d'aubaine et d'effets de substitution. L'échantillon étant représentatif, ces coefficients ont été généralisés à l'ensemble des entreprises exonérées pour estimer l'emploi créé par la mesure.

Source : « Aides publiques à l'embauche : quels effets sur l'emploi » - Premières synthèses, N°6, mai 1990, Ministère du Travail de l'Emploi et de la Formation Professionnelle.

### ● A quoi sert l'outil ?

Dans le contexte de l'évaluation, les enquêtes par questionnaire sont essentiellement utilisées pour collecter l'opinion des destinataires. Les questions posées peuvent être de nature descriptive (ex : par quel canal avez-vous été averti de l'existence de l'aide ?), de nature normative (ex : le service qui vous a été fourni répond-il à un besoin de première priorité pour votre organisation ?), ou encore de nature causale (ex : qu'auriez-vous fait en l'absence de l'intervention évaluée ?).

### ● Dans quels cas l'utiliser ?

L'avantage de l'enquête par questionnaire est de produire des résultats applicables à l'ensemble de la « population » observée, soit directement si l'enquête est exhaustive, soit par extrapolation à partir d'un échantillon représentatif. Cependant, la taille minimum d'un échantillon généralisable est de l'ordre de 1000 unités. L'enquête par questionnaire doit donc plutôt être utilisée si la population à observer est nombreuse et homogène (ex : stagiaires, entreprises, voyageurs, etc..).

Le questionnaire est le plus utile quand l'enquêteur a une idée précise et claire de ce qu'il/elle veut observer. Dans ce cas, les enquêtes les plus simples consistent en une série de réponses à des questions fermées où les répondants choisissent parmi des réponses prédéterminées. On peut cependant se trouver face à des cas où il est utile d'employer des questionnaires où les réponses peuvent être orientées, particulièrement lorsque l'enquêteur n'a pas au préalable une idée claire de l'opinion des répondants ou de la façon dont a fonctionné l'intervention. Dans ces cas, une combinaison de questions ouvertes et de questions fermées peut être utile. Il faut cependant noter que les questions ouvertes sont plus

complexes à analyser et plus coûteuses car elles exigent que l'enquêteur détermine, regroupe et codifie les réponses similaires, une fois que les données primaires ont été recueillies. Cet outil est particulièrement bien adapté pour produire une information descriptive et pour réaliser des classements dans différentes catégories (ex : comment les entreprises aidées se répartissent-elles en fonction de leur taille, de leur secteur d'activité, de leur ancienneté, de la nature de leur marché, etc.). Il peut être utilisé pour recueillir des informations aussi bien qualitatives que quantitatives. Cependant les données subjectives doivent être interprétées avec prudence quand l'évaluation cherche à démontrer des causalités de liens ou à fournir des informations normatives

Un même échantillon peut être interrogé à deux ou plusieurs périodes successives pour fournir des données longitudinales permettant de suivre les destinataires et d'obtenir des estimations sur la pérennité des effets. On utilise par exemple des enquêtes par panel pour évaluer la durabilité des résultats obtenus en matière d'emploi (ainsi, on pourra interroger les stagiaires avant ou après leur formation et à une ou plusieurs reprises par la suite).

### ● **Comment le mettre en œuvre ?**

Pour réaliser efficacement une enquête par questionnaire, quatre étapes sont à respecter :

**Etape 1. La conception du questionnaire :** Il convient en premier lieu de définir rigoureusement les objectifs de l'enquête et de concevoir les questions appropriées. Les questions devront être formulées clairement et d'une façon intéressante. Il est important de ne pas multiplier les questions secondaires qui pourraient brouiller les répondants et/ou les enquêteurs, et nuire ainsi à la qualité de l'enquête. Lorsqu'il sera nécessaire d'établir une comparaison avec d'autres données statistiques, les termes utilisés dans le questionnaire seront soigneusement choisis pour garantir la cohérence.

La plupart des questionnaires sont conçus à l'origine pour poser des questions fermées (choix entre un nombre limité de réponses), mais nombreux sont ceux qui posent également quelques questions ouvertes (comportant un espace vide que le répondant doit remplir lui-même), ou semi-ouvertes (liste d'options comportant une réponse « autre » permettant au répondant d'émettre une opinion différente de celles prévues par l'enquêteur). Dans ces deux derniers

cas cependant, les réponses doivent être codifiées après collecte des données, ce qui peut poser des problèmes d'interprétation et de comparabilité si les répondants soulèvent des points très différents. Ceci dit, il s'agit souvent d'un moyen très utile de s'assurer que des points qui auraient échappé au départ à l'enquêteur font bien l'objet d'une analyse s'ils se révèlent importants pour les destinataires du programme.

Il est essentiel de s'assurer que la formulation des questions n'influence pas les réponses. Les questions qui demandent au répondant de rapporter des événements, situations ou décisions passés sont généralement considérées comme problématiques, ainsi que les questions « à double négation ». L'importance de la formulation des questions dépend de la méthode utilisée. Ainsi, dans le cas d'une enquête envoyée par la poste ou toute autre enquête que le répondant complète lui-même, il est essentiel de s'assurer que les questions seront bien comprises. Ce point est moins important dans le cas d'enquêtes en face à face où l'enquêteur a la possibilité de clarifier la formulation ou d'aider le répondant qui ne comprendrait pas la question. Encore faut-il que l'enquêteur ait un niveau d'expérience et de compréhension suffisant pour le faire. Il est donc préférable de ne pas compter sur l'intervention de l'enquêteur, à moins qu'il ne soit très expérimenté, pour lever les ambiguïtés.

Les questions peuvent être nominales (ex : activité de votre organisation - agriculture, industrie, commerce, autre), ordinales (ex : qu'avez vous le plus apprécié dans cette formation : sa qualité, son prix, sa pertinence ?) ou numériques (ex : combien de jours de formation avez vous reçus ?).

Si les questions sont normatives ou causales, on aura recours à la triangulation. La triangulation consiste à poser plusieurs questions voisines qui permettent de mieux cerner les limites d'une réponse qui présente un risque de biais. Par exemple, les trois questions suivantes permettent d'estimer un effet d'aubaine : en l'absence de l'aide, auriez vous réalisé un investissement de même taille, un peu plus petit, beaucoup plus petit ? ... à la même date, un peu plus tard, beaucoup plus tard ? ... avec la même technologie, avec une technologie différente ?

**Etape 2. L'échantillonnage :** La population à enquêter doit être précisément délimitée (ex : toutes les entreprises de 1 à 50 salariés aidées au cours des 5 dernières années) et on doit recueillir les informations disponibles sur la structure de cette population. On doit vérifier si la population est connue (noms, adresses et numéros de téléphone) ou inconnue.

Lorsque l'on utilise un échantillon, celui-ci constitue un modèle réduit de l'ensemble de la population. Pour être représentatif sa taille doit être suffisante et répondre à des règles statistiques précises. Les indications qui suivent sont des règles extrêmement simplifiées qui doivent être vérifiées au cas par cas par un professionnel. La taille standard d'un échantillon représentatif est de l'ordre de 1 000 individus. La taille de l'échantillon ne dépend pas de celle de la population. Un échantillon de 1 000 personnes représentera aussi bien une population de 2 000 000 de personnes qu'une population de 6 000 personnes. Par contre, la taille de l'échantillon dépend fortement de la précision attendue. Il faut multiplier le nombre de questionnaires par 4 pour doubler la précision. Finalement, la taille de l'échantillon doit aussi tenir compte de l'hétérogénéité des réponses. Pour les questions les plus importantes, on doit faire une estimation anticipée de la proportion des répondants qui donneront telle ou telle réponse. Si on s'attend à ce que 20 à 80% des répondants choisissent une des réponses, un échantillon de 1 000 individus suffira pour que cette réponse soit valide. Si on s'attend à ce que 5% ou moins des répondants choisissent une des réponses, l'échantillon devra être plus grand. L'enquête fournit un chiffre valide pour une réponse si 100 répondants au moins ont choisi cette réponse.

Plusieurs techniques sont proposées dans le tableau qui suit pour construire un échantillon représentatif. Seules les deux dernières sont applicables si la population est peu ou pas connue.

**Encadré 24 - Les techniques d'échantillonnage**

Technique	Principe
<b>sélection aléatoire simple</b>	Chaque unité d'observation est tirée au sort. Elle a une chance égale de faire partie de l'échantillon et ne peut être sélectionnée qu'une seule fois
<b>sélection systématique</b>	On choisit 1 individu sur 10 ou sur 100, le premier individu étant choisi au hasard
<b>sélection aléatoire stratifiée</b>	Lorsque la population est peu homogène, elle est divisée en sous-groupes homogènes et un échantillon est ensuite pris dans chaque strate constituée
<b>sélection par grappes</b>	La population est divisée en sous-catégories ou grappes (entreprises aidées par ville,...) ; certaines grappes sont choisies aléatoirement puis tous les individus faisant partie d'une grappe sont interrogés
<b>par quota</b>	On construit un modèle réduit regroupant les caractéristiques connues de la population étudiée ; les enquêteurs sont ensuite chargés de trouver des individus correspondant à ces caractéristiques
<b>« boule de neige »</b>	A partir d'un échantillon comportant un nombre restreint de personnes, on adjoint des unités supplémentaires avec lesquelles les premières sont en relation

**Etape 3. Pré-test :** Un pré-test ou test-pilote du questionnaire, auprès d'un nombre limité de répondants, est indispensable ; il permettra de critiquer la forme et le contenu du questionnaire et d'en apprécier ainsi la pertinence. La rédaction finale du questionnaire sera ensuite réalisée. Ceci est particulièrement important dans le cas d'enquêtes où le répondant remplit lui-même le questionnaire ou dans le cas d'enquêtes à réponses complexes ou « filtres » (moyens de guider les répondants vers les questions qui correspondent le mieux à leur propre expérience : ainsi par exemple, s'assurer que seuls les usagers d'un service émettent un avis sur celui-ci et que ceux dont le jugement ne peut être fondé s'abstiennent).

**Etape 4. L'administration du questionnaire :** Les questionnaires peuvent être soit :

- envoyés par courrier ou par Internet ou distribués manuellement (par exemple, à l'occasion d'un événement particulier sur lequel l'enquêteur souhaite recueillir des opinions). Il s'agit alors d'enquêtes en auto-administration, ou

- complétés avec l'aide d'un enquêteur (par exemple par entretien téléphonique ou en face à face).

En cas d'auto-administration, il faut s'attendre à un taux de réponse faible (quelquefois 20% pour les réponses à un premier envoi). Le taux de réponse peut être augmenté par des relances et, le cas échéant, par une information dans la presse spécialisée. Dans le cas d'envoi par la poste, les réponses seront plus nombreuses si une enveloppe pré-imprimée et pré-affranchie a été jointe au questionnaire). Selon la difficulté du contexte, un bon taux de réponse varie entre 60 et 90%. En cas d'enquêtes répétées sur un même panel, on considère qu'il y a une perte inévitable de l'ordre de 10% par an.

Plus on aura exposé la finalité de l'enquête de façon à convaincre les répondants potentiels de l'importance de leur contribution, plus ils seront enclins à répondre. Quand c'est possible, il faudra garantir au répondant le traitement confidentiel de ses réponses. Les enquêteurs doivent être informés des « règles de conduite » et être capables de les expliquer clairement aux personnes interrogées. On pourra aussi prévoir d'encourager la participation et la rapidité de la réponse par des « offres ». De nombreuses enquêtes offrent la possibilité de participer à une loterie, etc.. Ceci est moins courant dans le secteur public, mais a été utilisé dans certains contextes. Dans certains cas, il peut être utile de proposer au répondant de revoir un condensé du rapport d'évaluation. Il est également recommandé d'adresser une lettre aux répondants pour les remercier de leur contribution. Un contrôle par sondage est généralement effectué après enquête pour vérifier que les entretiens ont bien eu lieu.

**Etape 5. L'encodage des données :** L'étape suivante est la codification et l'enregistrement des réponses, presque toujours sous forme informatique pour une analyse ultérieure. A ce stade, des redressements sont généralement effectués pour tenir compte des imperfections de l'échantillon, notamment pour corriger les biais introduits par les non-réponses.

**Etape 6. L'interprétation et la diffusion des résultats :** L'étape finale consiste à analyser la signification des résultats de l'enquête. Ceci implique une bonne compréhension des objectifs de l'enquête et ses limites (par exemple, les contraintes liées à la taille de l'échantillon, la fiabilité des réponses, etc..) Il est donc important que la phase d'analyse soit conduite par un enquêteur expérimenté qui appréhende bien ces problèmes et qui ait été impliqué dès le départ dans la conception de l'enquête. Pour que l'enquête soit utile, ses résultats doivent être présentés de façon à être

accessibles et « parlants » pour les décideurs. De nombreux logiciels spécialisés peuvent aider à réaliser cette étape. Par conséquent, il est généralement important de s'assurer que le rapport final fournit suffisamment de précisions sur la façon dont l'enquête a été menée (précisions qui apparaissent dans une annexe technique destinée à ceux qui veulent vérifier sa validité) mais que le texte du rapport lui-même se limite à une explication claire et nette des conclusions.

### ● **Application aux Fonds structurels**

L'enquête par questionnaire est adaptée à l'observation des résultats et des impacts d'une intervention qui touche une population de taille importante de façon suffisamment homogène. C'est donc probablement un outil à réserver à l'évaluation rétrospective d'une intervention simple, par exemple l'évaluation d'une action de formation destinée aux chômeurs. Cependant, cet outil peut également jouer un rôle important dans les analyses « avant-après » où il est utilisé pour établir au début du programme un point zéro destiné à mesurer la progression pendant et après la mise en œuvre de l'intervention.

### ● **Points forts et limites d'utilisation**

- L'enquête par questionnaire a l'avantage de produire une information structurée, quantifiée et généralisable, à condition que l'échantillon ait été bien construit. Elle est donc particulièrement intéressante pour quantifier des indicateurs, y compris des indicateurs de suivi.
- Cet outil est coûteux puisque pour fournir des informations qui sont statistiquement significatives, un échantillon doit compter plus de 1000 unités. Par contre, il permet de toucher une population étendue à un coût relativement restreint.
- Il est parfois difficile de connaître la population à interroger, spécialement si celle-ci est importante et mal cernée par le système de suivi. Dans ce cas, l'outil perd un de ses principaux avantages : la généralisabilité des observations.
- La nécessité de standardiser les réponses à travers des questions fermées peut conduire à une trop grande simplification. Pour cette raison, le questionnaire est habituellement réservé aux situations où les effets de l'intervention ont été préalablement identifiés et clairement compris. Par

opposition, les impacts d'une intervention très innovante seront mieux évalués par un questionnaire aux questions ouvertes et plusieurs autres techniques basées sur des méthodes autres que des enquêtes.

- La qualité d'une enquête par questionnaire peut être affaiblie si certaines réponses sont biaisées par la volonté des répondants qui sont généralement favorables à une intervention dont ils bénéficient.
- L'expérience révèle aussi que de nombreuses personnes répondent aux questions soit par politesse, soit par crainte de paraître ignorantes alors qu'elles ignorent en fait les réponses ou ne se sentent pas capables d'émettre un jugement. Pour être fiable, un questionnaire ne devrait par conséquent comprendre que des questions que tous les répondants interpréteront de la même façon et auxquelles ils pourront répondre avec exactitude.

#### ● Pour en savoir plus...

Javeau, C. (1992) *L'enquête par questionnaire, manuel à l'usage du praticien*, Edition de l'Université de Bruxelles.

Guide pratique en quinze étapes.

Dussaix, A.-M., Grosbras, J.-M. (1996), *Les sondages : principes et méthodes*. Collection Que sais-je ? Paris : Presses Universitaires de France.

Pratique et accessible, il contient les informations essentielles sur la technique et pratique du sondage.

Grangé, D., Lebart, L., ed. (1994), *Traitements statistiques des enquêtes*, Paris : Dunod.

Nombreuses informations techniques et statistiques, fournies sous forme accessible.

Fink, A. (1995), *The Survey Handbook*, London : Sage publications.

Guide pratique complet.

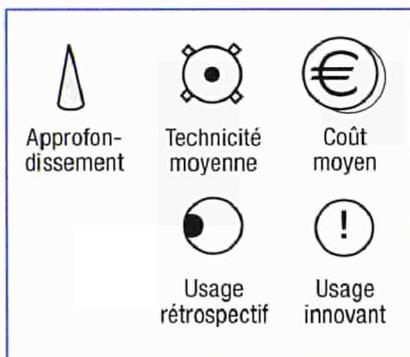
Henerson, M.E., Morris, L.L., Fitz-Gibbon, C.T. (1987) *How to measure Attitudes?*, London : Sage publications.

Traite de la construction et de l'utilisation d'échelles de mesure.



## ■ L'observation ethnographique

La technique de l'observation ethnographique consiste à observer l'activité quotidienne des acteurs qui mettent en œuvre le programme ou des destinataires qui entrent en contact avec le programme. Par exemple, l'enquêteur assiste au processus de sélection des projets ou il participe à un stage de formation. L'enquêteur cherche donc à observer et à comprendre la situation « de l'intérieur ».



Cette technique est particulièrement utile lorsque l'on doit observer une situation mal connue ou lorsque l'on soupçonne que la même situation est comprise très différemment selon que le point de vue est « extérieur » ou « intérieur ». L'observation ethnographique est la seule technique permettant d'observer des domaines où il est difficile de mener ouvertement des activités de recherche.

L'observation ethnographique est née, il y a plus d'un siècle, avec la science du même nom.

### Exemple : Evaluation d'un service public de logements sociaux

L'observation ethnographique a été appliquée à l'évaluation d'une politique de logement social en France. Il s'agissait d'analyser l'écart entre les objectifs nationaux de cette politique et sa mise en œuvre concrète sur le terrain par les fonctionnaires au contact avec le public.

Pendant plusieurs mois, le chargé d'évaluation a accompagné des équipes de fonctionnaires d'organismes gestionnaires de logements destinés à des ménages à faible revenu. En compagnie de ces fonctionnaires, il a participé aux visites des logements au moment de l'arrivée et du départ des locataires. Ces visites ont pour but de juger l'état du logement et de déterminer le montant des réparations qui incombent soit au locataire, soit à l'organisme gestionnaire.

La position de l'observateur lui a permis de se « fondre dans le décor » et d'observer une cinquantaine de situations dans lesquelles la politique nationale était concrètement appliquée par les fonctionnaires. Les situations observées ont été enregistrées et une vingtaine de séquences-types ont été sélectionnées. Ces séquences ont été rédigées en respectant la confidentialité, puis elles ont été

présentées aux acteurs de la politique dans le cadre de réunions organisées séparément à différents niveaux (fonctionnaires de terrain, responsables des organismes, responsables nationaux).

L'évaluation a permis de montrer comment les règles sont mises en œuvre par les fonctionnaires locaux, comment elles sont interprétées par les locataires, et comment ces derniers influencent leur application.

L'évaluation a mis en avant le fait que les locataires avaient des difficultés à comprendre la justification des règles. Au vu de ces résultats, plusieurs modifications ont été apportées : les directions des organismes ont révisé certaines règles et engagé en examen du recrutement et de la formation de leurs agents.

Source : Philippe Warin, « L'évaluation des usagers au cours des relations de service », Les annales de la recherche urbaine n°47 (1990).

### ● A quoi sert l'outil ?

L'observation ethnographique permet de réaliser une collecte d'information très approfondie sur quelques situations typiques de la mise en œuvre de l'intervention. Elle éclaire en détail les effets de l'intervention et l'influence du contexte. Elle privilégie les points de vue des acteurs clé et des destinataires.

### ● Dans quels cas l'utiliser ?

L'observation ethnographique a été utilisée pour comprendre le fonctionnement des politiques dans les domaines de l'éducation, de la santé (diagnostic, soins, ...), de la justice, de la police, de la recherche scientifique, des transports urbains, etc. Dans le cadre d'une évaluation, cette technique est particulièrement indiquée pour observer les processus d'interaction entre les administrations et leur public.

Dans l'exemple d'observation ethnographique présenté ci-dessus (l'évaluation d'un service public de logements sociaux), l'enquêteur a également effectué une analyse ethnométhodologique. L'éthnométhodologie est apparue dans les années 60 et s'est développée en Europe au début des années 80. Cette science étudie le raisonnement pratique par lequel des individus (fonctionnaires, destinataires, usagers, etc..) jugent quotidiennement les situations qu'ils rencontrent et les actions collectives dans lesquelles ils sont engagés. Si elle est utilisée correctement, cette technique ne perturbe personne et n'influence pas les comportements. On peut donc y avoir recours pour les comportements

spontanés de populations qui ne sont pas disposées à accepter le formalisme des enquêtes par questionnaire. C'est la seule technique applicable lorsqu'il y a de graves difficultés d'accès au terrain, ce qui peut arriver par exemple en cas de conflit au sein des organisations chargées de la mise en œuvre ou lorsque le comportement des destinataires est en partie illégal ou irrégulier. La technique est donc particulièrement intéressante si l'évaluation est lancée dans la perspective de modifier des règlements que l'on soupçonne d'être inefficaces.

En raison de son caractère inductif, l'observation ethnographique permet d'observer les résultats d'une intervention dont on connaît mal le fonctionnement. Elle est particulièrement utile lorsque l'on soupçonne que la réalité n'est pas vue de la même façon par ceux qui mettent en œuvre l'intervention et par le public qui la reçoit ou qui en bénéficie.

Par contre, la technique exige un délai important, alors même que les observations portent sur quelques situations très délimitées et qu'elles peuvent donc poser des problèmes de généralisation. L'outil convient donc plutôt à l'évaluation d'une intervention lourde et relativement homogène. Il exige également de la part de l'enquêteur d'être particulièrement habile à assimiler et à rendre compte fidèlement du comportement des acteurs clé. Il faut parfois un certain temps pour que l'enquêteur se « fonde dans le décor » et garantir ainsi un comportement spontané des participants.

### ● Comment le mettre en œuvre ?

L'observation est généralement mise en œuvre selon les étapes suivantes :

**Etape 1. Choix des situations d'observation :** Les situations à observer sont définies préalablement en fonction des préoccupations des commanditaires de l'évaluation. Il s'agit de situations d'interaction ou de négociation entre les acteurs publics et les destinataires de la politique évaluée.

**Etape 2. Observation :** L'observateur suit et enregistre le déroulement des interactions en s'attachant à perturber le moins possible le comportement des acteurs. Ce travail consiste à prendre des notes et à effectuer des enregistrements audiovisuels (les plus discrets possible).

Cette étape ne peut se limiter à une simple observation directe ; elle doit être complétée par une analyse de type institutionnel, afin d'identifier les champs de contraintes qui pèsent et influent sur les relations entre acteurs.

L'observateur doit recueillir un maximum d'informations, et capter une vue particulière, intime, de la situation donnée.

**Etape 3. Recherche et découpage de séquences-types :** Les événements observés au cours de la seconde étape sont analysés en termes de séquences caractéristiques. Chaque enregistrement est découpé comme on découperait un film en séquences.

L'observateur identifie les « assertions évaluatives », c'est-à-dire les phrases qui comportent un jugement de valeur explicite ou implicite. Les séquences-types et leur analyse sont concentrées sur ces assertions et révèlent la façon dont la politique est jugée sur le terrain. Ainsi employé, l'outil peut apporter un nouvel éclairage sur le bien fondé et sur l'efficacité de la politique.

**Etape 4. Analyse des séquences-types avec les acteurs:** Les séquences-types et les assertions sont réécrites ou modifiées de façon à être rendues anonymes. Elles sont alors transmises à des personnes représentatives de celles qui ont été observées afin de recueillir leurs commentaires et réactions. Cette étape permet de vérifier que l'on n'a pas créé de distorsion en sortant les séquences de leur contexte. Elle fournit, pour chaque séquence, des clefs d'interprétation reconnues et validées par la « communauté » étudiée.

### ● **Application aux Fonds structurels**

L'observation ethnographique a été peu utilisée, à notre connaissance, dans le contexte de l'évaluation des programmes socio-économiques. Cette technique est pourtant bien adaptée à certaines interventions des Fonds structurels, par exemple celles qui s'adressent à des publics difficiles à observer par les techniques d'enquête plus traditionnelles (ex : chômage de longue durée).

L'observation ethnographique est un outil générique qui traite à la fois la collecte, l'interprétation et le croisement des données. Elle partage ces caractéristiques avec la technique de l'étude de cas. Elle est donc particulièrement bien adaptée à l'analyse des effets d'une intervention innovante ou mal connue, et en particulier pour clarifier les facteurs concurrents qui ont une influence sur le succès ou l'échec apparent des interventions communautaires évaluées.

### ● Points forts et limites d'utilisation

- L'observation ethnographique permet de révéler l'écart entre la façon dont l'intervention publique est comprise en haut lieu et la façon dont elle est comprise sur le terrain ; elle met en avant l'interprétation qu'en font les individus en situation opérationnelle.
- L'observation est limitée à un petit nombre de situations. La généralisation n'est donc possible que si l'intervention est suffisamment homogène.
- Elle repose sur des informations spontanées, collectées par un observateur indépendant et expérimenté. La fiabilité de l'observation dépend beaucoup du savoir-faire professionnel de l'observateur-analyste.
- En contrepartie de ses avantages, l'observation ethnographique nécessite une préparation minutieuse pour permettre à l'observateur de s'insérer dans le contexte observé sans le perturber, ainsi qu'un temps de collecte important car il ne maîtrise pas le calendrier de ses observations.
- Cette technique permet de collecter des données dans des situations difficiles où les autres techniques d'enquête sont inapplicables.

### ● Pour en savoir plus...

Coulon, A. (1987), *L'ethnométhodologie*, Paris : PUF, Que sais-je ?

Approche à la fois théorique et pratique, illustrée par la présentation d'expériences.

Garfinkel, H. (1967), *Studies in ethnomethodology*, Englewood Cliffs, NJ : Prentice Hall.

Pour revenir aux sources de l'ethnométhodologie.

Jorgensen, D.L. (1989), *Participant Observation, a Methodology for Human Studies*, London : Sage Publications, Applied Social research methods Series, No 15.

Manuel complet sur une variante de l'observation ethnographique dans laquelle l'enquêteur se place personnellement dans la situation d'un acteur ou d'un destinataire du programme.



## Outils pour analyser les données

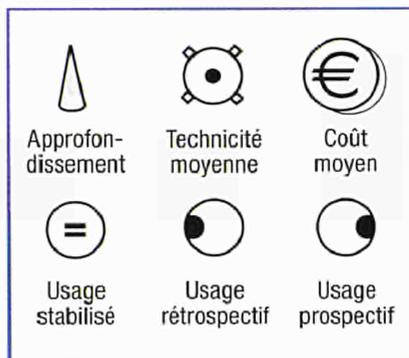
Dans la mesure où l'évaluation porte sur un domaine circonscrit, il est possible de réaliser une analyse approfondie des liens de causalité entre l'action et ses effets. Plusieurs outils sont susceptibles d'être utilisés dans ce contexte :

- L'analyse de la variance (outil non présenté dans cet ouvrage) et l'analyse factorielle permettent de découvrir des correspondances et des dissemblances au sein d'un échantillon d'observations, de constituer des typologies et d'identifier les facteurs exogènes qui sont associés à la production des impacts.
- L'enquête Delphi est classée dans cette section car elle a été conçue pour produire des estimations d'impacts. Elle est bien adaptée aux situations d'évaluation ex ante, pour lesquelles on doit s'appuyer sur les données secondaires. L'outil mobilise et analyse les données par le truchement d'experts. Il se rapproche donc du panel d'experts également adapté à la fonction d'analyse des données.
- Les groupes de comparaison permettent d'estimer les effets nets par différence entre un groupe de destinataires et un groupe de non-destinataires.
- L'analyse de régression permet d'estimer les effets nets et de déterminer si les liens de causalité entre l'intervention et ses effets sont statistiquement significatifs.
- L'étude de cas, l'entretien de groupe et l'observation ethnographique déjà cités comme outils d'observation, sont également des outils de croisement et d'analyse des données dont l'emploi est très souple. Il est même envisageable de fournir une estimation des effets (sous forme de fourchette d'effet maximum et minimum) à l'aide d'études de cas conduites sur une sélection de projets.



## ■ Analyse factorielle

L'analyse factorielle est une technique d'analyse statistique multidimensionnelle utilisée pour classer des volumes importants de données quantitatives concernant des groupes d'individus pour lesquels on dispose d'informations sur de nombreuses caractéristiques (variables). L'outil identifie les ressemblances ou les liens (facteurs) entre individus et entre variables. Il permet « d'échanger de l'information contre de la lisibilité », en produisant une représentation graphique simplifiée des données disponibles faisant apparaître les phénomènes à analyser.



Les origines de l'analyse factorielle remontent au début du siècle mais l'outil s'est développé récemment grâce aux avancées considérables du traitement informatique des données.

### Exemple : Les impacts sur l'emploi dans le secteur coopératif catalan

Cet exemple correspond à celui de l'Encadré 20.

L'utilisation de l'analyse factorielle peut être illustrée par un projet d'évaluation ex post, orienté vers les effets sur l'emploi (création et maintien de postes) d'une mesure de subvention dans le secteur des coopératives catalanes. Dans ce cas, l'évaluation était rendue plus complexe par la diversité des coopératives formant une population hétérogène. Le problème était double : d'une part, il fallait déterminer la pertinence des mesures prises par la région au regard de cette situation ; d'autre part, il fallait évaluer l'importance des impacts sur les coopératives ayant perçu une aide.

Etant donné la variété et la faible connaissance de la population ainsi que de celle des impacts, une analyse factorielle a été retenue pour sa capacité à faire émerger les principales caractéristiques de la population considérée (environ 3 000 coopératives).

La première étape de l'évaluation a consisté à envoyer un questionnaire à l'ensemble des coopératives. Il a été complété par des entretiens. Les données ainsi recueillies ont d'abord été analysées par des méthodes statistiques traditionnelles (recherche des principaux descripteurs de la population, réalisation de tableaux croisés). La deuxième étape de l'analyse des données a consisté à réaliser une analyse factorielle en vue d'isoler les variables susceptibles d'expliquer les variations de l'emploi dans les coopératives : y compris la taille de l'entité, l'âge, le secteur d'activité, le niveau technologique,.....

Source : Comitè Assessor para l'Estudi de l'Organització de l'Administració, Generalitat de Catalunya, 1995.

### ● A quoi sert l'outil ?

Entre autres finalités, l'analyse factorielle permet dans un cadre d'évaluation :

- d'analyser finement les réactions des destinataires et leurs caractéristiques et ainsi de mieux comprendre l'efficacité relative des interventions ;
- d'explorer les corrélations entre phénomènes et donc d'identifier d'éventuels facteurs concurrents qui ont pu amplifier ou réduire l'impact du programme,

### ● Dans quels cas l'utiliser ?

L'analyse factorielle est un outil de traitement des données qui est utilisé par les organismes statistiques nationaux, les instituts de sondage. Il présente l'avantage de traiter des données quantitatives de façon à permettre la construction de typologies. L'outil est donc très utile lorsque l'évaluation doit construire des classes au sein de la population concernée (par exemple : qualification, catégorie socio-professionnelle, criminalité, concentration urbaine, etc.).

Il est important de noter que les classements établis par analyse factorielle ne sont pas choisis au hasard par l'équipe d'évaluation. Au contraire, ils sont extraits d'un vaste ensemble de données et sont construits de façon à permettre la meilleure discrimination possible au sein de ces données.

### ● Comment le mettre en œuvre ?

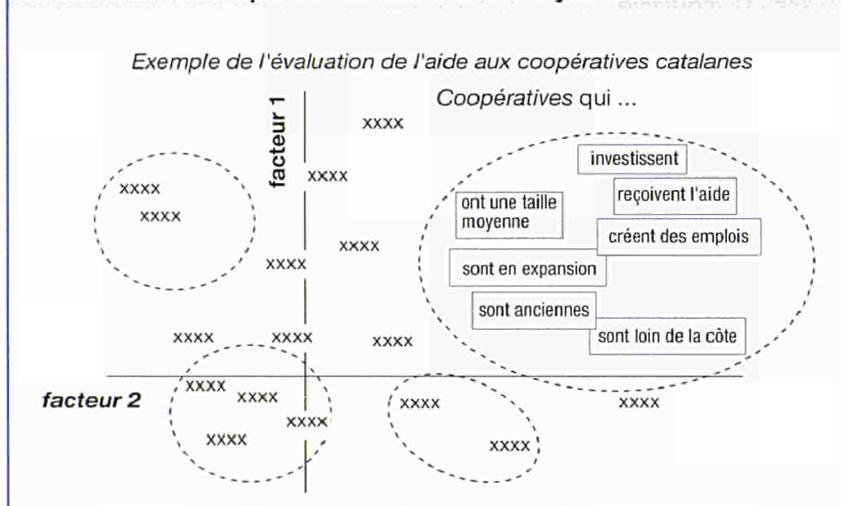
**Etape 1. Constitution d'un tableau de données :** les données traitées proviennent du suivi de l'information ou, plus souvent, d'une enquête par questionnaire. L'ensemble des données recueillies est regroupé dans une matrice de données, comme il en résulte de chaque sondage ou recensement. Une ligne correspond aux données relatives à une personne (une entreprise, etc.) et une colonne regroupe les valeurs d'une même variable pour différents individus.

**Etape 2 et 3. Choix d'un espace des individus et d'un espace des variables :** ce sont deux étapes relativement semblables. Dans les deux cas, il s'agit d'examiner l'échantillon, de rechercher les valeurs extrêmes ou erronées, d'extraire les variables et les groupes de personnes pertinentes pour l'analyse. C'est à ce niveau que l'on va orienter l'analyse vers les variables sur lesquelles l'étude doit se concentrer.

**Etape 4. Analyse multidimensionnelle :** l'analyse est une étape qui est facilitée aujourd'hui par l'existence de logiciels statistiques spécialisés. Il s'agit néanmoins d'une étape cruciale, notamment parce que plusieurs choix techniques doivent être faits en fonction des données disponibles et des résultats recherchés. Au cours de cette étape, le nombre d'observations sera réduit et les facteurs pertinents seront identifiés. Ils seront alors utilisés pour différencier la population.

**Etape 5. Projection dans un espace à deux dimensions (quelquefois trois):** les logiciels statistiques produisent une partie de leurs résultats sous forme d'une carte où l'on peut faire figurer les individus et/ou les variables.

### Encadré 25 - Exemple de résultat d'une analyse factorielle



**Etape 6. Interprétation des résultats :** l'analyse factorielle ne révèle pas la signification des facteurs ni celle des regroupements qui se forment sur la carte. L'analyste doit les interpréter et leur trouver un nom. L'idéal est d'associer les protagonistes clé de l'évaluation à cette étape car c'est pour eux une occasion de mieux comprendre le fonctionnement de l'intervention.

#### ● Application aux Fonds structurels

L'analyse factorielle peut être utilisée comme outil d'évaluation ex ante, par exemple pour identifier des typologies de la population cible en partant

de l'ensemble de la population. On peut ainsi s'en servir pour mettre en place des procédures de sélection de destinataires ou pour identifier la nature des problèmes socio-économiques auxquels doit faire face cette population.

On peut également s'en servir pour souligner les ressemblances et dissemblances au sein de groupes d'individus ou de zones géographiques. Ceci est particulièrement utile pour réaliser des analyses comparatives de développement régional ou pour évaluer les réactions des destinataires d'actions spécifiques.

Par contre, l'analyse factorielle ne peut servir à estimer les effets socio-économiques. Il s'agit avant tout d'un outil de comparaison des données et d'analyse inductive.

#### ● Points forts et limites d'utilisation

- Les analyses factorielles sont un outil de clarification de la matière première qui résulte d'enquêtes à grande échelle, car elles permettent de dégager les grandeurs pertinentes (quels sont les facteurs qui expliquent la différenciation ou la ressemblance d'un groupe de données ?).
- Les résultats d'une analyse factorielle sont très accessibles et se présentent sous une forme pédagogique (diagrammes) agrégeant un grand nombre de données.
- Si l'analyse factorielle permet d'illustrer l'existence de groupes de variables autour de certaines catégories de population, ce n'est pas un outil qui peut servir à fonder l'existence de causalités ou d'effets.
- L'analyse factorielle exige l'accès à une quantité considérable de données, dont la collecte peut s'avérer longue et coûteuse.
- Les techniques factorielles nécessitent des qualifications particulières et des précautions méthodologiques. Mal conduites, elles risquent de fournir des résultats tautologiques.

#### ● Pour en savoir plus

Bry, X. (1995), *Analyses factorielles simples*, Paris : Economica, 111p.

Ouvrage destiné en principe à tous les publics, il vaut mieux disposer de quelques notions de statistique.

Dazy, F et Le Barzic, J. F. (1996), *L'analyse des données évolutives, méthodes et applications*, Paris : Technip.

Forme écrite d'un séminaire de formation, qui requiert de bonnes bases statistiques.

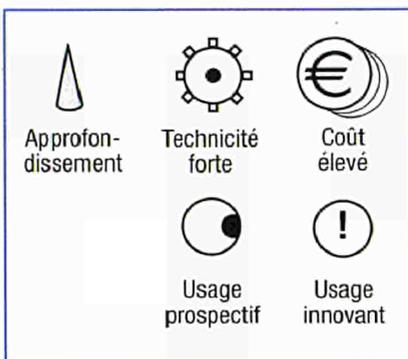
Grangé, D. et Lebart, L. (1994), *Traitements statistiques des enquêtes*, Paris : Dunod.

Ouvrage qui en vient rapidement aux points essentiels.



## ■ Enquête Delphi

Les évaluations ayant recours aux avis d'experts peuvent se faire dans un cadre ex ante. La technique de l'enquête Delphi en est une illustration ; elle se présente sous la forme d'une procédure de consultation anonyme et itérative d'experts par voie de sondage postal. L'absence de concertation préalable au sein du groupe permet une explicitation des divergences entre experts et donc de mieux comprendre la nature de l'expertise de chacun.



Cette technique est recommandée lorsque les questions posées sont simples (programme comportant peu d'objectifs, à caractère technique) et que l'on vise à établir une estimation quantitative des impacts potentiels d'une intervention isolée (par exemple : augmentation de la fiscalité, du prix de l'énergie,...).

### Exemple : Politique de vaccination contre le SIDA : une analyse de scénario par la méthode Delphi

La méthode Delphi a été utilisée pour examiner quelles seraient les réponses appropriées à l'introduction d'un premier vaccin préventif contre le SIDA en Suisse. Trente personnes concernées par cette stratégie de vaccination ont contribué de manière anonyme à cette étude. Cette étude s'est basée sur un scénario développé préalablement qui décrivait les caractéristiques d'un premier vaccin préventif partiellement efficace contre le SIDA.

La consultation Delphi s'est déroulée en trois cycles autour des questions suivantes :

- énumérer les objectifs à atteindre dans les cinq premières années,
- évaluer l'acceptabilité et la faisabilité de propositions concernant le processus de développement d'une stratégie de santé publique et la vaccination contre le SIDA,
- estimer l'utilisation potentielle du vaccin par différents groupes d'utilisateurs.

Ces résultats ont permis, d'une part d'apporter des enseignements importants sur les stratégies et les recommandations à développer dans le cadre des campagnes de prévention contre le SIDA et, d'autre part, de fournir un cadre de référence à la mise en place d'une future stratégie de vaccination contre le SIDA.

Source : Zuber, P.(1994) Introducing a first AIDS vaccine in Switzerland : A Policy Delphi Analysis. Lausanne : Institut de médecine sociale et préventive.

### ● A quoi sert l'outil ?

L'enquête Delphi, comme les autres démarches s'appuyant sur les experts, est utile pour produire, dans le cadre d'une évaluation, des recommandations dont la crédibilité est fonction de la légitimité des experts eux-mêmes. Cette approche peut être utilisée de manière prospective pour établir des hypothèses sur la réalisation d'événements et sur leurs implications socio-économiques.

Cette méthode permet notamment d'éclaircir l'évolution d'une situation, d'identifier les priorités, ou encore, d'élaborer des scénarios prospectifs.

### ● Dans quels cas l'utiliser ?

Cette technique est utilisable lorsqu'il existe une expertise importante sur le sujet, donc dans le cas de programmes qui ne sont pas innovants. Elle est en outre indiquée dans le cadre d'une évaluation ex ante s'il s'agit d'intervention publique à caractère technique. Elle a été ainsi très fréquemment utilisée dans le cadre des politiques énergétiques pour des études prospectives sur l'impact de modifications fiscales,....

On peut y avoir recours lorsqu'on craint que les experts s'influencent réciproquement. L'anonymat et l'absence d'interactions favorisent de plus une analyse en profondeur des raisonnements à l'origine des opinions exprimées.

### ● Comment le mettre en œuvre ?

La démarche consiste à interroger des experts à l'aide de questionnaires successifs afin de faire apparaître des convergences et dégager un éventuel consensus. Les phases principales de ce processus sont :

**Etape 1. Détermination et formulation des questions :** Un soin particulier doit être porté au choix et à la formulation des questions afin d'obtenir des informations utilisables.

**Etape 2. Sélection des experts.** Ceux-ci doivent avoir des connaissances spécifiques du sujet et être prêts à s'investir dans ce type de démarche. Le panel est généralement composé d'une cinquantaine d'individus.

**Etape 3. Elaboration du premier questionnaire et envoi aux experts :** Le premier questionnaire doit contenir un rappel de la nature de l'étude et comprendre deux ou trois questions semi-ouvertes ou ouvertes.

**Etape 4. Analyse des réponses au premier questionnaire** : Les réponses sont analysées afin de déterminer la tendance générale et les réponses extrêmes.

**Etape 5. Elaboration du second questionnaire et envoi aux experts** : A chaque expert informé des résultats du premier tour, il est demandé de fournir une nouvelle réponse et de la justifier si elle s'écarte de la tendance générale.

**Etape 6. Envoi d'un troisième questionnaire** : Ce questionnaire ne s'adresse qu'aux experts ayant maintenu des réponses « extrêmes ». Il est en outre demandé de critiquer les arguments des tenants de l'autre position extrême. La confrontation des avis joue un rôle modérateur et facilite l'apparition de convergence de points de vue.

Une convergence suffisante des opinions apparaît généralement lors de l'exploitation du quatrième questionnaire. Si tel n'est pas le cas, le cycle continue.

**Etape 7. Résumé du processus et rédaction du rapport final.**

### ● Application aux Fonds structurels

Dans le cas des Fonds structurels, l'enquête Delphi est conseillée pour obtenir des estimations macroéconomiques lorsque les phénomènes en cause sont peu complexes ; par exemple, pour quantifier l'impact d'un grand projet d'infrastructure. Elle peut être aussi envisagée pour préciser des relations de causes à effets potentielles dans le cas d'interventions innovantes.

C'est une technique particulièrement avantageuse lorsque l'on a affaire à un territoire très vaste : il n'y a pas de frais de déplacement des experts, uniquement des frais postaux.

### ● Points forts et limites d'utilisation

- Les conclusions issues de ce type de démarche permettent à l'inverse de la démarche par panel de mieux tenir compte des points de vue minoritaires. Elle offre de plus la possibilité de mieux comprendre les éléments sur lesquels se fondent les avis d'experts.
- Cette technique offre moins de souplesse d'utilisation car le temps nécessaire pour les différentes itérations est important (supérieur à trois mois).

- Comme dans le cas de la démarche par panel, la fiabilité de l'outil dépend de la subjectivité des avis d'experts, elle-même fonction du stock de connaissances dont ils disposent.

● **Pour en savoir plus...**

Godet, M. (1985), *Prospective et planification*. Paris : Economica, 119-125.

Dans le cadre de la première partie de cet ouvrage consacrée à la prospective, la méthode Delphi est présentée et illustrée par des exemples.

Nadeau, M-A. (1988). *L'évaluation de programme*. Laval, Québec : Presses de l'université de Laval. 349-352.

Manuel présentant les méthodes et les instruments utilisés dans le cadre de l'évaluation de programme. La méthode Delphi est exposée succinctement.

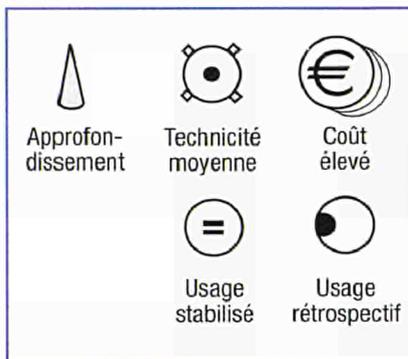
Witkin, B.R et Altschuld, J.W (1995), *Planning Conducting Needs Assessments*, Thousand Oaks : Sage. 193-203.

Recueil des méthodologies utilisables dans le cadre de l'évaluation des besoins, une partie du chapitre consacré aux techniques de groupe traite de la méthode Delphi.

## ■ Groupes de comparaison

La technique des groupes de comparaison consiste à comparer un « groupe participant » (les P.M.E. bénéficiant des services, par exemple) à un groupe similaire sur lequel l'intervention n'a pas eu d'effet. Dans l'idéal, les deux groupes comparés ne devraient différer que par leur participation ou leur non-participation au

programme. Dans ce cas, le groupe qui n'a pas été affecté par l'intervention est appelé « groupe de contrôle » et remplit la même fonction qu'un « contrôle » dans le domaine des sciences naturelles. Dans sa forme la plus extrême, la construction d'un groupe de contrôle fait appel à une sélection aléatoire des participants et des non-participants. En pratique cependant il est rare de pouvoir constituer des groupes qui soient identiques en tous points dans le domaine des sciences sociales, et en particulier dans le contexte des Fonds structurels européens. On a le plus souvent recours à des techniques plus pragmatiques qui sont décrites dans cette fiche sous le nom de « groupes de comparaison ». Un approfondissement est cependant proposé en fin de section pour traiter la technique du groupe de contrôle.



### Exemple : Estimation de l'impact d'un programme de formation et conseil aux entreprises au Royaume Uni

Cet exemple correspond à celui de Encadré 21.

Le programme anglais TEC (Training and Enterprise Council) a été évalué au moyen de la technique des groupes de comparaison. L'effet du programme a été estimé par comparaison « avant et après » d'entreprises utilisatrices des services du programme et un ensemble de binômes appariés d'entreprises n'ayant pas participé au programme TEC. Les binômes ont été constitués selon les critères suivants : le type d'activité de l'entreprise, la taille de l'entreprise en termes d'effectif, ainsi que son ancienneté et sa situation géographique. Un total de 180 paires d'entreprises a été interrogé.

Le rapport examine à la fois les difficultés pratiques et les difficultés théoriques de la technique des groupes de comparaison. La principale difficulté pratique consistait à assortir les paires. Les problèmes théoriques étaient quant à eux liés au fait que, même si les paires associaient des entreprises apparemment identiques, des

différences imperceptibles subsistaient entre elles et pourraient expliquer les différences de résultats sans lien avec le programme évalué.. Malgré ces difficultés, la technique des groupes de comparaison a permis de définir des normes pour d'autres facteurs régionaux ou nationaux ayant influencé les résultats des entreprises. Deux techniques de collecte de données ont été appliquées. L'une consistait en un questionnaire relatif à l'impact sur l'emploi du programme TEC. Afin d'éviter que les résultats ne soient faussés par des réponses « stratégiques » ou intéressées (par exemple, une réponse positive donnée dans l'espoir d'obtenir une aide supplémentaire), les entreprises interrogées devaient non seulement estimer les effets sur l'emploi des services du programme TEC, mais aussi donner des détails sur leur chiffre d'affaires et leurs effectifs. Ceci devait permettre d'établir des comparaisons objectives entre les paires.

Source : Cahier MEANS N°3, 41-42.

### ● A quoi sert l'outil ?

Le but de la technique des groupes de comparaison est d'estimer les effets nets d'une intervention publique, c'est-à-dire les effets directement imputables à cette intervention, et non pas ceux qui résultent de facteurs exogènes ou concurrents.

### ● Dans quels cas l'utiliser ?

La technique des groupes de comparaison s'emploie le plus souvent lorsque le programme offre des subventions ou des services à un groupe relativement homogène de destinataires, que ces destinataires ont déjà été atteints (situation rétrospective) et qu'il existe un groupe similaire n'ayant pas bénéficié du programme évalué.

Pour produire des résultats crédibles, les deux groupes comparés doivent être suffisamment similaires du point de vue des facteurs concurrents probables. Cela veut dire que l'on peut identifier et hiérarchiser ces facteurs avant de constituer les groupes à comparer. L'outil est donc réservé aux situations où l'on connaît déjà assez bien les mécanismes de production des effets.

L'outil des groupes de comparaison n'impose pas l'utilisation d'une technique de collecte de données particulière. En fait, il est possible d'appliquer plusieurs techniques de collecte de données (enquêtes par questionnaires, entretiens,...) et/ou des analyses secondaires.

## ● Comment le mettre en œuvre ?

**Etape 1. La construction des groupes de comparaison :** Il existe trois façons de construire des groupes de comparaison.

- La création de binômes appariés (*matching pairs*), en sélectionnant pour chaque individu du groupe participant au programme un individu non touché par le programme mais aussi ressemblant que possible selon les autres points. La difficulté consiste à trouver des individus qui ne soient pas concernés par le programme. Il est probable par exemple qu'une entreprise opérant sur le même marché qu'une entreprise bénéficiaire ait subi le contrecoup des aides reçues par l'un de ses concurrents. De même, il est probable qu'on dispose a priori de très peu d'informations sur les non-destinataires. La vérification des caractéristiques requises et la recherche d'individus correspondant à un profil spécifique étant des étapes souvent longues et difficiles, la création de binômes appariés peut se révéler onéreuse.
- La comparaison avec des individus connus du programme mais non participants (candidats ayant abandonné leur demande, candidats non sélectionnés, candidats sélectionnés ayant abandonné le programme en cours de route). Dans ce cas, le groupe de non-participants doit être ajusté pour ressembler, autant que possible, au groupe de participants. Il reste cependant toujours un risque lié aux raisons ayant entraîné l'abandon ou le rejet et qui peuvent affecter leur comportement dans son ensemble.
- La comparaison peut enfin se faire, de façon générique, avec un groupe de comparaison hypothétique reconstitué à partir de statistiques nationales. Pour cela, il faut que le programme n'ait touché qu'un petit groupe de participants de façon spécifique et que le reste du pays n'ait pas bénéficié de programmes similaires.

**Etape 2. La comparaison proprement dite :** Le principe consiste à comparer les changements observés dans les deux groupes pour en déduire, par différence, une estimation de l'effet net du programme.

Il existe différentes formes de comparaison présentant chacune des avantages et inconvénients, et des degrés de validité plus ou moins importants.

- La comparaison « **après seulement** » consiste à analyser les différences entre les deux groupes après obtention de la subvention ou du service. Elle présente l'avantage d'être simple. Par contre cette approche ne tient pas compte des différences qui existaient entre les deux groupes au départ.

	Participation au programme	Situation après le programme
Groupe participant	oui	X2
Groupe de comparaison	non	X'2
Estimation de l'effet	X2 - X'2	

- La comparaison « avant-après » permet de contourner l'inconvénient précédent en prenant en compte les différences éventuelles entre les deux groupes aussi bien avant qu'après l'intervention. Elle est donc axée sur l'amélioration relative pour les participants en comparaison des non-participants. Par contre, cela sous-entend que des données relatives à la situation avant l'intervention sont disponibles (par l'intermédiaire du suivi ou grâce à des données statistiques...), ce qui n'est pas toujours le cas. Par ailleurs, l'évaluation peut également débuter non pas avant mais en cours d'intervention, et dans ce cas les données relatives à la situation antérieure ne sont pas forcément disponibles. Enfin, il existe de nombreux programmes pour lesquels les participants et non-participants ne sont pas connus au départ et pour lesquels les informations relatives aux différences existant entre les deux groupes avant le programme ne sont pas disponibles. Dans toutes ces situations, les données concernant la situation « avant » doivent être collectées de façon rétrospective, ce qui ne les rend pas totalement fiables.

	Participation au programme	Situation avant le programme	Situation après le programme
Groupe participant	oui	X1	X2
Groupe de comparaison	non	X'1	X'2
Estimation d'effet	$(X2 - X1)$ $(X'2 - X'1)$		

- La comparaison de tendance avant-après est considérée comme plus fiable. Le comportement des deux groupes dans le temps est comparé pour observer dans quelle mesure les participants bénéficient d'une amélioration par rapport à leur situation antérieure nettement supérieure à celle des non participants.

## ● Application aux Fonds structurels

La technique des groupes de comparaison est adaptée aux évaluations ex post, pour lesquelles l'identification des impacts est primordiale. Il permet en particulier d'estimer les effets d'aubaine, c'est-à-dire les changements qui seraient survenus dans le groupe cible même en l'absence du programme et qui ne doivent donc pas être considérés comme un impact de ce dernier.

Cependant, dans le cadre des Fonds structurels, il est souvent difficile de déterminer des groupes de comparaison. D'une part, le public concerné doit être suffisamment homogène et, d'autre part, il faut qu'une partie significative des destinataires potentiels n'ait pas participé au programme (dans le cas contraire, on parle de « saturation » de la population exposée au programme).

Il est possible de comparer les comportements des participants avec ceux de groupes de comparaison sélectionnés hors des zones éligibles, mais peu d'études jusqu'à présent ont utilisé cette approche.

## ● Points forts et limites d'utilisation

- L'analyse d'un groupe de comparaison bien conçu fournit une estimation convaincante d'un ensemble d'effets du programme.
- L'un des problèmes posés par cet outil provient de la nécessité de bien prendre en compte, dès le début de l'étude, les facteurs susceptibles d'influer de manière déterminante sur la production des effets. Si l'on oublie des variables importantes (des facteurs souvent exogènes au programme) il devient difficile d'établir de fortes inférences causales et l'analyse perdra de sa valeur.
- Une autre limite de l'outil est due à son incapacité à expliquer les effets constatés. Il focalise sur la mesure des effets plutôt que sur les mécanismes qui ont produit ces effets (mise en œuvre, qualité des services fournis, réactions des participants). Le programme et ses effets sont donc traités comme une boîte noire. Cette limite peut être partiellement compensée par l'emploi d'autres outils (étude de cas par exemple) qui donnent à la comparaison une dimension qualitative.

## ● Pour en savoir plus...

Monnier, E. (1992), *Évaluation de l'action des pouvoirs publics*. Paris : Economica, 9-51.

Présentation largement illustrée des techniques comparatives.

Conseil du Trésor du Canada, Contrôleur Général (1991) *Méthodes d'évaluation des programmes*. Division des Communications, Conseil du Trésor du Canada.

Approche générale - voir le chapitre 3.1-3.35.

Matalon, B. (1988), *Décrire, Expliquer, Prévoir - démarches expérimentales et terrain*. Armand Colin - Collection U. 219-252.

Ouvrage plus élaboré et plus technique - voir la troisième partie. Par ailleurs, l'annexe 1 développe le cas particulier de la construction des groupes de comparaison par tirage au sort.

### ● **Approfondissement : les groupes de contrôle et le dispositif expérimental**

Pour éliminer tous risques de biais, on peut avoir recours à une affectation aléatoire des individus, familles ou entreprises à l'un ou l'autre des deux groupes. Cette pratique, dite dispositif expérimental, constitue dès lors le moyen le plus rigoureux d'établir l'existence des liens de causalité supposés entre l'intervention publique et les impacts escomptés. Elle suppose la comparaison entre un groupe exposé au programme, dit « expérimental », et un groupe strictement identique, le groupe « témoin », qui permet de constater ce qui se serait passé en l'absence du programme.

Le dispositif expérimental sert à vérifier la réalité des relations de cause à effets entre les changements socio-économiques observés et le programme évalué. Il permet de comparer empiriquement l'état actuel d'une situation sociale suite à la mise en oeuvre du programme avec ce qu'aurait été cette même situation sans le bénéfice du programme, et évaluer ainsi l'impact net du programme en « contrôlant » l'influence des facteurs concurrents. Il peut être considéré comme le moyen le plus rigoureux permettant d'établir des relations de causalité par rapport aux résultats d'un programme, car l'utilisation de groupes de comparaison statistiquement équivalents renforce la validité interne de la technique.

Les dispositifs expérimentaux peuvent être utilisés pour évaluer une grande diversité d'interventions publiques mais ont été principalement utilisés dans les programmes éducatifs (programmes contre l'usage et la consommation d'alcool, méfaits du tabac...). L'expérimentation peut être appropriée lorsqu'il est possible de réaliser une affectation aléatoire sans craindre un sentiment d'injustice sociale qui découle de l'exclusion du groupe témoin des bénéfices de l'intervention publique. Il est également justifié pour

l'évaluation de programmes innovants coûteux, pour lesquels les données les plus fiables possibles sont nécessaires.

#### ◆ *Mise en œuvre d'une expérimentation*

L'expérimentation repose, d'abord, sur l'affectation aléatoire des unités d'observation, réparties entre le groupe expérimental et le groupe témoin de façon à constituer des groupes statistiquement équivalents. Ainsi, la probabilité que les paramètres statistiques descriptifs de toute variable décrivant ces deux groupes soient semblables sera très grande. La deuxième étape consiste à appliquer le traitement uniquement au groupe expérimental (ce qui est la seule différence entre les deux groupes). Elle est suivie de la réalisation de mesures d'observations sur les deux groupes avant l'application du traitement (E (1) et E'(1)), puis après (E(2) et E'(2)). Enfin, la comparaison des écarts vérifiés sur les deux groupes entre état initial et état final fournit une mesure de l'impact du traitement, l'importance de la différence permettant de juger si l'impact est significatif ou non.

Le schéma ci-dessous résume la mise en œuvre du dispositif expérimental :

	mesure préalable	application du traitement	mesure in fine	variation des mesures
groupe expérimental	E (1)	X	E (2)	E (2) - E(1)
groupe témoin	E'(1)	-	E'(2)	E'(2) - E'(1)

#### ◆ *Limites de l'expérimentation*

Les dispositifs expérimentaux sont difficiles à mettre en œuvre. Souvent, l'affectation aléatoire ne peut être réalisée : il est parfois impossible de définir la variable permettant de constituer un groupe de comparaison, lorsque par exemple la population cible bénéficie en totalité du programme. Une sélection aléatoire des membres des groupes étudiés est rarement possible sur la totalité de la population cible, or le volontariat est source de biais dit « d'autosélection ». En outre, le recours à la création artificielle d'un groupe témoin nuit à la validité externe du dispositif expérimental. Quel que soit le mode de sélection des personnes, la validité externe peut être compromise. Les facteurs contextuels liés au choix des sites limitent également la représentativité de l'expérimentation. Autant d'éléments qui confirment que le dispositif expérimental réduit la capacité de généralisation des conclusions à l'ensemble du programme ou à d'autres contextes.

Un problème éthique et politique peut en outre se poser, si l'on considère que seules les personnes du groupe expérimental bénéficient du programme au détriment des individus du groupe témoin. Le dispositif expérimental nécessite par ailleurs d'importantes ressources financières, intellectuelles, institutionnelles, et la constitution des groupes statistiquement équivalents constitue en soit une tâche lourde et fastidieuse. Finalement, il ne permet pas de fournir une information sur l'incidence d'événements spontanés.

◆ **Comment utiliser l'outil pour l'évaluation ?**

Le dispositif expérimental est justifié pour les quelques cas où les relations causales sont peu nombreuses, simples, connues, dont l'objectif est de mesurer des impacts stables, et non d'expliquer des phénomènes complexes et mouvants.

Cet outil est donc peu adapté à l'évaluation des Fonds structurels et l'on ne trouve pas d'exemple en Europe d'application du dispositif expérimental dans ce cas. Les difficultés d'ordre pratique, relatives à la mise en œuvre de cet outil, et son utilité incertaine, peuvent expliquer cette absence. Les programmes des Fonds structurels sont appliqués sur une durée relativement longue. Aussi, les deux groupes, expérimental et témoin, pourraient être exposés à certaines variations, et des différences nettes se créeraient entre ces groupes.

● **Pour en savoir plus...**

Boruch, R. F. (1997), *Randomized Experiments for Planning and Evaluation*, Thousand Oaks, CA : Sage publications, 265 p.

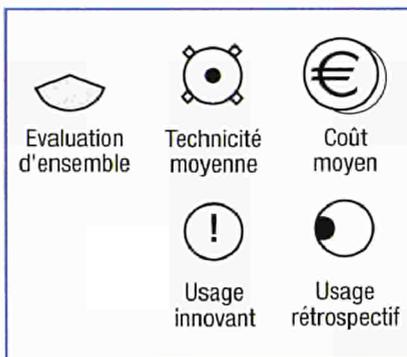
Un ouvrage qui milite de façon mesurée pour l'utilisation de la technique d'expérimentation.

## ■ Analyse de régression

L'analyse de régression permet d'obtenir une estimation quantitative de l'influence de plusieurs facteurs explicatifs (ex : l'intervention à évaluer, les facteurs concurrents) sur une variable à expliquer (ex : le résultat de l'intervention). Pour réaliser une analyse de régression, il faut se doter au préalable d'un modèle logique explicatif. L'outil calcule

ensuite les paramètres de ce modèle en les ajustant le mieux possible aux évolutions observées dans le passé. La technique de l'analyse de régression s'applique avec des variantes selon que les variables du modèle sont continues ou discrètes et selon que leurs relations sont linéaires ou non.

L'outil est employé depuis environ un siècle par les statisticiens, notamment en biologie. Il a été généralisé grâce au développement des microordinateurs et des logiciels statistiques qui permettent d'ajuster facilement des modèles complexes sur des données nombreuses.



### Exemple : L'évaluation des stages de formation FSE en Espagne

En 1995, les Départements d'analyse économique et d'économie appliquée de la Faculté de Sciences Economiques de la l'Université Autonome de Madrid ont conduit une évaluation des effets de certaines mesures de formation mises en place avec l'aide du Fond Social Européen. Cette évaluation devait répondre aux questions suivantes : Combien de participants ont trouvé un emploi après avoir suivi des cycles de formation ? Quel est le profil des participants qui ont trouvé un emploi ? Quelle est l'influence du contexte régional ?

Une enquête a été réalisée sur un échantillon de 13 127 stagiaires réels ou potentiels sur l'ensemble du territoire espagnol (hormis trois régions).

Un modèle explicatif a été construit avec les variables suivantes :

- Comme variables à expliquer (variables dépendantes) : le taux d'emploi et la durée de recherche de l'emploi ;
- Comme variables explicatives : le sexe, l'âge, le niveau d'études, le lieu de résidence, le type de chômage (longue durée, recherche d'un premier emploi), et l'expérience professionnelle ;
- Une variable explicative spéciale a été ajoutée : la participation au programme (durée de la formation FSE).

Ce modèle comprend des variables continues comme l'âge ou la durée de recherche d'emploi, ainsi que des variables discontinues comme le sexe ou le fait de trouver un emploi. Au moyen d'une analyse de régression, le modèle a été ajusté aux résultats d'une enquête par questionnaire.

L'analyse de régression a permis d'estimer l'impact du programme sur l'emploi des stagiaires. En outre, il a permis de produire des estimations supplémentaires telles que :

- Le fait d'être un homme améliore les chances de trouver un emploi de 5% ;
- La localisation régionale est importante : les participants des stages vivant aux Baléares ont environ 20% de plus de chances de trouver un emploi que ceux de la Galice ;
- Les personnes possédant un diplôme d'enseignement secondaire améliorent leurs chances de trouver un emploi de plus de 12% par rapport à celles qui ont un niveau d'études élémentaire.

Source : F. Saez-Fernandez, M. Toledo, "Education, Labour Market and Employment Policy - Empirical Evidence and Implications", 7th EALE Conference, Madrid 1995.

### ● A quoi sert l'outil ?

Lorsqu'elle est applicable, l'analyse de régression est une excellente technique d'estimation des effets. La technique sert alors à attribuer une part des changements observés au programme évalué (effet net) et une autre part aux facteurs explicatifs ou concurrents.

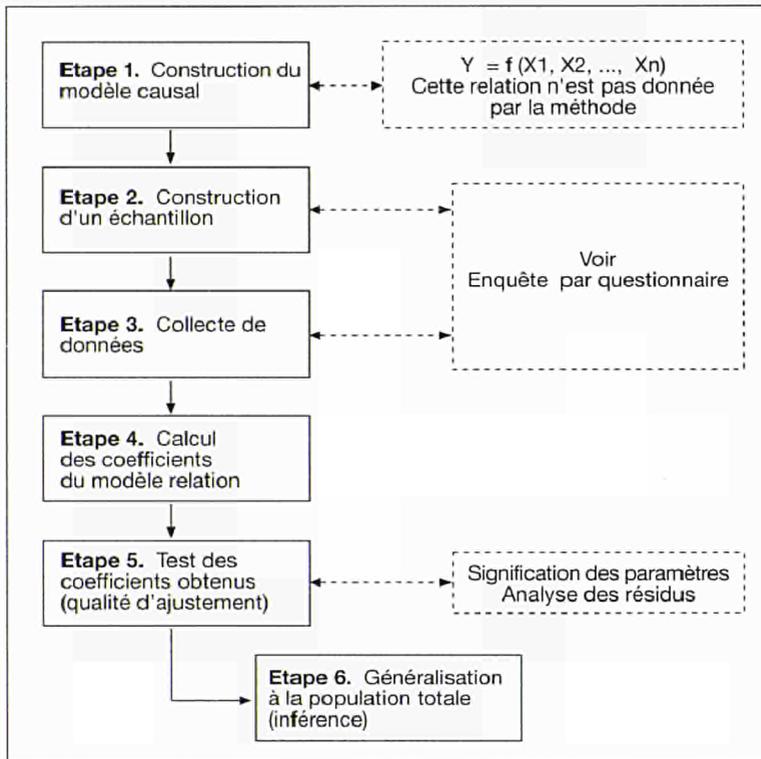
### ● Dans quel cas l'utiliser ?

L'outil n'est applicable que si l'on dispose au préalable d'un modèle explicatif crédible et robuste. Il doit donc être utilisé dans les cas où les effets de l'intervention sont bien identifiés et où le processus de production des effets est connu.

L'analyse de régression est typiquement utilisable lorsque l'on privilégie un des effets du programme évalué (par exemple le taux de placement).

Pour pouvoir appliquer l'outil, il faut disposer d'un volume d'observations quantitatives important : entre mille et dix mille unités (stagiaires, entreprises, exploitations,...). Les observations doivent être détaillées selon toutes les variables du modèle explicatif.

## ● Comment le mettre en œuvre ?

**Encadré 26 - Réaliser une analyse de régression**

**Etape 1. La construction du modèle causal** : La construction du modèle explicatif est une étape délicate qui doit être menée en concordance avec la réflexion sur la théorie d'action de l'intervention. Les variables peuvent être de plusieurs natures. Elles peuvent être spécialement créées, par exemple pour prendre en compte le fait qu'un individu ait bénéficié d'une aide ou non (variable muette - *dummy* - prenant les valeurs 0 ou 1). Une variable peut représenter une caractéristique observable (avoir un emploi ou non) ou inobservable (probabilité d'avoir un emploi). Le modèle peut supposer que telle ou telle variable évolue de façon linéaire, logarithmique ou probabiliste. Tous les modèles explicatifs sont construits sur la base d'un schéma de type :

$Y : f(X_0, X_1, X_2, \dots, X_n)$  dans lequel :

Y est le changement que le programme est principalement censé produire (ex : emploi des stagiaires)

f est une fonction de plusieurs variables explicatives

X<sub>0</sub> est une variable qui représente le fait qu'un individu bénéficie ou ne bénéficie pas de l'intervention

X<sub>1</sub>, ..., X<sub>n</sub> sont les autres variables susceptibles d'expliquer le changement

Un bon modèle doit sélectionner des variables peu nombreuses avec un fort pouvoir d'explication. Il ne doit pas oublier des facteurs importants. Il doit éviter les relations causales qui fonctionnent dans les deux sens, du type X<sub>1</sub> explique Y mais Y explique X<sub>1</sub> (causalité circulaire).

Si deux variables explicatives sont linéairement dépendantes et apportent la même information, il est inutile de les considérer toutes les deux. En effet, les phénomènes de colinéarité affaiblissent l'ajustement. Par exemple, si les femmes interrogées ont des durées de chômage préalable systématiquement plus longues que celles des hommes, il ne sera pas possible de départager l'influence des deux facteurs explicatifs : sexe et durée de chômage préalable.

Certaines formes d'analyse de régression (*path analysis*) permettent la séparation des effets directs et des effets indirects. Les effets indirects ont lieu par une ou plusieurs variables interposées. Vu la complexité qui caractérise la plupart des phénomènes socio-économiques, cette propriété est importante.

**Etape 2. La construction d'un échantillon :** Pour appliquer une régression multiple, il faut généralement un échantillon de grande taille (2 000 à 15 000 individus). Voir enquête par questionnaire pour la construction des échantillons.

**Etape 3. La collecte des données :** Les données proviennent généralement du système de suivi ou d'une enquête par questionnaire ou d'une combinaison des deux sources.

**Etape 4. Le calcul des coefficients :** Cette phase de l'analyse est largement facilitée par des logiciels statistiques d'un coût abordable et utilisables sur micro-ordinateur (ex : SAS, SPSS, SPAD, etc.).

**Etape 5. Test de l'ajustement effectué** : Lorsqu'on élabore et qu'on ajuste le modèle, on cherche à expliquer la plus grande part possible de la variabilité des changements observés. La partie qui reste inexpliquée représente les résidus. Ainsi, plus les résidus seront faibles, meilleure sera la qualité du modèle et de son ajustement. L'analyse des résidus est une étape très importante : c'est à ce niveau que l'on voit le degré d'adaptation du modèle aux phénomènes que l'on veut expliquer. C'est l'analyse des résidus qui permet de dire si l'outil a permis d'estimer les effets de façon valide ou non. Si l'on détecte des anomalies importantes, l'emploi de l'outil doit être considéré comme un échec et l'évaluation doit s'arrêter avant l'étape 6.

**Etape 6** : Cette dernière phase permet la généralisation des coefficients du modèle à l'ensemble de la population. Elle permet donc de produire l'estimation des effets.

### ● Application aux Fonds structurels

Dans le cadre des Fonds structurels, l'emploi de cette technique rencontre rapidement des limites contraignantes. La technique est lourde car elle suppose la connaissance quantitative de plusieurs milliers d'individus. Elle est donc également longue et chère et ne s'applique qu'à des interventions qui touchent un grand nombre de personnes.

### ● Points forts et limites d'utilisation

- L'utilisation de l'analyse de régression crée une occasion de préciser les hypothèses concernant la nature des effets (théorie de l'action), ainsi que les facteurs concurrents.
- Lorsqu'elle s'achève sur un succès (ajustement statistiquement valide), l'analyse de régression produit une estimation quantitative des effets nets et résout un des problèmes les plus difficiles d'une évaluation.
- Les observations doivent présenter des évolutions suffisamment contrastées pour permettre l'ajustement. Par exemple, si toutes les observations concernent la tranche d'âge 30-40 ans, on ne pourra pas estimer l'influence de l'âge sur l'emploi.
- L'analyse de régression risque d'aboutir à la conclusion qu'il existe un lien fort entre deux variables, alors qu'on a oublié d'estimer l'influence d'autres variables, pourtant plus significatives (cette erreur est nommée *data snooping*).

- Les relations entre les différentes variables expliquées et explicatives sont fréquemment circulaires (X explique Y et Y explique X). Dans ce cas, la méthode est inapplicable.
- La mise en œuvre de l'outil pourrait être longue et coûteuse.
- L'outil est d'un emploi très sensible : les erreurs commises au niveau de la collecte des données et du modèle explicatif se répercutent sur les conclusions sans qu'il soit facile de s'en apercevoir.

#### ● Pour en savoir plus

Mohr, L.B. (1995), *Impact analysis for program evaluation*, London : Sage Publications, 310 p.

Ouvrage qui aborde les principales techniques d'évaluation des programmes d'une façon claire, peu mathématisée.

Wonnacot, T.H., Wonnacot, R.J. (1984), *Statistique, Economie, Gestion, Sciences, Médecine*, Paris : Economica, 790 p.

Manuel qui applique les principes fondamentaux de la statistique aux sciences sociales. Ouvrage clair, illustré de nombreux exemples.

## Outils pour porter un jugement

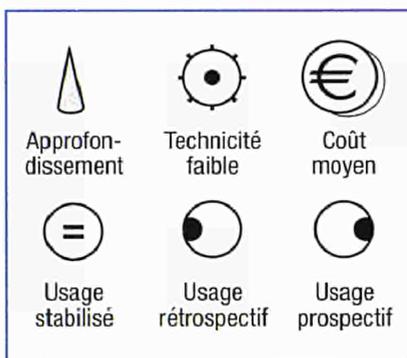
Dans le cadre d'une intervention approfondie, il devient possible d'employer tous les outils de jugement à la disposition des évaluateurs, et notamment les outils qui ont été développés pour l'évaluation des projets.

- L'**analyse coût-efficacité** juge l'efficacité de l'intervention par comparaison quantitative avec une ou plusieurs autres interventions. Elle privilégie un objectif jugé prioritaire. Elle est particulièrement adaptée à l'évaluation ex post.
- L'**étalonnage** (*benchmarking*) juge l'efficacité de l'intervention par comparaison qualitative avec une autre intervention reconnue comme excellente (étalon). L'étalonnage autorise la prise en compte et la synthèse de plusieurs effets.
- L'**analyse coût-avantages** juge l'efficacité de l'intervention en prenant en compte tous ses effets et en réalisant une synthèse de façon quantitative. L'analyse coût-avantages peut juger une intervention dans l'absolu, sans référence de comparaison à d'autres interventions. Elle est généralement employée ex ante, mais il est rare que toutes les conditions nécessaires à sa bonne application soient réunies.
- L'**analyse multicritère** (outil présenté précédemment) compare l'efficacité et l'efficacité de plusieurs projets ou interventions. L'outil prend en compte et synthétise plusieurs effets de façon quantitative et/ou qualitative. Cet outil permet de prendre en compte d'éventuelles différences entre les points de vue des partenaires. L'outil est généralement employé ex ante. Sa souplesse permet de l'utiliser dans de nombreuses situations.
- Le **panel d'experts** (outil présenté précédemment) est un outil très souple qui permet de produire un jugement synthétique, appuyé sur des données qualitatives et quantitatives, même incomplètes. Cet outil permet de prendre en compte d'éventuelles différences entre les points de vue des partenaires. Dans une moindre mesure, l'**enquête Delphi** (outil présenté précédemment) peut apporter une aide similaire pour la formulation du jugement.
- Finalement, le **vote coloré** (outil présenté précédemment) est un outil qui peut servir lors de la synthèse finale de l'évaluation pour valider les conclusions et les recommandations.



## ■ Analyse Coût-Efficacité

L'analyse coût-efficacité trouve son origine dans certaines formes d'évaluation qui se sont développées, notamment dans le domaine de la médecine et de la sécurité routière. Il s'agit d'une approche pragmatique qui se focalise directement sur l'effet principal d'un programme rapporté à son coût et qui exclue



volontairement d'autres effets potentiels. C'est donc une forme simplifiée d'évaluation « value for money » (VFM) dans laquelle la « valeur » du programme s'exprime en une seule dimension. L'analyse coût-efficacité se différencie de l'analyse coût-avantages par le fait que l'effet du programme s'exprime en quantités physiques (ex : nombre de morts évités) et non pas en valeur monétaire.

### Exemple : Création et maintien d'emplois au pays de Galles

Cet exemple correspond à celui de l'Encadré 3.

En 1989, la Cour des Comptes anglaise (NAO) a utilisé l'analyse coût-efficacité pour mener une évaluation des politiques de création et de maintien de l'emploi au Pays de Galles mises en œuvre sur les vingt dernières années par trois agences clés : le Welsh Office (WO) responsable du développement urbain, la Welsh Development Agency (WDA) pour le soutien au développement industriel et le Development Board for Rural Wales (DBRW) chargé du développement social et économique des régions Nord et Centre du Pays de Galles. L'évaluation a été mise en place conjointement par le NAO et le WO. Elle a duré 5 mois et a porté sur les effets de ces politiques sur un échantillonnage de 303 sociétés destinataires, lesquelles ont été interrogées sur ces effets.

Les résultats des entretiens ont été utilisés pour calculer le coût par emploi des mesures, sur la base du nombre de créations d'emplois prévues, du niveau de l'effet d'aubaine et de déplacement et des impacts bénéfiques sur les fournisseurs de ces sociétés. Les coûts globaux du programme ont alors été divisés par les bénéfices nets pour obtenir une mesure de coût-efficacité, c'est-à-dire :  $\text{Coût-Efficacité} = \frac{\text{Coût du programme}}{\text{(nombre brut d'emploi créé - effet d'aubaine - moins effet de déplacement + effet fournisseur)}}$ .

L'absence d'homogénéité dans la mise en œuvre des actions de développement régional au Pays de Galles a rendu difficile les comparaisons portant sur les coûts

par emploi créé. Cependant, le NAO a finalement conclu, bien que l'évaluation précédente ait surestimé les effets des programmes sur l'emploi parce que les niveaux des effets d'aubaine et de déplacement n'avaient pas bien été pris en compte, que les programmes avaient créé un nombre significatif de nouveaux emplois.

Source : S. Roberts and Ch. Pollitt, (1994) *Audit or Evaluation ? a National Audit Office VFM Study*, Public Administration.

### ● A quoi sert l'outil ?

L'analyse coût-efficacité sert à évaluer l'efficacité d'un programme au regard d'une seule mesure, quantifiable et considérée comme l'objectif prépondérant ou objectif clé du programme. Sur la base de ce critère, des comparaisons sont établies entre le programme évalué et des interventions similaires mises en œuvre dans d'autres régions ou des approches alternatives axées sur la création de résultats semblables dans la même région.

L'analyse coût-efficacité peut constituer la première étape d'une analyse d'étalonnage (*benchmarking*) axée sur des programmes ou des actions entreprises dans des contextes similaires mais qui sont apparemment d'un meilleur rapport coût/efficacité que le programme analysé.

### ● Dans quel cas l'utiliser ?

L'analyse coût-efficacité s'emploie aussi bien dans une évaluation ex ante que dans une évaluation ex post. Elle est particulièrement adaptée à l'évaluation de programmes dont les effets escomptés ont été clairement définis, sont relativement homogènes et/ou peuvent être mesurés en terme de résultat clé (par exemple le nombre d'emplois créés).

### ● Comment le mettre en œuvre ?

**Etape 1. Détermination de l'objectif du programme :** Cette première étape est celle de l'explication de la logique du programme. On répond aux questions suivantes : Quels sont les buts à atteindre ? Quelles sont les réalisations du programme ? Quels sont les impacts escomptés ? Par quels moyens pense-t-on atteindre ces résultats ? Parmi ceux-ci, quel est celui considéré comme primordial ?

**Etape 2. Définition du coût total du programme :** A ce stade, on additionne le coût de toutes les ressources publiques du programme afin

d'obtenir le coût total. En général, seules sont incluses les ressources directes qui ont une valeur monétaire bien définie. Les coûts moins tangibles, tels que le temps passé par le personnel, et en particulier, les coûts de supervision, ne sont normalement pas pris en compte. On valorise parfois le coût du programme en fonction des avantages qui auraient pu être obtenus en affectant l'argent public à d'autres fins (coût d'opportunité).

**Etape 3. Mesure de l'impact :** Cette étape est la plus difficile. De nombreuses études, y compris l'évaluation du NAO citée plus haut, utilisent des approches empiriques, basées sur la collecte de données primaires, pour recueillir des informations sur l'ampleur des effets positifs du programme en cours d'évaluation. Il est cependant possible aussi d'estimer le niveau d'impact sur la base des données secondaires et/ou la modélisation de la mise en œuvre du programme. Quelle que soit la méthode employée, il est important d'obtenir une vue exacte, non seulement des effets positifs mais aussi du niveau des effets d'aubaine, de déplacement, et autres effets d'éviction, etc. Certains évaluateurs ont estimé que l'on devrait également donner une indication des effets multiplicateurs et des effets secondaires non attendus, mais il est en réalité extrêmement difficile dans la pratique d'évaluer ces effets secondaires de façon précise ; La plupart des évaluations n'en ont donc pas tenu compte ou ont donné des estimations de leur niveau. Cette dernière solution est à l'évidence préférable étant donné l'importance des effets secondaires. On peut pallier le manque de précision dans l'estimation du niveau de ces impacts en ayant recours à des analyses de sensibilité qui indiquent le niveau probable de bénéfices.

**Etape 4. Calcul du ratio coût-efficacité :** La dernière phase est celle de la mise en œuvre du calcul qui aboutit au résultat final. Il s'agit d'une simple division dont la formule sera du type suivant :

Coût du programme / (effets bruts - effets d'aubaine et de déplacement, et le cas échéant + les effets multiplicateurs sur les fournisseurs des destinataires directs).

#### ● **Application aux Fonds structurels**

La technique coût-efficacité est difficile à appliquer aux programmes des Fonds structurels qui poursuivent souvent des objectifs multiples. De ce point de vue, l'analyse multicritère est souvent d'un emploi plus approprié. Elle peut cependant constituer un bon outil pour des évaluations

thématiques comme par exemple, celle des effets sur l'emploi, des effets sur la création de P.M.E., etc. ou pour comparer des approches destinées à produire des impacts similaires. Il est cependant important de s'assurer que la base de calcul et la fiabilité des données relatives aux coûts et aux bénéfices nets soient définies de façon homogène.

### ● **Points forts et limites d'utilisation**

- L'analyse doit être précédée d'une réflexion sur les objectifs du programme, sur l'identification de son objectif principal et sur un indicateur fiable du degré d'atteinte de cet objectif.
- Les analyses coût-efficacité ont tendance, par nature, à se focaliser sur les résultats directs à court et à moyen terme mais ignorent plus ou moins les impacts à plus long terme dont peuvent dépendre l'efficacité finale, ou l'absence d'efficacité, des programmes et des politiques. Par exemple, l'efficacité des programmes des Fonds structurels dépend de l'atteinte d'une combinaison d'objectifs multiples destinés à promouvoir, à l'intérieur d'une région, une croissance et un développement autonomes.
- L'analyse coût-efficacité peut fournir une base de comparaison entre programmes ayant un objectif similaire s'il est possible de déterminer très sûrement l'ampleur de l'impact sur cet objectif. Cependant, si des méthodes d'enquête différentes sont utilisées ou s'il y a des raisons de croire que la qualité des données diffère selon les programmes, le ratio coût-efficacité ne peut être appliqué pour établir des comparaisons.
- Comme l'analyse coût-efficacité repose sur une estimation de l'impact du programme par rapport à son objectif principal, elle présente l'autre avantage de produire des conclusions aisément compréhensibles et qui se concentrent sur les principales préoccupations des publics clés (y compris les politiques et les décideurs).
- Si l'analyse repose sur des liens spécifiques entre les entrées et les sorties relativement bien établis, l'outil permet de faciliter la description du fonctionnement réel des programmes. Ceci peut être utile pour affiner des politiques existantes ou pour améliorer l'efficacité des interventions futures.
- L'analyse coût-efficacité ne peut servir qu'à comparer des programmes dont la mise en œuvre est simple et dont les impacts sont du même type.

Or, ce cas de figure est beaucoup plus rare qu'on ne l'imagine. Ainsi, par exemple, même des programmes ciblés sur un objectif principal identique comme la création de postes, peuvent créer des opportunités d'emploi qui seront différentes qualitativement, en termes de longévité, sécurité, niveaux de rémunération, probabilité qu'ils soient accessibles à des personnes des régions ciblées, etc.

- L'établissement des liens de causalité de ce type nécessite de pouvoir obtenir des gestionnaires de programme et des destinataires des données détaillées. Lorsque celles-ci n'ont pas été recueillies dans le cadre du suivi courant des programmes/projets, l'évaluation devra avoir recours à une collecte laborieuse de données primaires.
- Dans l'idéal, les analyses coût-efficacité doivent donc être utilisées en combinaison avec d'autres techniques économiques pour analyser les impacts à plus long terme sur le PIB et la compétitivité au niveau régional.

#### ● Pour en savoir plus...

*Creating and Safeguarding Jobs in Wales, A Value for Money Audit* (1993), Brunel University, MEANS.

Une évaluation pilote du programme MEANS, 20 p.

Garrabé, M., *Ingénierie de l'évaluation économique*, Paris : Ellipses, 255 p. Ouvrage qui décrit les principaux outils d'évaluation économiques, parmi lesquels figure l'analyse coût-efficacité. Ouvrage accessible.

Gray, A.W. (1995), *EU Structural Funds and Other Public Sector Investments. A Guide to Evaluation Methods*, Dublin : Gill & Macmillan, 106 p.

Livre traitant des techniques d'évaluation économiques, avec une orientation particulière vers l'analyse coût-avantages. Contient une description complète de l'analyse coût-efficacité, de son articulation logique et des conditions de sa mise en œuvre. Très accessible.

Sanders, J.R (1994), *The Program Evaluation Standards, How to Assess Evaluations of Educational Programs*, Second Edition, London : Sage Publications, 222 p.

Un manuel dédié aux principales étapes de l'évaluation des programmes. Contient une description rapide de l'analyse coût-efficacité. Très accessible.

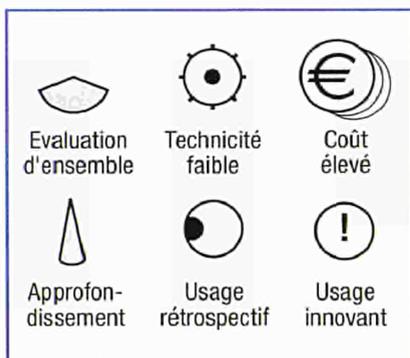


## ■ Etalonnage (Benchmarking)

L'étalonnage a été développé dans les années 80 par les grandes entreprises américaines du secteur électronique, soucieuses de combler leur retard en matière de compétitivité et de qualité face à la concurrence japonaise. Les principes de l'étalonnage reposent sur un échange d'informations

concernant l'organisation de certains segments de la production avec des entreprises réputées pour leur excellence dans un domaine donné.

Dans le cadre du « New Public Management », cette pratique a été transposée à la sphère publique, particulièrement à la gestion des services publics et des administrations municipales. Elle peut également être pertinente dans le cadre de l'évaluation concomitante d'interventions publiques.



### Exemple : l'évaluation de la gestion des villes

La fondation Bertelsmann, en Allemagne, a organisé un concours ouvert aux villes allemandes. Le concours jugeait la qualité des réformes administratives mises en œuvre par les villes candidates en fonction de trois critères.

Afin de se doter d'une norme de comparaison, la fondation a eu recours à un jury d'experts qui a défini les critères permettant de juger la qualité des réformes administratives d'une ville. Ce jury a ensuite identifié des cas de « meilleures pratiques », s'appliquant aux critères retenus, à savoir : trois villes situées respectivement en Nouvelle Zélande, au Royaume-Uni et en Suède. Il s'agit de villes reconnues pour l'originalité et l'efficacité de leurs réformes administratives. Chacune des villes identifiées apparaît comme étant « exemplaire » et sert de « benchmark » pour au moins un des critères définis par le jury.

Les critères sont les suivants : modalités de commercialisation des services publics, stratégie de qualité appliquée aux services publics, management stratégique concernant l'organisation des services, gestion financière.

Outre qu'elle a fourni une référence incontestable pour le jugement du concours, la technique de l'étalonnage a permis, de manière plus large, d'évaluer l'état d'avancement des réformes administratives dans un grand nombre de villes allemandes. Elle a permis d'évaluer le retard global pris par les villes allemandes dans ce domaine.

Source : Frieder Naschold, "Internal Modernisation, Competition, Fiscal Consolidation. International Experiences with Local Government Administrative Reform", WZB, 1996.

### ● A quoi sert l'outil ?

La finalité de l'étalonnage est de parvenir à identifier et à formaliser les forces et faiblesses d'une organisation en la comparant avec d'autres organisations réputées comme étant les meilleures dans le domaine dans lequel des progrès doivent être réalisés. Il s'agit d'une démarche de comparaison des performances qui peut se situer à des niveaux différents qui dépendent de l'étalon de comparaison choisi :

- L'étalonnage externe permet de comparer les différentes fonctions du processus de production d'une organisation aux fonctions similaires dans d'autres organisations.
- L'étalonnage interne permet d'étudier les meilleures solutions mises en œuvre (*best practice*) et de s'en servir pour améliorer les méthodes appliquées à l'intérieur de l'organisation. L'étalonnage interne s'applique aux divisions d'une même organisation, il constitue une démarche moins exigeante.

L'apport de l'étalonnage réside dans une décomposition de la logique productive en fonctions élémentaires qui feront l'objet de la comparaison. L'échange d'informations à propos de ces fonctions élémentaires devient donc possible entre des organisations très différentes.

Une telle démarche permet de prendre du recul par rapport aux pratiques en cours dans l'organisation, en les confrontant aux pratiques mises en œuvre ailleurs. On peut ainsi apprendre de l'expérience des autres. Praticqué à une plus grande échelle, l'étalonnage peut permettre la diffusion de bonnes pratiques.

Bien qu'ayant été conçu à l'origine pour être appliqué aux organisations, l'emploi de l'étalonnage peut être envisagé pour l'évaluation des mesures et projets.

### ● Dans quel cas l'utiliser ?

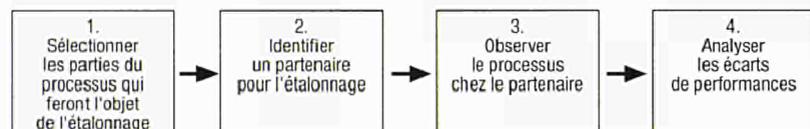
En tant qu'outil d'évaluation, l'étalonnage s'applique à l'évaluation des mesures. Il consiste à identifier des mesures (ou des projets) reconnues comme ayant fonctionné avec succès (*best practice*) et à s'en servir comme d'un étalon de comparaison pour la mesure (ou le projet) à évaluer.

Ceci permet de clarifier le fonctionnement du programme, en détaillant les différentes mesures qui le caractérisent, et d'en identifier les points forts et faibles.

Il peut servir à déterminer des *benchmarks* en termes d'efficacité et d'efficience : c'est-à-dire des programmes reconnus pour leur qualité et ainsi à mettre en œuvre les changements à effectuer dans le cadre du programme évalué, permettant à ce dernier de rattraper la performance des meilleurs programmes.

### ● Comment le mettre en œuvre ?

#### Encadré 27 - L'étalonnage en quatre étapes



**Etape 1.** Dans une phase préliminaire, il faut identifier la nature du problème rencontré et le représenter de façon claire. Un premier pas consiste dans la conceptualisation du processus que l'on doit analyser. L'étalonnage ne pourra porter que sur certaines des étapes du processus. Celles-ci devront faire l'objet d'un choix effectué en fonction de la nature du problème : ce sont les critères de l'étalonnage. Un deuxième pas consiste dans la construction d'indicateurs qui doivent permettre de rendre les critères d'évaluation mesurables.

**Etape 2.** Le choix de l'organisation avec laquelle la comparaison doit avoir lieu dépend des critères retenus pour l'étalonnage. Notons qu'il ne doit pas nécessairement s'agir d'une structure analogue, mais qu'elle peut être choisie en fonction de sa réputation dans le champ concerné par les critères de l'étalonnage. Mais l'étalonnage est basé sur une forme de réciprocité. Les deux partenaires doivent y trouver un intérêt. Si ce n'est pas le cas, l'étalonnage n'aura que peu d'intérêt.

**Etape 3.** La phase de l'étalonnage proprement dite se déroule à ce niveau. Elle consiste dans l'analyse approfondie des critères identifiés au niveau de l'organisation partenaire en vue d'établir une comparaison systématique des solutions mises en œuvre pour satisfaire les critères.

**Etape 4.** Cette analyse peut prendre des formes diverses : des contacts en face à face par des visites dans l'organisation, l'envoi d'un questionnaire, le recours à des consultants spécialisés, des panels composés de membres des organisations impliquées.

La technique de l'étalonnage comprend toujours ces quatre étapes, mais elle peut toutefois prendre diverses formes. Dans certains cas, l'étalonnage est utilisé simplement pour comparer les réalisations ou les impacts qui ont pu être obtenus par différentes organisations ou politiques. Dans le cas présent, l'accent est mis sur l'évaluation de l'efficacité relative et le résultat publié, par exemple, dans des tableaux de classement par ordre de performance. De plus en plus d'agences franchissent une étape supplémentaire et utilisent la technique de l'étalonnage comme un outil de diagnostic qui permet aux gestionnaires, non seulement d'observer les différences en terme de performance, mais aussi d'en analyser les raisons et ainsi de mieux appréhender les stratégies qui améliorent l'efficacité. Cette dernière approche est connue sous le nom d'étalonnage de processus. Elle nécessite l'analyse des différents styles de gestion, des méthodes organisationnelles et des contextes environnementaux des programmes ainsi qu'une connaissance de la manière dont les politiques sont mises en œuvre. Son intérêt réside dans son aptitude à identifier les bonnes pratiques et à les transposer ou à les modifier pour qu'elles puissent être mises en œuvre dans d'autres cas de figure.

Il est souvent admis que les comparaisons entre programmes similaires opérant dans des environnements comparables sont les plus parlantes. En effet, on peut s'attendre à ce que les performances de ces programmes soient largement comparables et que les gestionnaires aient avantage à adopter des pratiques similaires. Cependant, ce n'est pas toujours le cas. Certains gestionnaires de programme pensent qu'ils peuvent obtenir de nouveaux éclairages à partir de l'étalonnage réalisé sur des organisations très différentes des leurs. On constate ainsi de plus en plus fréquemment des échanges de bonnes pratiques entre les agences des secteurs privé et public. Dans le contexte des Fonds structurels, on pourrait trouver des exemples qui utilisent l'étalonnage de processus pour comparer le fonctionnement de programmes similaires dans différents Etats membres ou bien des exemples où l'on compare les programmes subventionnés par lue et les initiatives nationales dans les mêmes Etats membres.

### ● **Application aux Fonds structurels**

L'application de l'étalonnage au cadre des Fonds structurels consiste dans l'identification de projets reconnus comme relevant de bonnes pratiques et dans leur comparaison avec des projets mis en œuvre dans le cadre des Fonds structurels. Il peut s'agir de mesures mises en œuvre au sein d'organisations internationales, d'universités ou d'autres organisations publiques ou privées.

Il est également possible d'avoir recours à l'étalonnage lors de la préparation d'un nouveau cycle de Fonds structurels, ce qui peut se traduire par des modifications de nature réglementaire, stratégique et organisationnelle. L'outil peut également s'appliquer afin d'améliorer les systèmes d'évaluation eux mêmes.

### ● Points forts et limites d'utilisation

- L'étalonnage est une méthode souple, qui permet de formuler des réponses à des problèmes organisationnels, de mettre l'accent sur les points forts et faibles.
- La mise en œuvre de la démarche doit impliquer une équipe assez large pour permettre l'acceptation des résultats de l'étalonnage à l'intérieur de l'organisation, mais assez réduite pour conserver de la souplesse et de l'adaptabilité vis-à-vis des changements à effectuer. Dans le cadre de grandes institutions ceci peut constituer un facteur de blocage.
- La démarche doit être initiée de l'intérieur d'une organisation et être construite sur une base participative pour être efficace. S'il s'agit d'un travail formel, réalisé par des consultants extérieurs sans véritable participation interne, l'étalonnage sera forcément un exercice limité, dont l'impact au sein des organisations restera faible.
- L'identification d'un partenaire pour l'étalonnage peut être problématique dans le cadre de programmes publics, et il pourra être nécessaire d'avoir recours à des panels d'experts pour mieux identifier les expériences pertinentes (*best practices*).
- L'étalonnage est souvent basé sur le partage d'informations stratégiques, il peut donc constituer un risque pour l'organisation qui accepte de s'y livrer, s'il s'agit d'entreprises publiques ou de centres de recherche appartenant à des secteurs stratégiques sur un plan politique ou économique.

### ● Pour en savoir plus...

Balm, G. J. (1994) *Evaluer et améliorer ses performances. L'étalonnage*, AFNOR, Paris, 159 p.

Expose les étapes du *benchmarking* telles qu'elles sont entreprises par IBM aux Etats-Unis. Parfois un peu dense.

Chang, R. Y., Kelly, P. K.(1995) *Progresser avec le Benchmarking*, Paris : Les presses du management, 122 p.

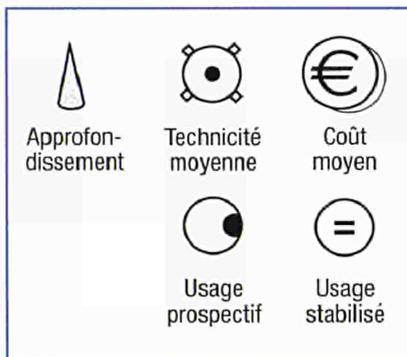
Décrit de façon narrative la mise en place d'un processus de *benchmarking* au sein d'une entreprise. Très simple et vivant.

Karlöf & Partners (1995), *Pratiquer le Benchmarking*, Paris : Les éditions d'organisation, Paris : 144 p.

Guide pratique, illustrant de façon vivante les différentes étapes du *benchmarking* du point de vue des entreprises.

## ■ Analyse Coût-avantages

L'analyse coût-avantages (ACA) (parfois hâtivement traduit de l'anglais par analyse coût-bénéfice) est un outil d'analyse des projets publics, qui a pour objectif de déterminer si la réalisation d'un projet est souhaitable du point de vue de l'ensemble de la société. Cette technique est basée sur la prise en compte et sur une



valorisation monétaire de tous les effets positifs et négatifs du projet au cours de sa durée de vie sur les différentes composantes de la société. Afin de valider l'opportunité du projet, il est important non seulement d'en apprécier la faisabilité mais également d'évaluer les options alternatives de manière à sélectionner la solution la plus appropriée pour la collectivité.

L'utilisation de l'analyse coût-avantages a été généralisée au cours des années 60 dans le cadre du PPBS (*planning and programming budgeting system*) aux Etats-Unis. En tant qu'outil d'aide à la décision, elle permet de justifier le choix d'un projet parmi un ensemble de possibilités. Quelques expériences d'utilisation systématique de l'ACA ont été conduites par exemple par la Banque Mondiale (années 70/80) et le gouvernement britannique (années 80). Aujourd'hui l'analyse coût-avantages ne joue de rôle significatif que dans le domaine de l'évaluation de projets de grande dimension tels que les grands travaux d'infrastructures.

### Exemple : Le train à haute vitesse du corridor Madrid-Seville : une évaluation par l'analyse coût-avantages

Le train espagnol à grande vitesse (AVE) relie Madrid à Séville depuis 1992. Lors de sa construction, plusieurs options techniques étaient possibles et permettaient le choix entre un train à haute vitesse intégrale nécessitant une infrastructure spécifique ou une technologie pendulaire nécessitant des investissements moins importants en matière d'infrastructure mais permettant des vitesses de circulation inférieures à celles de l'AVE.

L'analyse coût avantages a été appliquée rétrospectivement à ce projet en tenant compte de la structure des coûts et des tarifs constatés depuis 1987. Les coûts et avantages suivants ont été pris en compte :

- coût pour le contribuable espagnol et européen
- coût pour les passagers du train

- coût en termes de nuisances pour les riverains
- avantages en termes de temps gagné par les passagers du train (compté en équivalent monétaire).

Sachant que les coûts sont surtout concentrés au moment de la construction et que les avantages sont répartis sur une longue période, tous les coûts et avantages ont été ramenés à une même dimension temporelle en mesurant leur valeur nette présente.

L'étude a permis d'aboutir à une validation rétrospective du projet en montrant que ses avantages étaient globalement supérieurs à ses coûts.

Source : G. de RUS, Facultad de Ciencias Economicas de Tarifa, 1996.

### ● A quoi sert l'outil ?

L'analyse coût-avantages est principalement employée lorsque l'on souhaite comparer l'ensemble des impacts d'un projet public durant sa durée de vie et que ces impacts, même très hétérogènes, peuvent être mesurés en termes monétaires.

### ● Dans quel cas l'utiliser ?

L'analyse coût avantages est très répandue dans le cadre des procédures préalables à la construction d'infrastructures de transport. Elle est utilisée quand les impacts escomptés d'un projet sortent de la sphère privée, sont différenciés et concernent une population hétérogène. Elle est réalisée dans le cadre de projets où l'identification des destinataires et des perdants est possible et pour lesquels tous les effets, même hors marché, peuvent être identifiés et monétarisés.

### ● Comment le mettre en œuvre ?

**Etape 1. Définition du projet et de son coût :** Cette première phase sert à situer le projet dans son contexte. C'est lors de cette étape que l'on détermine les contraintes inhérentes à la réalisation du projet, que l'on examine les différentes variantes possibles et la durée de vie du projet. C'est la phase de détermination du coût, avec l'application éventuelle de prix fictifs (par exemple le prix fictif de la main d'œuvre peu qualifiée est nul s'il y a un fort taux de chômage).

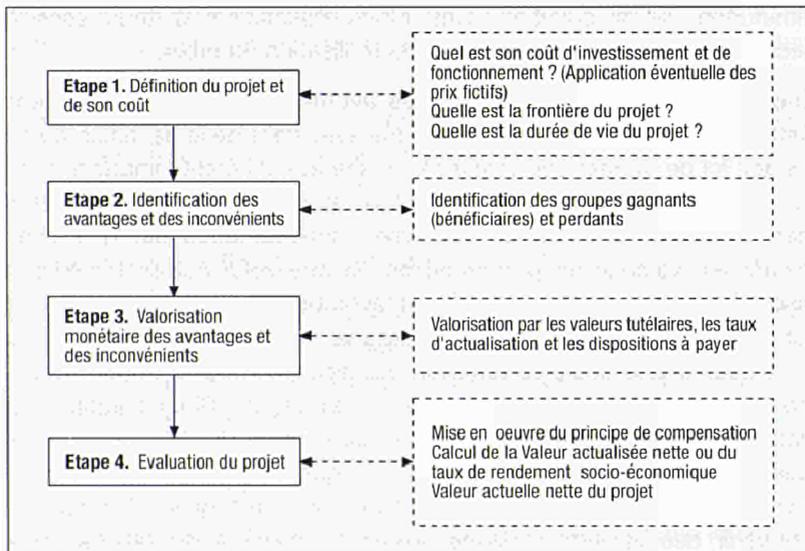
Quatre catégories de variables définissent l'ACA :

- les coûts (C) : il sera ici tenu compte non seulement des dépenses monétaires mais également des coûts sociaux engendrés par le projet

(par exemple, le coût réel de location d'un terrain mis à disposition sans frais par une institution publique à un utilisateur) ;

- les avantages (B) : tant les retombées strictement économiques que les incidences sociales du projet doivent être prises en considération (par exemple, réduction des effets de congestion suite à la construction d'une route à voie rapide) ;
- le temps (t) : tout investissement étant caractérisé par une durée de vie, il importe d'étendre l'analyse sur toute la période concernée par l'avenir du projet (par exemple, chacune des années durant lesquels un hôpital sera fonctionnel) ; le temps étant lui-même limité par la durée de vie de l'investissement égale à  $n$  (par exemple, 20 ans pour les infrastructures et 10 ans pour les investissements productifs).
- le taux d'actualisation ( $r$ ) : les coûts et les avantages étant répartis au cours du temps, il est indispensable d'en actualiser au présent les valeurs de manière à pouvoir globaliser les valeurs calculées (par exemple, l'appréciation d'un avantage à prix courant ne permet pas une comparaison correcte de deux périodes distinctes étant donné le phénomène d'inflation).

### Encadré 28 - Les étapes d'une analyse coût-avantages



**Etape 2. Identification et quantification physique des avantages et des inconvénients du projet :** Cette phase est celle de l'identification des avantages et des inconvénients du projet qui peuvent toucher des groupes de personnes différents (par exemple le temps gagné par les personnes qui empruntent une nouvelle ligne de train ou la quantité de nuisances subies par les riverains de l'infrastructure). L'ACA se différencie de l'évaluation de projets privés sur deux points fondamentaux :

- les coûts et les avantages de tous les membres de la société sont pris en compte et non pas seulement les dépenses et recettes monétaires de l'agence responsable ;
- le taux de rendement social peut différer de manière sensible du taux de rendement privé suite à la prise en compte des phénomènes d'externalité, ces derniers concernant les effets positifs ou négatifs que doit assumer la société comme conséquence de la réalisation du projet.

Les externalités concernent les effets positifs ou négatifs que doit assumer la société comme conséquence de la réalisation d'une activité économique donnée. Elles se réfèrent aux incidences positives et négatives que doit supporter la société sans que ces incidences ne soit prises en considération par les seules règles du marché. Par exemple, l'amélioration des infrastructures de transports se traduit par une réduction du risque d'accidents, des gains de temps, des coûts de transport moindre et une diminution de la pollution. Ces effets bénéficient à tous sans que nécessairement chacun contribue à la réalisation du projet.

**Etape 3. Valorisation monétaire des avantages et des inconvénients :** Cette phase ne doit pas être confondue avec celle de la définition du coût. Il s'agit ici de donner une valeur monétaire aux différents impacts (positifs ou négatifs du projet), les principaux étant en dehors de la sphère marchande. Le principe de l'analyse coût-avantages est que chaque impact est valorisé en proportion de ce que celui qui en bénéficie est disposé à payer pour l'obtenir (s'il est favorable) ou à payer pour l'éviter (s'il est défavorable). Par exemple, un passager est disposé à payer 100 euros pour gagner une heure, un riverain est disposé à payer 10000 euros de moins pour une maison exposée aux nuisances, etc. Si un avantage ou un inconvénient ne peut pas être évalué sous forme de disposition à payer (quelle est la valeur d'un biotope ?, quelle est la valeur d'une vie humaine ?, quelle est la valeur d'une connaissance acquise ?, etc.), on fait intervenir des valeurs tutélaires, fixées de manière discrétionnaire. Par valeur tutélaire, il faut entendre une valeur fixée par un processus politique

explicite (ou le plus souvent implicite). Les avantages et les inconvénients s'échelonnent au cours du temps et il faut pouvoir les mettre en rapport à des échéances différentes. Pour cela, on considère que le temps lui-même possède une valeur tutélaire : c'est le taux d'actualisation.

**Etape 4. Evaluation du projet :** une fois que l'ensemble des étapes précédentes ont été réalisées, on peut passer à l'évaluation proprement dite, qui permet de formuler un jugement sur le projet. C'est à ce stade que s'applique le principe de compensation. Si le projet est bon, les gagnants doivent pouvoir indemniser les perdants, ce qui se traduit par le principe suivant : la valorisation monétaire des inconvénients et des coûts ne doit pas excéder la valorisation monétaire des avantages. Ce principe est évidemment de nature abstraite, car la compensation n'a jamais lieu réellement. Les résultats de l'ACA pouvant varier sensiblement en fonction des hypothèses de travail, il importe de compléter l'appréciation du projet par une analyse de sensibilité. Une telle analyse permet de vérifier dans quelle mesure la rentabilité du projet est affectée par des variations des variables essentielles.

Deux indicateurs résument l'évaluation finale qui est faite du projet :

- la valeur actuelle nette de la réalisation du projet qui n'est autre que la résultante de la différence entre les bénéfices et les coûts socio-économiques de la réalisation du projet mesurés en tenant compte de la durée de vie du projet et ramenés à des valeurs comparables sur la base du taux d'actualisation ;
- le taux de rendement interne du projet qui porte sur la détermination du taux d'actualisation pour lequel la valeur nette présente du projet est nulle et qui permet d'apprécier l'utilité sociale du projet.

#### ● Application aux Fonds structurels

L'analyse coût-avantages est appliquée à l'évaluation des grands projets. Elle est une obligation réglementaire pour l'appréciation ex ante des grands projets européens. Elle est réalisée par les Etats membres qui la soumettent à la Commission. Elle constitue un préalable obligatoire à la décision de financement. L'analyse coût-avantages est rarement utilisée à l'extérieur de ce cadre.

Pour la période 1989-1993, les taux de rendement économique interne des grands projets financés par les Fonds structurels sont les suivants :

Secteurs	Taux moyen	Nombre de projets
Energie	13,7	5
Eau et environnement	21,7	21
Transport	19,2	63
Industrie	18,4	14
Autres services	18,7	5
<b>Total</b>	<b>19,3</b>	<b>108</b>

Source: *Guide de l'analyse des coûts et avantages des grands projets*, Commission européenne.

### ● Points forts et limites d'utilisation

- La mise en place d'une analyse coût-avantages oblige ses concepteurs à poser des questions essentielles telles que : quelles sont les personnes concernées ? Quels sont les avantages et les inconvénients du projet ? Quelle partie de la population y gagne ou perd et comment ? En ce sens, elle induit une clarification du projet.
- L'analyse coût-avantages oblige à poser la question des impacts d'un projet dans des termes clairs. Cependant, elle ne peut pas servir à l'analyse des impacts, qui reste subordonnée à d'autres techniques d'évaluation.
- Il n'est pas toujours possible d'effectuer de façon légitime et crédible la valorisation de certains avantages ou inconvénients (par exemple la destruction d'un biotope exceptionnel).
- L'analyse coût-avantages ne prend pas directement en compte les objectifs de la politique qui est conduite. Tous les avantages et inconvénients sont intégrés, qu'ils soient importants ou marginaux par rapport aux objectifs politiques.
- La complexité réelle d'un grand projet, le nombre de personnes concernées, la définition des impacts, des effets indirects, etc. rendent très difficile la quantification des avantages et des coûts.
- La mise en œuvre du principe de compensation n'est jamais effectuée dans la réalité, et l'on est loin des constructions théoriques pures (économie publique et économie du bien-être) qui servent de justification fondamentale à cette technique.

● **Pour en savoir plus...**

Gauthier, G et Thibaut, M (eds.) (1993), *L'analyse coût-avantages, défis et controverses*, Paris : Cetai, Economica, 526 p.

Réservé aux initiés : une compilation de textes émanant de différents spécialistes sur les avantages et les limites de l'analyse coût-avantages, notamment par leurs aspects épistémologiques.

Chervel, M. (1987), *Calculs économiques publics et planification. Les méthodes d'évaluation de projets*, Paris : Publisud, 268 p.

Une présentation détaillée de l'ensemble des techniques de calcul économique public. Les principaux aspects sont abordés et une classification est proposée.

Nas, T.F. (1996), *Cost benefit Analysis. Theory and application*, London : Sage Publications.

Manuel en anglais qui reprend tous les fondements microéconomiques de l'analyse coût-avantages et qui traite des études de cas : réservé aux initiés.

Diamond, D. and Spence, N. (1983), *Regional Policy Evaluation. A Methodological Review and the Scottish Example*, Gower International Library of Research and Practice, Aldershot, 170 p.

Une présentation des principaux outils d'évaluation issus de la théorie économique appliqués aux politiques de développement régional. Une approche critique de l'analyse coût-avantages, très accessible.

Commission Européenne (1997), *Guide de l'analyse Coût-Avantages des grands projets dans le cadre de la politique régionale communautaire*, DG XVI, Bruxelles, 90 p.

Guide destiné à l'analyse coût-avantages préalable à la mise en œuvre des grands projets décrivant les principales étapes de la méthode.

● **Approfondissement : l'apport de la technique à l'évaluation de projets**

Une fois le projet clairement identifié, ses objectifs définis, sa faisabilité établie et les options alternatives envisagées, l'équipe d'évaluation est en mesure d'aborder les différentes phases de quantification du projet : l'analyse financière, l'analyse coût-avantages et l'analyse de sensibilité.

Cette description des concepts est complétée par une présentation synthétique des constats établis sur base d'une analyse des grands projets cofinancés par les Fonds structurels.

#### ◆ *Analyse financière*

L'analyse financière du projet d'investissement nécessite que soit prise en compte la programmation de cet investissement et des recettes et coûts attendus de manière à apprécier au mieux les entrées et les sorties annuelles liées à la réalisation et à la gestion du projet. Cette approche permet d'évaluer la charge financière du projet et de s'assurer qu'il ne risque pas d'être arrêté à un moment quelconque suite à un manque de liquidités. C'est sur la base des résultats de cette analyse que sera effectuée l'ACA au travers de la prise en compte d'un certain nombre de composantes complémentaires de nature socio-économique permettant de mesurer l'impact social du projet.

Pour apprécier la valeur financière actuelle nette d'un projet, on tient compte des valeurs actualisées à un taux «  $r$  » des entrées (E) et des sorties (S) financières futures (en ce compris la valeur résiduelle de l'investissement en fin de période) :

$$VFAN = (E_0 - S_0) + (E_1 - S_1) / (1+r) + (E_2 - S_2) / (1+r)^2 + \dots + (S_n - S_n) / (1+r)^n$$

Le taux d'actualisation est représentatif du coût d'opportunité du capital, c'est-à-dire de la valeur économique du capital investi dans son utilisation alternative optimale. Le taux utilisé en pratique est souvent équivalent au taux de rendement à long terme du capital investi sur les marchés financiers. Le choix du taux d'actualisation est parfois difficile et peut apparaître comme arbitraire. Lorsque l'on compare plusieurs projets, il est préférable d'utiliser le taux de rendement financier interne «  $rr$  » qui n'est autre que la valeur de «  $r$  » qui annule l'équation suivante :

$$(E_0 - S_0) + (E_1 - S_1) / (1+r) + (E_2 - S_2) / (1+r)^2 + \dots + (E_n - S_n) / (1+r)^n = 0$$

Dans le cas des investissements publics, il arrive souvent que les taux de rendement financiers observés soient faibles, voire négatifs. Ceci s'explique par une multitude de facteurs parmi lesquels les distorsions de prix, les problèmes de tarification et les externalités. D'où la nécessité d'étendre l'analyse à d'autres critères que les seules variables financières pour juger de l'opportunité de réaliser un investissement public. Ceci implique de répertorier et d'évaluer monétairement l'ensemble des coûts et avantages du projet.

### ◆ **Analyse des coûts et avantages**

De manière similaire à l'analyse financière, on calculera la valeur économique actuelle nette en utilisant la formule suivante:

$$VEAN = (A_0 - C_0) + (A_1 - C_1) / (1+r) + (A_2 - C_2) / (1+r)^2 + \dots + (A_n - C_n) / (1+r)^n$$

dans laquelle  $A_i$  représente l'ensemble des avantages attendus en période  $i$  et  $C_i$  l'ensemble des coûts à couvrir au cours de cette même période. Si la valeur économique actuelle nette est supérieure à zéro, l'évaluation du projet est favorable. Le coefficient «  $r$  » représente le taux permettant d'actualiser l'ensemble des coûts et des bénéfices en fonction de l'année de référence définie dans ce cas comme l'année 0. L'analyse des grands dossiers européens de première et de seconde générations ont mis en évidence des taux d'actualisation variant de 3% à 10%. En règle générale, le taux conseillé est de 5%.

Toujours par analogie avec l'analyse financière, on peut calculer le taux de rendement économique interne du projet («  $rr$  ») qui n'est autre que la valeur de «  $r$  » qui annule la relation suivante :

$$(A_0 - C_0) + (A_1 - C_1) / (1+r) + (A_2 - C_2) / (1+r)^2 + \dots + (A_n - C_n) / (1+r)^n = 0$$

Tous les coûts et avantages d'un projet ne pouvant faire l'objet d'une monétarisation systématique, l'analyse sera complétée par une étude approfondie des aspects non monétaires liés à la réalisation du projet.

### ◆ **Analyse de sensibilité et des risques**

Selon les cas, les résultats de l'analyse coût-avantages peuvent être extrêmement sensibles à une modification des hypothèses de travail. L'analyse de sensibilité consiste à faire varier la valeur d'un certain nombre de variables clés de manière à apprécier l'impact d'une modification des hypothèses sur le taux de rendement du projet. Si les modifications apportées à une variable-clé se traduisent par un changement important du taux de rendement, il est indispensable de renforcer la fiabilité de la mesure obtenue pour cette variable.

Pas plus que les investissements privés les grands projets publics ne sont exempts d'incertitudes. L'analyse coût-avantages étant basée sur des hypothèses quant aux impacts attendus du projet, il est vital d'apprécier le degré de risque du projet. L'analyse du risque sera effectuée en associant des probabilités de réalisation à des valeurs alternatives des variables cruciales, en déterminant les taux de rendement économiques qui y sont

associés et en calculant alors la distribution de probabilité de ces taux de rendement. Ceci permet notamment d'évaluer la probabilité que le projet ait un taux de rendement économique inférieur au taux de rendement minimal requis. Cette analyse de risque permet aux décideurs de prendre leur décision finale en connaissance de cause quant aux risques encourus par le projet.

#### ◆ *Taux de rendement des projets européens*

Pour les périodes de programmation 1989-1993 et 1994-1999 correspondant aux programmes de première et deuxième génération issus de la réforme des Fonds structurels, l'horizon moyen (durée de vie économique du projet), le taux de rendement financier interne moyen et le taux de rendement économique interne moyen d'un ensemble de 400 grands projets cofinancés par les fonds structurels ont fait l'objet d'une analyse.

Secteurs	Horizon moyen	Nombre de projets (*)	Taux de rendement financier interne moyen	Nombre de projets (*)	Taux de rendement économique interne moyen	Nombre de projets (*)
Energie	24,7	9	7,0	6	12,9	6
Eau et environnement	29,1	47	-0,1	15	15,8	51
Transport	26,6	127	6,5	55	17,1	152
Industrie	8,8	96	19,0	68	18,4	14
Autres services	14,2	10	4,2	5	16,3	10
<b>Total</b>	20,1	289	11,5	149	16,8	233

(\*) Projets pour lesquels des données sont disponibles

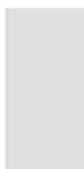
Source: *Guide de l'analyse des coûts et avantages des grands projets*, Commission des communautés européennes

Très instructif, ce tableau de synthèse permet d'épingler les aspects importants de l'apport de l'ACA à l'analyse des grands projets.

Premièrement, on observe que la durée de vie moyenne des projets varie sensiblement d'un domaine à l'autre. De 20 ans en moyenne, la durée de vie varie de 9 ans pour les infrastructures industrielles à 30 ans pour les infrastructures environnementales. L'appréciation de la durée de vie du projet est importante si l'on veut estimer correctement les impacts potentiels à moyen et long terme des investissements envisagés.

Deuxièmement, la mesure du taux de rendement financier interne permet de cerner les besoins de financement que nécessitera le projet tout au long de sa durée de vie. Cette analyse financière est également vitale pour la programmation des investissements. Alors que le rendement financier est élevé pour les infrastructures productives qui font en général l'objet de recettes substantielles (loyer ou prix de vente des terrains et installations, prix de vente des services...) et s'élève en moyenne à 19%, le taux de rendement financier des installations environnementales qui ne font que très rarement l'objet d'une tarification aux prix du marché est très faible (-0,1% en moyenne). Pour les projets qui se caractérisent par un taux de rendement financier faible, la programmation financière est d'autant plus importante que se pose la question de la viabilité du projet et de son incidence réelle sur les finances publiques.

Troisièmement, l'appréciation des coûts et avantages socio-économiques du projet permet d'évaluer sa valeur économique réelle et sa contribution au bien-être économique et social, objectif ultime de l'ACA. Un premier constat est que le taux de rendement économique interne diffère de manière significative du taux de rendement financier interne. Pour les projets cofinancés par les fonds structurels, le taux de rendement économique interne moyen varie sur une fourchette allant de 12,9% dans le secteur énergétique à 18,4% pour le secteur industriel. Ces chiffres mettent en évidence l'intérêt de la démarche coût-avantages qui permet de dépasser le cadre strict de l'analyse financière en tenant compte notamment des externalités que générera le projet.



## IV

### Recommandations pour sélectionner et assembler les outils

*Le repérage des outils les plus adaptés doit être fait dans chaque cas particulier d'évaluation. Cette partie apporte cependant une aide à ceux qui conçoivent les évaluations en décrivant quatre situations typiques : (1) question prospective sur un programme multisectoriel, (2) question rétrospective sur un programme multisectoriel, (3) approfondissement d'une question prospective et (4) approfondissement d'une question rétrospective. Dans chacune de ces quatre situations, les outils présentés dans cet ouvrage sont passés en revue en examinant leur plus ou moins bonne adaptation. L'ouvrage se termine par des recommandations applicables dans le cas où l'on assemble plusieurs outils au sein d'une même évaluation.*



## Guide pour repérer les outils adaptés

Le repérage des outils les plus adaptés doit être fait au cas par cas en fonction de la nature de l'objet à évaluer, et surtout en fonction des questions posées par le commanditaire. Il n'est donc pas question de donner des recettes générales. Il est cependant possible d'aider ceux qui conçoivent les évaluations en décrivant quelques situations typiques. Tout d'abord, l'objet à évaluer peut être l'ensemble d'un programme multisectoriel ou au contraire une question d'évaluation particulière qu'il s'agit d'approfondir.

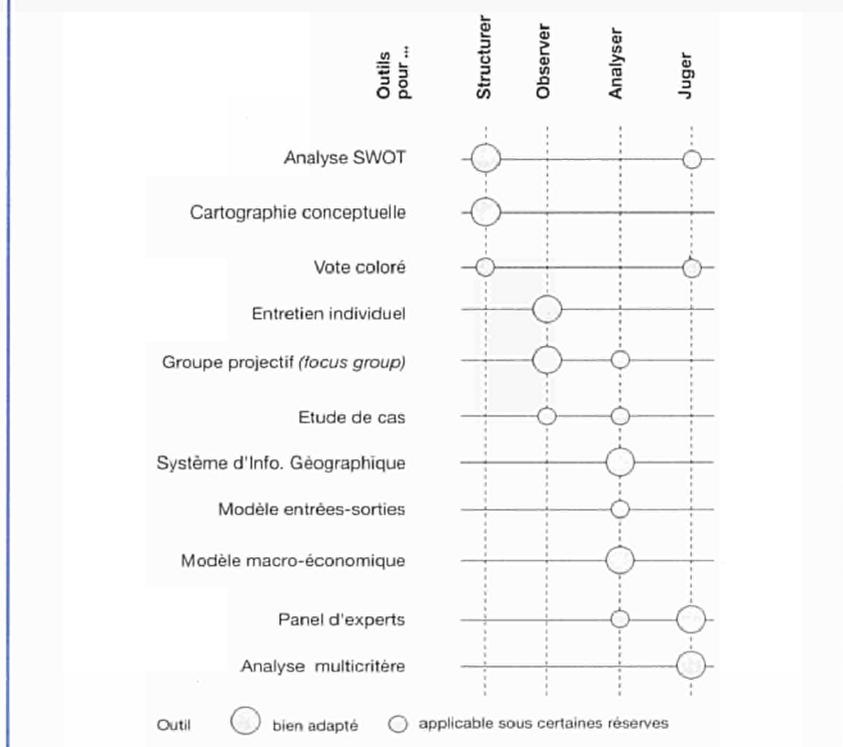
Du point de vue du choix des outils, il est également important de prendre en compte la dimension temporelle de la question posée. La question est-elle prospective ou rétrospective ? Une question est **rétrospective** si l'on peut y répondre en s'appuyant sur une observation des effets sur le terrain, c'est le cas de l'évaluation ex post. Elle est **prospective** dans le cas contraire, notamment pour l'évaluation ex ante. L'évaluation à mi-parcours peut comporter un mélange de questions prospectives (ex : quel est l'impact prévisionnel des projets qui ont déjà été sélectionnés ?) et de questions rétrospectives (ex : quels sont les premiers résultats obtenus sur le terrain ?). La combinaison de ces distinctions permet de définir les situations d'évaluation présentées ci-dessous.

<b>Encadré 29 - Situations types</b>		
	<b>question prospective</b>	<b>question rétrospective</b>
<b>évaluation d'ensemble d'un programme multisectoriel</b>	Situation 1  par exemple : évaluation ex ante ou intermédiaire d'un programme régional ou d'un CCA	Situation 2  par exemple : évaluation intermédiaire ou ex post d'un programme régional ou d'un CCA
<b>approfondissement d'une question évaluative</b>	Situation 3  par exemple : évaluation ex ante ou intermédiaire d'une mesure ou d'un programme sectoriel	Situation 4  par exemple : évaluation intermédiaire ou ex post d'une mesure ou d'un programme sectoriel

Le guide présenté ci-après propose donc une aide pour la sélection des outils dans ces quatre situations. Il est limité aux outils qui font l'objet d'une fiche technique dans cet ouvrage. Sa logique pourrait aisément être étendue à d'autres outils. Il ne recommande pas d'appliquer systématiquement tel outil à telle situation mais propose généralement plusieurs outils parmi lesquels on pourra choisir celui (ou ceux) qui est le mieux adapté à la situation considérée.

● **Outils pour l'évaluation d'ensemble d'un programme, question prospective**

**Encadré 30 - Evaluation d'ensemble d'un programme, question prospective**

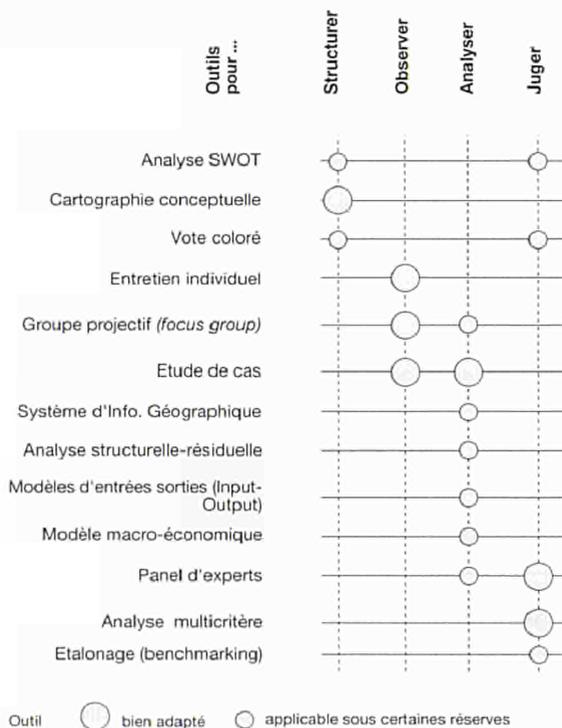


Dans cette situation, les grandes lignes du programme ont déjà été proposées. L'évaluation doit contribuer à ce que le programme projeté ait

des objectifs mieux justifiés, plus clairs, hiérarchisés et logiquement reliés aux actions prévues. Si nécessaire, l'évaluation peut comporter une observation de terrain pour vérifier la nature et l'importance des besoins. Elle peut également comporter une estimation prévisionnelle des effets, notamment pour aider à quantifier les objectifs. Une évaluation de ce type mérite d'être outillée si l'on dispose d'un délai et d'un budget suffisant.

● Outils pour l'évaluation d'ensemble d'un programme, question rétrospective

**Encadré 31 - Evaluation d'ensemble d'un programme, question rétrospective**

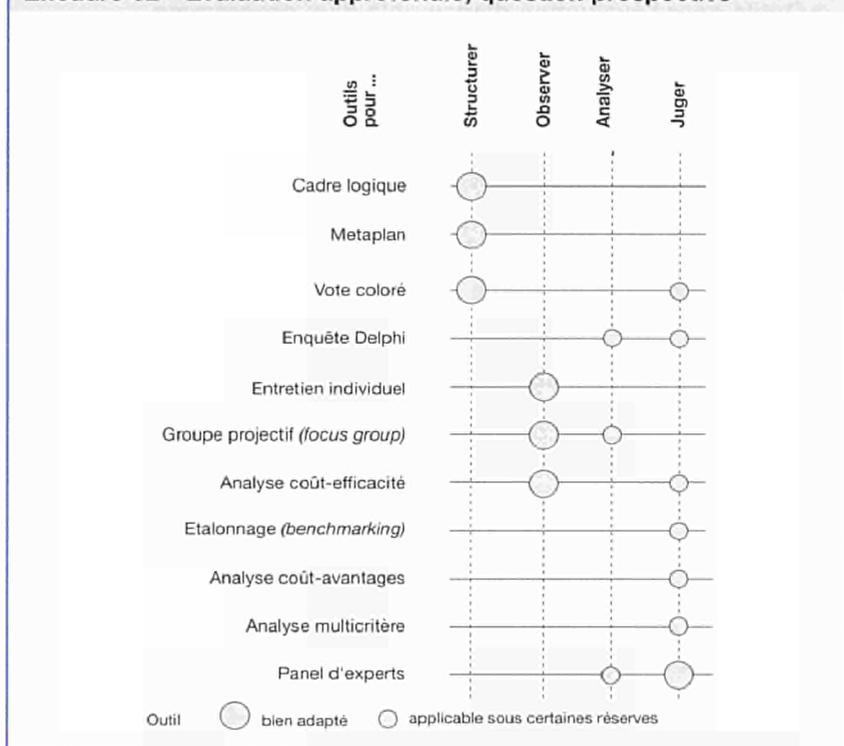


Dans cette situation, l'évaluation doit observer, analyser et juger les effets que le programme a produit sur le terrain. Ces effets sont nombreux et ils sont produits par des processus multiples qui diffèrent selon les

composantes du programme. En outre, l'évaluation peut avoir à tenir compte de différences de points de vue entre les partenaires. Cette situation présente des contraintes qui limitent fortement le choix des outils.

● Outils pour l'évaluation approfondie, question prospective

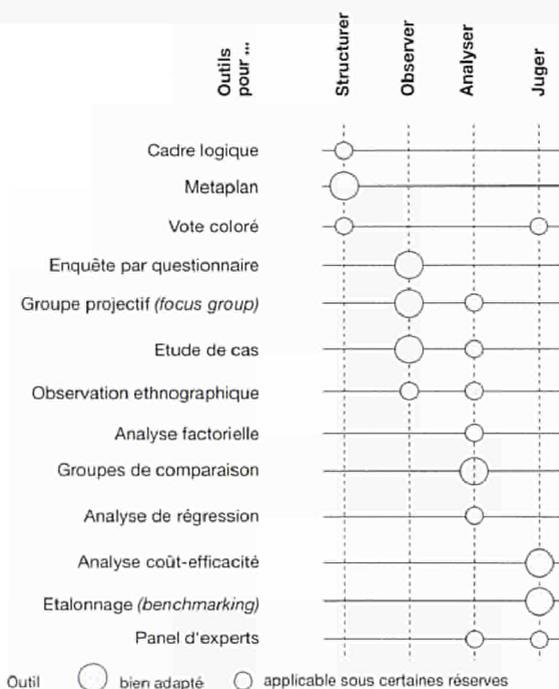
**Encadré 32 - Evaluation approfondie, question prospective**



Dans cette situation, on évalue de façon approfondie un aspect ou une partie d'un programme complexe. Cela ramène au cas d'une intervention simple, constituée d'actions similaires, s'adressant à un public relativement homogène et dotée d'un objectif principal. Les questions prospectives qui se posent généralement dans cette situation est celle de la faisabilité et de la pertinence de l'intervention projetée. Pour y répondre, l'évaluation doit aider à comprendre comment cette intervention va produire ses effets, quelle sera vraisemblablement l'ampleur de ces effets et dans quelle mesure l'intervention projetée répond à un réel besoin.

## ● Outils pour l'évaluation approfondie, question rétrospective

### Encadré 33 - Evaluation approfondie, question rétrospective



Dans cette situation, l'évaluation observe, analyse et juge les effets qui ont été obtenus sur le terrain. Ces effets sont peu nombreux et sont produits par des processus qu'il est possible de schématiser. De ce fait, il est possible de les étudier de façon approfondie. On peut également consacrer du temps à rechercher de bonnes références de comparaison pour juger si les effets obtenus sont suffisants ou non. La plupart des outils classiques de l'évaluation sont applicables à cette situation.



## Recommandations pour assembler plusieurs outils

### ● Démarche pour l'assemblage des outils

Le choix des outils et leur assemblage se font généralement par améliorations successives. Au départ, la nature de l'évaluation (programme multisectoriel ou question plus simple) et celle de la question posée (prospective ou rétrospective) suggèrent que certains outils pourraient être bien adaptés. Le guide qui précède peut aider à repérer ces outils.

A ce stade, on construit une méthode provisoire et on vérifie sa qualité à l'aide des questions suivantes :

- Les huit fonctions standard d'une évaluation sont-elle traitées ? Certaines fonctions sont-elles inutiles ? Certaines fonctions ont-elles été oubliées ?
- Quels outils ont été choisis ? Quelles fonctions remplissent ces outils ? Sont-ils bien adaptés aux fonctions à traiter ? Sont-ils appliqués conformément aux règles ?
- Est-ce que l'ensemble de la méthode est solide et homogène ?

Cette dernière question est peut-être la plus importante. On peut en effet comparer une méthode d'évaluation à une chaîne. Chaque fonction est un maillon de la chaîne et chaque maillon est d'autant plus solide que la fonction est traitée par un outil approprié. La métaphore de la « chaîne technique » souligne l'interdépendance des différents outils qui sont assemblés au sein de la méthode d'évaluation.

Un maillon est robuste si la fonction correspondante est traitée avec un niveau de technicité suffisant par rapport au risque de contestation des résultats de l'évaluation. Il n'est pas obligatoire qu'un maillon soit « outillé » pour être robuste. Cependant, si un maillon est outillé, sa solidité dépend du fait que l'outil est adapté à la question posée et à la situation.

La solidité de l'ensemble de la chaîne technique est celle de son maillon le plus faible. Ainsi, il est inutile d'employer un outil onéreux et sophistiqué pour estimer les effets si cet outil s'applique à des données collectées de façon contestable. La robustesse d'une méthode d'évaluation n'est que partiellement liée à la qualité des outils employés. Ce n'est pas tant la haute technicité du traitement de certaines fonctions qui fait une bonne évaluation, c'est plutôt le fait que toutes les fonctions sont traitées avec un niveau de technicité homogène. Quand on a construit une méthode

d'évaluation provisoire, on doit donc vérifier en priorité si elle présente des maillons faibles.

**Encadré 34 - La solidité d'une chaîne technique est celle de son maillon le plus faible**

Clarifier / hiérarchiser les effets et critères

Définir les instruments d'observation

Circonscrire le champ d'observation

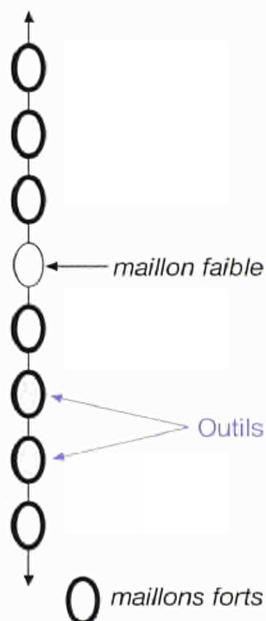
Collecter les données

Croiser les informations

Estimer les effets

Juger selon les différents critères

Formuler un jugement de synthèse



La construction de la méthode d'évaluation procède ensuite par des améliorations successives jusqu'à satisfaire les impératifs suivants :

- toutes les fonctions qui doivent être traitées le sont ;
- le ou les outils sélectionnés sont adaptés à la situation et à la fonction qu'ils remplissent ;
- la méthode construite est homogène et n'a pas de point faible ;
- la méthode choisie est compatible avec les contraintes de coût et de délai de l'évaluation.

On ne peut pas rechercher une robustesse maximum, ni choisir des outils nombreux et perfectionnés si les contraintes de temps et de budget sont trop sévères. Le réalisme d'une méthode impose que sa robustesse, mais

également son coût, soient proportionnés aux enjeux de l'évaluation pour ses clients. Tant que la méthode proposée ne répond pas à cette exigence de réalisme, il faut l'ajuster en supprimant des outils ou en les remplaçant par des outils plus légers à mettre en œuvre.

Un problème trop souvent oublié est qu'on ne construit pas une méthode pour chaque évaluation, mais bien une méthode pour chaque question au sein d'une évaluation. Supposons qu'une même évaluation contienne les deux questions suivantes :

- A. Le programme a-t-il eu un impact suffisant en termes de réduction du chômage dans la région ?
- B. Parmi les mesures du programme, quelles sont celles qui sont les plus efficaces pour la création d'emploi, et pourquoi ?

La question A devra être traitée par une estimation d'impact de toutes les mesures, au besoin avec l'aide d'un modèle économétrique. Par contre, la question B serait mieux traitée avec une méthode reposant sur l'observation et l'analyse de projets particulièrement réussis en termes de création d'emplois (études de cas). Il arrive qu'une même méthode puisse répondre simultanément à deux ou plusieurs questions, mais elle ne le fait généralement pas avec le même niveau de robustesse. En première approximation, on peut donc dire que le coût d'une évaluation augmente en même temps que le nombre de questions figurant dans les termes de référence (et auxquelles on donne une véritable réponse).

### ● Trois exemples d'assemblage d'outils

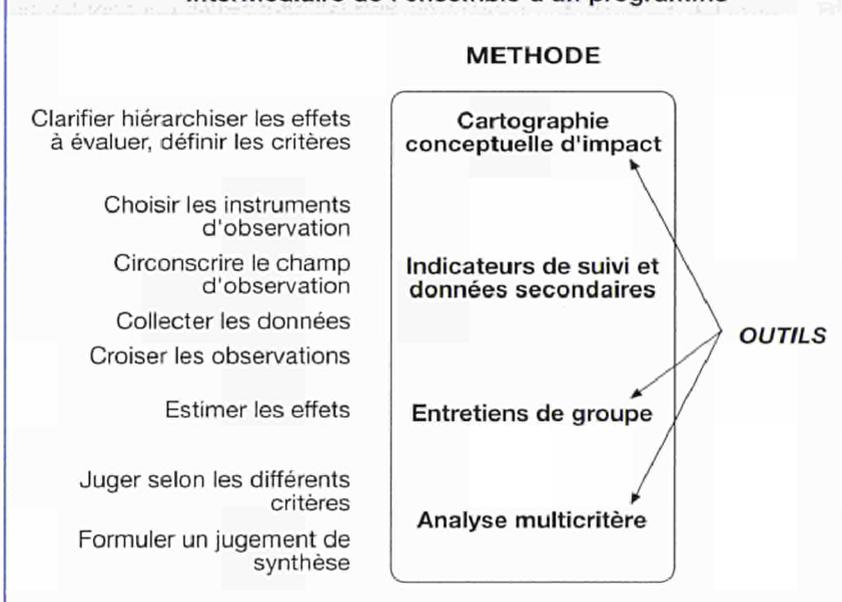
Les trois propositions qui suivent sont construites à partir d'exemples réels présentés dans les chapitres précédents.

Le premier exemple est l'évaluation d'ensemble du programme de développement régional du Hainaut qui est présenté à la page 129 ainsi que dans le Volume 4 de cette Collection consacré aux techniques innovantes développées dans le cadre du programme MEANS. Cette évaluation intermédiaire avait pour but de comparer l'efficacité des principales mesures du programme. Elle s'est terminée par une analyse multicritère qui a permis d'aboutir à un jugement synthétique. Ce jugement s'appuyait sur une estimation qualitative des effets de chaque mesure.

Pour renforcer la solidité des estimations d'impacts, on pourrait par exemple construire la méthode suivante :

- clarification des effets à évaluer avec l'aide d'une cartographie conceptuelle d'impacts ;
- rassemblement des indicateurs et des données secondaires disponibles sur chaque mesure évaluée ;
- organisation d'un ou plusieurs entretiens de groupe pour chacune des mesures évaluées ; estimation qualitative des impacts ;
- comparaison entre mesures et jugement de synthèse à l'aide d'une analyse multicritère.

**Encadré 35 - Exemple d'assemblage d'outils pour l'évaluation intermédiaire de l'ensemble d'un programme**

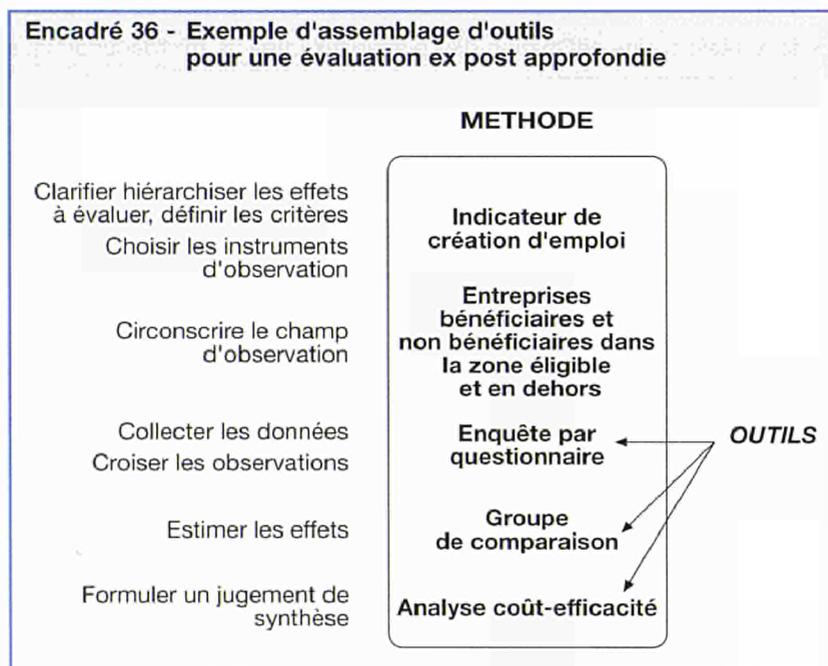


Le second exemple est construit à partir de l'évaluation de la création d'emploi au Pays de Galles décrite à la page 6 ainsi qu'à l'Encadré 3. Cette évaluation ex post s'attachait à analyser un impact de façon approfondie. Elle a comporté une enquête par questionnaire auprès des entreprises aidées qui a permis d'estimer les effets d'aubaine et de déplacement, ainsi qu'une analyse coût efficacité qui a permis de juger l'efficacité du programme par comparaison avec un programme écossais. Les maillons faibles de cette évaluation étaient probablement l'estimation des effets d'aubaine et la comparaison d'efficacité.

Pour renforcer l'évaluation sur ce point, on aurait pu améliorer la méthode de la façon suivante :

- définition du champ d'observation en trois groupes : les entreprises galloises aidées du programme, les entreprises non galloises aidées d'autres programmes considérés comme performants, les entreprises galloises non touchées par le programme ;
- collecte de données au moyen d'une enquête par questionnaire ;
- estimation des effets grâce aux groupes de comparaison ;
- analyse coût efficacité et jugement par rapport à d'autres programmes performants.

Cette méthode aurait évidemment été beaucoup plus onéreuse.



Le troisième exemple est construit à partir de l'évaluation ex ante de l'aide européenne à la République d'Irlande présentée à la page 39 dans l'Encadré 13. Cette évaluation s'efforçait d'estimer l'impact probable de l'aide sur le développement économique, notamment en termes de PIB par habitant. L'évaluation a comporté une première série d'estimations

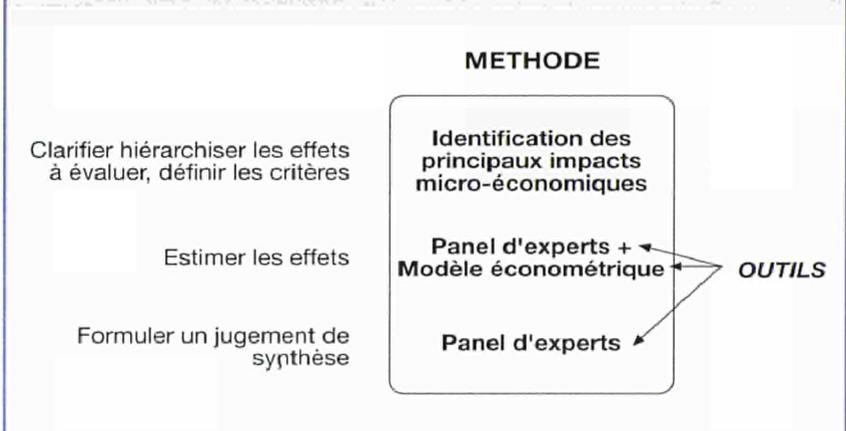
microéconomiques (impact sur la dépense publique, sur la productivité, sur les coûts de production). Ces estimations ont ensuite été introduites dans un modèle macroéconomique. Les estimations d'impacts microéconomiques sur la productivité et les coûts de production étaient certainement les maillons faibles de cette évaluation. Par ailleurs, l'évaluation s'est arrêtée à l'estimation des effets macroéconomiques et n'a pas formellement été jusqu'à juger si ces effets étaient suffisants ou insuffisants.

Pour améliorer l'homogénéité de la méthode d'évaluation, on aurait pu consacrer un peu moins de temps et d'argent au maniement du modèle économétrique, tout en renforçant les maillons faibles de la façon suivante :

- clarifier et hiérarchiser les impacts microéconomiques probables des différentes mesures, au besoin en s'appuyant sur une **matrice d'impacts** ;
- faire réaliser une estimation des principaux impacts microéconomiques par un **panel d'experts** ;
- estimer les conséquences économiques au niveau national au moyen d'un **modèle macroéconomique** comme cela a été fait en pratique ;
- réunir un **panel d'experts** et lui demander de conclure sur le fait que les impacts probables sont suffisant ou insuffisants.

Sans être beaucoup plus onéreuse, cette méthode aurait été à la fois plus crédible et plus conclusive.

### Encadré 37 - Exemple d'assemblage d'outils pour l'évaluation ex ante de l'ensemble d'un programme



## Conclusion

Le recours à un outil a de nombreux avantages en termes d'assurance sur le résultat obtenu, de sécurité sur le contenu de l'évaluation, de maîtrise des coûts, de comparabilité des résultats et de contrôle de qualité. Une évaluation plus technique est souvent plus simple à commander, plus facile à contrôler, plus crédible pour les lecteurs du rapport et plus prévisible pour le commanditaire. On comprend que tous ces avantages rendent les outils d'évaluation très attractifs.

Puisque la technicité est attirante, on a souvent tendance à survaloriser les outils employés en évaluation. Tout un travail d'évaluation se trouve parfois résumé par le nom d'un des outils mis en œuvre (ex : évaluation coût-avantages, évaluation macroéconomique) et qui prend une dimension emblématique. Il est cependant dangereux de survaloriser un outil en oubliant qu'il ne permet le plus souvent de remplir qu'une seule fonction d'évaluation, qu'il est adapté à un certain contexte, qu'il présente des contraintes spécifiques de disponibilité de données, qu'il a un coût et une durée de mise en œuvre incompressibles. Choisir un outil sans tenir compte de toutes ces dimensions risque d'entraîner des conséquences graves : absence de conclusions, dépassement de délai, conclusions pas crédibles, conclusions non-pertinentes, etc.

Cet ouvrage ne doit donc pas être considéré comme une publicité pour des évaluations les plus techniques possibles. L'ambition est plutôt d'adapter le niveau de technicité à l'importance stratégique des questions posées, ainsi qu'au budget et au délai alloué à l'évaluation. Cet ouvrage doit donc être considéré comme une aide à la maîtrise de la technicité des évaluations.



## A

- analyse coût-avantages 201
- analyse coût-efficacité 10, 24, 201, 205, 230
- analyse de faisabilité 68
- analyse de régression 21, 85, 136
- analyse de variance 176
- analyse factorielle 95, 136, 176, 177
- analyse multiforme 121, 129, 201, 238
- analyse qualitative 28
- analyse quantitative 28
- analyse structurelle réticulaire 65, 67
- analyse SWOT 22, 41, 43, 121
- approche 16
- approche narrative 10
- approche pragmatique 12

## B

- benchmarking déclinage 121

## C

- carte logique 42, 141, 143
- cartographie conceptuelle d'activités 41, 49, 141, 148
- check technique 235
- champ 27
- critère 26, 29

## D

- démarche 16
- démarche déductive 23
- démarche inductive 23
- descripteur 26
- désign 19
- donnée primaire 27
- donnée secondaire 27

## E

- état net 24
- évaluation Deshpai 176, 185, 201
- enquête par questionnaire 61, 126, 139, 167, 169, 229
- évaluation de groupe 31, 71, 163, 173, 233
- évaluation individuelle 61, 63, 123, 167
- évaluation 121, 201, 209
- étude de cas 61, 77, 85, 136, 157, 176
- évaluation à 360 degrés 31, 219
- évaluation approfondie 31, 229, 232, 233, 238
- évaluation d'ensemble 31, 229, 230, 231, 237, 238
- évaluation ex ante 31, 229, 238
- évaluation ex post 31, 229, 238
- évaluation formative 30
- évaluation intermédiaire 237
- évaluation rétrospective 30
- évaluation traditionnelle 32

## F

- fonction plane évaluation 8, 24, 258

## G

- groupe de comparaisons 65, 126, 173, 181, 229

## I

- indicateur 26
- intra-cultural analysis
- matrice d'analyse-synthèse 66

## L

- l'enquête par questionnaire 167

## Index



**A**

- analyse coût-avantages 201
- analyse coût-efficacité 19, 24, 201, 203, 239
- analyse de causalité 28
- analyse de régression 21, 85, 175, 195
- analyse de variance 175
- analyse factorielle 85, 138, 175, 177
- analyse multicritère 121, 129, 201, 238
- analyse qualitative 28
- analyse quantitative 28
- analyse structurelle-résiduelle 85, 93
- analyse SWOT 22, 41, 43, 121
- approche 18
- approche naturaliste 12
- approche pragmatique 12

**B**

- benchmarking étalonnage 121

**C**

- cadre logique 42, 141, 143
- cartographie conceptuelle d'impacts 41, 49, 141, 238
- chaîne technique 235
- champ 27
- critère 26, 29

**D**

- démarche 18
- démarche déductive 28
- démarche inductive 27
- descripteur 26
- design 18
- donnée primaire 27
- donnée secondaire 27

**E**

- effet net 28
- enquête Delphi 175, 183, 201
- enquête par questionnaire 61, 138, 139, 157, 159, 239
- entretien de groupe 61, 71, 157, 175, 238
- entretien individuel 61, 63, 139, 157
- étalonnage 121, 201, 209
- étude de cas 61, 77, 85, 139, 157, 175
- évaluation à mi-parcours 31, 229
- évaluation approfondie 31, 229, 232, 233, 238
- évaluation d'ensemble 31, 229, 230, 231, 237, 239
- évaluation ex ante 31, 229, 239
- évaluation ex post 31, 229, 238
- évaluation formative 30
- évaluation intermédiaire 237
- évaluation récapitulative 30
- évaluation thématique 32

**F**

- fonction (d'une évaluation) 18, 24, 235

**G**

- groupe de comparaison 85, 139, 175, 187, 239

**I**

- indicateur 26
- input-output analysis
- modèle d'entrée-sortie 86

**L**

- l'enquête par questionnaire 137

## M

- matrice d'impacts 41, 141, 240
- Metaplan® 20, 141, 151
- méthode 18, 235
- méthodologie 19
- modèle Beutel 105
- modèle d'entrée-sortie 86
- modèle HERMIN 30, 118
- modèle macroéconomique 30, 39, 86, 107, 240
- modèle QUEST 116

## O

- observation ethnographique 157, 169, 175
- outil 18
- outil générique 24
- outil spécialisé 24

## P

- panel d'experts 121, 123, 137, 201, 240
- public cible 32

## Q

- question évaluative 26

## S

- shift-share analysis
  - analyse structurelle résiduelle 85
- SIG
  - système d'information géographique 85
- structuration 26
- système d'information géographique 38, 85, 87

## T

- technique 18
- technique économique 12
- technique expérimentale 12

## V

- vote coloré 41, 121, 141, 201

## Remerciements

La préparation de ce Volume s'est étalée sur plusieurs années, de 1995 à 1998. Les travaux préparatoires ont comporté les étapes suivantes :

- Cinq études de cas ont été réalisées sur la base d'évaluations sélectionnées en raison de l'emploi de techniques d'évaluation reconnues. Cinq universitaires européens ont participé à ces études dont la synthèse a permis de dégager les règles d'utilisation des outils d'évaluation qui sont présentées dans ce Volume.
- Près de trois cents fonctionnaires de la Commission européenne ont participé à des séminaires de formation, comportant une journée spécialement consacrée aux outils d'évaluation. Le contenu de ces séminaires a progressivement évolué en fonction des attentes des participants. Ce Volume tient compte de ces interactions.
- Plusieurs experts ont participé à la rédaction et à la vérification des fiches.

La coordination des travaux de ce volume a été assurée par Miguel Benito Alonso (D.G. XVI.G.2 : Coordination de l'évaluation) et a impliqué plusieurs Directions Générales de la Commission Européenne, notamment les D.G. II, V, VI, XIV et XVI et en particulier Philippe Goybet, Anastassios Bougas, Martin Hallet, Andrea Mairate et Julius Op de Beke.

Ce Volume a été rédigé par le Centre Européen d'Expertise en Evaluation, notamment avec la participation de Kristian Colletis-Wahl, Olaf Rieper et Kevin Williams.

Les auteurs souhaitent remercier toutes celles et tous ceux qui ont contribué, soit directement, soit indirectement, à la production de ce Volume et en particulier les experts dont les noms suivent : Henri Capron (Université Libre de Bruxelles, Belgique), Dr. Basil Cracknell (consultant, Grande Bretagne), Prof. Massimo Florio (CSIL, Milan), Christophe Kellerhals (Evaluanda, Genève), Steve Martin (Warwick Business School, Warwick), Thanos Skouras (Athens University of Economics and Business), Evert Vedung (Uppsala Universitet).



Commission européenne

**Évaluer les programmes socio-économiques  
Principales techniques et outils d'évaluation**

Luxembourg: Office des publications officielles des Communautés européennes

1999 — 247 p. — 14,8 x 21 cm

ISBN 92-828-6627-0



**BELGIQUE/BELGIË**

**Jean De Lannoy**  
Avenue du Roi 302/Koninglaan 202  
B-1190 Bruxelles/Brussel  
Tel: (32-2) 538 43 08  
Fax: (32-2) 538 08 41  
E-mail: jean.de.lannoy@infoboard.be  
URL: http://www.jean-de.lannoy.be

**La librairie européenne/  
De Europese Boekhandel**  
Rue de la Loi 244/Velstraaat 244  
B-1040 Bruxelles/Brussel  
Tel: (32-2) 295 26 39  
Fax: (32-2) 735 06 60  
E-mail: mail@europop.be  
URL: http://www.europop.be

**Montaur beige/Belgisch Staatsblad**  
Rue de Louvain 40-42/Luuvenseweg 40-42  
B-1000 Bruxelles/Brussel  
Tel: (32-2) 552 22 11  
Fax: (32-2) 511 01 84

**DANMARK**

**J. H. Schultz Information A/S**  
Herstedvang 10-12  
DK-2620 Albertslund  
Tel: (45) 43 83 23 00  
Fax: (45) 43 83 19 69  
E-mail: schultz@schultz.dk  
URL: http://www.schultz.dk

**DEUTSCHLAND**

**Bundesanzeiger Verlag GmbH  
Vertriebsabteilung**  
Amsterdamer Straße 192  
D-50735 Köln  
Tel: (49-221) 97 66 80  
Fax: (49-221) 97 66 82 78  
E-Mail: vertrieb@bundesanzeiger.de  
URL: http://www.bundesanzeiger.de

**ΕΛΛΑΔΑ/GREECE**

**G. C. Eleftheroudakis SA**  
International Bookstore  
Pansipistimou 17  
GR-10564 Athens  
Tel: (30-1) 331 41 80/12/3/4/5  
Fax: (30-1) 323 96 21  
E-mail: ebooks@netor.gr

**ESPAÑA**

**Boletín Oficial del Estado**  
Trinidad, 27  
E-28017 Madrid  
Tel: (34) 915 38 21 11 (Líbroe),  
913 84 17 15 (Suscrip.)  
Fax: (34) 915 38 21 21 (Líbroe),  
913 84 17 14 (Suscrip.)  
E-mail: clientes@com.boe.es  
URL: http://www.boe.es

**Mundi Prensa Libros, SA**  
Castiello, 37  
E-28001 Madrid  
Tel: (34) 914 36 37 00  
Fax: (34) 915 75 39 98  
E-mail: lorena@mundiprensa.es  
URL: http://www.mundiprensa.es

**FRANCE**

**Journal officiel**  
Service des publications des CE  
26, rue Desaix  
F-75727 Paris Cedex 15  
Tel: (33) 1 40 56 77 31  
Fax: (33) 1 40 55 77 00  
URL: http://www.journal-officiel.gouv.fr

**IRELAND**

**Government Supplies Agency**  
Publications Section  
4-5 Harcourt Road  
Dublin 2  
Tel: (353-1) 661 31 11  
Fax: (353-1) 475 27 60

**ITALIA**

**l'Espresso SpA**  
Via Duca di Calabria, 1/1  
Casella postale 552  
I-50125 Firenze  
Tel: (39) 055 84 83 1  
Fax: (39) 055 64 12 57  
E-mail: licosa@l'eboc.it  
URL: http://www.l'eboc.it/licosa

**LUXEMBOURG**

**Messagerie du livre SARL**  
5, rue Raiffaens  
L-2411 Luxembourg  
Tel: (352) 40 10 20  
Fax: (352) 49 06 61  
E-mail: mail@mdl.lu  
URL: http://www.mdl.lu

**NETHERLAND**

**SDU Servicecentrum Uitgevers**  
Christoffel Plantinstraat 2  
Postbus 20014  
2500 EA Den Haag  
Tel: (31-70) 378 96 60  
Fax: (31-70) 378 97 63  
E-mail: sdu@sdu.nl  
URL: http://www.sdu.nl

**ÖSTERREICH**

**Manz'sche Verlags- und  
Universitätsbuchhandlung GmbH**  
Kohlmarkt 16  
A-1014 Wien  
Tel: (43-1) 53 16 11 00  
Fax: (43-1) 53 16 11 67  
E-Mail: bestellen@manz.co.at  
URL: http://www.manz.at/index.htm

**PORTUGAL**

**Distribuidora de Livros Bertrand Ld.**  
Grupo Bertrand, SA  
Rua das Terras dos Vales, 4-A  
Apartado 50037  
P-2700 Amadora  
Tel: (351-1) 495 90 50  
Fax: (351-1) 496 02 55  
Imprensa Nacional/Casa da Moeda, EP  
Rua Marquês SA da Boadisa, 16-A  
P-1050 Lisboa Codex  
Tel: (351-1) 353 03 99  
Fax: (351-1) 353 02 94  
E-mail: del.incm@mail.telepac.pt  
URL: http://www.incm.pt

**SUOMI/FINLAND**

**Akateeminen Kirjakauppa/  
Akademiska Bokhandeln**  
Keskuskatu 1/Centrålgatan 1  
PL/PB 128  
FI-00101 Helsinki/Helsingfors  
P/n:n (358-9) 121 44 18  
F/fax (358-9) 121 44 35  
Sähköposti: akatila@akateeminen.com  
URL: http://www.akateeminen.com

**SVENIGE**

**BTJ AB**  
Traktörvägen 11  
S-221 82 Lund  
Tel: (46-46) 18 00 00  
Fax: (46-46) 30 79 47  
E-post: btjpub@btj.se  
URL: http://www.btj.se

**UNITED KINGDOM**

**The Stationery Office Ltd**  
International Sales Agency  
51 Nine Elms Lane  
London SW8 5DR  
Tel: (44-171) 873 90 90  
Fax: (44-171) 873 54 63  
E-mail: enq.enquiries@hso.co.uk  
URL: http://www.tb-stationery-office.co.uk

**ISLAND**

**Bokabud Lausar Blöndal**  
Skólavörðslög, 2  
IS-101 Reykjavík  
Tel: (354) 551 56 50  
Fax: (354) 552 55 80

**NORGE**

**Sveits Norge AS**  
Ostenveien 18  
Boks 6512 Etterstad  
N-0606 Oslo  
Tel: (47-22) 97 45 00  
Fax: (47-22) 97 45 45

**SCHWEIZ/SUISSE/SVIZZERA**

**Euro Info Center Schweiz**  
c/o OSEC  
Stämpfenbachstraße 65  
PF 492  
CH-8005 Zürich  
Tel: (41-1) 365 53 15  
Fax: (41-1) 365 54 11  
E-mail: eics@osec.ch  
URL: http://www.osec.ch/eics

**BALGARIA**

**Europress Euromedia Ltd**  
59, bul Vitosha  
BG-1000 Sofia  
Tel: (359-2) 980 37 66  
Fax: (359-2) 980 42 30  
E-mail: Milena@mboc.bg

**ČESKÁ REPUBLIKA**

**ŠSIS**  
NIS-prodávna  
Javnickova 22  
CZ-130 00 Praha 3  
Tel: (420-2) 24 23 14 86  
Fax: (420-2) 24 23 11 14  
E-mail: nikpso@dec.nis.cz  
URL: http://usisr.cz

**CYPRUS**

**Cyprus Chamber of Commerce and Industry**  
PO Box 1455  
CY-1509 Nicosia  
Tel: (357-2) 66 95 00  
Fax: (357-2) 66 10 44  
E-mail: demetrap@ccc.org.cy

**ESTI**

**Eesti Kaubandus- ja Tööstuskoja (Estonian Chamber of Commerce and Industry)**  
Toom-Kooli 17  
EE-0001 Tallinn  
Tel: (372) 646 90 44  
Fax: (372) 646 02 45  
E-mail: erio@koda.ee  
URL: http://www.koda.ee

**HRVATSKA**

**Mediatrade Ltd**  
Pavla Hatza 1  
HR-10000 Zagreb  
Tel: (385-1) 481 94 11  
Fax: (385-1) 481 94 11

**MAGYARORSZÁG**

**Euro Info Service**  
Európa Ház  
Marsóczy utca  
PO Box 475  
H-1396 Budapest 62  
Tel: (36-1) 350 80 35  
Fax: (36-1) 350 82 30  
E-mail: euroinfo@mail.matav.hu  
URL: http://www.euroinfo.hu/index.htm

**MALTA**

**Miller Distributors Ltd**  
Malta International Airport  
PO Box 25  
Luqa LQA 05  
Tel: (356) 66 44 88  
Fax: (356) 67 67 99  
E-mail: gwirth@usa.net

**POLSKA**

**Asa Polonia**  
Krakowska Przedmieście 7  
Skw. pocztowa 1001  
PL-01-050 Warszawa  
Tel: (48-22) 826 12 01  
Fax: (48-22) 826 62 40  
E-mail: ars\_pok@bry.hsn.com.pl

**ROMÂNIA**

**Euromedia**  
Str: G-ral Beldiuc Nr 41  
RO-70419 Bucuresti  
Tel: (40-1) 315 44 03  
Fax: (40-1) 314 22 86

**ROSSIYA**

**CEEC**  
50-letiya Otkrybrya Av. 9  
117312 Moscow  
Tel: (7-095) 135 52 27  
Fax: (7-095) 135 52 27

**SLOVAKIA**

**Centrum VTI SR**  
Nám. Slobody, 19  
SK-81223 Bratislava  
Tel: (421-7) 54 41 83 64  
Fax: (421-7) 54 41 83 64  
E-mail: europ@ibti.sltk.stuba.sk  
URL: http://www.sltk.stuba.sk

**SLOVENIA**

**Gospodarski Vestnik**  
Dunajska cesta 5  
SI-1000 Ljubljana  
Tel: (386) 613 09 16 40  
Fax: (386) 613 09 16 45  
E-mail: europ@gvestnik.si  
URL: http://www.gvestnik.si

**TÜRKIYE**

**Dünya İntelöl AS**  
100. Yıl Mahallesi 34440  
TR-80050 Bağcılar-İstanbul  
Tel: (90-212) 629 46 89  
Fax: (90-212) 629 46 27  
E-mail: intelol@dunya-gazete.com.tr

**AUSTRALIA**

**Hunter Publications**  
PO Box 404  
3067 Abbotsford, Victoria  
Tel: (61-3) 94 17 53 61  
Fax: (61-3) 94 19 71 54  
E-mail: jpdaves@ozemail.com.au

**CANADA**

**Les éditions La Liberté Inc.**  
3020, chemin Sainte-Foy  
Q1A 3V3 Sainte-Foy, Québec  
Tel: (1-418) 658 37 63  
Fax: (1-800) 567 54 49  
E-mail: liberte@mediom.qc.ca

**Renouf Publishing Co. Ltd**

5369 Chemin Canolek Road Unit 1  
K1J 9J3 Ottawa, Ontario  
Tel: (1-613) 276 26 65  
Fax: (1-613) 745 76 60  
E-mail: orderdept@renoufbooks.com  
URL: http://www.renoufbooks.com

**ISRAEL**

**ROY International**  
41, Mishmar Hayarden Street  
PO Box 12056  
61130 Tel Aviv  
Tel: (972-3) 649 94 69  
Fax: (972-3) 646 60 30  
E-mail: roy@nebivon.net.il  
URL: http://www.royint.co.il

**Sub-agent for the Palestinian Authority:**

**Index Information Services**  
PO Box 19502  
Jerusalem  
Tel: (972-2) 627 16 34  
Fax: (972-2) 82 17 19

**JAPAN**

**PSI-Japan**  
Asahi Sanbancho Plaza #206  
7-1 Sanbancho, Chiyoda-ku  
Tokyo 102  
Tel: (81-3) 32 34 69 21  
Fax: (81-3) 32 34 69 15  
E-mail: books@psi-japan.co.jp  
URL: http://www.psi-japan.com

**MALAYSIA**

**EBIC Malaysia**  
Level 7, Wisma Hong Leong  
18 Jalan Perak  
50450 Kuala Lumpur  
Tel: (60-3) 262 92 98  
Fax: (60-3) 262 61 98  
E-mail: ebocuk@mnl.net.my

**MEXICO**

**Mundi Prensa Mexico, SA de CV**  
Rio Pánuco No 141  
Colonia Guadalupe  
MX-06500 Mexico, DF  
Tel: (52-5) 533 56 58  
Fax: (52-5) 514 67 96  
E-mail: 1015452361@compuserve.com

**PHILIPPINES**

**EBIC Philippines**  
19th Floor, PS Bank Tower  
Sen. Gil J. Puyat Ave. cor. Tondo St.  
Makati City  
Metro Manila  
Tel: (63-2) 759 66 80  
Fax: (63-2) 759 66 80  
E-mail: ebicppcm@globe.com.ph  
URL: http://www.ebicpp.com

**SRI LANKA**

**EBIC Sri Lanka**  
Trans Asia Hotel  
11 S Sirchittampalam  
A. Gardiner Mawatha  
Colombo 2  
Tel: (94-1) 874 71 50 78  
Fax: (94-1) 44 87 79  
E-mail: ebicsl@imim.com

**THAILAND**

**EBIC Thailand**  
29 Vanich Building, 8th Floor  
Sao Chulalongkorn  
Ploenchit  
10330 Bangkok  
Tel: (66-2) 855 06 27  
Fax: (66-2) 855 06 28  
E-mail: ebicthk@kscsi5.th.com  
URL: http://www.ebicthk.org

**UNITED STATES OF AMERICA**

**Berman Associates**  
4511-F Aspenway Drive  
Lanham MD20706  
Tel: (1-800) 274 44 47 (toll free telephone)  
Fax: (1-800) 865 34 50 (toll free fax)  
E-mail: query@berman.com  
URL: http://www.berman.com

**ANDERE LÄNDER/OTHER COUNTRIES/ AUTRES PAYS**

**Bitte wenden Sie sich an ein Büro Ihrer Wahl/Please contact the sales office of your choice/Veuillez vous adresser au bureau de vente de votre choix**

**Office for Official Publications of the European Communities**

2, rue Mercier  
L-2985 Luxembourg  
Tel: (352) 29 29 42 65  
Fax: (352) 29 29 42 75  
E-mail: info.info@opoca.ccc.be  
URL: http://eur-op.eu.int

## La collection MEANS : “Évaluer les programmes socio-économiques”

Cette collection en six volumes met à la disposition d'un public élargi les enseignements du programme MEANS, lancé par la Commission européenne en vue d'améliorer et de promouvoir les méthodes d'évaluation. Cette publication répond au besoin de plus en plus manifeste de disposer de références solides pour traiter de façon efficace les problèmes d'évaluation rencontrés dans la mise en œuvre de l'action publique en général et en particulier dans le domaine des Fonds Structurels.

De ce point de vue, cet ouvrage constitue un guide méthodologique original qui fournit des solutions à des problèmes techniques et organisationnels rencontrés par l'évaluation auxquels la littérature existante apporte peu de solutions.

La collection s'adresse non seulement aux évaluateurs professionnels pour qui elle constituera un vade-mecum précieux, mais aussi aux non-spécialistes et en particulier aux autorités de gestion et aux membres des comités de suivi des programmes structurels.

Chaque volume est conçu de manière à être lu indépendamment des autres, mais tous sont complémentaires et traitent conjointement la totalité des aspects de l'évaluation des programmes socio-économiques :

- I. Conception et conduite d'une évaluation
- II. Choix et utilisation des indicateurs pour le suivi et l'évaluation
- III. Principales techniques et outils d'évaluation
- IV. Solutions techniques pour évaluer dans un cadre de partenariat
- V. Évaluation transversale des impacts sur l'environnement, l'emploi et les autres priorités d'intervention
- VI. Glossaire de 300 concepts et termes techniques



OFFICE DES PUBLICATIONS OFFICIELLES  
DES COMMUNAUTÉS EUROPÉENNES

L-2985 Luxembourg