



Fonds structurels communautaires

Évaluer les programmes socio-économiques

**Solutions techniques pour évaluer
dans un cadre de partenariat**

**Collection
MEANS**

Volume 4

Commission européenne

Évaluer les programmes socio-économiques

Solutions techniques pour évaluer dans un cadre de partenariat

Volume 4

*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 Structuels. Elle permettra aux responsables de l'évaluation de développer leurs

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

Témoin 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

Introduction	9
I Dessiner la carte conceptuelle des impacts	15
Une solution pour évaluer des mesures différentes à l'aide de critères communs	17
Mode d'emploi de la cartographie conceptuelle des impacts	29
Utiliser la cartographie conceptuelle des impacts	45
II Construire et utiliser des échelles de notation	51
Une solution pour évaluer des mesures différentes sous une forme homogène	53
Mode d'emploi des échelles de notation	59
Utiliser les notes pour synthétiser l'évaluation des mesures	71
III Analyser le programme à l'aide d'une matrice des impacts croisés	75
Une solution pour analyser la synergie au sein du programme	77
Mode d'emploi de la matrice des impacts croisés	85
Faire une vérification empirique des principales synergies identifiées	91
IV Appliquer l'analyse multicritère dans un cadre de partenariat	97
Une solution pour juger la réussite d'ensemble d'un programme	99
Mode d'emploi de l'analyse multicritère	105
Utiliser l'analyse pour produire des conclusions utiles	123
Conclusion	125
Annexe	133
Index	137



Les programmes de développement socio-économiques font souvent l'objet de co-financements et s'inscrivent donc dans un cadre partenarial. C'est notamment le cas pour les programmes des Fonds structurels. Il en résulte des orientations particulières pour leur évaluation, non seulement parce que le partage de la responsabilité accroît le nombre des protagonistes, mais aussi parce qu'il a tendance à multiplier le nombre des objectifs et à diversifier les domaines d'intervention.

Introduction

Jusqu'à une date récente, l'évaluation s'est principalement développée dans le cadre d'interventions financées par une seule institution dans un domaine précis. De ce fait, les évaluateurs manquaient de savoir-faire nécessaires pour traiter des programmes partenariaux, multi-objectifs et multi-secteurs. Les équipes d'évaluation ont fondé sur à concevoir les programmes complexes comme des séries d'activités perçues, traitant successivement chacune de ses composantes isolées, mesurées au séparé. Cette démarche conduit cependant à des difficultés au moment où il faut produire des conclusions synthétiques et faire des recommandations stratégiques à l'échelle de l'ensemble du programme. En appliquant les méthodes d'évaluation traditionnelles, on parvient souvent à réaliser des réallocations budgétaires car on ne fait pas comparer les différents composants du programme à l'aune de critères d'évaluation communs et acceptables pour tous les acteurs.

Au cours des dernières années, plusieurs équipes d'évaluation travaillent dans le domaine des Fonds structurels se sont distinguées en appliquant des innovations méthodologiques dans des cas d'évaluation « multi-acteurs et multi-secteurs ». Les évaluations pilotes réalisées dans le cadre du programme MIFAS ont également permis de mettre au point de nouvelles solutions opérationnelles aux échelles de projets. En s'appuyant sur l'expérience accumulée au cours des dernières années, ce Volume rassemble quatre solutions techniques nouvelles qui sont présentées à la disposition des professionnels, avec en outre la mise à disposition de fiches de ces techniques méthodologiques conceptuelles des modèles et outils multiformes, figurant également dans le Volume 3 de la collection, qui décrit les outils d'évaluation dans le secteur d'aide aux collectivités locales. C'est-à-dire de façon plus synthétique dans le présent Volume, mais en précisant que l'on trouve dans le présent Volume.

La première technique présentée est la cartographie conceptuelle des impacts. Cet outil peut être employé pour définir les principaux impacts



Les programmes de développement socio-économique font souvent l'objet de co-financements et s'inscrivent donc dans un cadre partenarial. C'est notamment la règle pour les programmes des Fonds structurels. Il en résulte des contraintes particulières pour leur évaluation, non seulement parce que le partage de la décision augmente le nombre des protagonistes, mais aussi parce qu'il a tendance à multiplier le nombre des objectifs et à diversifier les domaines d'intervention.

Jusqu'à une date récente, l'évaluation s'est principalement développée dans le cadre d'interventions financées par une seule institution dans un domaine précis. De ce fait, les évaluateurs manquent du savoir-faire nécessaire pour traiter le cas d'un programme partenarial, multi-objectifs et multi-secteurs. Les équipes d'évaluation ont tendance à considérer les programmes complexes comme des séries d'analyses partielles, traitant successivement chacune de ses composantes (axes, mesures ou actions). Cette démarche conduit cependant à des difficultés au moment où il faut produire des conclusions synthétiques et faire des recommandations stratégiques à l'échelle de l'ensemble du programme. En appliquant les méthodes d'évaluation traditionnelles, on parvient rarement à justifier des réallocations budgétaires car on ne sait pas comparer les différentes composantes du programme à l'aune de critères d'évaluation communs et acceptables pour tous les décideurs.

Au cours des dernières années, plusieurs équipes d'évaluation travaillant dans le domaine des Fonds structurels se sont distinguées en apportant des innovations méthodologiques dans des cas d'évaluation « multi-objectifs et multi-secteurs ». Les évaluations pilotes réalisées dans le cadre du programme MEANS ont également permis de mettre au point plusieurs solutions opérationnelles applicables à ce problème. En s'appuyant sur l'expérience constituée au cours des dernières années, ce Volume rassemble quatre solutions techniques nouvelles qui sont maintenant à la disposition des professionnels et il en donne le mode d'emploi détaillé. Deux de ces techniques (cartographie conceptuelle des impacts et analyse multicritère) figurent également dans le Volume 3 de la collection, qui décrit les outils d'évaluation dans un esprit d'aide aux choix méthodologiques, c'est-à-dire de façon plus synthétique sans fournir tous les conseils de mise en œuvre que l'on trouve dans le présent Volume.

La première technique présentée est la **cartographie conceptuelle des impacts**. Cet outil peut être employé pour définir les principaux impacts

intermédiaires du programme de façon commune aux différents protagonistes. La cartographie conceptuelle propose des critères d'évaluation applicables, de façon transversale, à l'ensemble d'un programme multi-objectifs et multi-secteurs. Elle fournit également une description précise de ces impacts, utilisable pour construire des échelles de notation et des indicateurs. A ce jour, il n'existe pas de manuel pour la mise en œuvre de cette technique dans le contexte de l'évaluation.

La seconde technique présentée est celle de la construction d'échelles de notation. La notation est de plus en plus utilisée, notamment dans les pays d'Europe du Nord, pour la sélection des projets. L'originalité de ce Volume réside dans le fait que la notation est appliquée à la synthèse des conclusions obtenues au niveau des mesures, et que la notation est directement reliée aux principaux critères d'évaluation du programme.

Au-delà de l'évaluation de chaque mesure, on doit aussi s'intéresser à l'interaction des mesures entre elles. On peut en effet s'attendre à ce que l'impact d'un ensemble cohérent de mesures soit plus grand que celui des mêmes mesures mises en œuvre séparément. Le troisième outil présenté dans ce Volume apporte une aide à l'évaluation de ces effets de synergie. Il s'agit de la **matrice des impacts croisés** qui est apparue comme un élément clé, commun aux démarches imaginées et testées par plusieurs équipes européennes dont les travaux ont été rassemblés dans le cadre du programme MEANS.

Enfin, chacun des co-financeurs souhaite retrouver la trace de ses priorités dans le rapport d'évaluation. L'**analyse multicritère**, quatrième des techniques présentées, permet de fonder des conclusions de synthèse en respectant les attentes de chaque décideur. L'analyse multicritère est typiquement employée pour l'évaluation ex ante en vue de classer ou de trier des projets concurrents ou des variantes d'un même projet. En appliquant l'outil à l'évaluation d'ensemble d'un programme de développement socio-économique, ce Volume propose une utilisation inhabituelle. Parmi les multiples variantes techniques de l'analyse multicritère, ce Volume présente celle qui est la mieux adaptée au contexte des programmes cofinancés.

Les quatre outils qui font la substance de ce Volume sont présentés avec leur mode d'emploi détaillé. Pour chaque outil, on s'est efforcé de présenter plusieurs variantes permettant une mise en œuvre très simple ou impliquant, au contraire, des procédures plus techniques et plus rigoureuses. Le lecteur

constatera que, dans leurs variantes les plus légères, les propositions contenues dans ce Volume ne sont rien d'autre que du « bon sens organisé » et qu'elles rejoignent nombre de pratiques actuelles.

Il était intéressant de réunir ces quatre techniques dans le même Volume car elles ont de nombreux points communs : forte participation des partenaires du programme, capacité à traiter des informations quantitatives et qualitatives, prise en compte de plusieurs points de vue, etc. Les quatre techniques peuvent d'ailleurs être réunies au sein d'une même méthode d'évaluation. La conclusion du Volume présente deux propositions pratiques permettant de réaliser cet assemblage.

Le public visé par ce Volume est d'abord constitué des équipes d'évaluation (c'est le seul Volume de la collection qui leur est prioritairement destiné). Le Volume s'adresse également aux responsables d'évaluation qui désirent approfondir leurs connaissances. Par extension, il est susceptible d'intéresser ceux qui travaillent dans des situations d'évaluation multi-décideurs hors du domaine des Fonds structurels.

Plusieurs modes de lecture de ce Volume sont possibles, comme le propose l'encadré ci-après.

Encadré 1 - Utilisations possibles de ce Volume

Si vous voulez préparer l'évaluation d'ensemble d'un programme socio-économique en faisant appel à tout ou partie des quatre techniques,

- Lisez la conclusion (voir page 125), puis reportez vous, autant que de besoin, aux différentes parties du Volume.

Si vous voulez étudier la cohérence d'un programme, réfléchir à ses objectifs implicites, trouver des critères d'évaluation transversaux pour toutes les mesures,

- Lisez la partie consacrée à la cartographie conceptuelle des impacts (voir page 15) ou simplement son mode d'emploi (voir page 29).

Si vous voulez évaluer toutes les mesures d'un programme avec une même logique homogène et transversale,

- Lisez la partie consacrée à la construction et à l'utilisation des échelles de notation (voir page 51) ou simplement le mode d'emploi correspondant (voir page 59).

Si vous voulez analyser la synergie au sein du programme et comprendre comment les différentes mesures renforcent mutuellement leurs impacts,

- Lisez la partie consacrée à la matrice des impacts croisés (voir page 75) ou simplement son mode d'emploi (voir page 85).

Si vous voulez conclure l'évaluation d'ensemble d'un programme, classer ou comparer les mesures qui le composent en fonction de leur plus ou moins grande réussite, prendre en compte tous les critères qui sont importants pour les partenaires,

- Lisez la partie consacrée à l'analyse multicritère (voir page 97) ou simplement son mode d'emploi (voir page 105).

Dessiner la carte conceptuelle des impacts

La cartographie conceptuelle des impacts, telle qu'elle est présentée dans ce chapitre, est une adaptation de la technique dite de « concept-mapping » utilisée depuis une dizaine d'années dans le domaine de l'évaluation. L'outil est employé pour définir les principaux impacts intermédiaires du programme en recherchant la plus grande transversalité possible entre les différentes mesures. La carte conceptuelle des impacts permet de clarifier les principaux objectifs du programme (explicites et implicites) et de fonder l'évaluation sur des critères clairs et acceptés par tous les partenaires. Elle fournit également une description précise des impacts attendus, utilisable de façon opérationnelle dans le cadre d'une enquête de terrain, par exemple pour construire des échelles de notation et des indicateurs.



Une solution pour évaluer des mesures différentes à l'aide de critères communs

■ Evaluation d'ensemble d'un programme

Les programmes financés par la politique européenne de cohésion visent tous à accélérer le développement socio-économique dans telle ou telle de ses dimensions (revenu, emploi, qualité de vie, etc.). Le principal impact est donc généralement attendu au niveau macro-économique ou macro-social. Pour obtenir cet impact, la plupart des programmes adoptent des stratégies multisectorielles, comprenant des mesures diversifiées autant par leur nature (équipement, conseil, formation, communication, etc.) que par leur domaine (transports, recherche, tourisme, etc.).

Dans le cadre de la politique européenne de cohésion économique et sociale, on a recours à une typologie d'intervention en trois grands domaines d'intervention, à savoir : l'amélioration des infrastructures de base (transport, communication, équipements publics en général), le soutien à l'environnement productif comprenant les aides aux acteurs économiques (pour accroître, par exemple, leur maîtrise de la qualité, leur capacité d'exportation, leur insertion dans des réseaux d'innovation, etc.), et enfin, le développement des ressources humaines (par des actions de formation essentiellement).

Ce large éventail de domaines d'intervention hétérogènes est fréquent en matière de programmes socio-économiques. Toutefois, cette diversité atteint, pour les programmes financés par les Fonds structurels, une ampleur rarement égalée, et qui pourrait conduire à la confusion si les programmes n'étaient pas fortement structurés. Le mode de structuration qui a été adopté comporte trois niveaux principaux auxquels s'ajoutent un ou deux niveaux intermédiaires.

- Le niveau d'ensemble du programme est associé à l'objectif global, par exemple le développement économique ou l'emploi. Ce niveau comprend généralement quelques axes prioritaires qui déclinent l'objectif global dans ses principales dimensions stratégiques ;
- Le niveau des mesures (une dizaine à plusieurs dizaines), correspond à l'unité de base de la gestion du programme, chaque mesure faisant l'objet d'un dispositif de gestion particulier. La mesure correspond en

- Le niveau des projets (souvent plusieurs centaines) est l'unité de mise en œuvre sur le terrain. Chaque projet est le point d'articulation entre le programme et ses destinataires.

internationaux plus faciles pour les entreprises régionales ou un meilleur accès des touristes aux sites touristiques). L'existence même du programme présuppose que tous ces mécanismes convergent vers l'objectif global de développement. Il n'en reste pas moins que cette convergence se fait par des cheminements nombreux et très ramifiés.

La réalisation d'une évaluation de programme pourra prendre des formes très différentes selon la façon dont on veut analyser ces impacts. Dans cette Collection, on a choisi de présenter les trois situations suivantes :

- L'évaluation d'ensemble qui se donne pour contrainte d'examiner tous les impacts de toutes les mesures.
- L'analyse plus approfondie des impacts macro-économiques.
- L'analyse plus approfondie d'un impact particulier, par exemple au niveau d'une mesure.

Ce Volume a été préparé dans la perspective de la première situation, c'est-à-dire l'évaluation d'ensemble de tous les impacts d'un programme multisectoriel. On se reportera aux Volumes 1 et 3 de la Collection pour une introduction aux techniques d'évaluation applicables aux deux autres situations.

■ Synthèse d'évaluations spécifiques

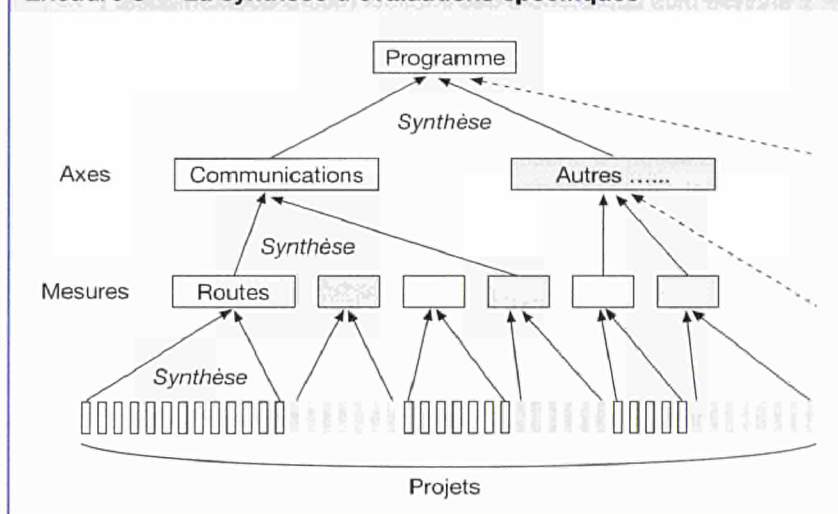
Chaque mesure s'évalue naturellement en fonction de critères qui lui sont propres. Dans l'Encadré 2 par exemple, la mesure d'amélioration des routes sera évaluée, sur la base de ses résultats et de ses impacts spécifiques, en fonction du gain de temps potentiel ou réel pour les véhicules utilisateurs. Comme cela se fait souvent, on peut supposer que l'équipe d'évaluation estimera le nombre total d'heures de transport économisées annuellement par les véhicules utilisateurs. Ce nombre d'heures peut être rapporté au coût de la mesure d'amélioration des routes de façon à permettre des comparaisons avec des programmes routiers dans d'autres régions similaires.

Les indicateurs qui viennent d'être cités sont des indicateurs spécifiques à un domaine d'intervention (les routes), de même que la démarche d'évaluation suivie est celle d'une évaluation spécifique au secteur routier. On peut conduire des évaluations spécifiques pour toutes sortes de mesures : voies ferrées, ports, aéroports, mais aussi pour des secteurs d'intervention aussi divers que la formation des chômeurs de longue durée,

la réhabilitation de quartiers dégradés ou la mise en place de réseaux d'innovation. On trouvera dans le Volume 2 de la Collection un grand nombre d'exemples d'interventions, accompagnés d'une description de leurs résultats et impacts spécifiques, et de propositions d'indicateurs adaptés.

Lorsqu'une équipe d'évaluation doit réaliser l'évaluation d'ensemble d'un programme, une pratique courante consiste à réaliser des évaluations spécifiques, mesure par mesure, et à synthétiser ensuite les résultats au niveau des axes, puis du programme, comme le schématise l'Encadré 3.

Encadré 3 – La synthèse d'évaluations spécifiques



Cette procédure d'évaluation « de bas en haut » ne présente pas de problème majeur quand il s'agit d'évaluer une mesure en faisant la synthèse des évaluations des projets qui la composent. En effet, tous les projets d'une même mesure se ressemblant, on peut procéder par agrégation, calcul de moyenne, comparaison des projets entre eux, identification des projets exemplaires, etc.

L'évaluation par synthèse peut, au contraire, présenter de graves faiblesses méthodologiques si elle porte sur des mesures hétérogènes évaluées de façon spécifique. A la limite, chaque évaluation spécifique utilise ses propres indicateurs, analyse des impacts distincts et applique des critères de jugement différents. Si la situation se présente ainsi, l'évaluation d'ensemble

d'un programme ne peut être que la juxtaposition d'évaluations partielles des mesures. La notion même de synthèse n'a plus de sens. On arrive à produire un rapport d'évaluation qui consacre un long chapitre à chaque mesure, mais dont la conclusion d'ensemble tient en quelques lignes au contenu très pauvre ou très fragile. Dans une telle situation, il n'est pas possible d'aboutir à des recommandations intéressant réellement l'ensemble du programme. Par exemple, il n'est pas possible de justifier une recommandation de réallocation budgétaire entre mesures car on n'a pas d'argument pour dire qu'une mesure est plus efficace ou plus réussie qu'une autre.

■ **Evaluation selon des critères transversaux**

Pour produire de véritables conclusions synthétiques, il faut disposer de critères d'évaluation qui permettent de comparer la valeur des mesures les unes par rapport aux autres. Un élément de comparaison fréquemment utilisé est celui du rythme de mise en œuvre, qui se traduit par exemple dans un indicateur de taux de consommation budgétaire ou de taux d'avancement des réalisations. Cet élément de comparaison joue un rôle important dans le cadre du suivi. Il constitue trop souvent le principal argument de réallocation budgétaire quand l'équipe d'évaluation ne parvient pas à fonder ses recommandations sur des critères communs d'efficacité, d'efficience ou de pertinence.

Exemple : Comparaison du taux d'atteinte des objectifs

Une mesure a pour objectif de requalifier 2 000 salariés menacés de perdre leur emploi, et une autre mesure vise à créer 500 nouveaux contacts internationaux pour des PME régionales. A la date de l'évaluation à mi-parcours, la première mesure a atteint 53% de son objectif et la seconde seulement 25%. Peut-on pour autant dire que la seconde mesure est un échec relatif ? En réalité, la première mesure est la continuation d'une intervention ancienne dont les objectifs n'ont pas véritablement changé, tandis que la seconde mesure est innovante et ambitieuse. Le rapprochement des deux taux ne peut donc pas donner d'argument pour comparer l'efficacité des deux mesures.

Pour renforcer les évaluations, on peut envisager de comparer les mesures en fonction du taux d'atteinte de leurs objectifs spécifiques (voir l'exemple ci-dessus). Le fait qu'une mesure atteigne rapidement ses objectifs peut être un argument en faveur de son efficacité ou de sa réussite mais cet argument est forcément fragile. Il suffit que les objectifs aient été fixés de

manière très prudente ou au contraire très ambitieuse, pour fausser les conclusions. L'évaluation du degré d'atteinte des objectifs renvoie à une question d'évaluation *ex ante* : peut-on dire que les objectifs de toutes les mesures ont été fixés avec le même degré d'ambition ? Cette question renvoie elle-même à une autre question fondamentale traitée dans cet ouvrage : au nom de quel critère commun peut-on comparer le degré d'ambition des différentes mesures d'un même programme ?

Dans le cadre de l'évaluation d'ensemble d'un programme, ce Volume cherche à promouvoir un véritable travail de synthèse fondé sur l'examen de la contribution des différentes mesures au développement socio-économique, et non pas sur leur rythme d'exécution.

■ Contribution à l'atteinte de l'objectif global du programme

Exemple : Objectif global et objectifs intermédiaires

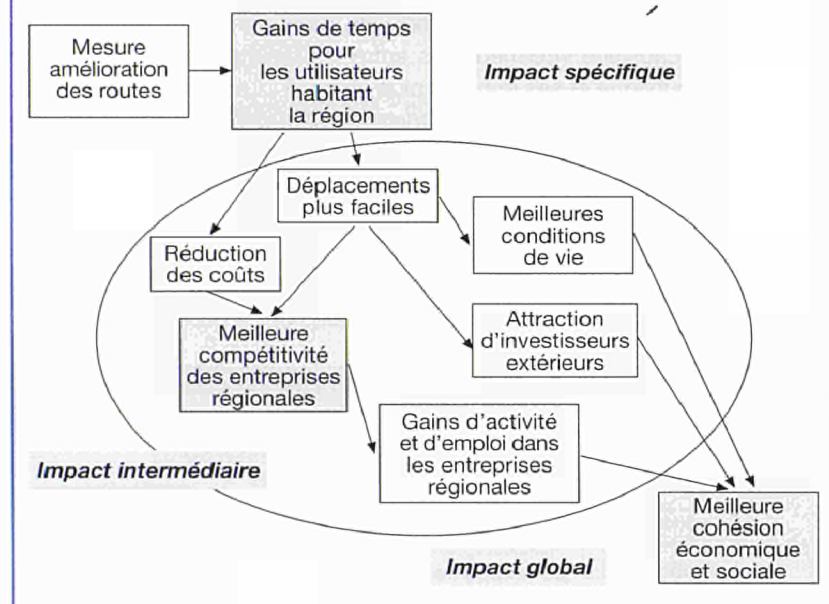
La province belge du Hainaut reçoit des aides de l'Union Européenne dans le cadre du programme de développement dont la priorité est l'adaptation économique de la province. L'objectif global fixé au début du programme en 1994 est le redressement du PIB/habitant en vue d'atteindre un niveau de 79% de la moyenne communautaire en 1999.

Le programme contient quatre axes d'intervention, chacun d'entre eux étant subdivisé en mesures. L'axe prioritaire 1 a pour objectif de relancer l'activité économique. Il s'agit tout à la fois de contribuer à la modernisation du tissu économique et au développement endogène de la région, de renforcer ses capacités en matière de recherche et développement technologique, d'offrir aux entreprises un environnement favorable à leur installation et de transformer en emplois la croissance économique résultant de ces efforts. L'objectif de l'axe 2 est de renforcer l'attractivité de la région. La province du Hainaut porte encore les traces d'un intense passé industriel. Rendre la zone attractive, assurer un cadre de vie de qualité et préserver l'environnement constituent dès lors un défi tout particulier. L'axe prioritaire 3 a pour but de compléter les infrastructures du transport dans une perspective d'intégration par rapport aux autres mesures du programme. L'objectif de l'axe 4 est de donner sa chance à chacun. Les problèmes du marché du travail dans le Hainaut justifient la mise en œuvre d'une politique d'emploi qui cherche à promouvoir le développement d'une économie à la fois compétitive et solidaire.

Chaque programme se doit de contribuer à l'objectif global de cohésion économique et sociale qui est poursuivi par l'Union Européenne. L'objectif global d'un programme s'exprime lui-même en termes de développement

socio-économique, avec une pondération plus ou moins importante accordée à tel ou tel aspect de ce développement : diversification sectorielle, compétitivité, intégration sociale, durabilité de la croissance, respect de l'environnement, etc. Comme le montre l'exemple précédent, l'objectif global décrit l'impact ultime que le programme cherche à obtenir. Dans le cadre d'un programme conçu de façon cohérente, chacune des mesures contribue à la production de l'impact global du programme.

Encadré 4 – Complexité des impacts et niveaux d'objectifs



On pourrait vouloir juger les mesures du programme en fonction de leur contribution à l'atteinte de l'objectif global, mais des obstacles pratiques rendent cette façon de faire impossible. Comme le montre l'encadré 4, l'impact d'une mesure donnée sur la cohésion économique et sociale passe par des mécanismes trop nombreux et trop complexes pour faire l'objet d'une évaluation robuste. Certains de ces mécanismes font d'ailleurs appel à des raisonnements macroéconomiques. Une équipe d'évaluation peut certainement apporter des connaissances utiles sur ces mécanismes, mais elle pourra rarement les analyser de façon différenciée, mesure par mesure, avec suffisamment de robustesse pour justifier des recommandations

fortes, de type réallocations budgétaires. En pratique, l'observation empirique des impacts doit s'arrêter à un niveau intermédiaire, dépassant souvent les destinataires directs, mais n'atteignant jamais l'ensemble de l'économie régionale. Au-delà de ce niveau intermédiaire, on travaille en utilisant des coefficients d'extrapolation ou des modèles économétriques, techniques qui fournissent des résultats utiles mais qui ne permettent pas de départager les mesures en fonction de leur efficacité relative.

■ Repérage des principaux impacts intermédiaires

Plusieurs techniques ont été développées pour permettre l'identification des liens de causalité intermédiaires entre les résultats spécifiques d'une intervention publique et ses impacts plus globaux. La plus simple et la plus connue est celle du cadre logique. Originellement conçue pour l'évaluation ex ante des projets dans les pays en développement, cet outil donne une présentation structurée de l'intervention évaluée et de ses effets attendus. En appliquant le cadre logique à l'exemple cité dans l'encadré 2, on obtient un tableau qui énonce les attentes en matière de réalisations (en bas à gauche de l'encadré 5) ainsi que pour les principales étapes logiques séparant les réalisations de l'impact global (en haut à gauche).

Encadré 5 – Classement des objectifs à l'aide du cadre logique

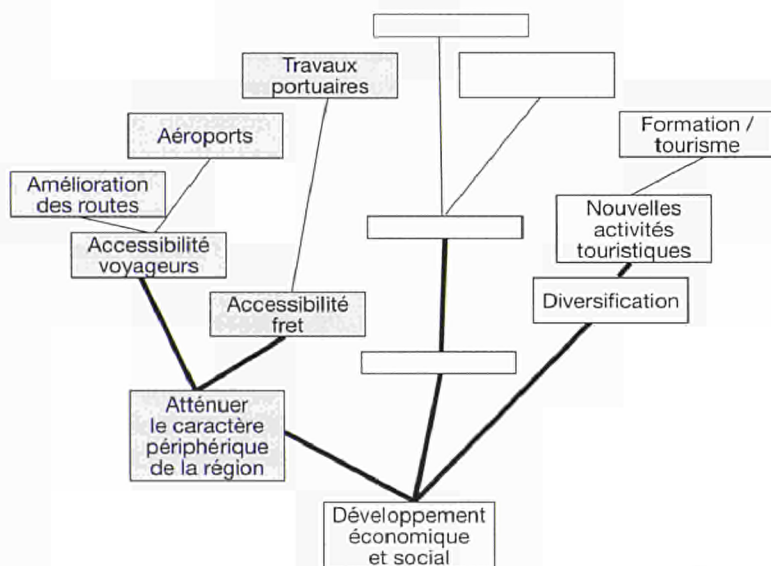
Projet de route entre A et B – voir Encadré 2.

	Enoncé des attentes	Indices de succès vérifiables	Conditions et risques
Impact global	Développement économique et social		
Impact spécifique	Economie pour les entreprises régionales		
Résultat	Gain de temps sur parcours-type		
Réalisation	Amélioration des routes		

A chaque étape logique, l'équipe d'évaluation mentionne les moyens prévus ou à prévoir pour vérifier l'atteinte des objectifs (indices de succès vérifiables). Elle définit en même temps ce qui doit être considéré comme un succès ou un échec (réalisations, résultats et impacts suffisants/insuffisants). En obligeant à remplir cette seconde colonne de la matrice avec des indicateurs ou des descriptions qualitatives précises, la technique apporte souvent une clarification utile qui permet de revenir à une description plus précise de la première colonne. La troisième colonne intitulée « conditions et risques » établit une liste des facteurs exogènes qui peuvent faciliter ou contrarier le passage d'une étape à l'étape suivante. On trouvera plus de détails sur la technique du « cadre logique » dans le Volume 3 de la Collection.

La structure du cadre logique, comportant seulement deux niveaux d'impact, ne permet pas de bien représenter les impacts intermédiaires. Un second outil, mieux adapté à la complexité des programmes socio-économiques, est celui de l'arbre des objectifs. Comme le montre

Encadré 6 – Arbre des objectifs



Dans la pratique, l'arbre des objectifs est dessiné sous une forme couchée avec les mesures à gauche et l'objectif global à droite pour faciliter sa présentation

L'Encadré 6, cette technique permet de déterminer la contribution logique de chacun des objectifs de niveau plus spécifique à un objectif de niveau plus global, et ainsi de suite jusqu'à l'objectif global du programme constitué par le tronc de l'arbre.

Par rapport au cadre logique, l'arbre des objectifs apporte deux éléments intéressants : d'une part, il permet de ramifier les branches en autant de rameaux qu'il y a de mesures dans le programme, et d'autre part, il permet de multiplier les niveaux intermédiaires pour mieux rendre compte de la complexité des mécanismes de production des impacts.

L'arbre des objectifs fait également apparaître des objectifs intermédiaires communs à plusieurs mesures. Dans l'exemple de l'Encadré 6, on voit que l'objectif « atténuer le caractère périphérique de la région » peut servir de référence commune pour juger l'efficacité des trois mesures situées en haut à gauche de l'arbre. En fait, les critères communs à plusieurs mesures apparaissent à tous les embranchements de l'arbre. Cependant, chaque embranchement crée une possibilité de comparaison limitée aux branches qui en dépendent. Pour comparer des mesures situées à droite et à gauche de l'arbre, il faudrait redescendre jusqu'au tronc, or on a vu que c'est généralement impossible. Du fait de sa construction en arborescence, la technique a donc un faible potentiel pour l'évaluation synthétique et transversale des multiples mesures d'un programme.

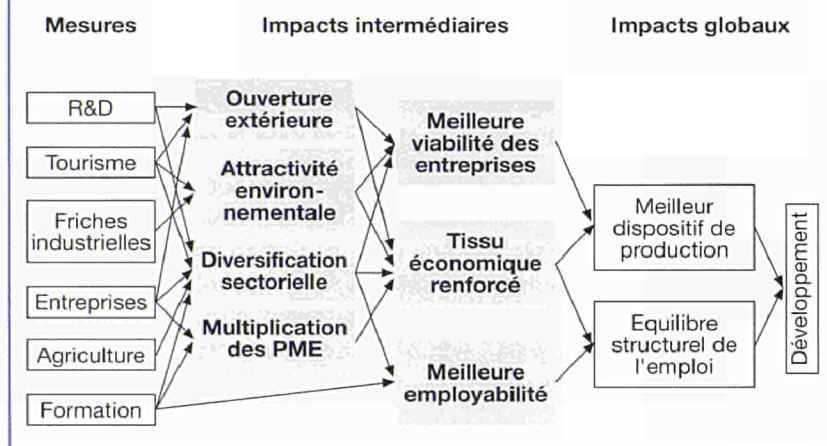
■ Recherche d'impacts intermédiaires transversaux

Pour progresser vers une évaluation véritablement synthétique, on doit abandonner l'image de l'arbre pour passer à l'image d'un réseau d'objectifs. L'outil adéquat est le diagramme logique des impacts qui permet de considérer une multiplicité de relations logiques entre les mesures et les impacts intermédiaires. Ainsi, dans l'exemple de l'Encadré 7 ci-après, l'impact intermédiaire « viabilité des entreprises » est indirectement relié à toutes les mesures du programme, ce qui permet d'envisager des comparaisons entre toutes les mesures sur la base de leur contribution à la viabilité des entreprises.

Le diagramme logique fait apparaître une série d'impacts intermédiaires qui vont permettre de fonder l'évaluation d'ensemble du programme. Parmi les impacts intermédiaires, il conviendra de rechercher ceux qui répondent aux deux exigences suivantes :

- Ne pas être trop globaux pour permettre une analyse empirique à partir d'observations de terrain ;
- Être connectés au plus grand nombre possible de mesures pour permettre des comparaisons et des agrégations au sein du programme.

Encadré 7 - Diagramme logique des impacts attendus



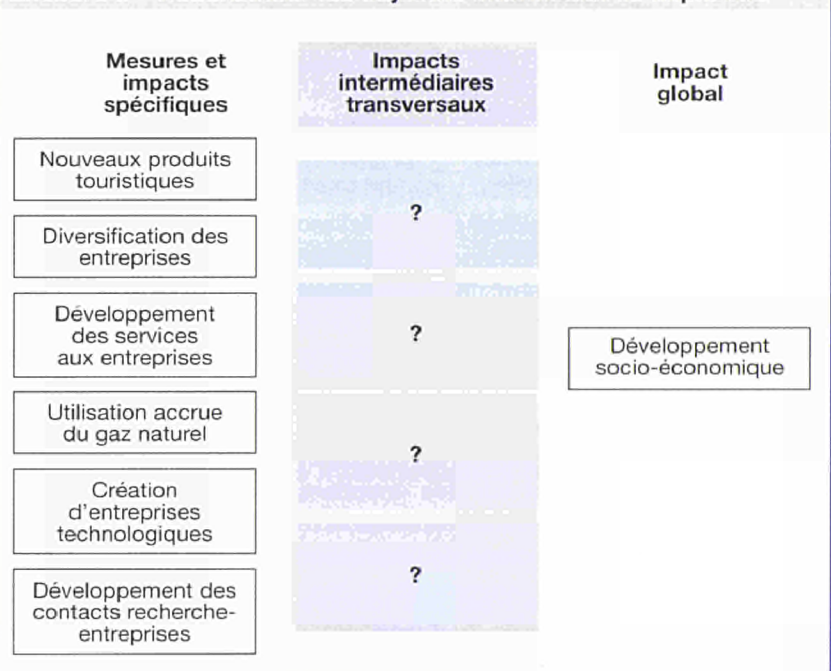
Le diagramme logique des impacts peut être construit de plusieurs façons. Une première démarche consiste à construire l'arbre des objectifs en s'appuyant sur la documentation du programme, puis à transformer l'arbre en réseau en s'appuyant sur des entretiens avec les gestionnaires du programme et avec quelques opérateurs. Le diagramme est ensuite présenté au groupe de pilotage de l'évaluation pour être discuté et validé.

Cette première démarche n'est pas toujours suffisante, en particulier si l'équipe d'évaluation constate des ambiguïtés ou des incohérences dans les objectifs intermédiaires. Dans le passé, il est arrivé que l'équipe d'évaluation invoque de tels problèmes pour déclarer le programme inévaluable dans son ensemble et pour arrêter son travail au stade des évaluations spécifiques des mesures. Dans cette situation où les objectifs intermédiaires ne sont pas définis assez clairement, l'équipe d'évaluation ne devrait pas renoncer, mais plutôt s'engager dans la voie d'une reconstitution des objectifs implicites.

Dans cette perspective, la démarche consiste à construire le diagramme logique des impacts, non pas à partir de l'arbre des objectifs, mais à partir

des impacts attendus par les gestionnaires et les opérateurs du programme. Pour ce faire, l'équipe d'évaluation commence par schématiser les mesures, leurs résultats et leurs impacts spécifiques (voir les cadres blancs à gauche de l'Encadré 8) ainsi que l'impact global (cadre de droite). La démarche consiste ensuite à identifier et décrire de façon précise une série d'impacts intermédiaires attendus (cadres intermédiaires colorés). Ceux-ci doivent permettre de connecter les impacts spécifiques et globaux par l'intermédiaire d'un réseau de relations logiques le plus dense possible.

Encadré 8 – Reconstituer les objectifs intermédiaires implicites



Cette démarche a fait l'objet d'une évaluation pilote dans le cadre du programme MEANS en 1996. Le travail faisait partie de l'évaluation à mi-parcours du programme de développement socio-économique de la région de Valencia, (Espagne) cofinancé par les autorités régionales, le gouvernement espagnol et la Commission européenne. L'équipe d'évaluation a employé la technique de la cartographie conceptuelle des impacts dont le mode d'emploi est donné ci-après.

Mode d'emploi de la cartographie conceptuelle des impacts

La cartographie conceptuelle d'impacts est une adaptation de la technique dite de « *concept-mapping* » utilisée depuis une dizaine d'années dans le domaine de l'évaluation. La technique a d'abord été utilisée comme aide à la définition de stratégies ou à la réorganisation des entreprises avant d'être appliquée à la conception de politiques publiques, notamment de programmes sociaux. L'application développée dans le cadre du programme MEANS a été spécialement conçue pour préciser les objectifs intermédiaires explicites et implicites d'un programme de développement socio-économique et pour faciliter le choix des indicateurs associés.

L'outil est fondé sur une réflexion collective et privilégie la diversité et la richesse des points de vue. Les idées de départ sont énoncées et regroupées par les participants eux-mêmes, de façon intuitive. Un logiciel de traitement de données qualitatives synthétise les représentations individuelles sous la forme d'une image graphique collective (la carte conceptuelle). Cette technique évite que l'équipe d'évaluation impose sa propre structure logique qui risquerait de rester étrangère aux participants.

■ Constituer un ou plusieurs groupes de travail

Exemple : Réflexion collective avec un grand nombre de participants

La technique de la carte conceptuelle a été utilisée en Irlande du Nord pour évaluer un programme de réconciliation entre les communautés catholique et protestante. La question évaluative était la suivante : « quels types d'actions de réconciliation contribuent le mieux à améliorer la compréhension mutuelle entre les communautés ? ». Grâce à une réflexion collective impliquant 200 personnes, un grand nombre de propositions ont été énoncées (exemple : « soutien de la langue irlandaise comme faisant partie intégrale de notre héritage culturel »).

Au cours d'une réunion plus restreinte, les propositions ont été regroupées par thème. Chaque participant a classé les propositions en piles cohérentes et les a notées de 1 à 5 selon leur pertinence. Les regroupements effectués par les partenaires ont été synthétisés sous forme de carte conceptuelle. Le résultat a permis de regrouper les propositions en six classes qui ont reçu une dénomination (exemple : « traditions culturelles ») et qui ont servi à articuler la suite de la réflexion.

La technique peut être mise en œuvre au sein du groupe de pilotage de l'évaluation, à condition que celui-ci comprenne au moins une dizaine de

personnes. Elle peut aussi être appliquée dans le cadre d'une série de groupes de travail. Dans ce cas, elle peut impliquer jusqu'à deux cents personnes comme dans l'exemple ci-dessus.

Pour constituer des groupes de travail, on peut regrouper les participants par catégorie, par exemple : un groupe de gestionnaires du programme, un groupe de représentants des autorités locales, un groupe de chefs d'entreprise, etc.

Exemple : Groupes de travail par catégorie de participants

La technique de la cartographie conceptuelle a été appliquée aux Etats-Unis dans le cadre d'un programme de soins infirmiers. Les objectifs de l'étude étaient, d'une part de donner des informations aux gestionnaires sur les pratiques de soins apportés, et d'autre part, de faire état des théories et des résultats de recherche en matière de soins. Trois groupes distincts ont été formés : un groupe d'infirmiers, un groupe de patients et un groupe de théoriciens des soins infirmiers. Pour chacun des groupes, on a appliqué parallèlement la technique de la cartographie conceptuelle.

Si l'on choisit de constituer plusieurs groupes de travail, une solution est de regrouper les différents acteurs suivant les domaines d'intervention du programme ; par exemple : un groupe pour les interventions dans le domaine des infrastructures, un groupe pour les interventions dans le domaine de l'environnement productif et un troisième groupe pour les interventions dans le domaine des ressources humaines.

Il est souhaitable que chaque groupe compte entre 10 et 15 personnes. Les groupes doivent être composés de façon à ce que leurs participants aient le même niveau de connaissance du programme, de sa logique et de ses impacts. Enfin les participants doivent s'engager à être présents de façon continue lors des réunions du groupe. Il faut prévoir deux à trois réunions échelonnées sur une à deux semaines. Lorsque les circonstances le permettent, une partie du travail peut être organisée sous forme de réunion virtuelle par échange de courrier électronique.

**■ Repérer les impacts attendus
dans les documents du programme**

Pour commencer son travail, l'équipe d'évaluation consulte tous les documents disponibles et susceptibles de faire apparaître les impacts

potentiels du programme. En premier lieu, elle s'appuie sur une analyse du document de programmation. Les objectifs sont inventoriés et on identifie ceux qui sont associés à une description des impacts attendus. Les intitulés des mesures, des sous-programmes ou des axes prioritaires peuvent aussi servir au repérage des impacts attendus. Les constats, les faits ou les motifs évoqués, comme la raison d'être du programme, peuvent également inspirer l'équipe d'évaluation pour définir ce que le programme est censé produire.

Au-delà des documents officiels de programmation, l'équipe d'évaluation peut étendre son analyse aux travaux de préparation du programme et notamment au plan établi par l'Etat membre, aux analyses stratégiques (SWOT), ainsi qu'aux justifications apportées pour l'éligibilité du territoire ou du public cible compte tenu des dispositions réglementaires régissant les Fonds structurels. Le Cadre Communautaire d'Appui, lorsqu'il en existe un, constitue également une source d'information utile. Les rapports d'évaluation précédents sont également utilisables ainsi que les informations issues des documents de gestion et de suivi, par exemple les fiches techniques décrivant les critères de sélection des projets. On étudie également les bilans et les tableaux de bord, ainsi que les comptes-rendus des réunions de suivi, les communiqués de presse et les déclarations officielles à propos du programme et des bénéfices qu'on en attend.

L'analyse des documents aboutit à une liste d'extraits décrivant ou évoquant les effets que le programme doit produire ou est supposé produire. Si plusieurs groupes ont été constitués pour couvrir différents domaines d'intervention, l'équipe d'évaluation établit une liste par groupe en rassemblant tous les extraits susceptibles de concerner les différents groupes, un même extrait pouvant concerner plusieurs groupes.

■ Faire exprimer les impacts attendus par les acteurs du programme

Dans une seconde étape, l'équipe d'évaluation réalise une enquête qualitative auprès de dix à vingt participants des groupes de travail. Le but de l'enquête est de révéler les objectifs implicites des gestionnaires et des opérateurs qui mettent le programme en œuvre. Les questions de base sont du type :

- Parmi les interventions du programme, quelles sont celles que vous connaissez le mieux ?

- Pour les interventions que vous connaissez le mieux, pouvez-vous dire quels sont les publics touchés directement ou indirectement ?
- De quelle nature sont les effets du programme sur ces publics ?
- Par quels mécanismes ces effets se produisent-ils ?
- Les effets produits doivent-ils être considérés comme positifs, désirables ?

Pour obtenir le point de vue des acteurs sur les effets du programme, l'équipe d'évaluation procédera par entretiens individuels téléphoniques ou en face à face. L'entretien de groupe (*focus group*) est également envisageable. Il prendra alors la forme d'une première réunion du groupe de travail. Par contre, l'entretien ne doit pas être directif, ni faire appel à un questionnaire. Il s'agit en effet de faire émerger le maximum d'idées possibles, en respectant complètement les mots et les phrases des personnes interrogées plutôt que de leur demander de se ranger dans des cadres préétablis.

Les impacts sont ceux que les acteurs du programme s'attendent à constater à court ou moyen terme. Ils peuvent être escomptés ou non dans les documents du programme. Ils sont généralement favorables, mais l'équipe d'évaluation n'hésitera pas à retenir aussi ceux qui sont défavorables. Par exemple, les deux impacts suivants pourront être retenus : « les entreprises aidées augmentent leur productivité » et « des emplois sont perdus dans les entreprises aidées du fait des gains de productivité ». L'expérience montre que les participants n'évoquent pas facilement les impacts défavorables. L'équipe d'évaluation ne doit donc pas hésiter à insister sur l'intérêt d'examiner également ce type d'impacts. On doit aussi veiller à ce que les personnes interrogées ne mentionnent pas uniquement des impacts « faciles » à atteindre.

L'analyse des entretiens aboutit à une liste de phrases, extraites du discours des personnes interrogées, et décrivant les effets du programme, tels qu'ils sont attendus par ses acteurs. Si plusieurs groupes ont été constitués, on constitue pour chaque groupe la liste des phrases susceptibles de l'intéresser.

■ Etablir une liste homogène d'énoncés d'impacts

Tous les extraits de textes et de discours qui ont été réunis dans les deux étapes précédentes doivent être réécrits sous une forme homogène par

l'équipe d'évaluation. Il s'agit de reformuler l'ensemble des éléments recueillis sous forme de courtes phrases appelées « énoncés d'impact ». Un énoncé d'impact prend toujours une des formes suivantes : tel public change son comportement de telle façon, telle variable économique est modifiée de telle façon, tel élément de l'environnement évolue de telle façon. Par exemple, si un objectif du programme est une « diminution de l'exode rural des jeunes », il donnera lieu à l'écriture de l'énoncé suivant : « plus de jeunes s'installent dans les bourgs ruraux ». L'énoncé sous-entend que le changement ou l'évolution sont dus au programme, mais il n'est pas indispensable de faire référence à cette relation causale à chaque fois. Par ailleurs, l'énoncé décrit la nature de l'impact attendu et non pas son niveau. Par exemple, si un impact est énoncé sous la forme « plus de jeunes s'installent dans les bourgs ruraux », différentes mesures pourront avoir différents niveaux de réussite de ce point de vue : « un nombre significatif de jeunes se sont installés durablement » ou « quelques jeunes se sont installés de façon exceptionnelle » ou encore « aucun jeune ne s'est installé ». Ces niveaux de réussite seront évoqués plus loin sous l'appellation de « descripteurs d'impact » (voir page 66).

Un énoncé s'applique toujours à un seul impact. Par exemple, si un objectif du programme est « accroître le revenu des exploitants agricoles par la diversification des activités », il donnera lieu à l'écriture de deux énoncés d'impact : (1) « les activités des exploitants agricoles se diversifient » et (2) « le revenu des exploitants agricoles s'accroît ».

L'équipe d'évaluation doit garder à l'esprit l'objectif de l'exercice qui est d'inventorier les impacts intermédiaires. A ce stade, il convient donc d'éliminer les énoncés d'impacts trop spécifiques qui ne peuvent manifestement concerner qu'une mesure et une seule. De même, on pourra écarter les impacts trop globaux lorsqu'on estime qu'ils ne pourront pas être reliés, de façon empirique, aux différentes mesures du programme.

A l'issue de cette étape, l'équipe d'évaluation fait un premier tri pour éliminer les énoncés qui font manifestement double emploi et qui sont écrits dans des termes quasiment identiques. Elle obtient une liste d'énoncés (voir exemple suivant) qui ne doit pas dépasser la centaine. Si la liste est trop longue, on écarte les impacts qui concernent trop spécifiquement telle ou telle mesure. Si plusieurs groupes de travail ont été constitués, on établit une liste pour chaque groupe.

Exemple : Impacts énoncés par les gestionnaires d'une mesure d'aide directe aux PME

Les PME disposent de meilleurs systèmes de gestion

La qualification des employés des entreprises artisanales s'accroît

L'offre touristique se diversifie

Un plus grand nombre d'entreprises respectent les normes environnementales

Les entreprises remplacent des sources d'énergie dérivées du pétrole par le gaz naturel

Les Centres de Recherche développent des projets de R&D adaptés aux besoins de l'industrie régionale

Les entreprises améliorent leur processus de production

Les entreprises augmentent leurs ventes à l'exportation

La situation financière des PME s'améliore

Des emplois sont créés dans les zones prioritaires

Les entreprises réduisent leur production de déchets

■ Valider les énoncés

L'étape qui suit est réalisée dans le cadre d'une réunion du groupe de pilotage ou de plusieurs réunions s'il y a plusieurs groupes de travail. Dans ce second cas, chaque groupe examine la liste des énoncés d'impact qui a été préparée à son intention.

L'animateur de la séance commence par :

- exposer les principes de la technique
- préciser la logique qui a régi le regroupement des participants
- expliquer ce que l'on entend par énoncé d'impact et comment les énoncés ont été établis
- présenter la liste des énoncés

La liste fait ensuite l'objet d'une discussion approfondie qui porte notamment sur les points suivants :

- Y a-t-il contradiction ou non entre les énoncés qui proviennent des documents de programmation, de gestion et de suivi et ceux qui proviennent des entretiens ?
- Certains énoncés peuvent-ils être réécrits de façon à s'appliquer à un plus grand nombre de mesures ?

- Les énoncés sont-ils bien compris de la même façon par tous les participants ? Sinon, comment les préciser ?
- Peut-on fusionner des énoncés dont la rédaction est proche et qui n'ajoutent rien en terme d'impact ?
- Doit-on ajouter des énoncés importants ?

■ Ponderer les énoncés

Après cette première réunion, l'équipe d'évaluation finalise la liste des énoncés. Les énoncés sont répartis de façon aléatoire en veillant à ne pas constituer des successions d'énoncés proches. L'équipe d'évaluation imprime pour chaque participant une liste des énoncés numérotés.

Une deuxième réunion du groupe de travail a lieu peu de temps après la première. S'il y a plusieurs groupes de travail, chacun d'entre eux tient deux réunions successives et rapprochées. Le premier objectif de la réunion consiste à hiérarchiser les impacts attendus en fonction de leur importance pour la réussite d'ensemble du programme. La réunion se déroule sur la base des consignes suivantes :

- Distribuer à chaque participant la liste des énoncés.
- Demander aux participants d'attribuer à chaque énoncé une note en fonction de l'importance de l'impact correspondant pour la réussite du programme (1=impact peu important, ..., 4=impact essentiel). Cette notation vaut aussi pour les impacts défavorables (4=impact très pénalisant, ..., 1=impact peu pénalisant). Les notes sont inscrites sur la liste des énoncés d'impacts. La liste des énoncés d'impacts est collectée.

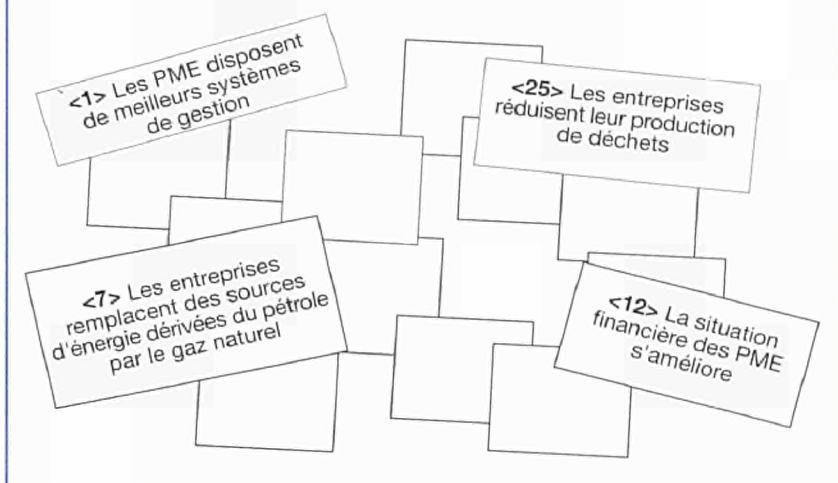
■ Regrouper les impacts conceptuellement proches

Le travail collectif se poursuit par un regroupement des énoncés d'impact en quelques grandes classes. C'est ici qu'intervient la technique de la cartographie conceptuelle proprement dite. L'équipe d'évaluation imprime pour chaque participant un jeu de cartons ou bistrots en faisant figurer sur la même face de chaque carton un énoncé et son numéro.

La réunion se poursuit sur la base des consignes suivantes :

- Distribuer à chaque participant la série complète des énoncés d'impact.

Encadré 9 – Distribution des énoncés d'impact



- Demander aux participants de regrouper les cartons en plusieurs piles en fonction de leur proximité conceptuelle. Chacun des membres du groupe de travail va ainsi procéder à des regroupements d'impacts qui relèvent pour lui d'une même « classe » d'impacts. La consigne est la suivante : « Mettez dans une même pile les impacts qui vous semblent proches ». La constitution des piles est très libre. La seule contrainte est que tous les cartons ne doivent pas se trouver dans une seule pile et qu'il ne doit pas y avoir autant de piles que de cartons.
- On demande aux participants d'expliciter, au dos du premier carton de chaque pile, la logique de constitution de la pile.
- On récolte les piles de chacun des participants, on agrafe chacune des piles et on les place dans une enveloppe au nom du participant.

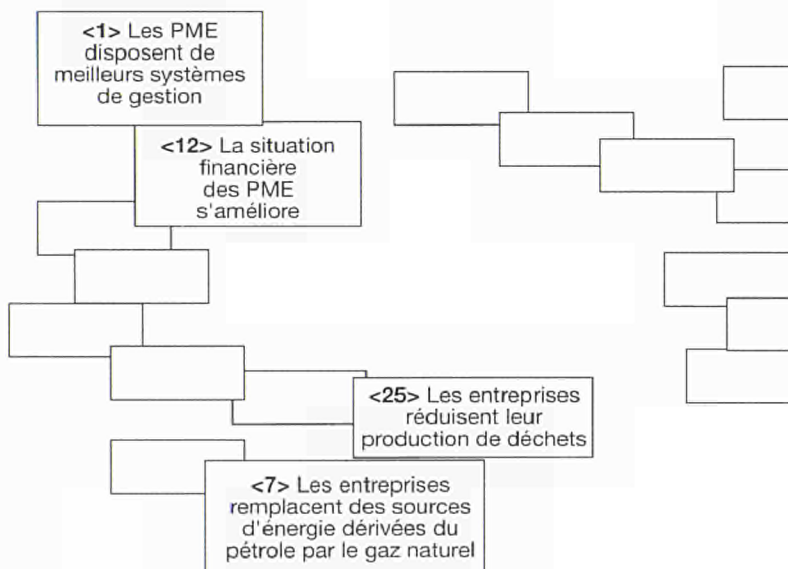
Les piles produites par chacun des membres du groupe font l'objet d'une analyse statistique. Il s'agit de constituer des classes d'impact rendant compte le plus fidèlement possible des regroupements effectués par les participants. La technique consiste à compter le nombre de fois que deux impacts ont été mutuellement associés et à considérer ce nombre comme un indice de proximité conceptuelle entre ces deux impacts, la distance conceptuelle étant au contraire égale au nombre de fois où les deux impacts n'ont pas été associés. On construit ensuite une matrice carrée comportant tous les énoncés en ligne et en colonne, chaque cellule faisant apparaître la

distance conceptuelle de deux impacts. Cette matrice est traitée par une analyse statistique multidimensionnelle qui produit deux résultats :

- Une carte conceptuelle sur laquelle les impacts sont positionnés et qui reflète le mieux possible leurs proximités et leurs distances conceptuelles (voir encadré ci-dessous).
- Une classification hiérarchique montrant les énoncés qui sont conceptuellement les plus proches (voir encadré 11).

L'ensemble de la démarche peut être conduit à l'aide d'un logiciel d'analyse statistique et d'un micro-ordinateur, sachant qu'il existe plusieurs logiciels standards sur le marché. Pour faciliter la démarche, on peut également utiliser le logiciel qui a été développé spécialement pour le « *concept mapping* » (voir références bibliographiques W.M.K. Trochim, 1989).

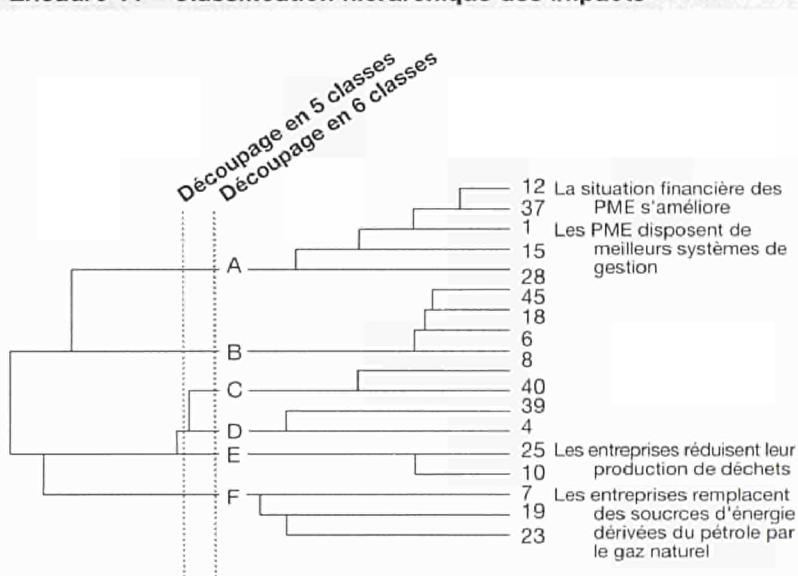
Encadré 10 – Carte conceptuelle des impacts



A partir de cette carte, il est possible de poursuivre l'analyse statistique en formant un certain nombre de classes d'impacts. On utilise pour cela la classification hiérarchique en coupant l'arborescence en un point donné.

Dans l'exemple ci-dessous, deux découpages sont proposés, le premier donnant naissance à cinq classes d'impacts, le second à six (classes A, B, C, D, E, F). Dans cet exemple, il vaut mieux retenir le découpage en cinq classes car l'encadré montre que si l'on regroupe les classes C et D, alors on devrait également étendre ce regroupement à la classe E. Une fois que les classes ont été découpées, il est possible de calculer un indice de liaison pour chaque énoncé. L'indice de liaison varie entre 0 et 1, une valeur proche de 1 signifie que l'énoncé a joué un rôle important dans la construction de la classe d'impacts à laquelle il appartient, une valeur proche de 0 signifie au contraire que l'énoncé n'est pas relié étroitement à sa classe.

Encadré 11 – Classification hiérarchique des impacts



Dans le cadre de l'évaluation pilote réalisée à Valencia, le groupe « infrastructures » a retenu trois classes d'impacts, le groupe « aides directes aux entreprises » en a retenu huit et le groupe « soutien à l'environnement productif » sept.

Si les participants sont peu nombreux et si le nombre d'énoncés d'impact est inférieur à une trentaine, on peut procéder sans faire appel à un logiciel d'analyse statistique, par exemple en utilisant une technique de discussion en groupe de type Metaplan®. Cette technique est décrite dans le

Volume 3 de la Collection. Elle aboutit à un regroupement visuel des énoncés d'impact, réalisé dans le cadre d'une discussion collective, ainsi qu'à une pondération de ces énoncés. Cette approche simplifiée a été utilisée dans le cadre de l'évaluation à mi-parcours d'un programme de développement socio-économique comme le montre l'exemple suivant.

Exemple : Utilisation de la technique METAPLAN® pour obtenir les classes d'impacts

Dans le cadre de l'évaluation à mi-parcours d'un programme de reconversion industrielle en France, la première réunion du groupe de pilotage avait comme principal objectif de produire une liste des impacts attendus.

La préparation du plan de discussion et l'animation de la séance ont été pris en charge par des consultants extérieurs à l'aide de l'outil METAPLAN®. Les participants (13 personnes) étaient des gestionnaires de mesures et des opérateurs.

La séance a duré environ deux heures. Elle a fait ressortir les principaux impacts attendus, puis les a regroupés en cinq classes dénommées comme suit :

- A - Requalification de la main d'œuvre
- B - Adaptation des formations
- C - Attrait du territoire
- D - Reconversion/Diversification
- E - Création d'emplois d'initiative locale

■ Valider les classes d'impacts obtenues

Dans le processus de cartographie conceptuelle, chaque groupe de travail se réunit une nouvelle et dernière fois pour discuter et valider les classes d'impacts qui apparaissent sur la carte. L'équipe d'évaluation aura au préalable établi et imprimé la carte, ainsi qu'une liste d'énoncés regroupés.

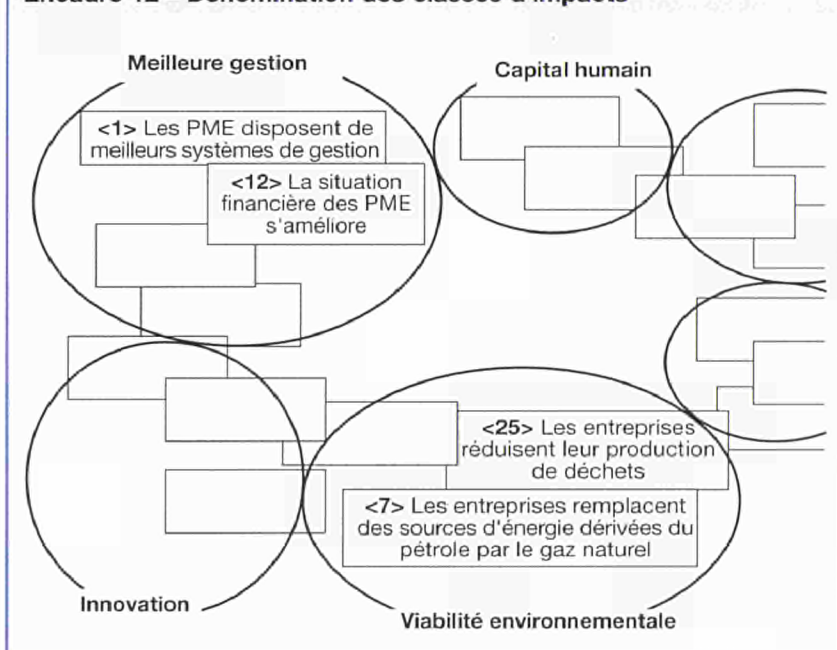
Pour chaque classe successivement, les participants procèdent de la façon suivante :

- Parcourir les impacts faisant partie de la classe et vérifier leur signification.
- Reclassez, si cela paraît pertinent, les énoncés d'impact qui ont de faibles indices de liaison.
- Repérer, au sein de la classe d'impacts, ceux qui ont reçu la plus forte pondération de la part des participants et ceux qui ont le plus fort indice de liaison.

- Discuter de la logique de constitution des classes, en s'aidant au besoin des indications écrites par les participants au moment de la constitution des piles.
- Chercher un mot ou une expression qui synthétise l'ensemble des énoncés d'impact regroupés dans la classe, et en particulier ceux qui ont reçu une forte pondération et qui ont le plus fort indice de liaison.

La discussion de chaque classe d'impacts conduit généralement le groupe de travail à choisir un nom un peu abstrait comme le montre l'encadré ci-dessous. Certains pourront alors s'inquiéter du fait qu'après un intense travail collectif, le résultat final reste imprécis et peu clair. Un tel constat serait d'autant plus gênant dans le cas d'un travail entrepris pour clarifier des objectifs considérés comme trop vagues. En réalité, la procédure ne se prête pas à une telle critique. Si le nom d'une classe d'impacts est abstrait, les énoncés qui composent la classe sont au contraire très clairs et suffisamment précis pour constituer la base de l'évaluation.

Encadré 12 – Dénomination des classes d'impacts



■ Travail avec deux niveaux de transversalité

Une variante intéressante a été choisie dans le cadre de l'évaluation conduite à Valencia. Elle consiste à travailler en deux temps. Dans un premier temps, trois groupes de travail ont été constitués avec, dans chaque groupe, une dizaine de responsables locaux (gestionnaires des mesures et opérateurs). Dans un deuxième temps, un groupe « transversal » a été réuni avec les représentants de la Commission, de l'administration nationale et de l'administration régionale (voir l'Encadré 13).

Encadré 13 – Les groupes de travail constitués à Valencia (Espagne)	
Groupes sectoriels	
Groupe A « Aides directes aux entreprises »	12 gestionnaires de 4 institutions administratives (tourisme, industrie & commerce, finances, économie et entreprises)
Groupe B « Soutien à l'environnement productif »	11 gestionnaires de 6 institutions (commerce, formation professionnelle, tourisme, économie,...)
Groupe C « Infrastructures »	11 gestionnaires de 5 institutions (travaux publics, transports, industrie, qualité environnementale...)
Groupe transversal	
Groupe « Responsables centraux »	Représentants de la Communauté Autonome de Valencia, représentants de l'Etat espagnol et rapporteurs de la Commission européenne

Par exemple, le groupe de travail « Aides directes aux entreprises » comprenait des gestionnaires et des opérateurs concernés par les mesures suivantes :

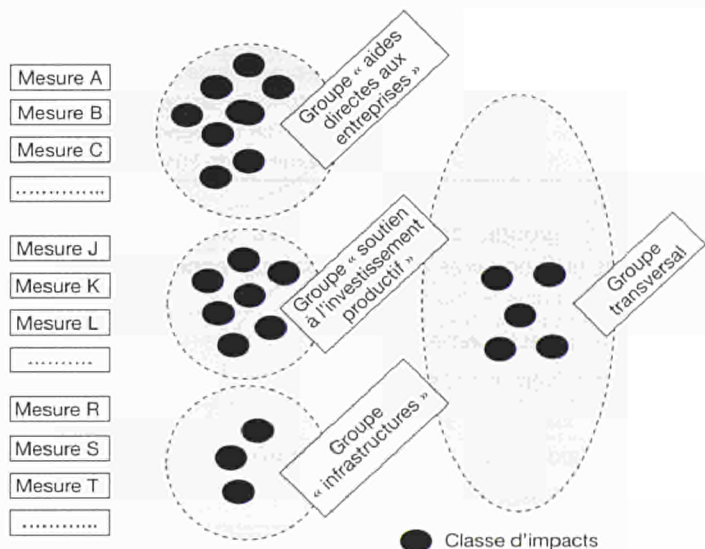
- Plan d'aide à l'investissement ;
- Modernisation des établissements et des systèmes productifs de l'artisanat ;
- Programme d'aides pour les zones prioritaires ;
- Fonds de capital risque pour les PME ;
- Soutien aux économies d'énergie, à la diversification et/ou à l'emploi des sources alternatives d'énergie par les entreprises ;

- Plan de mise en valeur des services aux entreprises ;
- Plan technologique ;
- Soutien à l'amélioration de l'offre touristique ;
- Soutien aux entreprises pour l'adaptation aux normes environnementales.

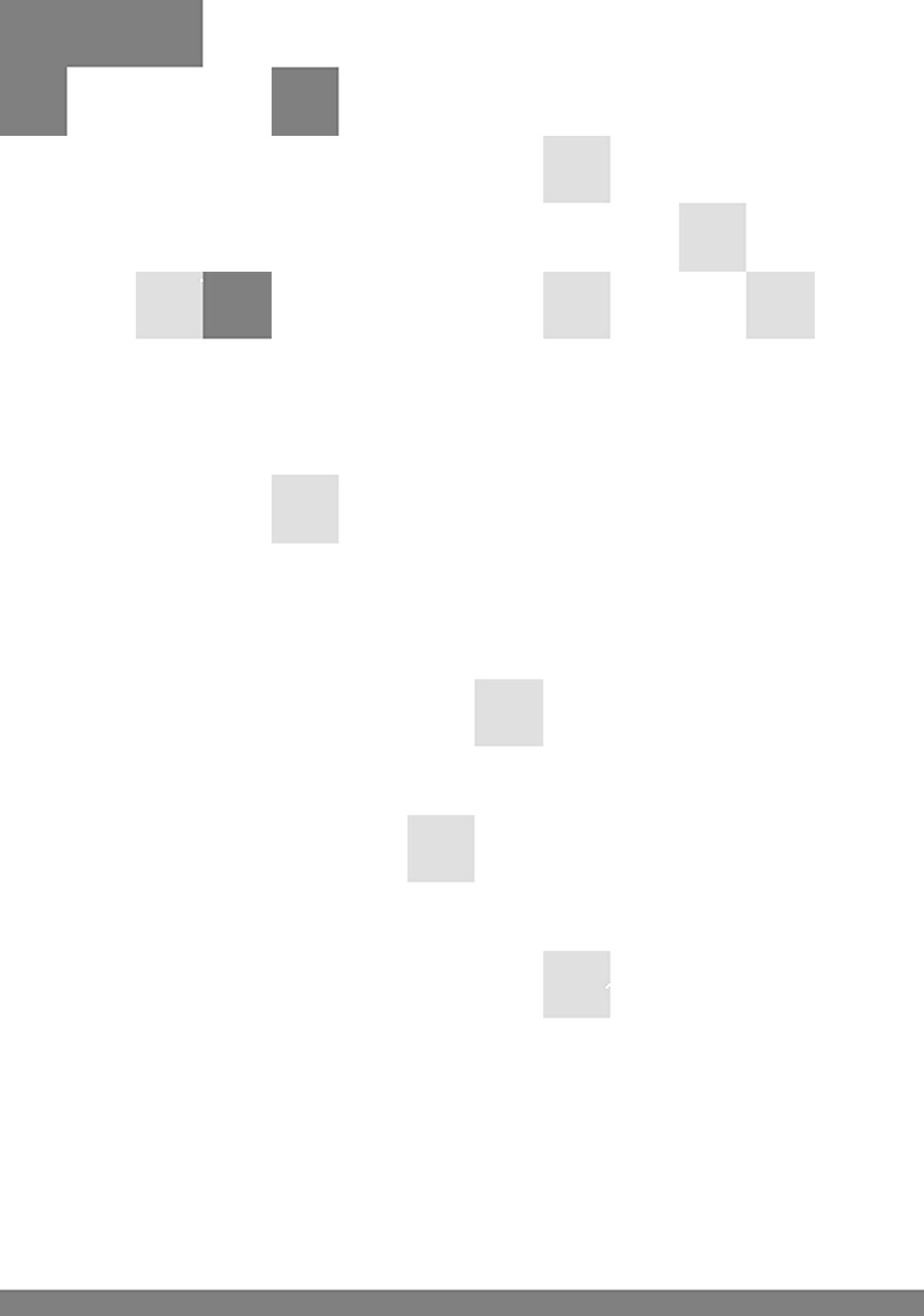
Chacun des trois groupes sectoriels a établi sa propre liste d'énoncés d'impact. Ces énoncés ont été systématiquement séparés en deux niveaux, le premier niveau étant plus proche des impacts spécifiques des mesures et le deuxième niveau étant plus proche de l'impact global du programme.

Au niveau inférieur (plus proche des impacts spécifiques des mesures), on a classé par exemple les énoncés suivants : « le nombre d'initiatives entrepreneuriales a augmenté », « les nouvelles PME disposent de meilleurs systèmes de gestion », « la qualification des employés des entreprises artisanales s'est élevée ». Au niveau supérieur (plus proche de l'impact global du programme), on a classé par exemple les énoncés suivants : « le nombre d'emplois au sein des entreprises régionales a augmenté », « les zones qui ont bénéficié des aides n'ont pas souffert d'émigration de la population ».

Encadré 14 – Deux niveaux de classes d'impacts



Chacun des trois groupes sectoriels a travaillé uniquement sur les énoncés d'impact de niveau inférieur. Tous les énoncés d'impact de niveau supérieur ont été réunis en une seule liste et traités par le quatrième groupe. Les classes d'impacts ainsi constituées se sont donc organisées en deux niveaux comme le montre l'Encadré 14.



Utiliser la cartographie conceptuelle des impacts

■ Construction des critères d'évaluation du programme

La réflexion collective sur laquelle se fonde la cartographie conceptuelle des impacts privilégie la diversité et la richesse des points de vue. Cette démarche qui tient compte de l'opinion des différents partenaires permet de produire des critères d'évaluation acceptés par les différents partenaires d'un programme. La médiation informatique apparaît comme un atout majeur pour parvenir à l'acceptation du référentiel commun d'évaluation par les différents partenaires. Cette technique évite que l'équipe d'évaluation impose sa propre structure logique qui risquerait de rester étrangère aux participants.

Encadré 15 – Critère d'évaluation commun aux différentes mesures

Classe d'impacts	Viabilité environnementale
Critère d'évaluation	Efficacité des mesures en termes de viabilité environnementale.
Question évaluative	Quelle a été la contribution des différentes mesures à la viabilité environnementale des unités économiques concernées par le programme ?
Exemple de méthode d'évaluation	Enquête auprès des unités économiques ciblées par les différentes mesures : entreprises, agents touristiques, exploitations agricoles, artisans, etc. Enquête par questionnaire, par entretiens ou par étude de cas. Les questions posées font référence aux énoncés des impacts qui appartiennent à la classe de la viabilité environnementale, c'est à dire : « Y a-t-il eu réduction de la production de déchets ? », « Y a-t-il eu transfert de consommation d'énergie en faveur du gaz naturel ? », etc.

L'intérêt essentiel de la cartographie conceptuelle est de permettre la construction de critères d'efficacité qui pourront être appliqués de façon transversale à plusieurs mesures. Par exemple, en revenant à l'étape de dénomination des classes (voir Encadré 12), l'équipe d'évaluation pourra utiliser la classe « viabilité environnementale » de la façon indiquée dans

l'Encadré 15. A travers cet exemple on voit qu'une démarche très simple conduit successivement d'une classe d'impacts à un critère d'évaluation, puis à une question évaluative, puis à une méthode d'enquête qui permettra de répondre à la question de façon empirique et pragmatique. Au-delà, il est possible d'utiliser les énoncés constituant chaque classe pour construire des échelles de notation.

L'évaluation vérifiera que les classes d'impacts sont réellement transversales, c'est-à-dire qu'elles s'appliquent à un grand nombre de mesures. La démarche peut être considérée comme un succès si chaque classe d'impacts est potentiellement connectée à un grand nombre de mesures. Plus les connexions sont nombreuses, plus l'évaluation pourra faire des comparaisons et des agrégations entre les mesures et plus il sera possible de réaliser une évaluation synthétique du programme. Il est à noter que la technique de la cartographie conceptuelle favorise de multiples connexions à cause du mode de construction de classes d'impacts. Rappelons en effet que ces classes comprennent tous les impacts intermédiaires attendus par les gestionnaires des différentes mesures, et que ces impacts ont fait l'objet d'un travail de regroupement systématique.

■ Construction d'indicateurs génériques

Un indicateur générique est un indicateur qui permet d'effectuer des quantifications comparables sur plusieurs mesures de nature différente au sein d'un même programme (voir Volume 2 de la Collection). Un tel indicateur permet l'agrégation des données au sein du programme sous forme de somme ou de moyenne. On citera comme exemples d'indicateurs génériques le taux de consommation budgétaire et le taux de réalisation. Certains indicateurs génériques ne s'appliquent qu'à une partie du programme : nombre d'innovations (pour toutes les mesures qui favorisent le développement technologique), proportion des PME parmi les destinataires (pour toutes les mesures qui bénéficient à des entreprises).

Plus les mesures d'un programme sont hétérogènes, plus il est difficile de définir des indicateurs véritablement génériques. Si l'on réalise une cartographie conceptuelle, les classes d'impacts intermédiaires constituent une très bonne aide pour créer ces indicateurs génériques. Il est même possible de consacrer une partie de la dernière réunion du groupe de travail à cet exercice. La discussion portera successivement sur chaque classe d'impacts et procédera comme suit :

- Parcourir les énoncés d'impact constituant la classe.

- Demander aux participants de trouver des indicateurs associés à ces énoncés.
- Redéfinir les indicateurs proposés pour qu'ils soient les plus génériques possibles. Par exemple, un indicateur d'« emplois créés dans les entreprises aidées » pourra être remplacé par un indicateur plus générique d'« emplois créés dans les unités économiques aidées ».
- Privilégier les indicateurs associés aux impacts qui ont reçu les plus importantes pondérations.
- Retenir les indicateurs qu'il sera possible de renseigner pour le plus grand nombre possible de mesures.
- 2 à 5 indicateurs pourront être proposés pour chaque classe d'impacts.
- Du fait de l'importance de la question et du temps nécessaire pour une discussion approfondie des indicateurs, il est conseillé d'aborder le sujet après une pause conséquente, voire de prévoir une demi-journée de travail à cet effet.

Exemple : Indicateurs associés aux classes d'impacts

Figurent ci-dessous quelques indicateurs d'impact intermédiaire correspondant aux critères génériques définis par le groupe central dans le cadre de l'évaluation pilote MEANS. Rappelons que cette évaluation a porté sur 25 mesures du programme de développement socio-économique de la région de Valencia en Espagne pour la période 1994-99.

Classe d'impacts	Indicateur générique	Source d'information	
		Programme	Contexte
Meilleur aménagement du territoire	Fréquentation des transports publics urbains Nombre de voyageurs par km des transports publics	Opérateurs de transports aidés Opérateurs de transports aidés	Statistiques du Ministère des Transports Statistiques du Ministère des Transports
Meilleur environnement productif	Elévation du niveau d'études de la population active Densité d'établissements bancaires	Opérateurs des mesures de formation et d'éducation	Institut Statistique National Institut Statistique National

■ Une technique adaptée au partenariat

La cartographie conceptuelle des impacts permet d'engager une dynamique de groupe. Cette dynamique présente un intérêt évident dans une démarche d'évaluation : l'évaluation sera d'autant mieux acceptée par les parties prenantes du programme que ces dernières auront collectivement participé à la définition du référentiel commun d'évaluation. Par ailleurs, les échanges entre les gestionnaires ne peuvent que renforcer la connaissance mutuelle des actions et favoriser ainsi les effets de synergie. La participation des gestionnaires au processus favorise une meilleure prise de conscience par ces derniers des effets que sont censées produire les interventions, au-delà du stade des réalisations.

La possibilité de mettre en œuvre la méthode et d'obtenir les résultats en un laps de temps relativement court, de l'ordre d'une semaine pour un cycle de réunions, rend la cartographie conceptuelle des impacts compatible avec le calendrier d'évaluation. La cartographie conceptuelle des impacts présente cependant des limites qui peuvent restreindre l'utilité de la technique dans certaines circonstances.

Il apparaît ainsi que cette méthode mobilise un grand nombre de gestionnaires ou de responsables sur plusieurs jours. Même si certaines réunions sont remplacées par des échanges de courrier, un tel projet nécessite que du temps soit consacré par de nombreux fonctionnaires. Outre les problèmes pratiques liés à la gestion des agendas, une telle mobilisation d'agents publics occasionne un coût pour la collectivité. Cette mobilisation n'apparaît justifiée que dans la mesure où d'importants enjeux motivent l'évaluation.

Il est également clair que la constitution des groupes est sensible. En cas de conflit, si tous les partenaires n'acceptent pas de participer au processus, ou s'il n'est pas possible de parvenir à une représentation équilibrée, la mise en œuvre de la technique est compromise. La difficulté rencontrée pour constituer les groupes est aussi forte lorsqu'il s'agit d'impliquer au processus des représentants des publics cibles. L'identification, comme l'équilibre de la représentation, apparaissent dans ce cas relativement complexes.

Enfin, il faut tenir compte du jeu des acteurs dans le cadre des réflexions de groupe. La cartographie conceptuelle des impacts permet d'éviter qu'un leader au sein du groupe influence la constitution des piles ou la notation des énoncés. En revanche, cette influence est possible lors de la

production collective des énoncés, et lors de la discussion de la carte des impacts. Il revient donc à l'animateur des réunions de veiller à ce que les résultats soient bien ceux de l'ensemble du groupe. Il est bon que l'animateur maîtrise les techniques d'animation de groupe afin que le jeu des acteurs ne biaise pas les résultats de l'exercice.

Sachant que la mise en œuvre de la technique nécessite une bonne connaissance du domaine de l'intervention évaluée, ainsi que des connaissances statistiques suffisantes, cela peut inciter à ce que le projet soit réalisé par une équipe pluridisciplinaire.



II Conclusion et Synthèse

Construire et utiliser des échelles de notation

Les systèmes de notation sont de plus en plus utilisés dans le cadre de la sélection des projets, notamment au Royaume-Uni. La notation a l'avantage de prendre en compte, avec une grande souplesse, toutes les informations quantitatives et qualitatives qui permettent de juger selon un critère donné. Son domaine d'emploi est donc plus large que celui des indicateurs qui sont difficilement applicables à certains critères à cause de leur caractère quantitatif. L'originalité de la démarche proposée dans ce chapitre réside dans ce que la notation est utilisée pour l'évaluation des mesures et non pas pour la sélection des projets. Ce chapitre se relie au précédent car il est possible de construire une échelle de notation pour chaque classe d'impacts. Il se relie également au chapitre IV car l'analyse multicritère s'appuie le plus souvent sur des évaluations d'impact réalisées à l'aide d'échelles de notation.



Une solution pour évaluer des mesures différentes sous une forme homogène

■ Conclusions qualitatives sur les impacts des mesures

L'évaluation de l'ensemble des mesures d'un programme se donne pour contrainte d'examiner tous les impacts de toutes les mesures. Dans le cas d'un programme socio-économique qui comprendrait par exemple une quarantaine de mesures à évaluer selon cinq grandes classes d'impacts (compétitivité, emploi, égalité des chances, environnement, qualité de vie), l'équipe d'évaluation devrait se prononcer sur pas moins de 200 impacts. Ce simple chiffre montre qu'il ne sera pas possible d'analyser systématiquement tous ces impacts de façon approfondie. L'évaluation d'ensemble d'un programme peut s'accompagner tout au plus d'approfondissements concernant quelques impacts particulièrement sensibles.

L'étude approfondie d'un impact donné suppose une observation de terrain. Les techniques de base employées pour ce type d'analyse sont, par exemple, les groupes de comparaison et l'analyse de régression. L'emploi de ces techniques nécessite du temps et des moyens importants, et imposent que le champ d'observation soit limité à une seule catégorie de réalisations et à un seul public. L'approfondissement porte généralement sur un mécanisme d'impact de niveau microéconomique ou microsocial. L'analyse d'un impact est seulement possible à une date suffisamment tardive pour que celui-ci ait eu le temps de se matérialiser (voir le Volume 1 de la Collection).

Toutes ces contraintes font qu'une véritable analyse des impacts ne peut pas être étendue à l'ensemble d'un programme multisectoriel. Dans cette situation, on doit donc privilégier des approches applicables de façon pragmatique, par exemple l'utilisation de coefficients d'extrapolation pour estimer les impacts de façon forfaitaire. Ainsi, un coefficient d'extrapolation permettra d'estimer le nombre d'emplois temporaires créés par million d'euros d'investissement en infrastructure ou la valeur ajoutée annuelle générée par million d'euros d'investissement dans les PME. Un coefficient peut être déduit d'une enquête approfondie conduite dans le cadre de la même évaluation ou d'une évaluation précédente. Un coefficient peut aussi être emprunté à une évaluation ou à une recherche portant sur des

interventions similaires, à conditions que les conclusions soient considérées comme généralisables. Un coefficient peut également être établi pour les besoins d'une évaluation par un ou plusieurs experts.

Dans une situation d'évaluation à mi-parcours ou dans une évaluation ex post, il est possible d'utiliser les données de suivi des réalisations comme base d'extrapolation. Ainsi par exemple, on se fondera sur le nombre de lits d'un équipement touristique récemment construit, pour estimer le nombre d'emplois créés à l'aide d'un ratio de nombre d'emplois par lit. Dans certains cas, il est demandé aux destinataires de faire une déclaration sur les impacts au moment où ils sont en contact avec les opérateurs. Ils doivent par exemple indiquer le nombre d'emplois créés au moment du dernier paiement ou le nombre d'emplois à créer au moment de leur demande d'aide. Dans ce cas il est recommandé d'appliquer un coefficient correcteur pour tenir compte des effets d'aubaine. Ce coefficient peut varier de 0,8 dans les meilleurs cas, à moins de 0,3 pour certaines interventions où les effets d'aubaine sont particulièrement importants.

L'estimation des impacts par extrapolation n'est possible qu'à certaines conditions : (1) il faut disposer d'une base d'extrapolation sous forme d'indicateurs de suivi des réalisations et des résultats, (2) il faut que l'impact à estimer soit défini de façon précise et associé à un indicateur et (3) il faut que la littérature ou les évaluations précédentes fournissent des ratios de type « impact / réalisation » et que ces ratios soient considérés comme généralisables. Ces conditions sont relativement sévères et il est rare qu'elles soient réunies pour tous les impacts et toutes les réalisations. La situation la plus fréquente est donc celle où l'équipe d'évaluation ne peut pas réaliser une estimation chiffrée de tous les impacts pour toutes les mesures. S'il est exigé que l'évaluation donne une vue d'ensemble des effets du programme, alors elle devra se contenter d'apporter des conclusions qualitatives pour une partie au moins des impacts.

Même lorsqu'une démarche d'extrapolation permet de produire des conclusions quantitatives, il convient encore de nuancer les estimations d'impacts par des interprétations plus qualitatives. En particulier, l'équipe d'évaluation doit expliquer et justifier les coefficients d'extrapolation qu'elle a choisis, en considérant l'ensemble des éléments de contexte qui conditionnent les processus de mise en œuvre du programme, qui caractérisent les publics concernés et qui ont une influence sur les performances obtenues. Les professionnels de l'évaluation reconnaissent généralement que les indicateurs sont seulement une des bases parmi

celles qui permettent de formuler leurs conclusions. En particulier, les indicateurs doivent être interprétés à la lumière de toutes les informations qualitatives accessibles et pertinentes.

Pour conclure ce rapide tour d'horizon de l'évaluation des impacts, on est conduit à reconnaître qu'elle débouche souvent sur des conclusions qualitatives.

■ Présentation synoptique des conclusions qualitatives

En rédigeant des conclusions qualitatives sur les impacts d'une mesure donnée, l'équipe d'évaluation se place naturellement dans la perspective et dans le contexte de cette mesure. Les conclusions ont donc tendance à être rédigées avec une logique et un vocabulaire spécifique à telle ou telle mesure comme le montre l'exemple suivant. Ce fait crée des problèmes pour rédiger le chapitre de synthèse de l'évaluation à l'échelle du programme. On constate ainsi que des rapports d'évaluation sont organisés avec une section par mesure contenant des conclusions solides et argumentées, alors que la section de synthèse finale est souvent à la fois courte, pauvre et fragile.

Exemple : Conclusions qualitatives hétérogènes pour différentes mesures

Conclusion sur l'impact de la mesure A (formation destinée aux chômeurs de longue durée) : « Sur la base de l'examen du parcours post-formation de quelques stagiaires considérés comme typiques, on estime que l'employabilité a été améliorée dans la majorité des cas, avec des résultats meilleurs dans les métiers du tourisme et des résultats moins bons dans les métiers de l'environnement (ces derniers ne représentent qu'une petite partie des formations). Sur la base d'évaluations de programmes comparables, le taux de placement peut être estimé entre 30 et 50%. On a identifié un impact positif en terme de réorientation des femmes vers des emplois non menacés et considérés comme masculins. Il existe cependant un risque important d'apparition d'effet de substitution et d'écroulement. »

Conclusion sur l'impact de la mesure B (Aide à la création d'entreprise) : « Sur la base d'évaluations de programmes comparables, le taux de survie peut être estimé entre 20 et 30%. Il semble que le taux de survie soit supérieur pour les entreprises créées dans le secteur de l'environnement dans lequel se situent 35% des entreprises aidées. Par ailleurs on considère que la majorité des créateurs qui échouent ont cependant amélioré significativement leur employabilité. Il est à noter que 45% des créations aidées sont le fait de femmes, alors que cette proportion n'est que de 30% dans l'ensemble du pays. La différence peut être mise à l'actif du programme. »

Pour faciliter la préparation d'un chapitre de synthèse finale, il est intéressant de récapituler toutes les conclusions qualitatives sous la forme d'un tableau synoptique organisé en fonction des différentes mesures et des principales classes d'impact comme le montre l'Encadré ci-dessous.

Encadré 16 – Tableau synoptique qualitatif des mesures et de leurs impacts

Mesures	Impact sur l'emploi	Impact sur l'égalité des chances	Impact sur l'environnement
Formation destinée aux chômeurs de longue durée	L'employabilité a été améliorée dans la majorité des cas. Sur la base d'évaluations de programmes comparables, le taux de placement peut être estimé entre 30 et 50%. Il existe un risque important d'apparition d'effet de substitution et d'écrémage.	La formation a eu un impact positif en terme de réorientation des femmes vers des emplois non menacés et considérés comme masculins.	Une petite partie des formations ont abouti à des qualifications dans le domaine des métiers de l'environnement. Cependant, très peu de stagiaires ont trouvé un emploi correspondant à ces qualifications.
Aide à la création d'entreprise	Sur la base d'évaluations de programmes comparables, le taux de survie peut être estimé entre 20 et 30%. On considère que la majorité des créateurs qui échouent ont cependant amélioré significativement leur employabilité.	45% des créations aidées sont le fait de femmes, alors que cette proportion n'est que de 30% dans l'ensemble du pays. La différence peut être mise à l'actif du programme.	35% des entreprises créées l'ont été dans le domaine de l'environnement. Il semble que leur taux de survie soit supérieur au taux moyen.
Etc.....			

Le tableau synoptique constitue une aide pour l'équipe d'évaluation mais il lui reste beaucoup à faire pour produire de véritables conclusions synthétiques. En théorie, trois démarches au moins sont à entreprendre :

- En premier lieu, il convient de produire des conclusions de synthèse pour chaque classe d'impacts. Par exemple, l'équipe d'évaluation pourra identifier les mesures qui sont les plus efficaces en termes d'impact sur l'emploi, ainsi que les mesures qui ont un effet négligeable ou éventuellement dommageable sur ce plan, et ainsi de suite pour les autres impacts. Cette première démarche est développée dans les pages qui suivent consacrées aux échelles de notation.
- Ensuite, l'équipe d'évaluation doit faire une analyse des synergies entre les mesures. Par exemple, on pourra montrer que les deux mesures de l'Encadré 16 ont des impacts qui se renforcent mutuellement dans le domaine de l'emploi et de l'environnement. Cette seconde phase est développée dans le chapitre III, en relation avec la technique de la matrice des impacts croisés.
- Enfin, il convient de faire une synthèse générale couvrant les principaux impacts et la synergie entre les mesures. A l'issue de cette troisième phase, l'équipe d'évaluation devrait être capable de classer les mesures en fonction de leur contribution à l'efficacité globale du programme. Ce point sera développé dans le chapitre IV, en relation avec la technique de l'analyse multicritère.

Le présent chapitre est consacré à la première de ces trois démarches, c'est-à-dire à la rédaction de conclusions synthétiques portant séparément sur les différentes classes d'impacts. L'esprit de la démarche est de formuler des jugements comparatifs entre les différentes mesures, de façon transversale à l'ensemble du programme. La technique proposée est celle de l'échelle de notation (*scoring scale*). Cette technique produit des conclusions qui se prêtent naturellement à des comparaisons, à des agrégations, à des calculs de moyennes. Bien que les échelles de notation ne soient pas comparables, d'une classe d'impacts à l'autre, la technique apporte cependant une aide directe à la rédaction de la synthèse finale de l'évaluation.



Mode d'emploi des échelles de notation

Les échelles de notation sont utilisées dans de nombreux domaines, souvent de façon spontanée et peu formalisée. De ce fait, elles sont parfois critiquées pour leur caractère subjectif. Il existe cependant des cas où la notation est établie de façon très rigoureuse et où elle joue un rôle reconnu par tous. Il suffit de citer l'exemple des agences internationales de notation qui diffusent des appréciations sur la fiabilité des emprunteurs, sans parler du rôle économique joué par les notes décernées dans les guides gastronomiques. Dans les domaines des politiques d'éducation, de santé ou d'intégration sociale, il existe des centaines d'échelles de notation qui ont été développées dans le cadre de travaux de recherche et qui sont souvent reprises par des travaux d'évaluation.

A ce jour, la notation reste peu utilisée pour l'évaluation des politiques de développement socio-économique. Une exception notable est cependant celle de l'aide au développement. Dans ce domaine, la technique est employée par de nombreux bailleurs de fonds depuis la seconde moitié des années 1980 (voir l'exemple cité page 63). Les pages qui suivent proposent un mode d'emploi pour l'usage de cette technique dans le cadre des programmes socio-économiques de l'Union européenne. Ce mode d'emploi se décline en trois variantes selon le moment où l'échelle de notation est établie :

Variante 1 : l'équipe d'évaluation a déjà formulé des conclusions qualitatives pour chaque mesure, l'échelle de notation aide à formuler une conclusion de synthèse au niveau du programme.

Variante 2 : l'équipe d'évaluation a déjà collecté des informations qualitatives sur les effets des mesures, mais elle n'a pas encore rédigé ses conclusions, l'échelle de notation aide à formuler des conclusions, mesure par mesure, d'une façon homogène et facile à synthétiser.

Variante 3 : l'équipe d'évaluation commence son travail, l'échelle de notation aide à collecter les données sous une forme qui permettra de formuler des conclusions homogènes et faciles à synthétiser.

■ Echelles pour la synthèse de conclusions qualitatives

Dans la première variante, l'équipe d'évaluation a déjà formulé ses conclusions. Celles-ci se présentent sous la forme du tableau synoptique

de l'Encadré 16, chaque colonne concerne une classe d'impacts et réunit les conclusions rédigées pour toutes les mesures à propos de cette classe d'impacts. L'Encadré 17 est un extrait du tableau synoptique concernant l'impact sur l'égalité des chances.

Encadré 17 – Présentation synoptique des conclusions relatives à un impact

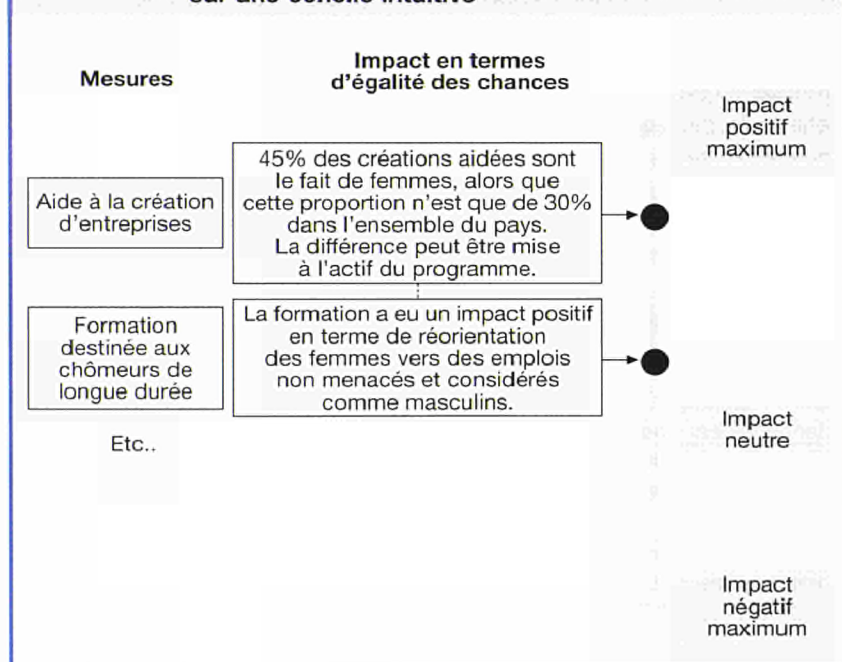
Mesure		Description de l'impact sur l'égalité des chances
1	Formation destinée aux chômeurs de longue durée	La formation a eu un impact positif en terme de réorientation des femmes vers des emplois non menacés et considérés comme masculins.
2	Aide à la création d'entreprise	45% des créations aidées sont le fait de femmes, alors que cette proportion n'est que de 30% dans l'ensemble du pays. La différence peut être mise à l'actif du programme.
..	Autres mesures...	Etc....

Les deux impacts décrits dans l'Encadré 17 sont-ils comparables ? Peut-on dire, en lisant ces conclusions, que la mesure n°1 est la plus efficace en termes d'impact sur l'égalité des chances ? Arrivé à cette phase de son travail, l'équipe d'évaluation constate que les impacts ne sont pas directement comparables et qu'elle peut difficilement aller plus loin par ses propres moyens en direction d'une conclusion de synthèse. Une voie possible, pour formuler des jugements comparatifs, consiste à faire exprimer un avis par le groupe de pilotage ou par un comité d'experts. La démarche est la suivante :

- Organiser un séminaire d'une demi-journée avec un panel d'experts ou avec le groupe de pilotage. De préférence il s'agira d'un groupe de pilotage élargi en forme d'instance d'évaluation, comme proposé dans le Volume 1 de la Collection.
- Présenter aux participants les conclusions de l'évaluation mesure par mesure. Vérifier la crédibilité des conclusions et les valider.
- Présenter le tableau synoptique des conclusions (voir Encadré 16).
- Examiner la première colonne du tableau synoptique et demander à chaque participant de situer intuitivement les impacts de chaque mesure sur une échelle comme dans l'Encadré 18. Comparer les classements réalisés par les participants, les discuter et les faire converger autant que possible.

- Poursuivre le classement avec les autres colonnes du tableau synoptique.
- Après le séminaire, convertir les classements effectués en notes échelonnées de -10 (impact négatif maximum) à + 10 (impact positif maximum) en passant par 0 (impact neutre). Convertir le tableau synoptique de l'Encadré 16 en tableau de notes (matrice de notation des impacts). La construction des échelles de notation autorise des comparaisons au sein d'une même colonne (par exemple, telle mesure est mieux notée qu'une autre). Par contre, les notes ne peuvent pas se comparer d'une colonne à l'autre (par exemple, une note 5 pour l'impact sur l'emploi n'est pas comparable à une note 5 pour l'impact sur l'environnement).

Encadré 18 – Classement des impacts obtenus sur une échelle intuitive



Dans le cadre d'une évaluation pilote du programme MEANS, une procédure techniquement plus rigoureuse a été proposée et appliquée. Dans cette variante, il n'est pas demandé aux participants de placer intuitivement les impacts sur une échelle de notation comme dans

L'Encadré ci-dessus. On leur demande plutôt d'exprimer leurs préférences entre tous les impacts pris deux à deux. Leurs opinions sont recueillies à travers une série de questions du type de celle-ci : « laquelle de ces deux situations préférez-vous :

- une situation où l'on constate un impact positif en terme de réorientation des femmes vers des emplois non menacés et considérés comme masculins,.....ou..
- une situation où l'on constate que des femmes ont créé leur entreprise dans une proportion supérieure à celle qui s'observe dans l'ensemble du pays (45% des créations aidées sont le fait de femmes au lieu de 30%)..... ? - En ce qui concerne la situation que vous préférez, votre préférence est-elle très forte, forte, faible ? ».

Le questionnement se poursuit pour tous les impacts de la première colonne comparés deux à deux, puis de même pour les autres colonnes. L'ensemble des réponses est traité à l'aide d'un logiciel informatique qui attribue à chaque impact une note comprise entre -10 et +10 de façon à refléter le plus fidèlement possible les préférences des participants. Le logiciel utilisé est dénommé « Macbeth ». Il a été créé par une équipe de l'Université de Lisbonne (voir annexe bibliographique Bana e Costa 1994). Ce logiciel est également évoqué plus loin (voir page 115) dans le cadre d'une des variantes de l'analyse multicritère.

Encadré 19 – Matrice de notation des impacts

Mesures	Impact sur l'emploi	Impact sur l'égalité des chances	Impact sur l'environnement
Formation destinée aux chômeurs de longue durée	+4	+3	0
Aide à la création d'entreprise	+8	+7	+5
Etc.....			

Quelle que soit la procédure retenue, le résultat final est constitué par une série d'échelles de notation (une par classe d'impacts). Chaque mesure se voit attribuer une note sur chaque échelle. On peut rassembler toutes les notes dans une matrice de notation des impacts (voir Encadré 19) qui complète le tableau synoptique des conclusions qualitatives (voir

Encadré 16) et qui a la même structure. La combinaison des conclusions qualitatives et des notes permet de formuler des conclusions transversales à l'échelle du programme pour chaque classe d'impacts.

■ Echelles pour la formulation de conclusions structurées

Dans la mesure du possible, il est préférable de construire des échelles de notation un peu plus tôt dans le processus d'évaluation, c'est-à-dire avant la formulation des conclusions mesure par mesure. De cette façon, l'équipe d'évaluation pourra utiliser les échelles de notation pour rédiger des conclusions d'une façon plus structurée et plus comparable d'une mesure à l'autre. Lorsqu'elles sont utilisées en évaluation, les échelles de notation sont généralement conçues dans cet esprit.

Dans certains cas (voir l'exemple de la Banque Mondiale ci-dessous), des échelles de notation standard sont imposées à toutes les évaluations qui s'inscrivent dans le cadre d'une même politique. Cette formule est difficilement applicable aux programmes socio-économiques financés par l'Union européenne. En effet, les différents programmes étant cofinancés en partenariat avec différentes autorités nationales et/ou régionales, il ne paraîtrait pas légitime de standardiser la façon de juger leurs impacts.

Exemple : Les échelles de notation utilisées par la Banque mondiale

La Banque Mondiale réalise environ 200 évaluations de projets tous les ans. Toutes ces évaluations comportent des conclusions qualitatives et des conclusions sous forme de notes, structurées en fonction d'un système standard. Les échelles de notation intègrent cinq dimensions de la performance des projets : l'impact économique, la durabilité, l'impact sur le développement institutionnel, la performance financière de la banque et celle de l'emprunteur. Tous les projets sont évalués par leur responsable (auto-évaluation) au moment de leur achèvement selon une notation en 4 échelons pour chaque dimension (« très satisfaisant », « satisfaisant », « insatisfaisant », « très insatisfaisant »). Dans un certain nombre de cas, et notamment en cas de doutes sur l'auto-évaluation, une deuxième notation est réalisée par le Département d'Évaluation des Opérations de la Banque (OED). Cette seconde notation fait toujours l'objet d'une discussion avec les responsables du projet. Si les jugements ne convergent pas (le cas est très rare), c'est à une commission d'arbitrage qu'il revient de fixer les notes définitives. La Banque Mondiale ne calcule pas de note globale pondérée pour synthétiser toutes les dimensions de la performance mais d'autres organismes le font (ex : Banque Asiatique pour le Développement).

Dans d'autres cas, l'évaluation est réalisée par un panel d'experts qui construit ses propres échelles de notation (voir l'exemple britannique ci-dessous). Dans cette démarche, les membres du panel reçoivent toutes les informations qui ont été collectées et rassemblées à propos des différentes mesures. Ils formulent à la fois des conclusions qualitatives et des conclusions sous forme de notes. Pour cela, ils se mettent d'accord sur une ou plusieurs échelles de notation qui sont établies lors de la première réunion du panel. Chaque panel construit son propre système de notation. Il est à noter que les panels d'experts sont peu employés dans l'évaluation des programmes socio-économiques de l'Union européenne.

Exemple : La notation par des panels d'experts

La notation par des panels d'experts est une formule employée depuis une dizaine d'année pour évaluer la performance des centres de recherche universitaires au Royaume-Uni. Dans ce pays, les travaux de tous les centres de recherche font l'objet à la fois d'un suivi quantitatif et d'évaluations périodiques par des panels d'experts. Dans leur évaluation, les experts prennent en compte les indicateurs (ex : nombre de publications) mais aussi des éléments purement qualitatifs (ex : qualité des quatre meilleures publications de chaque chercheur). L'évaluation débouche sur un classement des équipes de recherche (notées sur une échelle de 1 à 5). L'attribution des fonds publics est directement liée à ce classement.

Dans le domaine de la politique européenne de cohésion, une pratique intéressante est celle qui est employée au Royaume-Uni pour sélectionner les projets candidats aux financements des programmes. L'exemple qui suit présente la technique adoptée dans la région des Highlands and Islands. Cet exemple est intéressant car les échelles de notation sont construites en partenariat et sont utilisées pendant toute la vie du programme. Beaucoup de programmes britanniques adoptent d'ailleurs des échelles de notation similaires.

Exemple : Echelles de notation pour la sélection des projets

Dans le cadre du programme de développement régional des Highlands and Islands, deux cycles d'appel à projets ont lieu chaque année. Toutes les demandes de financement sont d'abord examinées en fonction de leur éligibilité. Les projets considérés comme éligibles font l'objet d'une notation selon huit échelles communes à l'ensemble du programme (liens avec les objectifs, efficacité probable, respect de l'environnement, effet de levier, innovation, liens avec d'autres projets, intervention du secteur privé et soutien des partenaires locaux). D'autres échelles sont également employées mais sont plus spécifiques à certaines mesures.

Chaque projet est noté sur chacune des échelles. Un système de pondération vient compléter la procédure pour permettre d'attribuer une note globale pondérée à chaque projet. Après confrontation des demandes de financement et des ressources disponibles, une barre d'admission est établie et une liste de projets à financer est recommandée. La décision finale revient aux responsables du programme.

Les échelles de notation étant utilisées de façon fréquente et satisfaisante pour la sélection des projets, plusieurs recommandations ont été faites pour étendre l'emploi de cette technique à l'évaluation des programmes. Les propositions suggèrent généralement de créer des échelles concernant les principaux impacts du programme et de s'en servir pour noter, non plus les projets, mais les mesures. Une démarche de ce type a été appliquée à l'évaluation transversale de plus de cent mesures de développement socio-économique en République d'Irlande. Cette démarche a permis de construire des comparaisons entre mesures et elle a abouti à des recommandations de réallocations budgétaires dont une partie significative a été acceptée.

Exemple : Echelles de notation pour évaluer la pertinence des mesures

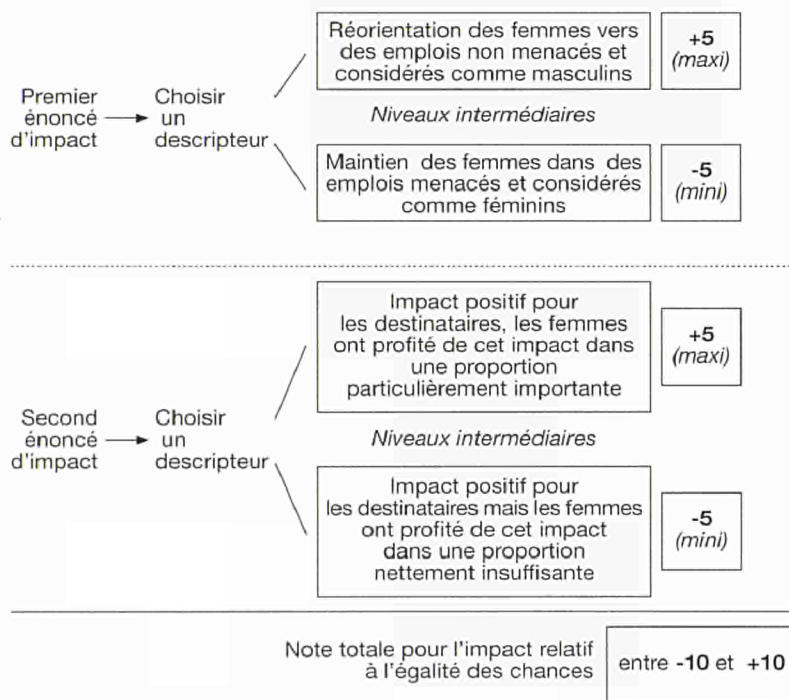
L'évaluation à mi-parcours d'un ensemble de programmes socio-économiques en République d'Irlande a porté sur plus d'une centaine de mesures. Pour chaque mesure, l'équipe d'évaluation a recherché s'il était possible de justifier son financement par des fonds publics. Quatre catégories de mesures ont été établies, correspondant à quatre formes de justification. Pour chaque catégorie l'équipe d'évaluation a construit une échelle de notation permettant d'apprécier plusieurs degrés de justification. Chaque échelle a été centrée sur une mesure typique (« anchor measure ») pouvant servir de point de comparaison pour les autres mesures. Toutes les mesures ont ensuite été notées en s'appuyant notamment sur des données qualitatives collectées précédemment. La notation a permis d'identifier une vingtaine de mesures sensibles, susceptibles de faire l'objet de décisions budgétaires importantes. Ces mesures sensibles ont fait l'objet d'une évaluation plus approfondie. Finalement, l'équipe d'évaluation a recommandé l'abandon ou le renforcement d'une série de mesures. Dans un nombre de cas significatif, les décisions prises ont été en ligne avec les recommandations.

L'exemple qui précède s'applique à une évaluation de la pertinence des mesures. On peut s'en inspirer très directement pour réaliser des échelles de notation des impacts. La procédure aboutit à créer des échelles de notation du type de celle qui est présentée dans l'Encadré 20. Le mode d'emploi est le suivant :

- L'équipe d'évaluation fait une première récapitulation rapide des informations collectées et des analyses réalisées. Elle identifie quelques classes d'impacts, par exemple impact sur l'emploi, impact sur l'égalité des chances, impact sur l'environnement, etc.
- Pour chaque classe d'impacts, l'équipe d'évaluation sélectionne des phrases qui décrivent les différents aspects des impacts constatés, par exemple : « les femmes qui ont participé à telle mesure ont bénéficié d'une réorientation vers des emplois non menacés et considérés comme masculins » ou « telle mesure a eu un impact positif pour ses destinataires et les femmes ont profité de cet impact dans une proportion particulièrement importante ». L'équipe d'évaluation réécrit ces phrases de la façon la plus générale possible afin qu'elles puissent s'appliquer à la plupart des mesures. Chaque phrase décrit précisément et clairement une situation considérée comme une réussite en terme d'obtention d'un impact. Ces phrases sont appelées « énoncés d'impact ». Pour une construction plus rigoureuse des énoncés d'impact et des classes d'impacts, on peut avoir recours à la cartographie conceptuelle des impacts.
- Chaque énoncé décrit une situation considérée comme réussie du point de vue de l'obtention d'un impact. En réalité, certaines situations feront apparaître des réussites partielles, ou une absence d'impact, ou même des contre-performances. Pour pouvoir traiter tous les cas de figure, l'équipe d'évaluation doit donc décrire plusieurs niveaux de réussite possibles. On peut ainsi construire des échelles d'impact comportant plusieurs niveaux entre 0 (pas d'impact) et 10 (excellent impact). Les phrases décrivant les différents niveaux d'impact sont appelées « descripteurs ». Pour chaque énoncé d'impact, il convient de rédiger plusieurs descripteurs pour des niveaux d'impact positifs, mais il convient aussi de prévoir des niveaux d'impacts négatifs.
- L'équipe d'évaluation rassemble ensuite tous les descripteurs concernant une classe d'impact sur une même échelle comme cela apparaît dans l'exemple de l'Encadré 20. Si une même échelle fait référence à plusieurs énoncés, les plages de variation de chaque énoncé sont diminuées (par exemple de $-10/+10$ à $-5/+5$) de façon à ce que le total des notes d'une même fiche soit toujours compris entre -10 (cas où l'on choisit tous les descripteurs les plus défavorables) à $+10$ (cas où l'on choisit tous les descripteurs les plus favorables).
- On organise ensuite un séminaire d'une demi-journée avec un panel d'experts ou avec le groupe de pilotage élargi.

- L'équipe d'évaluation présente ses travaux aux participants en commençant par la première classe d'impacts et en continuant jusqu'à la dernière. Pour chaque classe d'impacts, tous les énoncés et les descripteurs sont discutés et améliorés jusqu'à ce qu'ils soient compris sans ambiguïté. L'équipe d'évaluation présente sa proposition d'échelle de notation. Cette échelle est discutée par les participants et améliorée jusqu'à refléter correctement leur opinion.

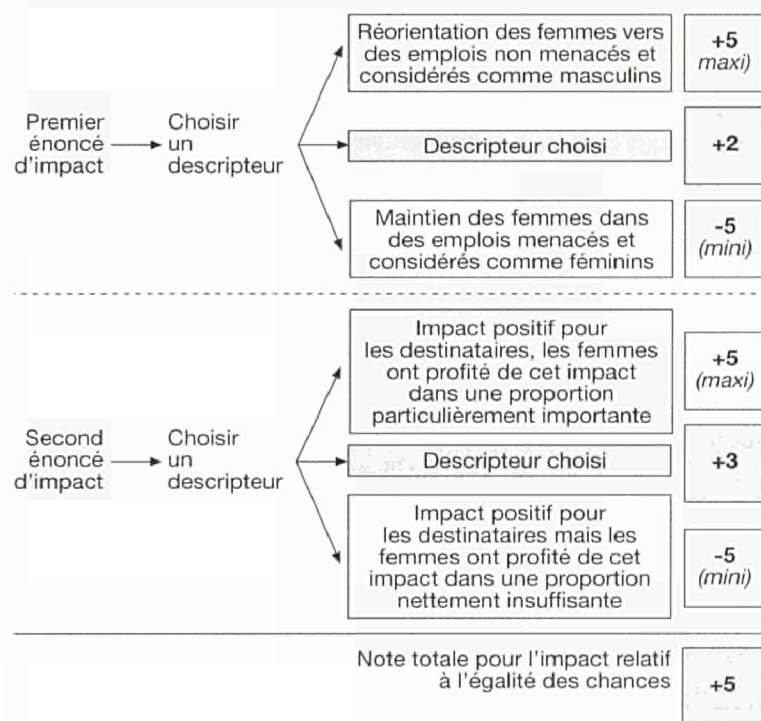
Encadré 20 – Echelle de notation constituée avec des descripteurs d'impact



- Après le séminaire, l'équipe d'évaluation fait une seconde récapitulation des informations collectées et de ses analyses. Pour chaque mesure, elle choisit sur les échelles de notation le ou les descripteurs qui lui paraissent refléter le mieux ses impacts. Cela lui permet d'établir la notation des impacts de la mesure (voir Encadré 21). Les descripteurs

choisis n'ont pas besoin de s'ajuster très exactement aux conclusions que l'équipe d'évaluation pense devoir dégager de ses observations. En effet, au-delà de la notation, l'équipe rédige aussi des conclusions qualitatives complètes, ce qui lui donne toute latitude pour apporter des nuances. Le principal intérêt des échelles de notation est de fournir un cadre pour étudier toutes les mesures selon une structure homogène qui facilitera la rédaction des conclusions de synthèse à l'échelle du programme.

Encadré 21 – Evaluation d'une mesure à l'aide d'une échelle de notation relative à un impact attendu



A l'issue de son travail, l'équipe d'évaluation obtient un tableau synoptique des conclusions qualitatives et une matrice de notation des impacts, dans une forme exactement similaire à celles qui ont été décrites précédemment (voir Encadrés 16 et 19). On verra plus loin comment cette matrice est utilisée.

■ Passage des échelles de notation aux indicateurs

Si de multiples échelles de notation sont construites pour les besoins de l'évaluation d'ensemble d'un programme, et si elles sont utilisées une seule fois, il convient de les construire de façon légère, c'est-à-dire avec un petit nombre de descripteurs par échelle. On peut considérer qu'un nombre raisonnable se situe entre quatre et huit descripteurs.

On peut cependant envisager d'utiliser la même échelle de notation à plusieurs reprises. C'est le cas, par exemple, dans les situations suivantes :

- Une échelle de notation a été créée pour évaluer les mesures à mi-parcours et on veut ensuite s'en servir dans le cadre de la sélection des projets.
- Une échelle de notation a été créée pour évaluer les mesures à mi-parcours et on souhaite la réutiliser dans le cadre de l'évaluation ex post.
- Une échelle de notation a été créée dans le cadre de l'évaluation d'un programme et on souhaite la réutiliser dans le cadre de l'évaluation d'un autre programme.

Encadré 22 – Affinement et harmonisation des échelles de notation

	Beaucoup	Un peu	Pas du tout
Quelle est la proportion de femmes parmi les participants?	3	1,5	0
Dans quelle proportion les femmes participantes ont-elles acquis une qualification dans un métier considéré comme masculin ?	2	1	0
Dans quelle proportion les femmes participantes ont-elles augmenté leur employabilité ?	3	1,5	0
Dans quelle proportion les femmes participantes ont-elles trouvé un emploi non menacé correspondant à leur qualification après un an ?	2	1	0
Note totale pour l'impact relatif à l'égalité des chances	= 5,5		

Si une échelle de notation doit être utilisée plusieurs fois, il est justifié de la construire de façon plus précise. L'échelle de l'Encadré 21 pourra par

exemple passer de quatre à douze descripteurs pour aboutir à une échelle affinée comme dans l'Encadré 22.

Une étape supplémentaire peut être franchie dans la mesure où l'échelle de notation a déjà été appliquée au moins deux ou trois fois. A ce stade, il est possible de la valider, de stabiliser les définitions et de s'en servir dans le cadre d'une enquête par questionnaire pour quantifier des indicateurs harmonisés et permettant des comparaisons entre différentes mesures ou différents programmes, comme cela est montré dans l'encadré ci-dessous.

Encadré 23 – Passage des échelles de notation aux indicateurs

	Mesure évaluée	Moyenne des mesures comparables
Proportion de femmes parmi les participants	72%	45%
Proportion des femmes participantes qui ont acquis une qualification dans un métier considéré comme masculin	5%	23%
Proportion des femmes participantes qui ont augmenté leur employabilité	52%	55%
Proportion des femmes participantes qui ont trouvé un emploi non menacé correspondant à leur qualification après un an	25%	29%

Utiliser les notes pour synthétiser l'évaluation des mesures

■ Vue d'ensemble des impacts d'un programme

Encadré 24 – Exemple de matrice de notation des impacts dans le cas d'un programme de développement urbain

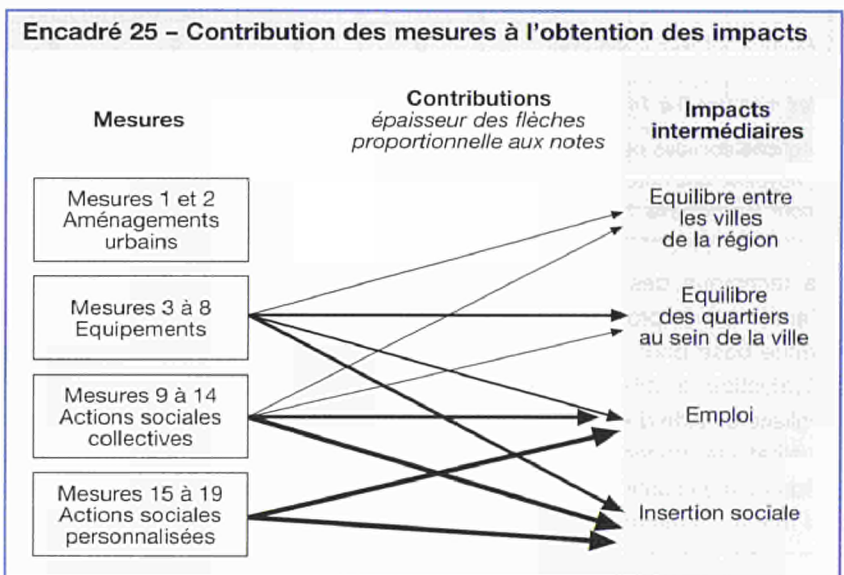
Contribution des mesures à la production des impacts attendus (note entre 0 et 10)	Insertion sociale	Emploi	Equilibre des quartiers au sein de la ville	Equilibre entre les villes de la région
Aménagements urbains (moyenne des notes pour les mesures 1 et 2)	1	1	1	2
Equipements (moyenne des notes pour les mesures 3 à 8)	5	4	4	3
Actions sociales collectives (moyenne des notes pour les mesures 9 à 14)	6	5	3	3
Actions sociales personnalisées (moyenne des notes pour les mesures 15 à 19)	6	6	2	2

La technique des échelles de notation a l'avantage de donner une vue d'ensemble du programme sous une forme très résumée et qui constitue une bonne base pour la rédaction des conclusions synthétiques. Par exemple, l'évaluation à mi-parcours d'un programme de développement urbain, réalisée à l'aide d'études de cas, a abouti à de très nombreuses conclusions qualitatives concernant les impacts. Pour synthétiser les conclusions, l'équipe d'évaluation a construit et utilisé des échelles de notation. L'Encadré 24 montre comment l'ensemble des conclusions qualitatives a été synthétisé sous forme d'une matrice de notation des impacts. Dans une telle matrice, les mesures (ou ici les groupes de mesures), peuvent se comparer colonne par colonne. On dira par exemple que telle mesure réussit mieux qu'une autre en termes d'impact sur l'emploi mais qu'elle réussit moins bien pour un

autre critère. Par contre, cette matrice ne permet pas des additions ou des comparaisons entre colonnes. Une note 5 dans la première colonne ne se compare pas, et peut encore moins s'additionner, avec une note 5 dans la dernière colonne. Pour faire des raisonnements ou des jugements qui combinent plusieurs colonnes, il faut employer la technique d'analyse multicritère (voir chapitre IV de ce Volume).

■ Jugement de la cohérence du programme au vu des impacts obtenus

Dans l'exemple du programme de développement urbain présenté dans l'Encadré 24, la matrice de notation des impacts a également été présentée sous forme de diagramme afin de montrer visuellement les impacts et leur complémentarité. Ce diagramme qui apparaît dans l'Encadré ci-dessous constitue un progrès important par rapport à celui qui peut être réalisé au début de l'évaluation (voir Encadré 7 – Diagramme logique des impacts) en ce sens qu'il montre l'importance des impacts obtenus et non plus la simple logique des impacts attendus.



Ce diagramme a l'avantage de faire apparaître visuellement la cohérence du programme. On voit par exemple que certaines mesures contribuent

très peu à la réussite d'ensemble du programme. On voit de même que certains impacts sont obtenus de façon importante et complémentaire par plusieurs mesures. Au contraire, la faiblesse de certains impacts laisse penser que plusieurs objectifs ne sont pas atteints de façon satisfaisante, alors même qu'ils sont considérés comme importants par plusieurs partenaires du programme.



Le lien entre programmation et évaluation

III

Analyser le programme à l'aide d'une matrice des impacts croisés

Au-delà de la réussite de chaque mesure, l'évaluation doit aussi s'intéresser à l'interaction des mesures entre elles. On peut en effet s'attendre à ce que l'impact d'un ensemble cohérent de mesures soit plus grand que celui des mêmes mesures mises en œuvre séparément. Le troisième outil présenté dans ce Volume apporte une aide à l'évaluation de ces effets de synergie. Il s'agit de la matrice des impacts croisés. Cet outil a été formalisé et testé dans le cadre des travaux du programme MEANS, à partir d'une synthèse entre des démarches proposées par plusieurs équipes européennes.



Une solution pour analyser la synergie au sein du programme

■ Liens entre programmation et synergie

Un programme est composé d'un ensemble de projets qui sont tous orientés vers le même objectif. On peut imaginer un programme constitué d'une juxtaposition de projets indépendants dont les effets s'additionneraient purement et simplement. Dans une telle situation, la programmation serait un exercice purement formel. En effet, l'objectif visé s'atteindrait aussi bien en gérant les projets indépendamment les uns des autres. Si la programmation apporte une valeur ajoutée, c'est bien parce que les projets se renforcent mutuellement, parce qu'ils entrent en synergie. La recherche de la synergie est une des raisons d'être de la programmation.

Encadré 26 - QUBIST, un outil d'analyse ex ante des synergies entre projets

En 1986, la Commission a fait réaliser par la Queen's University of Belfast le développement d'un outil informatisé d'aide à la sélection des projets. Cet outil a été construit dans le contexte de la préparation d'un programme de développement intégré en Irlande du Nord.

L'outil s'applique à tous les projets susceptibles d'être inscrits dans le programme. Pour chaque projet examiné, on commence par expliciter sa contribution probable aux objectifs du programme (une pondération est attribuée aux objectifs en fonction de leur importance). On dénombre ensuite les relations probables entre le projet examiné et les autres projets (synergies).

Les résultats de ces examens sont rassemblés dans une matrice de synergie qui permet de calculer la valeur synergétique de chaque projet (1 point par relation identifiée). On classe ensuite les projets selon le meilleur rapport entre leur coût et leur potentiel synergétique.

La sélection finale des projets prend en compte ce classement établi sur la base du critère de synergie ainsi que la contribution potentielle des projets aux objectifs du programme.

Dès la réforme des Fonds structurels de 1988, les notions de programmation et de synergie sont apparues comme inséparables. L'accent a été mis sur la nécessité de rechercher une meilleure efficacité par une programmation intégrée. « Un programme intégré cherche à obtenir un effet de synergie entre les différentes mesures et à assurer la convergence des effets des

différents partenaires ayant des projets et des responsabilités différents »
(Extrait des règlements des Fonds structurels de 1988).

Les caractéristiques d'un programme intégré sont doubles. Il s'agit d'une part, d'obtenir une meilleure complémentarité des projets et donc d'optimiser l'allocation des ressources. Il s'agit d'autre part, de renforcer mutuellement les interventions cofinancées par les différentes institutions qui agissent pour atteindre le même objectif.

Dès l'apparition du principe de programmation, la Commission a cherché à mettre au point des outils d'évaluation de la synergie. Un premier travail technique a été réalisé en 1986 avec la mise au point d'une technique intitulée QUBIST (voir Encadré 26). Cette technique était destinée à évaluer le potentiel de synergie des projets au stade de leur sélection. Elle est restée peu utilisée en tant que telle.

Pour accompagner le développement de la programmation et de l'évaluation, la Commission a souhaité poursuivre le développement des outils d'analyse de la synergie dans le cadre du programme MEANS. Un séminaire d'experts a été organisé à Bruxelles en 1993 dans le but de confronter les meilleures expériences existantes. Le séminaire a produit un rapport méthodologique et des propositions techniques.

A l'occasion de l'évaluation ex post des programmes d'Objectif 2 de la période 1989-1993, la Commission a décidé d'introduire une question relative à la synergie dans ses termes de références et de remettre les propositions techniques du programme MEANS aux équipes d'évaluation.

Plusieurs équipes d'évaluation ont pu tester, à cette occasion, la faisabilité de ces propositions techniques. Ce test a conduit à des améliorations et aménagements pris en compte dans les pages suivantes.

■ Définition de la synergie et des notions connexes

Le terme « synergie » est souvent employé et fait partie du vocabulaire courant dans beaucoup de domaines, si bien que sa signification est parfois floue. Ainsi une entreprise va s'allier avec une autre dans le but de développer des synergies. En pharmacologie, il y a synergie si deux substances administrées conjointement produisent un effet différent de celui qui résulte de leur injection séparée. C'est cette définition qui est retenue dans cet ouvrage, à savoir qu'il y a synergie lorsque plusieurs interventions publiques (ou plusieurs composantes d'une intervention)

produisent ensemble un impact supérieur à l'addition des impacts qu'elles produiraient isolément (ex : l'intervention finance l'agrandissement d'un aéroport qui facilite le remplissage des équipements touristiques, également financés par l'intervention).

Si rien n'est précisé, la synergie fait référence à des impacts positifs. On peut cependant évoquer des phénomènes de renforcement d'impacts négatifs, de synergie négative ou d'antisynergie (ex : l'intervention subventionne la diversification des entreprises alors qu'une politique régionale favorise le renforcement de l'activité dominante).

Le fait qu'un effet de synergie puisse apparaître au sein d'un programme peut d'abord s'expliquer par la notion de masse critique : une impulsion suffisamment importante est nécessaire pour franchir un certain seuil dans l'impact du programme. En deçà de ce seuil, rien ne se passe, et au delà, les effets s'enchaînent. Cela s'apparente à un effet de boule de neige. L'Encadré qui suit montre que la synergie peut résulter de nombreux autres mécanismes. Dans tous les cas de figure, la synergie est une notion dynamique : elle est le produit d'un processus dans lequel des investissements de différentes natures entrent en interaction sur un territoire donné.

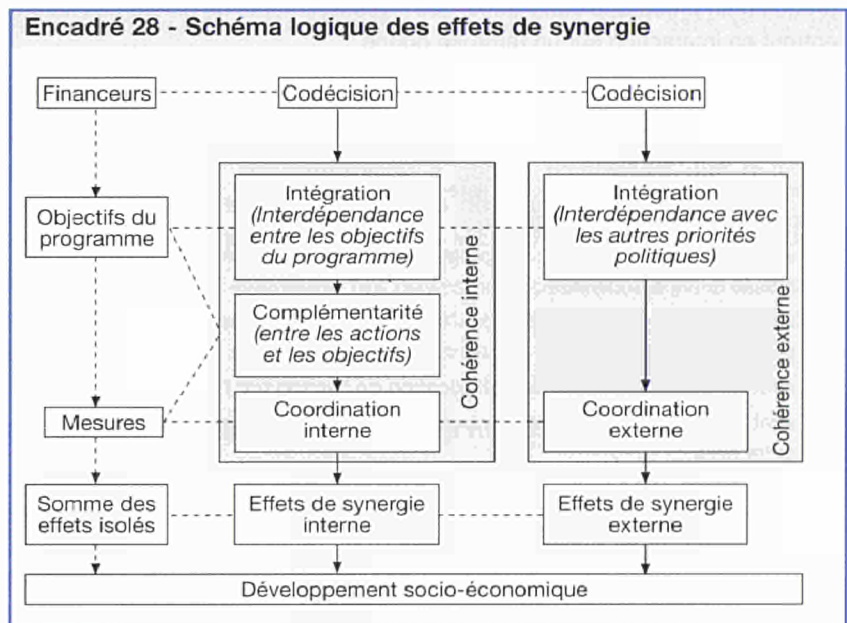
Encadré 27 - Les effets dynamiques de programmation

Sous le titre de *Program Induced Dynamic Effects*, une étude de Cambridge Policy Consultants, propose une liste des mécanismes qui sont susceptibles de produire des effets de synergie :

- effet de masse critique : un projet isolé n'aurait aucun impact faute d'une masse critique suffisante ;
- effet de taille : abaissement des coûts du fait de la présence de plusieurs projets au même endroit ;
- effet de démonstration : démultiplication de l'action par l'exemple ;
- effet d'ajustement : meilleure coordination des différentes actions publiques entre elles ;
- effet de capacité : une action publique isolée n'aurait aucun impact faute d'une masse critique suffisante ;
- effet de productivité : les économies d'échelle sur l'action publique elle-même améliorent sa productivité ;
- effet de déblocage : la réalisation d'un projet débloque une contrainte pesant sur un autre projet ;
- effet d'entraînement : la réalisation d'un projet incite à réaliser un autre projet ;

Dans le cas des programmes socio-économiques, la synergie peut être obtenue par différents leviers : cofinancements, partenariat, articulation de projets complémentaires sur un même territoire, convergence de plusieurs aides vers les mêmes destinataires, circulation de l'information entre opérateurs, etc. Dans tous les cas, la synergie est le produit de mécanismes complexes qui se forment au sein d'un système et qui se concrétisent dans le moyen terme. La synergie peut d'ailleurs être interne ou externe selon qu'elle intervient au sein du système constitué par le programme où qu'elle s'étend également aux effets d'autres interventions publiques.

Le schéma ci-dessous représente les mécanismes de production des effets de synergie. La synergie interne est représentée au centre du schéma, la synergie externe apparaît à droite, tandis que la colonne de gauche représente une situation imaginaire (situation contrefactuelle), où les mesures du programme seraient indépendantes les unes des autres et également indépendantes par rapport aux autres interventions publiques.



La synergie interne est le produit de la cohérence interne qui est constituée de trois éléments :

- L'interdépendance des objectifs (les objectifs sont reliés les uns aux autres) ;
- La complémentarité entre les mesures et les objectifs (les mesures visent les mêmes objectifs) ;
- La coordination entre les mesures (mesures sont choisies et mises en œuvre de façon à ce que leurs effets puissent se rencontrer dans le temps et dans l'espace).

Ces éléments ne sont pas les seuls à être nécessaires pour obtenir des effets de synergie, mais ils sont essentiels pour l'analyse. La synergie interne peut prendre des formes très diverses comme le montrent les exemples suivants :

- Le développement d'un complexe touristique est favorisé par la construction d'une route d'accès au littoral, l'un et l'autre subventionnés par le programme (renforcement mutuel).
- Les aides aux laboratoires universitaires et celles concernant le développement technologique dans les entreprises permettent une mise en réseau des connaissances et favorisent des innovations qui se traduisent par de nouvelles technologies industrielles (combinaison d'effets de masse critique et de boule de neige).
- La combinaison d'aides à différents niveaux d'une filière agricole permet de minimiser les coûts (effet de taille).

A cette notion de la synergie s'oppose le phénomène inverse, qu'on peut qualifier de synergie négative :

- La construction d'un complexe sidérurgique en bord de mer et d'un port de plaisance côte à côte (gêne mutuelle).
- L'attribution d'aides à différents niveaux d'une même filière agricole se faisant concurrence entre elles (effet de dispersion).

La deuxième catégorie de synergie, dite synergie externe, s'applique au renforcement ou à l'affaiblissement mutuel des effets du programme et des effets d'autres interventions. Par exemple, on recherche une synergie entre l'aide apportée aux PME par les programmes des Fonds structurels et les prêts qui leur sont accordés par la Banque Européenne d'Investissement. Un autre exemple moins heureux serait celui d'un programme d'insertion des chômeurs financé par l'Union européenne et qui recruterait une partie de ses participants aux dépens de programmes de formations financés par les autorités nationales. Une telle synergie négative a été dénommée « effet de cannibalisme » par certaines équipes d'évaluation.

L'étude de la synergie externe n'est pas toujours compatible avec les ressources budgétaires de l'évaluation. Dans le cadre d'une évaluation d'ensemble, il est cependant conseillé d'inventorier rapidement les principales politiques européennes, nationales et régionales qui sont voisines du champ d'évaluation afin de se faire une idée des synergies externes potentielles. Cette démarche peut se révéler très pertinente, sans nécessiter un temps d'analyse important.

■ Démarche d'évaluation de la synergie en trois phases

La synergie est rarement évaluée, car elle ne figure pas parmi les objectifs explicites des programmes, bien qu'il s'agisse d'un mécanisme important pour l'atteinte de ces objectifs. L'évaluation de la synergie est souvent limitée, dans le cadre d'une démarche a contrario, au constat d'un manque de coordination ou de gaspillages liés à des superpositions d'interventions. L'évaluation des effets de synergie est également restreinte par la limitation des ressources disponibles. Le temps et le budget de l'évaluation sont souvent mobilisés par d'autres thèmes considérés comme plus prioritaires tels que l'emploi ou l'environnement.

Il convient aussi de noter quelques véritables difficultés techniques. D'abord, les interactions qui favorisent la synergie sont extrêmement nombreuses et ne peuvent pas être toutes examinées. Il faut également prendre en compte l'absence de données fiables permettant de discriminer l'impact de l'ensemble des mesures d'un programme de celui qu'auraient produit ces mêmes mesures si elles avaient été mises en œuvre séparément.

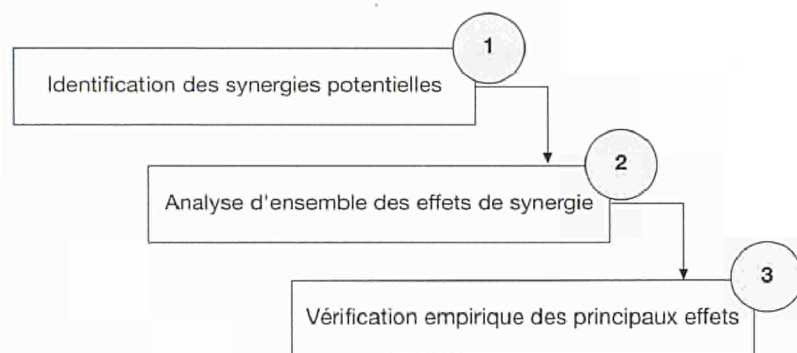
La solution proposée dans cet ouvrage repose sur l'utilisation d'un outil appelé matrice des impacts croisés. Cette solution a été conçue de manière à pouvoir être appliquée sans accroître fortement le volume du travail consacré à l'évaluation d'ensemble d'un programme. Elle se présente plutôt comme un complément d'analyse des informations qui devront inévitablement être collectées par ailleurs.

En testant empiriquement cet outil, il est apparu que son application devait rester réaliste et flexible. C'est pourquoi l'utilisation de la matrice des impacts croisés a été décomposée en trois phases à adapter en fonction du temps et du budget disponibles pour l'évaluation :

- une phase de construction de la matrice qui permet l'identification des synergies potentielles ;

- une phase d'analyse d'ensemble des effets de synergie à l'aide de la matrice ;
- une phase de vérification empirique des principaux effets, sélectionnés sur la base de la matrice.

Encadré 29 - Une démarche pour évaluer la synergie interne



Seules les deux premières phases sont réellement indispensables à l'évaluateur confronté à un agenda serré. Elles offrent de bonnes garanties pour une analyse des synergies au sein du programme considéré. La troisième phase apporte quant à elle des informations en provenance du terrain qui permettent de confirmer ou d'infirmer les hypothèses émises par les gestionnaires lors de la seconde phase. Elle est vivement recommandée dans le cas où la question de la synergie doit être approfondie et/ou lorsque les deux premières phases n'ont pas donné de conclusion crédible.



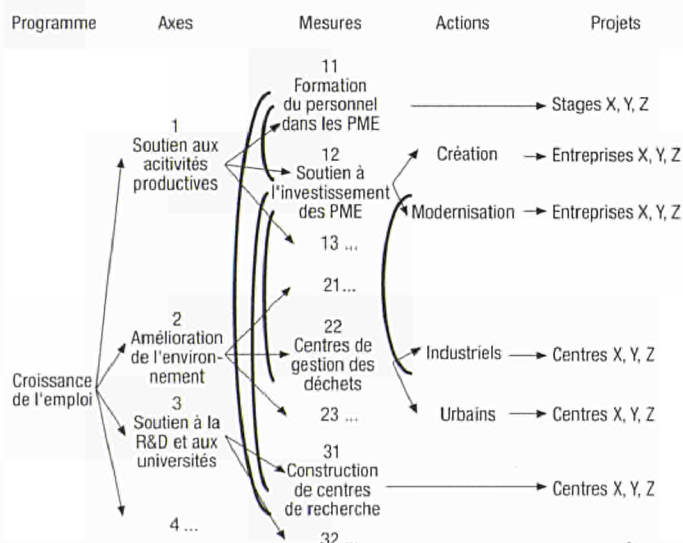
Mode d'emploi de la matrice des impacts croisés

La construction et la validation de la matrice d'impacts croisés s'effectuent en plusieurs étapes successives : élaboration de la structure de la matrice, notation des effets de synergie, validation avec les gestionnaires, calcul des coefficients de synergie potentielle, présentation d'ensemble.

■ Choix du niveau d'analyse de la synergie

Selon la structure du programme concerné, il sera plus pertinent d'analyser les synergies entre les axes, entre les mesures, entre les actions ou entre les projets. Le niveau d'analyse choisi dépend évidemment du nombre des composantes du programme à chaque niveau. Certains programmes ne comportent que quelques projets, ce qui permet d'analyser rapidement la synergie à leur niveau. Si le nombre de projets est très important, on pourra préférer analyser la synergie au niveau des mesures.

Encadré 30 - Choix du niveau d'analyse de la synergie dans un arbre des objectifs



Les traits en sur-épaisseur indiquent les synergies potentielles les plus évidentes.

Le choix du niveau d'analyse peut être fait en revenant à l'arbre des objectifs (voir Encadré 6 – Arbre des objectifs). Sur cette présentation schématique du programme, on peut en effet dessiner les liens de synergie potentielle qui paraissent les plus évidents comme indiqué dans l'Encadré 30.

Dans l'exemple du programme catalan d'Objectif 2, le niveau d'analyse le plus approprié semble être celui des mesures car c'est à ce niveau qu'on a repéré le plus grand nombre de synergies potentielles.

■ Les interactions entre mesures

Une fois le niveau d'analyse choisi, la matrice des impacts croisés est construite avec autant de lignes et de colonnes qu'il existe de composantes du programme au niveau choisi. Les cellules de la matrice permettent de « cocher » les synergies repérées comme indiqué dans l'encadré suivant. La matrice étant considérée comme symétrique à ce stade, on n'en remplira que la moitié.

Encadré 31 - Construction de la matrice des impacts croisés

Axes		I. Création et développement d'activités productives			II. Protection et amélioration de l'environnement			III. Investissement en R&D & formation		IV. Amélioration du réseau des télécommunications		
	mesures	11	12	13	21	22	23	31	32	41	42	43
I. Création et développement d'activités productives	1.1											
	1.2	X										
	1.3		X									
II. Protection et amélioration de l'environnement	2.1	X										
	2.2											
	2.3											
III. Investissement R&D & formation	3.1			X								
	3.2		X	X								
IV. Amélioration du réseau des télécommunications	4.1		X									
	4.2											
	4.3											

Cette approche reste cependant peu nuancée et purement théorique. C'est pourquoi il convient d'aller plus loin en qualifiant de manière plus précise

l'importance des effets de synergie attendus et en validant cette analyse au cours d'entretiens avec les gestionnaires du programme.

Les développements qui suivent constituent le cœur de la méthode : il s'agit du minimum à réaliser pour procéder à une analyse raisonnable des effets de synergie. Cette phase requérant des entretiens, il convient de les prévoir assez tôt dans l'agenda de l'évaluation.

■ La notation des effets de synergie

L'équipe d'évaluation effectue un rapprochement des mesures deux à deux pour repérer des synergies éventuelles. Lorsqu'un effet de synergie lui paraît possible, elle choisit une valeur sur une échelle variant en fonction de l'importance de l'effet :

- 2 pour un effet de synergie particulièrement important
- 1 pour un effet de synergie plus faible
- -1 même échelle appliquée à une synergie négative
- -2 même échelle appliquée à une synergie négative

Encadré 32 - Notation des synergies potentielles

Axes		I. Création et développement d'activités productives			II. Protection et amélioration de l'environnement			III. Investissement en R&D & formation		IV. Amélioration du réseau des télécommunications		
	mesures	11	12	13	21	22	23	31	32	41	42	43
I. Création et développement d'activités productives	1.1											
	1.2	(2)										
	1.3	(2)	(2)									
II. Protection et amélioration de l'environnement	2.1	2	(1)									
	2.2	1	(1)	-1								
	2.3	(2)	(1)	(1)	(2)	(1)						
III. Investissement R&D & formation	3.1	(1)	(1)	(2)			(1)					
	3.2	(2)	(2)	(2)			(1)	(1)				
IV. Amélioration du réseau des télécommunications	4.1	(1)	(2)									
	4.2	(2)	(2)					(1)				
	4.3	(1)										

Si l'équipe d'évaluation anticipe des effets de synergie plus ou moins importants, elle peut aussi estimer que certains effets sont très probables et que d'autres sont très aléatoires. Elle pourra noter les synergies incertaines par des parenthèses.

A ce stade, la matrice d'impacts croisés représente une synthèse quantifiée des hypothèses de synergie identifiées par l'équipe d'évaluation. Bien que la matrice des impacts croisés soit d'allure quantitative, elle ne contient que des appréciations subjectives. Les commentaires et remarques de l'équipe d'évaluation sont donc des éléments essentiels qui méritent d'être conservés.

Il arrive que des synergies ne soient pas réciproques entre des mesures prises deux à deux. Par exemple, une route d'accès au littoral est indispensable au succès d'un complexe touristique, mais en revanche le développement touristique escompté grâce à la construction de la route peut se réaliser en partie sans que le complexe touristique soit créé. On parle dans ce cas de synergie asymétrique, c'est-à-dire qu'un effet positif important existe si la mesure A s'ajoute à la mesure B, alors que l'effet est moindre si la mesure B s'ajoute à la mesure A. On signale cette particularité dans le commentaire qualitatif.

Lors du remplissage de la matrice, les valeurs de chaque côté de la diagonale sont identiques en cas de synergie symétrique. Il suffit de compléter la moitié de la matrice, l'autre partie (en grisé) étant identique. Si l'on souhaite mentionner les cas de synergie asymétrique (relation d'interdépendance non réciproque), il conviendra de les inscrire de l'autre côté de la diagonale.

Exemple : Cas de synergie asymétrique relevée en Catalogne dans un programme de reconversion économique

Entre autres exemples de synergie asymétrique, l'équipe d'évaluation a constaté que « la formation des chômeurs et des travailleurs issus du secteur de l'environnement et de la recherche-développement a un effet fortement positif sur les efforts de développement et de modernisation des PME. Par contre, le soutien au développement et à la modernisation des PME a pour sa part une influence positive comparativement plus faible sur la formation des chômeurs et des travailleurs issus du secteur de l'environnement et de la recherche-développement ».

■ La validation avec les gestionnaires

Une fois la matrice remplie, l'équipe d'évaluation organise des entretiens ou une réunion avec les principaux gestionnaires du programme. Le but est de

discuter et de valider les hypothèses de synergie présentées dans la matrice.

Pour chaque questionnaire de synergie, les explications et les arguments retenus par les gestionnaires sont répertoriés sous forme de commentaires qualitatifs annexés à la matrice avec une indication de la probabilité pour que cette synergie apparaisse (voir Encadré suivant).

Le commentaire détaillé est très important, car en cas de temps limité, il permet de mener une analyse qualitative déjà porteuse de résultats et qui peut être répercutée sur la mise en œuvre du programme concerné.

Encadré 33 - Commentaires sur les synergies potentielles

Liens entre les mesures	Effets des interactions	Avis des gestionnaires	Probabilité
11 & 12	Un personnel mieux formé facilite la décision de moderniser une PME	oui	forte
21 & 23	La formation favorise le choix de méthodes de production non polluantes	non	faible
..	...	...	...

■ Les coefficients de synergie potentielle

Une fois les hypothèses de synergies validées avec les gestionnaires, on peut finaliser la matrice des impacts croisés par le calcul des coefficients de synergie (C_{s+}). Ils sont calculés pour chaque mesure et pour l'ensemble du programme. La formule est la suivante :

$$C_{s+} = \frac{\text{Somme des notes positives}}{(\text{Nombre de notes positives}) \times 2}$$

Si toutes les synergies potentielles entre les mesures recevaient la note maximum, le coefficient serait égal à 1,00. Le coefficient serait égal à 0,00 pour un programme ne comportant aucune synergie positive. S'il existe des

synergies négatives, on calcule également le coefficient de synergie négative (C_{s-}) :

$$C_{s-} = \frac{\text{Somme des notes négatives}}{(\text{Nombre de notes négatives}) \times 2}$$

Pour faciliter le calcul de ces coefficients, il convient d'ajouter six colonnes à la matrice des impacts croisés comme indiqué dans l'Encadré suivant. Les colonnes "s+" et "s-" font respectivement la somme des valeurs de synergie positive et négative, tandis que les colonnes "n+" et "n-" comptent le nombre de notes positives et négatives présentes dans la matrice. Les colonnes Cs+ et Cs- calculent les coefficients synthétiques de synergie positive et négative.

Encadré 34 - Calcul des coefficients synthétiques de synergie

Annexe 04 - Calcul des coefficients synthétiques de synergie																				
		I.			II.			III.			IV.			Somme		Nombre		Coefficient		
		1.1	1.2	1.3	2.1	2.2	2.3	3.1	3.2	4.1	4.2	4.3	s+	s-	n+	n-	Cs+	Cs-		
I.	1.1		2	2	2	1	1		2				10		6			0,83		
	1.2	2		1		1			1	1	1	1	8		7			0,57		
	1.3	2	2				1		2				7		4			0,87		
II.	2.1	2		1		-1	2						5	-1	3	1		0,83 -0,5		
	2.2	1	1		-1		1						3	-1	3	1		0,50 -0,5		
	2.3	2	1	1	2	1			1				8		6			0,66		
III.	3.1	1	1	2			1		1				6		5			0,60		
	3.2	2	2	2			1	1					8		5			0,80		
IV.	4.1	1	2										3		2			0,75		
	4.2	2	2					1					5		3			0,83		
	4.3	1	1										2		2			0,50		
T.													65	-2	46	2			0,7	-0,5

Le calcul des coefficients synthétiques peut être étendu à l'ensemble du programme en faisant la moyenne de tous les coefficients non-nuls (coefficient synthétique de synergie positive = 0,7 et coefficient synthétique de synergie négative = -0,5). On doit cependant noter que l'intérêt de ce calcul est limité au cas où le même travail a été fait pour d'autres programmes par des équipes d'évaluation adoptant les mêmes modes de notation. Ce dernier cas étant peu probable, les coefficients de synergie auront une plus grande utilité en tant que critère pour un jugement comparatif des mesures dans le cadre d'une évaluation d'ensemble du programme.

Faire une vérification empirique des principales synergies identifiées

■ Présentation de l'évaluation d'ensemble de la synergie

Sur la base des calculs de coefficients complétés par les commentaires, l'équipe d'évaluation proposera des recommandations pour améliorer ou déclencher des effets de synergie au sein du programme évalué.

Dans l'exemple ci-dessus, inspiré de l'évaluation du programme d'Objectif 2 en Catalogne, il apparaît donc que les mesures 1.1, 1.3, 3.2, et 4.2 sont essentielles à la réussite du programme compte tenu de leurs effets synergétiques. À l'inverse, la disparition de la mesure 4.3 ne devrait pas affecter significativement l'efficacité des autres interventions. La question se poserait pour 1.2 qui a un coefficient de synergie assez faible (0.57), mais qui entre en synergie avec beaucoup d'autres mesures (7). On ne peut conclure dans ce cas que 1.2 a un faible potentiel de synergie car ses interactions, bien que de faible importance, sont nombreuses. La mesure 2.1 qui combine des effets de synergie positifs et négatifs doit également être examinée de façon nuancée. En fait, il convient toujours de nuancer la lecture des coefficients en fonction du nombre de synergies (colonnes "s" et "n" du tableau).

Pour une approche globale d'un programme, on peut ainsi se servir des coefficients de synergie positive et négative comme deux critères d'évaluation parmi d'autres. Dans l'optique d'une analyse multicritère telle que décrite dans le chapitre suivant, on peut intégrer ces deux colonnes de résultats dans la matrice multicritère. Les notes concernant la synergie côtoient alors d'autres critères relatifs à l'impact sur l'environnement, l'emploi, etc.

■ Identification des synergies à évaluer

Cette phase n'est à réaliser que si les synergies précédemment repérées présentent un intérêt prépondérant pour la continuation ou la réussite du programme. Elle a l'avantage de transformer les hypothèses formulées lors des phases précédentes en une véritable évaluation des effets de synergie, tels qu'ils sont observables dans la réalité. Elle nécessite d'étendre les entretiens aux destinataires du programme (public-cible), ce qui suppose

de pouvoir les identifier et les rencontrer pour obtenir des informations pertinentes. Pour un bon déroulement de cette phase, il faut disposer de ressources suffisantes pour les entretiens et d'un agenda pas trop « serré » avant la rédaction du rapport final de l'évaluation.

La matrice des impacts croisés peut servir également à repérer les interactions les plus critiques entre les mesures prises deux à deux. Une interaction est considérée comme critique si elle fait l'objet d'une hypothèse de forte synergie (positive ou négative) ou si elle fait l'objet d'hypothèses très divergentes entre les personnes interrogées (mentionnées dans les commentaires qualitatifs). Il y a également lieu de considérer comme critique l'interaction entre des mesures dont les gestionnaires se coordonnent particulièrement bien ou mal. Par exemple, si une compétition entre partenaires d'un programme vient contrarier la réussite des mesures pourtant programmées dans le but de créer des synergies.

Lorsqu'une interaction est considérée comme critique, on ouvre une « fenêtre », au format le plus adapté, dans la matrice des impacts croisés. Les hypothèses de synergie seront alors développées et vérifiées de façon fine à l'intérieur de ces quelques fenêtres.

Encadré 35 - Ouverture de « fenêtres » pour la vérification empirique de la synergie

		I.			II.			III.		IV.		
		1.1	1.2	1.3	2.1	2.2	2.3	3.1	3.2	4.1	4.2	4.3
I.	1.1											
	1.2	(2)										
	1.3	(2)	(2)									
II.	2.1	2		(1)								
	2.2	1	(1)		-1							
	2.3	(2)	(1)	(1)	(2)	(1)						
III.	3.1	(1)	(1)	(2)			(1)					
	3.2	(2)	(2)	(2)			(1)	(1)				
IV.	4.1	(1)	(2)									
	4.2	(2)	(2)					(1)		0		
	4.3	(1)	(1)							0	0	

Le commanditaire de l'évaluation est appelé à donner son avis sur le nombre et l'emplacement des fenêtres à ouvrir, puisque cela influence la durée et le coût de l'opération. Dans l'Encadré 35, on a considéré que les synergies à

l'intérieur des axes (en grisé) ne sont pas prioritaires et on a privilégié les zones contenant les valeurs les plus élevées (note 2) et/ou les plus contestées (note entre parenthèses). Cela a conduit à ouvrir trois fenêtres pour la vérification empirique de la synergie.

■ Vérification empirique des effets de synergie

La vérification empirique de la synergie nécessite d'effectuer des entretiens avec les destinataires des mesures concernées. Un destinataire est un individu, une entreprise ou une institution qui a été touchée d'une quelconque manière par le programme, directement ou indirectement.

Les « fenêtres » ouvertes à l'étape précédente peuvent concerner :

- des individus, des entreprises ou des institutions figurant au sein d'un même groupe cible ;
- une série d'individus / entreprises / institutions appartenant à un même territoire ;
- ou encore, une série d'entreprises appartenant à un même réseau (filière de production ou secteur d'activité).

Dans tous les cas, on doit rechercher des destinataires qui ont été touchés ou qui ont pu être touchés par les mesures étudiées.

L'équipe d'évaluation tente par cette opération de confronter les hypothèses de synergie, faites avec l'aide des questionnaires à ce qui peut être observé empiriquement. En fonction du nombre de destinataires, l'équipe d'évaluation réalise des entretiens individuels ou un entretien de groupe, voire une succession d'entretiens de groupe lorsque cela est possible. On se reportera au Volume 3 de cette collection consacré aux outils d'évaluation pour en savoir plus sur les entretiens de groupe.

Le contenu des entretiens passera en revue les hypothèses faites lors de la construction de la matrice des impacts croisés. L'intérêt des entretiens est d'ordre qualitatif ; il faut donc essentiellement relever l'argumentation des bénéficiaires.

De façon exceptionnelle, on peut tenter une analyse approfondie de la synergie entre les mesures concernées par une « fenêtre ». Cela suppose de multiplier les entretiens de façon à recouper les informations ou de réaliser une série d'études de cas. Cette opération permet de comprendre en profondeur la nature et les mécanismes de tel ou tel effet de synergie.

Elle est onéreuse et doit être prévue très tôt dans le déroulement de l'évaluation. Elle peut être particulièrement utile si les conclusions obtenues sont généralisables à d'autres programmes (validité externe).

■ Consolidation de l'analyse et recommandations

L'équipe d'évaluation réalisera une comparaison systématique entre les estimations des questionnaires et ses observations de terrain. Cette comparaison sera restituée aux gestionnaires des mesures et aux responsables du programme sous forme qualitative.

Il s'agit ici d'expliquer les divergences d'interprétation, mais aussi de valider les différents types de synergies qui apparaissent dans le programme et d'en connaître la nature. A cet égard, il est important pour les gestionnaires d'obtenir cette information le plus tôt possible afin qu'ils puissent procéder aux aménagements éventuels concernant la mise en œuvre.

La matrice des impacts croisés est corrigée pour tenir compte des vérifications empiriques si cela est jugé utile. On peut ensuite réitérer le calcul des coefficients synthétiques de synergie mesure par mesure.

L'encadré qui suit présente, domaine par domaine les effets de synergie qui ont été relevés le plus fréquemment par les équipes d'évaluation des programmes financés par les Fonds structurels dans les zones de reconversion industrielle.

Encadré 36 - Principaux effets de synergie constatés dans les évaluations

Domaine	Aide aux entreprises	Tourisme	Infrastructure	Formation	Technologie
Aide aux entreprises	-				
Tourisme	●	-			
Infrastructure	●	○	-		
Formation	●	●	○	-	
Technologie	●	○	●	●	-
Légende : Effets de synergie ● Forts ● Moyens ○ Faibles					

Les évaluations qui ont abordé la question de la synergie, surtout au sein des programmes d'Objectif 2, ont révélé des expériences intéressantes de

coordination décentralisée. Par exemple, dans les régions de Twente et Groningen-Drenthe (NL), on a utilisé des groupes de travail pour promouvoir l'intégration entre les mesures de différents axes. Dans la région de Bremen (R.F.A.), un « Fonds de Cohérence » a été adopté comme instrument pour inciter les opérateurs à se coordonner. Dans la région Nord-Pas-de-Calais (F), une « Structure Centrale d'Observation » agit dans la même optique.



Une solution pour juger la réussite d'ensemble d'un programme

Tenir compte de tous les critères pertinents pour favoriser le bon impact

IV Appliquer l'analyse multicritère dans un cadre de partenariat

L'analyse multicritère permet de fonder des conclusions de synthèse en tenant compte, comme son nom l'indique, de tous les critères qui ont été considérés comme pertinents par l'équipe d'évaluation. Cette capacité est particulièrement utile dans un cadre d'évaluation en partenariat car elle permet de respecter les attentes de chaque décideur. L'analyse multicritère est traditionnellement employée en tant qu'outil d'aide à la décision. Il existe de multiples variantes techniques de cet outil, dont les principales sont décrites dans ce chapitre. En faisant un choix parmi ces variantes et en utilisant l'analyse comme outil d'aide au jugement, ce Volume propose une utilisation nouvelle de la technique. Il s'agit de fonder des conclusions classant les mesures d'un programme en fonction de leur réussite, et selon une procédure acceptable pour tous les partenaires. Pour mettre en œuvre cette technique, l'équipe d'évaluation doit disposer de conclusions complètes sur toutes les mesures et sur leurs principaux impacts, si possible résumées à l'aide d'échelles de notation.



Une solution pour juger la réussite d'ensemble d'un programme

■ Tenir compte de tous les critères pertinents pour fonder le jugement

Au moment de conclure son travail, l'équipe d'évaluation doit se prononcer sur les mérites respectifs des différentes mesures, et en particulier sur leur contribution à la réussite globale du programme. Si l'évaluation ne parvient pas jusqu'à ce stade, alors elle ne peut pas déboucher sur des recommandations d'un niveau suffisamment stratégique en termes d'amélioration ou de réorganisation du programme. Pourtant, l'expérience montre que les équipes d'évaluation hésitent à formuler de telles conclusions, notamment parce que cela implique un jugement sur la valeur relative des résultats et des impacts obtenus.

Exemple : Conclusions comportant un jugement mal justifié

Parmi les conclusions d'un rapport d'évaluation, on peut lire les deux jugements suivants :

- « la mesure A est hautement appréciée par les professionnels »
- « la mesure B a un coût par emploi créé très élevé »

Dans cet exemple, l'équipe d'évaluation termine son rapport de façon conclusive, ce qui témoigne d'une qualité qui ne se rencontre pas en toutes circonstances. De plus, ses jugements sont étayés par des éléments factuels, ce qui est également un point positif. Cependant, on peut regretter que ces conclusions ne soient pas mieux justifiées. Par exemple, en ce qui concerne la mesure A, le rapport n'explique pas pourquoi on se contente du point de vue des professionnels. En ce qui concerne la mesure B, on aimerait savoir pourquoi, et par rapport à quelle norme, on peut conclure que le coût par emploi créé est très élevé. On aimerait également savoir si la mesure a d'autres effets positifs susceptibles de contrebalancer son coût élevé par emploi créé.

Si l'équipe d'évaluation ou le groupe de pilotage sont réticents à comparer les mérites des différentes mesures, c'est souvent par crainte de formuler un tel jugement de valeur, et surtout par crainte que ce jugement soit considéré comme trop subjectif. Pourtant, il est normal qu'une évaluation parvienne à formuler des jugements de valeur, le mot évaluation n'est-il pas lui-même apparenté au mot valeur ? Dans cette perspective, l'équipe d'évaluation doit s'efforcer de fonder son jugement sur des critères

explicites et acceptés. Comme le montre l'exemple qui précède, des progrès restent à réaliser pour une meilleure justification des conclusions.

Si l'équipe d'évaluation n'explique pas suffisamment ses jugements, cela fragilise l'ensemble du rapport. Toute la crédibilité de l'évaluation repose alors sur la confiance que l'on porte à l'impartialité de l'équipe d'évaluation, laquelle peut toujours être mise en doute (au besoin avec un peu de mauvaise foi). Pour éviter ce danger, tout en ne renonçant pas à produire de véritables conclusions, l'équipe d'évaluation peut s'aider d'un certain nombre de techniques qui permettent d'explicitier ses jugements et qui la mettent à l'abri des principales critiques. L'Encadré 37 présente quatre de ces techniques et leurs principales caractéristiques.

Encadré 37 – Techniques d'aide au jugement

	Analyse coût-efficacité	Enquête de satisfaction	Analyse coût-avantages	Analyse multicritère
Nombre de critères d'efficacité	Un seul	Nombreux	Nombreux	Nombreux
Qui décide de l'importance respective des critères ?	(pas nécessaire)	Les destinataires	L'équipe d'évaluation, avec le point de vue des publics gagnants ou perdants	Le groupe de pilotage de l'évaluation
Les critères sont-ils explicités ?	Oui	Non	Pas totalement	Oui

La technique la plus simple consiste à décider que le jugement fera référence à un impact prioritaire, généralement lié à l'objectif global du programme. On fait alors intervenir deux critères : l'obtention de l'impact prioritaire et les ressources mobilisées. L'évaluation prend alors la forme d'une évaluation coût-efficacité ou *value for money*. Par exemple, elle compare les différentes mesures du programme en fonction du coût par emploi créé. Dans le cadre des programmes socio-économiques de l'Union européenne, il est difficile de se contenter d'une technique aussi simplificatrice. En réalité, il est difficile d'estimer la contribution de toutes les mesures à l'atteinte d'un objectif global (ex : développement socio-

économique) car celui ci est à la fois trop vague et trop lointain (voir page 22). On est plus souvent dans une situation où les mérites d'une mesure se jugent par rapport à plusieurs critères (ex : compétitivité, diversification, emploi, environnement, etc.). L'équipe d'évaluation doit être attentive à ne pas oublier un critère pertinent sous peine de rendre les conclusions inacceptables pour l'un ou l'autre des lecteurs du rapport. Il faut donc avoir recours à des techniques qui permettent de juger en tenant compte de plusieurs critères.

Une première forme de jugement multicritère, souvent choisie implicitement, consiste à faire juger la mesure par ses destinataires. Lorsqu'une équipe d'évaluation conclut, après enquête auprès des entreprises aidées, que « la mesure est hautement appréciée par les professionnels », cela revient à dire que les professionnels interrogés ont identifié tous ses avantages et inconvénients et les ont synthétisés au moment de répondre à l'enquêteur. Après extrapolation à partir de l'échantillon enquêté, on obtient bien un jugement multicritère mais son défaut est d'être totalement opaque. Il est impossible de retrouver la trace de ce qui a compté dans les jugements individuels. Son autre défaut est de s'appuyer seulement sur les destinataires directs, sachant que leur jugement peut être biaisé par complaisance à l'égard de l'administration qui leur a apporté une aide.

Pour bien structurer les jugements en prenant en compte tous ses impacts, les manuels d'évaluation proposent également la technique de l'analyse coût-avantages. Selon cette technique, chaque impact est identifié et estimé de façon quantitative par l'équipe d'évaluation, puis il est pondéré à l'aide d'une valeur monétaire (appelée dans certains cas prix fictif). Cette valeur représente ce que la société est prête à payer ou plutôt ce que les personnes concernées sont prêtes à payer, pour obtenir l'impact. On valorise également les impacts négatifs par la disposition à payer pour les éviter.

L'analyse coût-avantages est une méthode qui aboutit en principe à un jugement tranché en forme de « oui ou non » sans avoir besoin de trouver une norme de comparaison. Le jugement est positif si les avantages actualisés pour les publics gagnants sont plus grands que les inconvénients actualisés pour les publics perdants. En pratique, cette simplicité est rarement mise à profit parce que trop de projets ou d'interventions réussissent le test de l'analyse coût-avantages et que leur nombre dépasse les possibilités de financement public. Un bilan coût-avantages positif n'est donc pas suffisant pour porter un jugement définitif sur une intervention isolée.

L'emploi de l'analyse coût-avantages est souvent délicat parce que de nombreux impacts sont difficiles à monétariser de façon incontestable (impacts sur l'environnement notamment). De plus, la technique est extrêmement exigeante en termes de quantification des impacts, ce qui la rend difficile à appliquer à un vaste programme de développement socio-économique. En pratique, son emploi n'est véritablement indiqué que pour certaines évaluations ex ante de grands projets.

Au contraire des trois approches précédentes (coût-efficacité, enquête de satisfaction et coût-avantages), la technique de l'analyse multicritère est une approche qui fait expliciter les critères de jugement par le groupe de pilotage de l'évaluation, ce qui peut constituer un avantage dans un contexte de partenariat. La technique n'est pas très exigeante en termes de quantification des impacts. Elle prend facilement en compte des conclusions qualitatives, pour peu qu'elles aient été structurées à l'aide d'échelles de notation.

Encadré 38 – Un atelier européen sur les techniques d'évaluations adaptées au contexte du partenariat

Les évaluations des programmes socio-économiques de l'Union européenne sont organisées et réalisées dans un contexte de partenariat entre la Commission et les autorités nationales et/ou régionales. Dans le cadre du programme MEANS, un atelier a réuni à plusieurs reprises une vingtaine d'experts et de praticiens européens dans le but d'identifier les difficultés de l'évaluation partenariale et de faire des propositions techniques pour les traiter.

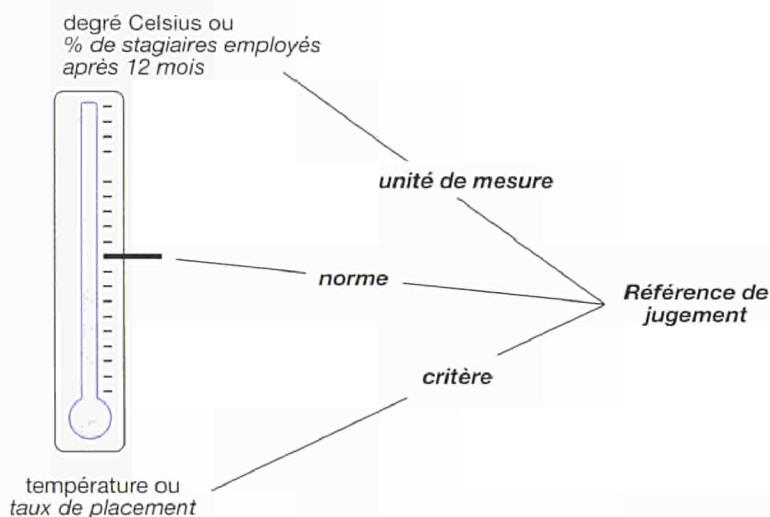
Les participants ont reconnu que des divergences d'intérêts apparaissent fréquemment lorsque l'on procède en commun à la clarification des objectifs, à la construction d'indicateurs ad hoc et à l'établissement de références de comparaison. De nombreuses techniques ont été évoquées au cours des trois réunions de l'atelier : diagramme logique des impacts, échelles de notation, analyse multicritère, SWOT, Metaplan®, vote coloré. Ces techniques sont généralement simples, au point que certaines ont été accusées d'être enfantines. Leur point fort est cependant de structurer la réflexion des partenaires et de gérer leurs désaccords sans mettre en danger leurs relations de travail.

En effet, l'évaluation partenariale n'échappe pas aux jeux de pouvoir et au fait que les participants ont tendance à parler le langage officiel de leur institution. Si le groupe de pilotage est élargi aux principaux partenaires du programme, on constate souvent qu'il est victime d'une sorte de « rituel de courtoisie » qui conduit à gommer les différences de point de vue, et à censurer certaines questions sensibles. Des techniques telles que l'analyse multicritère permettent de faire exprimer des désaccords au sein du groupe de pilotage, ce qui évite les phénomènes d'autocensure et augmente la pertinence des conclusions. Par contre, la technique canalise les discussions au sein du groupe de pilotage en évitant les risques de conflit.

■ Critères d'évaluation : définition et exemples

Pour bien justifier ses conclusions, l'équipe d'évaluation explicite ses références de jugement, c'est-à-dire qu'elle décrit ses principaux critères, chacun d'entre eux étant éventuellement associé à une unité de mesure et à une norme à atteindre, comme le symbolise l'exemple du thermomètre présenté ci-dessous.

Encadré 39 – Référence de jugement, norme et critère



Les principaux critères d'évaluation sont la pertinence, l'efficacité, l'efficience et l'utilité. La pertinence juge si les objectifs correspondent aux véritables besoins de l'économie et de la société. L'efficience juge si les résultats et les impacts ont été obtenus à un coût raisonnable, ou s'ils avaient pu être obtenus à moindre coût. L'efficacité juge si les résultats et les impacts obtenus permettent d'atteindre les objectifs explicites. Dans les pages qui suivent, on étendra la notion d'efficacité à l'atteinte des objectifs implicites, ce qui donne au terme une signification proche de celle du quatrième critère, l'utilité, qui juge l'ensemble des résultats et des impacts obtenus, attendus ou imprévus, au regard des besoins socio-économiques.

Dans les pages qui suivent, l'analyse multicritère est employée en privilégiant l'efficacité entendue au sens large, c'est-à-dire l'obtention des

impacts attendus. L'analyse est dite « multicritère » parce que les partenaires attendent plusieurs impacts, chacun d'entre eux étant considéré comme un critère de réussite pour l'analyse du programme.

Mode d'emploi de l'analyse multicritère

La problématique de choix multicritère est très ancienne. Elle a fait l'objet de développements techniques importants depuis les années 1960 dans le cadre de recherches sur l'aide à la décision initiées par des mathématiciens, spécialistes de la recherche opérationnelle. Son développement a longtemps été porté par deux « écoles » séparées, l'une ayant son centre en France, l'autre aux Etats-Unis. Les deux courants ont maintenant tendance à se rejoindre.

L'analyse multicritère permet d'éclairer une décision complexe en aidant à comparer toutes les solutions envisagées et en tenant compte de tous les points de vue jugés pertinents. En tant qu'outil d'aide à la décision, l'analyse multicritère est couramment utilisée dans le secteur privé, par exemple lorsqu'une entreprise choisit un site pour sa localisation ou dans le secteur public, par exemple lorsqu'une collectivité choisit un mode de stockage de ses déchets spéciaux.

Dans le domaine de l'évaluation des interventions publiques, l'analyse multicritère a été peu utilisée, si ce n'est dans l'évaluation ex ante des projets d'infrastructure de transport. Les pages qui suivent proposent une utilisation de la technique, non plus dans un contexte d'aide à la décision concernant un projet, mais sous la forme d'une aide au jugement concernant un programme. Il s'agit de juger la réussite des différentes mesures d'un programme de développement socio-économique et d'aboutir à des conclusions synthétiques, en prenant en compte les principaux critères pertinents pour le groupe de pilotage.

Dans une telle situation, l'emploi de l'analyse multicritère constitue une originalité. Le mode d'emploi qui suit doit cependant être considéré comme opérationnel dans la mesure où il prend en compte les résultats de plusieurs tests en vraie grandeur. Le déroulement de la procédure est décrit en cinq étapes, plusieurs d'entre elles pouvant faire l'objet de variantes.

■ Structuration de l'analyse

● Choix des objets à comparer

En principe, la procédure qui est décrite ci-dessous va permettre de comparer entre elles les mesures d'un programme. On doit cependant

s'interroger sur l'intérêt éventuel de travailler à un autre niveau en décomposant certaines mesures (niveau souvent appelé « action ») ou au contraire, en regroupant des mesures similaires. Le choix ne doit pas conduire à comparer un trop petit nombre d'objets (deux ou trois groupes de mesures par exemple) car la formalisation multicritère ne se justifierait pas dans ce cas. Le nombre d'objets à comparer peut au contraire être aussi grand que l'on veut (50 mesures par exemple, ou plus) car l'analyse multicritère ne comporte pas de limite dans ce sens.

On peut être conduit à limiter le nombre de mesures analysées, non pas pour des contraintes liées à la technique d'aide au jugement, mais pour des raisons de collecte de données. Telle qu'elle est décrite ci-dessous, l'analyse s'appuie en effet sur une matrice de notation des impacts, matrice dont la construction peut être simplifiée si on choisit d'analyser seulement les mesures les plus sensibles. La décision à prendre est de nature coût / utilité. S'il s'agit d'éclairer une décision de réallocation budgétaire ou une nouvelle programmation, la logique voudrait que l'on s'intéresse en priorité aux mesures les plus importantes en terme de budget ou aux mesures plus innovantes ou à celles dont l'efficacité est mise en doute.

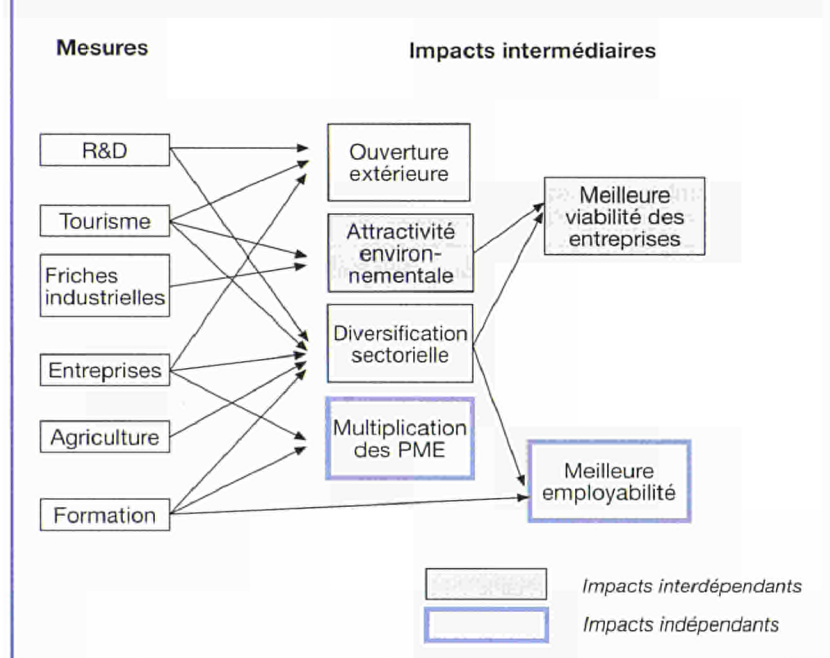
● **Choix des critères**

Dans le cas des programmes socio-économiques de l'Union européenne, la réussite d'une mesure se juge normalement par sa contribution à l'atteinte des objectifs intermédiaires affichés dans les documents de programmation. Les grandes priorités politiques de l'Union européenne (ex : environnement, égalité des chances) font également partie des critères de jugement. Une variante consiste à s'appuyer plutôt sur les objectifs implicites du programme, tels qu'ils ont été reconstitués par le groupe de pilotage ou par des groupes de travail élargis, par exemple avec l'aide de la cartographie conceptuelle des impacts (chapitre I).

C'est cette seconde option qui est présentée ci-après. A l'aide d'un diagramme logique des impacts (voir Encadré 7 – Diagramme logique des impacts), et en accord avec le groupe de pilotage, l'équipe d'évaluation définit la liste des critères, chaque critère étant dérivé d'un impact intermédiaire. A ce stade, l'équipe d'évaluation doit veiller à ce que les critères choisis soient logiquement indépendants les uns des autres. Ainsi, dans l'Encadré 40, l'impact intitulé « meilleure viabilité des entreprises » est logiquement relié à l'impact intitulé « attractivité environnementale »

qui le précède. L'analyse multicritère devra choisir de juger les mesures selon l'un ou l'autre de ces critères, mais pas sur les deux. Par contre, le diagramme montre que l'attractivité environnementale et l'amélioration de l'employabilité sont considérées comme des impacts indépendants. Tous deux pourront donc être choisis comme critères de jugement.

Encadré 40 – Indépendance et interdépendance des critères d'évaluation



Pour faciliter la formulation des conclusions et leur compréhension, on évitera également de mélanger des critères de nature trop différente. Si l'évaluation porte principalement sur l'efficacité du programme, l'analyse multicritère s'intéressera aux résultats et impacts obtenus par les différentes mesures (c'est l'option prise dans ce chapitre). Si l'évaluation porte principalement sur la pertinence du programme, on peut également réaliser une analyse multicritère, mais elle s'intéressera alors aux principales forces et faiblesses de l'économie régionale et à la façon dont les différentes mesures s'appuient sur les forces ou compensent les faiblesses.

La synergie entre les impacts des différentes mesures constitue un cas particulier. Il est logique d'en faire un critère de jugement à part entière si l'on pense que certaines mesures sont indispensables pour permettre la réussite d'autres mesures. Si l'étude des effets de synergie a été réalisée dans l'esprit du chapitre III de ce Volume, il est possible d'utiliser une matrice des impacts croisés et en particulier les coefficients de synergie pour prendre en compte ce critère dans la formulation du jugement de synthèse sur les mesures.

Contrairement au nombre de mesures à comparer, qui peut être potentiellement très grand, le nombre de critères doit être, quant à lui, inférieur à une limite raisonnable. L'expérience a montré que les participants d'un groupe de pilotage ou d'un groupe de travail ne peuvent pas s'exprimer valablement sur un nombre de critères supérieur à huit. C'est ce nombre qui a été retenu dans l'exemple suivant.

Exemple : Critères retenus pour l'évaluation d'un programme socio-économique en Belgique

Dans le cadre d'une évaluation pilote, une analyse multicritère a été réalisée en impliquant un groupe de pilotage élargi à une dizaine de personnes. Huit critères ont été choisis et le groupe de pilotage leur a donné les dénominations et les significations suivantes :

- Diversification sectorielle (la mesure réoriente l'activité économique de la région vers des marchés porteurs).
- Renforcement des PME (La mesure renforce la part des petites ou moyennes entreprises dans l'activité économique de la région).
- Ouverture extérieure (La mesure renforce les exportations hors de la région).
- Renforcement des services aux entreprises (La mesure permet à plus d'entreprises de se fournir en services dans la région).
- Meilleure insertion environnementale (La mesure réduit le nombre d'entreprises régionales dont le développement ou la survie est contrarié par des problèmes d'environnement).
- Meilleure répartition territoriale (La mesure réoriente l'activité économique de la région vers les zones les plus défavorisées).
- Meilleure employabilité (La mesure améliore l'employabilité des personnes de la région qui travaillent ou cherchent un emploi).
- Synergie (La mesure contribue à la réussite d'autres mesures du programme).

● Matrice d'évaluation multicritère

Une fois choisis les critères de jugement et les mesures à évaluer (ou les groupes de mesures, ou parties de mesures), l'équipe d'évaluation construit une matrice d'évaluation multicritère. Cette matrice n'est rien d'autre qu'un tableau comportant autant de colonnes que de critères et autant de lignes que de mesures à comparer. Chaque cellule représente l'évaluation d'une mesure pour un critère. L'analyse multicritère impose de faire une évaluation de toutes les mesures pour tous les critères (aucune cellule ne doit être vide), mais elle n'impose pas que toutes les évaluations aient la même forme. Comme le montre l'Encadré 41, la technique peut prendre en compte un mélange de critères quantitatifs exprimés par des indicateurs, de critères qualitatifs exprimés par des descripteurs, et de critères intermédiaires exprimés part des notes.

Encadré 41 – Combinaison de critères qualitatifs et quantitatifs

Critère Modalités d'évaluation du critère	Diversification (% des entreprises aidées intervenant dans des secteurs en forte croissance)	Employabilité (descripteurs d'impact)	Environnement (note d'impact entre 0 et 10)
Mesures			
Aide à l'investissement	72%	<i>Impact neutre</i>	3
Formation en entreprise	21%	<i>Augmentation significative de l'employabilité pour les stagiaires déjà qualifiés, impact neutre pour les autres</i>	1
Friches industrielles	52%	<i>Impact neutre</i>	8

La façon la plus pragmatique de remplir la matrice d'évaluation multicritère est de structurer toutes les conclusions de l'évaluation, qu'elles soient quantitatives ou qualitatives, à l'aide d'échelles de notation construites à cet effet par l'équipe d'évaluation (voir page 63). La matrice d'évaluation multicritère est alors équivalente à la matrice de notation des impacts décrite dans l'Encadré 19.

Avant de poursuivre l'analyse multicritère, l'équipe d'évaluation doit vérifier si ses premières conclusions vont permettre de comparer les mesures entre elles de façon satisfaisante. En choisissant ses critères, l'équipe d'évaluation a déjà veillé à ce qu'ils s'appliquent au plus grand nombre possible de mesures. Cependant, il faut encore que la majorité des mesures aient produit des impacts au regard de la majorité des critères. En d'autres termes, il faut que la matrice de notation des impacts ne comporte pas trop d'impacts neutres, nuls ou insignifiants. A la limite, si les seules notes non nulles sont situées dans la diagonale (comme dans l'Encadré 42), cela veut dire que les mesures évaluées n'ont aucun point en commun et que leurs critères d'évaluation sont spécifiques. L'analyse multicritère doit alors être abandonnée.

Encadré 42 – Cas de critères d'évaluation trop spécifiques

Critère	Diversification	Employabilité	Environnement
Modalités d'évaluation du critère	(note d'impact) entre 0 et 10)	(note d'impact) entre 0 et 10)	(note d'impact) entre 0 et 10)
Mesures			
Aide à l'investissement	7	0	0
Formation en entreprise	0	5	0
Friches industrielles	0	0	8

■ Choix de la procédure d'analyse

Deux possibilités principales s'offrent à l'équipe d'évaluation pour comparer les mérites des différentes mesures : l'analyse multicritère par compensation ou par surclassement. La variante la plus connue est la première, qui consiste à attribuer une pondération à chaque critère, puis à calculer une note globale pour chaque mesure en faisant la moyenne arithmétique pondérée des notes attribuées à cette mesure pour les différents critères. Cette variante s'appelle « compensatoire » parce que le calcul de la moyenne pondérée permet une compensation entre critères. Par exemple, une mesure qui a un très mauvais impact sur l'environnement pourra malgré tout obtenir une bonne note globale pondérée si son impact sur l'employabilité a été jugé excellent.

Dans la variante du surclassement, on ne calcule pas une note globale car on considère que les critères ne sont pas tous commensurables. Dans ce

cas, l'analyse repose sur des comparaisons multiples du type : « est-ce que la mesure A surclasse la mesure B du point de vue du critère d'environnement ? », « est-ce que la mesure A surclasse la mesure B du point de vue du critère d'employabilité ? », etc. On peut répondre à ces questions, par oui ou par non (critère franc) ou nuancée, auquel cas on introduit la notion de préférence faible et de critère à seuil. L'analyse fait toutes les comparaisons possibles et présente une synthèse du type : « la mesure A est au moins aussi bonne que la mesure B relativement à une majorité de critères (cas de concordance), sans être nettement trop mauvaise relativement aux autres critères (cas de discordance) ».

Ce type d'analyse intègre souvent une protection pour éviter de juger favorablement une mesure qui serait catastrophique du point de vue d'un critère donné. A cet effet, on détermine un seuil de veto pour chaque critère. Tout comme on peut établir des préférences nuancées entre critères, il est possible de nuancer le seuil de veto en introduisant un veto progressif. L'introduction d'un seuil de veto différencie fortement la logique du surclassement et celle de la compensation. S'il existe un seuil de veto, un très mauvais impact sur l'environnement interdira de considérer la mesure comme bonne, même si son impact sur l'employabilité a été jugé excellent par ailleurs.

Le surclassement a l'avantage de mieux refléter la nature des relations entre institutions publiques. En effet, quand on dit que « la mesure A surclasse la mesure B pour la majorité des critères et elle n'est pas exagérément surclassée pour aucun des critères », on adopte un mode de raisonnement très familier pour un responsable politique. Chacun sait en effet qu'une décision publique est plus facile à prendre si la majorité des protagonistes sont d'accord et s'il n'y a pas un groupe de pression violemment opposé. On peut d'ailleurs constater qu'il s'établit souvent une correspondance entre les critères et les protagonistes de l'évaluation. Dans le cas où le groupe de pilotage est élargi à une dizaine de partenaires, il n'est pas rare qu'un participant s'identifie fortement au critère « environnement » ou un autre au critère « emploi ». Dans cette situation la variante du surclassement est probablement de nature à mieux refléter le processus collectif de formulation du jugement au sein du groupe de pilotage.

D'un point de vue technique, la variante de compensation est très simple à mettre en oeuvre avec n'importe quel logiciel tableur. La variante de surclassement nécessite au contraire l'emploi de logiciels spécialisés qui

diffèrent selon les types d'analyse, dont les plus connus sont présentés dans l'Encadré ci-dessous :

Encadré 43 – Variantes d'analyse multicritère utilisant le surclassement

Les principales variantes de l'analyse multicritère utilisant le surclassement sont les suivantes :

ELECTRE I – Cette variante fonctionne avec un indice de concordance et un indice de discordance, présentés sous forme de notes. Un seuil de discordance (veto) est introduit pour tous les critères. Les surclassements et les seuils de veto sont francs. Le logiciel traite une situation où il s'agit de choisir la ou les meilleures mesures par exemple une situation où l'on cherche à identifier les meilleures pratiques.

ELECTRE TRI – Cette variante permet de trier les mesures en différentes catégories, par exemple les mesures les plus réussies, les mesures qui n'ont pas d'impact significatif et les mesures intermédiaires.

ELECTRE II produit un classement entre les mesures depuis la plus réussie jusqu'à la moins réussie. Les surclassements et les seuils de veto sont francs.

ELECTRE III réalise également un classement mais introduit des relations de surclassement floues.

PROMETHEE, utilise seulement un indice de concordance et introduit des surclassements progressifs.

Pour en savoir plus, voir annexe bibliographique : Vincke 1989

Le surclassement n'aboutit pas toujours à des conclusions tranchées. Par exemple, si l'analyse produit un tri, elle montrera que certaines bonnes mesures surclassent nettement les autres, que d'autres mesures plutôt mauvaises sont fréquemment surclassées et qu'un troisième groupe de mesures sont inclassables.

Au contraire, si l'analyse repose sur une compensation, elle est toujours conclusive. Dans la suite de ce chapitre, on a opté pour la variante utilisant la compensation, à la fois pour sa simplicité technique et pour sa capacité à conclure de façon tranchée. En effet, dans une logique d'évaluation d'ensemble du programme et dans une perspective de réallocation budgétaire, on considère que les lecteurs du rapport d'évaluation ont besoin d'obtenir des conclusions sur toutes les mesures évaluées. La formule de la note globale pondérée est adaptée à cette situation, mais elle correspond à une application simplificatrice de l'analyse multicritère. Pour une application concrète, il est prudent de vérifier avec chaque membre du groupe de pilotage qu'aucun problème majeur ne justifierait un système de veto.

■ Choix des assesseurs

L'analyse multicritère constitue une aide au jugement pour l'équipe d'évaluation. Elle lui permet d'explicitier ses jugements et de rédiger ses conclusions finales en se mettant à l'abri des principales critiques concernant sa subjectivité. En réalité, la technique produit un jugement fondé sur des éléments factuels et objectifs concernant les impacts, mais aussi sur la prise en compte impartiale des points de vue et les préférences des principaux partenaires.

L'analyse multicritère est une technique qui ne peut pas être mise en œuvre sans la participation effective des partenaires du programme. Lorsqu'elle utilise cette technique, l'équipe d'évaluation doit avoir des entretiens individuels ou des entretiens de groupe, avec un certain nombre de personnes porteuses des points de vue considérés comme les plus pertinents pour juger les mesures. Ces personnes seront désignées ci-dessous sous le nom d'assesseurs.

Une première formule consiste à considérer comme assesseurs les membres du groupe de pilotage de l'évaluation. Il faut pour cela que le groupe de pilotage ait été élargi suffisamment pour refléter les principaux points de vue en présence. En effet, même si la commande de l'évaluation est presque toujours le fait d'un des partenaires, il serait dangereux de considérer l'évaluation comme un exercice unilatéral commandé par un seul décideur pour son usage exclusif. Pour cette raison, la pratique conduit souvent à instituer un groupe de pilotage qui permet l'expression de plusieurs points de vue. L'élargissement du groupe de pilotage à une dizaine de protagonistes présente de nombreux avantages : les personnes invitées font entendre les attentes et les craintes qui s'expriment à leur niveau, elles améliorent la pertinence des questions posées, et leur présence crédibilise l'évaluation. Une plus grande diversité au sein du groupe de pilotage met en présence un plus grand nombre de points de vue, facilite leur confrontation, et permet d'élever le degré de consensus. Chaque participant apporte ses questions, son système de valeurs, ses critères de jugement et ses normes. Sous réserve de savoir tirer parti de cette pluralité, l'évaluation aboutit à des conclusions plus riches, plus crédibles et plus utiles. La prise en compte des attentes de plusieurs partenaires permet une meilleure acceptation des résultats et augmente l'intérêt qui sera porté à ces résultats. On trouvera dans le Volume 1 de la Collection, une série de recommandations pour constituer et faire fonctionner un groupe de pilotage élargi.

Si le groupe de pilotage a été élargi au-delà de dix personnes, toutes ne pourront pas participer efficacement à l'analyse multicritère. Un groupe de six à dix assesseurs sera alors choisi au sein du groupe de pilotage. Comme on le verra ci-dessous, l'équipe d'évaluation pourra être amenée à travailler avec les assesseurs de façon collective ou par entretiens individuels.

■ Expression des préférences entre critères

L'étape d'expression des préférences est destinée à établir les pondérations des différents critères. Cette pondération est établie en tenant compte des préférences des assesseurs.

● Expression directe des préférences des assesseurs

Une première variante consiste à réunir les assesseurs et à leur demander d'exprimer leurs préférences sous forme d'une pondération attribuée à chaque critère. On peut procéder par un vote par répartition de points. La discussion peut également être conduite au moyen de plusieurs réunions successives comme dans l'exemple ci-dessous.

Exemple : Création d'un système de notation-pondération en Angleterre

Dans le cadre du programme de développement socio-économique de la région North-East England pour la période 1991-93, un petit groupe de dix personnes a mis au point un système de sélection des projets. Le groupe de travail s'est réuni à plusieurs reprises pendant trois mois. Il comprenait des représentants des principales autorités responsables du programme et de sa mise en œuvre. Il a établi la liste des critères et le poids de chaque critère ainsi que des échelles de notation pour les différents critères.

Certaines réunions ont été conflictuelles et le travail a traversé une période de crise. Les membres du groupe sont toutefois parvenus à un accord, sachant qu'il existait dans la région un excellent esprit de partenariat et une unanimité en faveur de la politique de développement régional. Après que le groupe de travail soit arrivé à un accord, toutes les agences impliquées dans le programme ont été consultées et les autorités responsables du programme ont validé le système qui a été ensuite appliqué à la sélection de tous les projets et qui a servi de modèle pour la plupart des autres programmes britanniques.

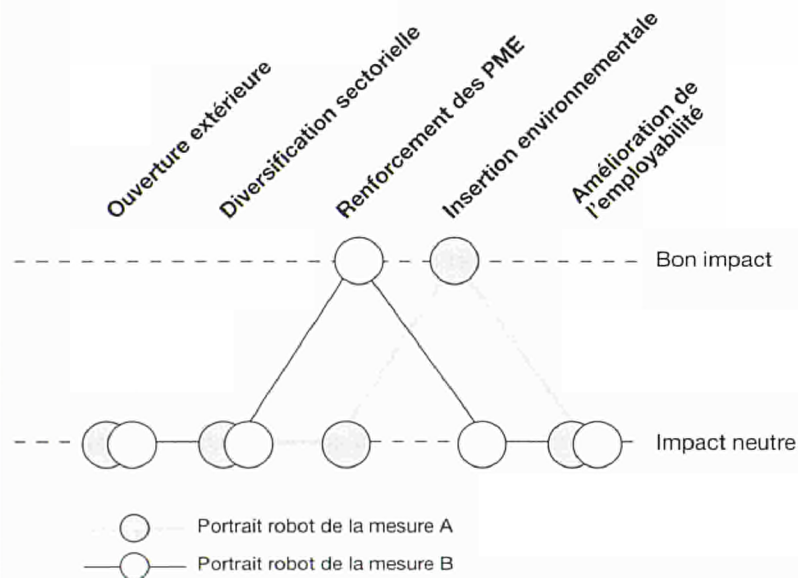
● Révélation des préférences par classement de portraits robots

Une autre formule consiste à présenter aux assesseurs des « portraits-robots » de mesures ou de projets décrits de telle façon qu'ils révèlent les

préférences entre les critères. On procédera par exemple en proposant aux assesseurs de comparer deux mesures « théoriques » A et B ayant des impacts schématisés comme dans l'Encadré ci-dessous.

La mesure théorique A sera décrite comme une mesure dont l'impact est neutre, c'est-à-dire nul, pour tous les critères, à l'exception du critère « insertion environnementale » pour lequel son impact est bon, c'est-à-dire que la mesure a contribué de façon systématique à améliorer la viabilité environnementale des activités de ses destinataires. La seconde mesure théorique B sera décrite comme une mesure dont l'impact est neutre, c'est-à-dire nul, pour tous les critères, à l'exception du critère « renforcement des PME » pour lequel son impact est bon, c'est-à-dire que la mesure a touché des PME, dans une large proportion, et a contribué de façon décisive à améliorer leur compétitivité.

Encadré 44 – Portraits-robots de deux mesures privilégiant deux critères



Entre ces deux portraits robots, les assesseurs doivent choisir celui qu'il préfère et, si possible, ils doivent dire si leur préférence est faible, moyenne, forte ou très forte. L'exercice est répété pour tous les couples de portraits

robots. L'ensemble des réponses est traité à l'aide d'un logiciel informatique qui attribue à chaque impact une pondération en % de telle façon que le total des pondérations soit égal à 100%. Le logiciel utilisé a été présenté par ailleurs dans le chapitre consacré aux échelles de notation (voir page 62).

● Révélation des préférences par classement de projets réels

Le choix proposé aux assessseurs dans la variante précédente pourra avoir l'inconvénient d'apparaître comme artificiel. Pour éviter ce problème, il est préférable de demander aux assessseurs de se prononcer sur leurs préférences entre des projets réels. Cette variante a été appliquée dans le cadre de l'évaluation d'un programme français de développement urbain (voir l'Encadré 24). C'est celle qui doit être privilégiée par l'équipe d'évaluation dans la mesure du possible. La procédure est présentée dans l'exemple qui suit.

Exemple : Classement de projets réels par les assessseurs

Un programme français de développement urbain, co-financé au niveau national et régional, a fait l'objet d'une évaluation à mi-parcours. Huit critères d'évaluation ont été choisis. L'équipe d'évaluation a réalisé une enquête par études de cas, portant sur une vingtaine de projets déjà réalisés. Chaque projet a fait l'objet d'une monographie décrivant ses impacts selon les différents critères. L'équipe d'évaluation a eu ensuite des entretiens individuels et confidentiels d'environ deux heures avec six assessseurs désignés par le groupe de pilotage. Ces entretiens ont pris la forme suivante :

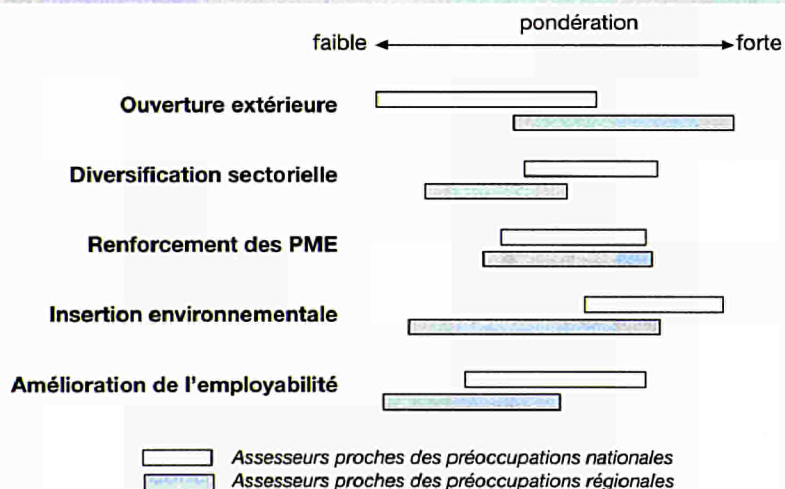
- L'équipe d'évaluation a montré à l'assesseur plusieurs monographies de projets dont il avait une bonne connaissance personnelle. Il a été vérifié que les monographies reflétaient fidèlement la réalité aux yeux des assessseurs.
- L'assesseur a ensuite pris connaissance de trois ou quatre monographies de projets qu'il ne connaissait pas. Les projets étaient choisis de façon à ce que les uns soient plutôt réussis du point de vue de certains critères et les autres soient plutôt réussis du point de vue d'autres critères. Les monographies ont généralement été considérées comme suffisamment précises pour permettre de juger les projets.
- L'assesseur a alors été invité à classer ces trois ou quatre projets de façon intuitive en les rangeant depuis le meilleur (le plus digne d'être financé par le programme) jusqu'au moins bon.
- L'équipe d'évaluation a débattu avec l'assesseur les raisons de son classement et en a déduit une pondération entre les critères qui a été discutée et validée par l'assesseur. Il s'est confirmé que les assessseurs avaient des points de vue très différents sur l'importance à donner aux différents critères.

● Différences de points de vue entre assessseurs

La révélation des préférences des assessseurs peut se faire en réunion ou en entretien individuel. Dans le premier cas, on en déduit un seul et unique système de pondération entre les critères. Dans le second cas, l'équipe d'évaluation et le groupe de pilotage peuvent décider d'établir des pondérations moyennes, ce qui a pour effet de gommer les différences de points de vue entre assessseurs.

Une variante consiste, au contraire, à conserver les pondérations individuelles de chaque assessseur. Dans ce cas, les pondérations réalisées individuellement par un assessseur doivent être traitées en préservant son anonymat. Il peut alors être intéressant de présenter les pondérations de façon séparée pour telle ou telle catégorie d'assesseurs. Dans l'exemple de l'Encadré ci-dessous, les pondérations ont été présentées de façon séparée pour deux catégories d'assesseurs, ceux qui ont déclaré se sentir plus proche des préoccupations nationales et ceux qui ont déclaré se sentir plus proche des préoccupations régionales.

Encadré 45 – Différences de pondérations entre les assessseurs



■ Classement multicritère des mesures

● Note globale et classement des mesures

Arrivé à ce stade de son travail, l'équipe d'évaluation dispose de tous les éléments pour calculer les notes globales pondérées des différentes mesures. On considère en effet que les résultats et impacts de chaque mesure ont été évalués au regard des mêmes critères, que l'ensemble de ces évaluations ont été présentées sous forme de notes dans une matrice de notation des impacts, et que l'on dispose d'un système de pondération exprimant les préférences moyennes des assesseurs pour tel ou tel critère.

Encadré 46 – Calcul des notes globales des mesures

Contribution des mesures à la production des impacts attendus (note entre 0 et 10)	Insertion sociale	Emploi	Equilibre des quartiers au sein de la ville	Equilibre entre les villes de la région	Note moyenne pondérée
Aménagements urbains (moyenne des notes pour les mesures 1 et 2)	1	1	1	2	1,2
Equipements (moyenne des notes pour les mesures 3 à 8)	5	4	4	3	4,1
Actions sociales collectives (moyenne des notes pour les mesures 9 à 14)	6	5	3	3	4,4
Actions sociales personnalisées (moyenne des notes pour les mesures 15 à 19)	6	6	2	2	4,2
Coefficient de pondération des critères	35%	20%	20%	25%	100%

Le calcul de la note globale se fait alors en multipliant chaque note élémentaire par sa pondération et en additionnant les notes élémentaires pondérées, comme dans l'exemple de l'Encadré 46 qui reprend la matrice de notation des impacts de l'Encadré 24.

A partir des notes moyennes pondérées, l'équipe d'évaluation peut classer les mesures par ordre de contribution à la réussite d'ensemble du programme. Ce jugement de synthèse sur l'efficacité des mesures peut être considéré comme solide et impartial à la condition que l'équipe d'évaluation ait pris les précautions nécessaires et qui sont présentées ci dessous.

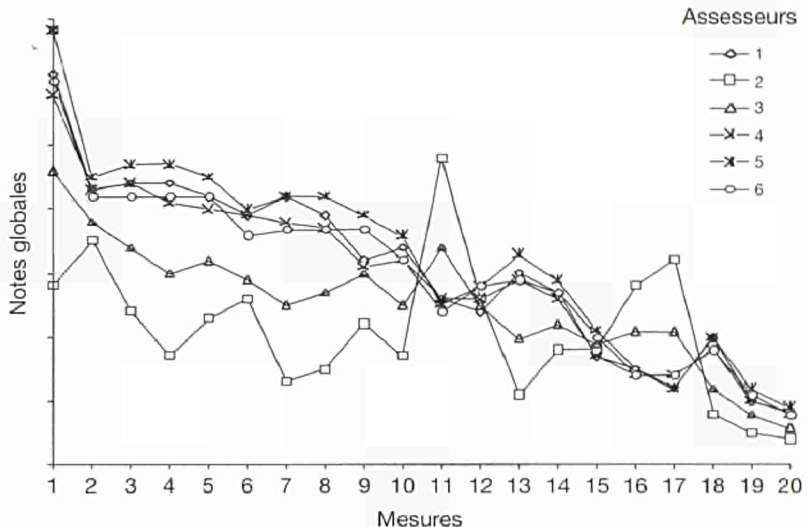
- Faire valider les critères d'évaluation par le groupe de pilotage.
- Présenter et faire valider les conclusions relatives aux impacts de chaque mesure, ainsi que la matrice de notation des impacts qui les résume.
- Etablir les coefficients de pondération des critères avec l'aide d'assesseurs choisis avec l'accord du groupe de pilotage.
- Vérifier avec le groupe de pilotage s'il y lieu d'instituer des seuils de veto sur certains critères, auquel cas il conviendra de préférer une procédure de surclassement.

Si, au contraire, l'équipe d'évaluation effectue elle-même le choix et la pondération des critères, sans interaction avec les partenaires, l'impartialité des résultats en souffrira et l'intérêt de l'analyse multicritère sera moindre.

● **Analyse multicritère-multijuge**

Si la révélation des préférences entre les critères a été faite grâce à des entretiens individuels avec les assesseurs, et si ces entretiens ont montré des divergences de points de vue notables entre les assesseurs, il est possible d'établir plusieurs classements des mesures. Sur la même matrice de notation des impacts, l'équipe d'évaluation appliquera différents systèmes de pondération (un jeu de pondérations différent pour chaque assesseur). Il en résultera des différences entre les notes globales pondérées, et donc des différences de classement, telle mesure pouvant être considérée comme une réussite du point de vue d'un assesseur et comme un échec du point de vue d'un autre. L'Encadré 47 montre comment vingt mesures ont été notées par six assesseurs différents en utilisant les pondérations selon une formule appelée « analyse multicritère-multijuge ».

Encadré 47– Classement multicritère-multijuge des mesures selon leur contribution à la réussite du programme



Dans cet exemple, le graphique laisse apparaître des zones de consensus et des zones de dissension. Pour certaines mesures, il existe en effet un consensus sur le fait qu'elles sont des réussites (2, 12 par exemple) ou qu'elles n'en sont pas (19, 20 par exemple). Par contre, certaines mesures donnent lieu à des jugements hétérogènes au sein du groupe des assesseurs (1, 4, 11 par exemple).

L'évaluation multicritère peut donc être organisée en vue de produire une conclusion synthétique unique à la fin de l'évaluation, ou au contraire, en vue de produire des conclusions adaptées aux préférences et aux priorités de plusieurs partenaires différents. Dans le cas des programmes socio-économiques de l'Union européenne, les différents niveaux de partenariat (européen, national régional) peuvent être concernés. Chacun de ces niveaux est en effet légitime pour établir ses propres priorités et à exprimer ses propres préférences entre critères.

L'objectif de l'analyse multijuge consiste donc à tenir compte de ces priorités différentes dans la restitution des résultats de l'évaluation. Elle est bien adaptée au mode de fonctionnement du partenariat en ce sens qu'elle dessine des zones consensuelles dans lesquelles les partenaires sont

d'accord sur le classement des mesures, et des zones de dissension qui révèlent des mesures considérées comme des réussites pour les unes et des échecs pour les autres. L'expérience a montré que les conclusions consensuelles sont généralement les plus nombreuses. Cela s'explique par le fait que les différentes pondérations s'appliquent à la même matrice de notation des impacts. Une mesure ayant de mauvaises notes pour tous les critères ne pourra donc jamais avoir une bonne note globale pondérée quelles que soient les différences de priorités entre partenaires. Les différences de points de vue entre partenaires ne peuvent donc pas contredire fortement les conclusions qui résultent de l'observation empirique si celles-ci montrent que certaines mesures font réellement partie des meilleures pratiques et que d'autres posent de réels problèmes d'efficacité. L'expérience montre aussi que les partenaires sont beaucoup plus disposés à accepter les conclusions du rapport si l'équipe d'évaluation a recueilli leur point de vue avec soin et si elle a pris la peine de présenter ses conclusions en tenant compte de leurs préférences.



Utiliser l'analyse pour produire des conclusions utiles

● Recommandations de réallocations budgétaires

Dans le cadre de l'évaluation des programmes socio-économiques, l'évaluation pourra orienter les décisions de réallocations budgétaires, soit en cours de programme, soit dans le cadre de la préparation du programme suivant. Les principales décisions de réallocations budgétaires sont en effet prises au niveau de la mesure. Les mesures jugées les moins réussies devront être réexaminées, soit dans le sens d'une réduction de leur budget, soit dans le sens d'une réorganisation tendant à les rendre plus efficaces. Les mesures les mieux classées pourront faire l'objet de recommandations d'augmentation budgétaire si cela est pertinent.

● Diffusion des meilleures pratiques

L'analyse multicritère permet d'identifier les succès du programme et les mesures les plus efficaces. La communication et la diffusion de ces conclusions au sein des acteurs du programme sont intéressantes. Il est possible de présenter les mesures jugées comme les plus réussies (meilleures pratiques) au cours de séminaires et de diffuser des brochures ou des monographies décrivant ces mesures. La diffusion peut également être étendue aux médias si les autorités gestionnaires du programme souhaitent montrer au public l'usage qui a été fait des ressources budgétaires du programme.

Il est également possible de publier des exemples concrets de mesures réussies en vue d'informer les gestionnaires de mesures similaires financées dans le cadre d'autres régions. En effet, l'échange d'informations sur les pratiques identifiées comme les meilleures, est un excellent moyen d'augmenter à terme la performance de la politique européenne de cohésion.

● Rétroaction sur les méthodes de sélection des projets

Le choix des critères d'évaluation, leur définition précise et leur pondération constituent un apport intéressant de l'analyse multicritère. Ce travail permet de constituer une description claire, complète et cohérente des intentions et des priorités des partenaires du programme. Il est ensuite possible

d'utiliser ces résultats pour diffuser des messages clairs aux gestionnaires des mesures et aux opérateurs.

Les résultats de l'analyse multicritère pourront aussi servir à renforcer le processus de sélection des projets. En effet, il est relativement facile de transférer les critères, les échelles de notation, et les pondérations vers le système de sélection des projets si celui-ci est organisé également sur le principe de la notation-pondération (voir l'exemple décrit à la page 64). En rapprochant la logique de la sélection des projets et la logique de l'évaluation des mesures, on a de meilleures chances de susciter l'émergence et le financement de projets contribuant bien aux priorités fixées dans le cadre du programme. Pour être globalement réussi un programme doit sélectionner ses projets sur des critères les plus proches possibles de ses propres critères d'efficacité.

● **Approfondissement des analyses**

Si l'analyse multicritère a été conduite dans sa variante « multijuges ». Il convient de savoir également utiliser ses non-conclusions, à savoir les cas de mesures jugées favorablement par certains et défavorablement par d'autres. Dans ce cas, le groupe de pilotage peut recommander aux autorités responsables du programme de lancer un approfondissement de l'évaluation pour mieux comprendre les raisons qui ont conduit à un désaccord sur ces mesures : Le désaccord tient-il à une insuffisante définition des critères ? Les critères prennent-ils bien en compte les objectifs de la politique, donc les points de vue de l'ensemble des partenaires de manière équitable ?

■ Combinaison des quatre solutions techniques

Pour récapituler les quatre solutions techniques on peut schématiser la démarche d'évaluation en quatre étapes structurées comme indiqué dans l'Encadré 46. L'équipe d'évaluation commence par analyser les documents de programmation puis réalise une concertation avec les partenaires. Elle choisit ses indicateurs puis définit les modalités de collecte des données. Elle réalise les indicateurs quantitatifs et qualitatifs puis analyse les données et les impacts (phase 4). À l'issue de ces quatre étapes, l'équipe d'évaluation réalise

Conclusion

Cet ouvrage a présenté quatre solutions techniques adaptées à l'évaluation d'ensemble d'un programme de développement socio-économique. Ces quatre solutions sont construites dans le même esprit. Elles donnent une grande importance à l'interaction entre l'équipe d'évaluation et les partenaires réunis au sein d'un groupe de pilotage élargi. Elles sont structurées et formalisées sans toutefois être trop exigeantes en termes de données quantitatives. Elles peuvent aisément être combinées entre elles comme cela est montré en conclusion.

Encadré 46 : Synthèse des quatre solutions techniques

1- Analyse des documents de programmation	2- Concertation avec les partenaires	3- Définition des indicateurs et des modalités de collecte	4- Réalisation des indicateurs et analyse des données
- Document de programmation - Document de planification - Document de suivi - Document de reporting	- Document de concertation - Document de concertation - Document de concertation - Document de concertation	- Document de concertation - Document de concertation - Document de concertation - Document de concertation	- Document de concertation - Document de concertation - Document de concertation - Document de concertation



● Combinaison des quatre solutions techniques

Pour récapituler les quatre solutions techniques on peut schématiser la démarche d'évaluation en quatre étapes standard comme indiqué dans l'Encadré 48. L'équipe d'évaluation commence par examiner les documents de programmation pour repérer les impacts attendus. Au besoin, elle clarifie ces impacts avec l'aide du groupe de pilotage (étape 1). Il faut ensuite choisir des indicateurs quantitatifs ou des descripteurs qualitatifs pour étudier les impacts (étape 2). A l'aide de ces instruments, l'équipe d'évaluation réalise une observation empirique des résultats et des impacts sur le terrain (étape 3). L'évaluation s'achève par un jugement synthétique sur le programme, jugement qui prend en compte les différents impacts obtenus et l'importance relative de ces impacts pour les différents partenaires (étape 4).

Dans le cas d'une évaluation ex ante, l'étape 3 est remplacée par une simple estimation prévisionnelle des impacts, sur la base des connaissances disponibles.

Encadré 48 – Enchaînement des quatre solutions techniques

1 Clarifier les principaux impacts attendus	2 Choisir des indicateurs et des descripteurs adaptés à chaque impact	3 Rassembler des informations et les analyser pour estimer les impacts	4 Porter un jugement synthétique au regard des différents impacts obtenus
Dessiner la carte conceptuelle des impacts	Construire des échelles de notation	Evaluer les mesures par observation empirique des résultats et impacts sur le terrain	Appliquer l'analyse multicritère
		Analyser la synergie entre mesures à l'aide d'une matrice des impacts croisés	Conclure l'évaluation et faire des recommandations

Encadré 49 – Etude empirique des impacts au moyen d'études de cas parallèles

Une étude de cas examine un projet individuel (le cas) et aboutit à une monographie décrivant ce projet en quelques pages. L'enquêteur rassemble tous les documents disponibles sur le projet et sur sa gestion. Il visite le site et interroge le gestionnaire de la mesure dont le projet fait partie, ainsi que l'opérateur responsable du projet. Il réalise aussi des entretiens individuels avec plusieurs destinataires directs et indirects. Chaque étude de cas représente une dépense non négligeable (2 000 à 5 000 euros) mais le nombre de cas étudiés est suffisamment réduit pour contenir le coût total de l'évaluation dans une limite raisonnable.

Une enquête par étude de cas parallèles porte typiquement sur dix à trente projets, observés et analysés avec la même grille d'enquête. Une telle enquête permet d'obtenir des informations généralisables sur les motivations des destinataires, sur leurs véritables besoins, sur les raisons de leur participation au programme, sur leur façon de juger la qualité du programme, sur les mécanismes d'apparition des impacts, sur l'existence d'effets imprévus, etc. L'enquête par étude de cas étant essentiellement qualitative, sa principale fonction est de mieux faire comprendre le fonctionnement du programme.

Les études de cas peuvent également produire des estimations d'impact quantitatives ou qualitatives, ou encore structurées à l'aide d'échelles de notation. Dans ce cas, chaque monographie produit des conclusions sur les impacts du projet étudié et de lui seul. Le nombre limité d'études de cas ne permet pas de procéder à des extrapolations de caractère statistique.

Il est cependant possible d'utiliser une enquête par étude de cas pour produire des conclusions systématiques sur les résultats et impacts des mesures d'un programme socio-économique. Par exemple, on pourra sélectionner un ou plusieurs cas considérés comme « typiques » de chaque mesure. Les conclusions obtenues seront utilisées pour estimer l'ordre de grandeur des impacts, mesure par mesure. Plus intéressante est sans doute la démarche qui consiste à étudier deux projets par mesure, un projet considéré comme très réussi par les gestionnaires de la mesure et un cas de relatif échec. De cette façon on obtient une fourchette d'estimation des effets, mesure par mesure. Les conclusions formulées par l'équipe d'évaluation pourront donc utilement s'appuyer sur une enquête par études de cas parallèles, sachant que les autres sources d'informations ne doivent pas être négligées (données de suivi sur les réalisations, opinions des gestionnaires de mesure, analyses d'impacts de mesures comparables, etc.).

Pour une présentation plus complète de la technique de l'étude de cas, on se reportera au Volume 3 de la Collection.

Pour réaliser l'évaluation d'ensemble d'un programme dans un cadre de partenariat, il est possible d'articuler les différentes solutions techniques présentées dans ce Volume, de la façon qui est présentée dans l'Encadré 48. Ce schéma fait cependant apparaître que les quatre solutions techniques ne suffisent pas à assurer la réalisation complète de l'évaluation. En particulier, elles ne touchent pas au cœur de l'évaluation, c'est-à-dire à l'observation empirique des résultats et des impacts produits par les mesures sur le terrain (étape encadrée en gras dans le schéma). Aucune évaluation n'a de sens si cette fonction n'est pas remplie. Différentes stratégies peuvent être adoptées pour réaliser l'observation empirique des résultats et des impacts, sachant que la difficulté majeure consiste ici à observer plusieurs classes d'impacts pour de nombreuses mesures qui ne s'adressent pas aux mêmes publics. Chaque équipe d'évaluation doit choisir la meilleure démarche possible compte tenu des caractéristiques du programme et de ses publics. L'Encadré 49 montre une des solutions possibles, basée sur le principe des études de cas parallèles.

● Une démarche d'ensemble privilégiant la rapidité

S'agissant de combiner entre elles les solutions techniques présentées dans ce Volume, plusieurs démarches sont possibles. On a choisi de présenter en premier lieu une démarche qui privilégie la facilité de mise en œuvre de l'évaluation. Cette démarche peut se dérouler comme suit :

- Constituer un groupe de pilotage élargi (voir pages 29 et 113).
- Repérer les impacts attendus dans les documents de programmation (voir page 30).
- Définir cinq à sept classes d'impacts en utilisant la technique METAPLAN® selon la procédure présentée page 39. Cette étape nécessite une première réunion du groupe de pilotage, éventuellement combinée avec la validation du plan de travail de l'évaluation.
- Réaliser l'évaluation des impacts des mesures (cette étape est importante et difficile, mais elle est commune à tous les types d'évaluation).
- Rassembler toutes les conclusions qualitatives et quantitatives concernant les impacts des mesures sous forme d'un tableau synoptique (voir Encadré 16).
- Transformer le tableau synoptique des conclusions en matrice de notation des impacts selon la procédure simplifiée qui est présentée en page 59. Cette étape nécessite une seconde réunion du groupe de

pilotage, éventuellement combinée avec la présentation des travaux d'enquête réalisés par l'équipe d'évaluation.

- Construire la matrice des impacts croisés, faire l'analyse d'ensemble des effets de synergie potentiels et produire les coefficients de synergie (voir page 83 et suivantes).
- Faire établir les pondérations des critères par le groupe de pilotage selon la procédure simplifiée qui est présentée en page 114. Cette étape nécessite une réunion du groupe de pilotage, éventuellement combinée avec la présentation d'une première série de conclusions par l'équipe d'évaluation.
- Calculer des notes globales pondérées pour toutes les mesures (voir page 118) et les utiliser pour fonder des conclusions d'ensemble et des recommandations de réallocations de ressources.

Comme le montre l'Encadré 50, l'ensemble de cette démarche n'impose pas un travail beaucoup plus volumineux que celui qui s'impose de toutes façons pour la réalisation d'une évaluation de bonne qualité.

Encadré 50 - Temps nécessaire pour mettre en oeuvre une démarche simplifiée

Temps nécessaire (selon la complexité du programme):	Jours min.	Jours max.
Constitution du groupe de pilotage élargi	(hors évaluation)	(hors évaluation)
Identification des impacts attendus et constitution des classes	4	15
Réalisation de l'évaluation des mesures	(comme dans une autre évaluation)	(comme dans une autre évaluation)
Tableau synoptique des conclusions et matrice de notation des impacts	3	6
Identification des synergies potentielles et analyse d'ensemble des effets de synergie	4	15
Faire la pondération des critères et le calcul des notes pondérées	3	6
Produire le rapport d'évaluation, les conclusions de synthèse et les recommandations	(comme dans une autre évaluation)	(comme dans une autre évaluation)
Temps supplémentaire par rapport à une autre évaluation	14	42

● Une démarche d'ensemble privilégiant la rigueur technique

La seconde démarche proposée privilégie l'emploi des quatre techniques dans leur forme la plus rigoureuse. Cette démarche s'applique à une évaluation à mi-parcours ou à une évaluation ex post. Elle se déroule de la manière suivante :

- Constituer plusieurs groupes de travail sectoriels, ainsi qu'un groupe de pilotage élargi qui fonctionnera comme groupe transversal (voir pages 29, 41 et 113).
- Dessiner la carte conceptuelle des impacts (voir pages 30 à 41).
- Pour chaque classe d'impacts, définir une échelle de notation structurée (voir Encadrés 20 à 22) en s'appuyant sur les énoncés d'impact qui composant cette classe (voir Encadré 12).
- Construire la matrice des impacts croisés et étudier la synergie entre les mesures (voir pages 85 à 90) et identifier les synergies à étudier de façon plus approfondie.
- Sélectionner un ou plusieurs projets typiques de chaque mesure ainsi que des projets porteurs d'importantes synergies potentielles.
- Réaliser une enquête par étude de cas parallèles portant sur tous les projets sélectionnés. Présenter les conclusions de chaque étude de cas sous forme d'une série de notes en utilisant les échelles de notation construites pour l'évaluation (voir Encadré 20).
- Rédiger des conclusions qualitatives mesure par mesure en se basant sur les études de cas et sur toutes les autres informations disponibles et pertinentes. Résumer toutes les conclusions sous forme d'une matrice de notation des impacts (voir Encadré 19).
- Réaliser l'analyse multicritère dans sa variante multijuge (voir pages 97 à 106).
- Conclure l'évaluation et procéder aux recommandations de réallocations budgétaires ou de réorientations stratégiques (voir page 123).

Il n'est pas réaliste d'indiquer des éléments de coût pour la mise en œuvre de cette démarche plus technique, car elle doit faire l'objet d'une adaptation fine au contexte et à la nature du programme évalué. Il faut cependant noter qu'elle implique une série de réunions avec les partenaires du programme sous forme de travail en groupe, de séminaires du groupe de pilotage ou d'entretiens individuels avec les assessseurs choisis pour l'analyse multicritère. La démarche comporte également une observation

empirique des effets des mesures. A ce titre, elle implique des visites sur le terrain et des entretiens avec des destinataires.

Au total, le souci de privilégier la qualité technique conduit à un calendrier d'évaluation qui s'étend sur au moins huit mois. Il est donc clair que cette démarche est plus lourde que celle qui est adoptée à ce jour par de nombreuses évaluations. Cet investissement supplémentaire dans l'évaluation peut être justifié si le « retour sur investissement » est suffisant. Rappelons les résultats qui peuvent être attendus d'une telle évaluation et qui ne sont généralement pas atteints de façon satisfaisante à ce jour :

- Des conclusions et des recommandations qui proposent et qui justifient des réallocations budgétaires et des changements de priorités entre les mesures du programme.
- Des conclusions et des recommandations fondées non pas sur les taux de consommation budgétaire et les taux de réalisation, mais sur un véritable examen des impacts et sur la contribution des mesures à la réussite globale du programme.
- Des conclusions et des recommandations qui tiennent compte, de façon impartiale, de tous les points de vue en présence, dans le cadre du partenariat.

Références bibliographiques

- Bana e Costa, C.A. et Vansnick, J.C. (1994) 'MARBETH - An interactive path towards the construction of the cardinal value functions', in *International Transactions on Operations Research*, 1(3), pp 459-530.
- Guba, E.G., Lincoln, Y.S. (1989) *Fourth Generation Evaluation*, Newbury Park, Sage Publications.
- Hemerson, M.E., Morris, L.L., Pitt-Gibbert, G.I. (1987) *How to Measure attitudes 2*, London : Sage publications.
- Knott, C. (1995) 'Concept mapping in policy evaluation: a research review of Community efforts in Northern Ireland', in *Evaluation, the International Journal of Theory, Research and Practice* 1(1), 65-79.
- Miller, D. C. (1981) 'Guide for the selection and construction of social scales and indices', in *Handbook of Research Design and Social Measurement*, London : Sage publications, 174-189.
- Monnier, J. (1992) *Évaluations de l'action des pouvoirs publics*, Paris : Economica.
- OECD - DAC Expert Group on Aid evaluation (1998) *The Role of Rating Systems in Aid Management: Experiences of DAC Member Organisations*, Paris : OECD (non publié).
- Scharif, A. (1989) *Décider sur plusieurs critères*, Lausanne : Presses polytechniques et Universitaires romandes.
- Ray, B. (1997) 'Un chaînon manquant en Recherche Opérationnelle - Aide à la Décision : les conclusions robustes', *Genève de L'IMSADE*, 144, Paris : Université Dauphine.
- Teulemonde, J., Fontaine, C., Laudry, E., Vincke, Ph. (1995) 'Evaluation in partnership: Practical suggestions for improving their quality', in *Evaluation, the International Journal of Theory, Research and Practice*, 4(2).
- Techuan, W.M.K. (1989), 'An introduction to concept mapping for planning and evaluation', in *Evaluation and Program Planning*, 12, 1-16.
- Vincke, Ph. (1989) *L'aide multicritère à la décision*, Bruxelles : Editions de l'Université de Bruxelles. Traduction anglaise (1997), *Multicriteria Decision-aid*, Chichester : Wiley.

Annexe



Références bibliographiques

- Bana e Costa, C.A. et Vansnick, J.C. (1994) 'MACBETH - An interactive path towards the construction of the cardinal value functions', in : *International Transactions in Operations Research*, 1(4), pp 489-500.
- Guba, E.G., Lincoln, Y.S. (1989) *Fourth Generation Evaluation*, Newbury Park: Sage Publications.
- Hemerson, M.E., Morris, L.L., Fitz-Gibbon, G.T. (1987) *How to Measure attitudes ?*, London : Sage publications.
- Knox, C. (1995) 'Concept mapping in policy evaluation: a research review of Community relations in Northern Ireland', in : *Evaluation, the International Journal of Theory, Research and Practice* 1(1), 65-79.
- Miller, D. C. (1991) 'Guide for the selection and construction of social scales and indices', in *Handbook of Research Design and Social Measurement*, London : Sage publications, 174-188.
- Monnier, E. (1992) *Evaluations de l'action des pouvoirs publics*, Paris : Economica.
- OECD - DAC Expert Group on Aid evaluation (1996) *The Role of Rating Systems in Aid Management Experiences of DAC Member Organisations*, Paris : OECD (non publié).
- Schärlig, A., (1985) *Décider sur plusieurs critères*, Lausanne : Presses polytechniques et universitaires romande.
- Roy, B. (1997) "Un chaînon manquant en Recherche Opérationnelle - Aide à la Décision : les conclusions robustes", *Cahiers du LAMSADE*, 144, Paris : Université Dauphine.
- Toulemonde, J., Fontaine, C., Laudren, E., Vincke, Ph. (1998) 'Evaluation in partnership: Practical suggestions for improving their quality', in : *Evaluation, the International Journal of Theory, Research and Practice*, 4(2).
- Trochim, W.M.K. (1989), 'An introduction to concept mapping for planning and evaluation', in : *Evaluation and Program Planning*, 12, 1-16.
- Vincke, Ph ., (1989) *L'aide multicritère à la décision*, Bruxelles : Editions de l'Université de Bruxelles., Traduction anglaise (1992), *Multicriteria Decision-aid*, Chichester : Wiley.

- Yin, R.K., Kaftarian, S.J., Jacobs, N.F. (1996) 'Empowerment Evaluation at Federal and Local Levels : Dealing with quality', FETTERMANN & al. *Empowerment evaluation*, London: Sage, pp 188-207.
- Yoon, K. P. (1995) *Multiple Attribute Decision Making*, London: Sage publications.
- Young, F.WW., Harris, D.G. (1994) 'Multidimensional Scaling', Norusis M.J. *SPSS Profesional Statistics* : SPSS Inc.

A

action 18, 25, 106
 analyse coût-avantages 101
 analyse coût-efficacité 100
 analyse de régression 53
 analyse multicritère 12, 57, 107, 131
 arbre des objectifs 25, 27, 36
 axe 17, 22, 85, 93

B

besoin 103

C

carte logique 24
 cartographie conceptuelle des impacts 11, 29, 106, 131
 classement 64, 77, 112, 119, 120
 coefficient d'extrapolation 53
 cohérence interne 80
 cohésion économique et sociale 17, 72
 commanditaire 32
 comparaison 70
 compensation 110
 complémentarité 72, 73, 78, 81
 coordination 61
 crédibilité 60, 100
 critère 21, 183

D

détermination 65, 67, 69, 109, 127
 destinataire 18, 24, 54, 91, 93
 destinataire direct 101
 destinataire des impacts 72
 données qualitatives 29

E

évaluation 101

Index

évaluation externe 54
 effet de contrebalancement 51
 efficacité 103
 efficience 102
 effets des changes 106
 effet d'impact 13, 68
 enquête par questionnaire 45
 évaluation de groupe 32
 évaluation individuelle 32, 95, 113, 114, 119
 équipe d'évaluation 25, 32, 43, 83, 130, 132
 étude de cas 45, 71, 128, 131
 évaluation à mi-parcours 21, 28, 54
 évaluation approfondie 53
 évaluation d'ensemble 12, 14, 19, 20, 22, 53, 82, 125
 évaluation ex ante 12, 22, 24, 105
 évaluation ex post 54
 évaluation post-mortem 102

F

financier 12

G

gestionnaire 27, 28, 31, 41, 83, 85, 87
 groupe de comparaison 53
 groupe de pilotage 29, 50, 68, 102, 105, 106, 111, 112, 124, 129, 131

H

harmonisation 70

I

impact 18
 impact global 24, 26
 impact intermédiaire 28, 100
 impact spécifique 18



A

action 18, 85, 106
 analyse coût-avantages 101
 analyse coût-efficacité 100
 analyse de régression 53
 analyse multicritère 12, 57, 102, 105, 131
 arbre des objectifs 25, 27, 86
 axe 17, 22, 85, 93

B

besoin 103

C

cadre logique 24
 cartographie conceptuelle des impacts 11, 29, 106, 131
 classement 64, 77, 112, 119, 121
 coefficient d'extrapolation 53
 cohérence interne 80
 cohésion économique et sociale 17, 22
 commanditaire 92
 comparabilité 70
 compensation 110
 complémentarité 72, 73, 78, 81
 coordination 81
 crédibilité 60, 100
 critère 21, 103

D

descripteur 66, 67, 69, 109, 127
 destinataire 18, 24, 54, 91, 93
 destinataire direct 101
 diagramme logique des impacts 72
 donnée qualitative 29

E

échantillon 101

effet d'aubaine 54
 effet de cannibalisme 81
 efficacité 103
 efficience 103
 égalité des chances 106
 énoncé d'impact 33, 66
 enquête par questionnaire 45
 entretien de groupe 32
 entretien individuel 32, 93, 113, 114, 119
 équipe d'évaluation 25, 32, 53, 63, 110, 117
 étude de cas 45, 71, 128, 131
 évaluation à mi-parcours 21, 28, 54
 évaluation approfondie 53
 évaluation d'ensemble 12, 14, 19, 20, 22, 53, 82, 125
 évaluation ex ante 12, 22, 24, 105
 évaluation ex post 54
 évaluation partenariale 102

F

financeur 12

G

gestionnaire 27, 28, 31, 41, 83, 88, 93
 groupe de comparaison 53
 groupe de pilotage 29, 60, 66, 102, 105, 106, 111, 113, 124, 129, 131

H

harmonisation 70

I

impact 18
 impact global 24, 28
 impact intermédiaire 26, 106
 impact spécifique 19

impartialité 100, 119
indicateur 19, 20, 21, 25, 46, 54, 70,
102, 109, 127
instance d'évaluation 60
interdépendance 81

M

masse critique 79
matrice de notation des impacts 61,
62, 68, 71, 118
matrice des impacts croisés 57, 85,
86, 131
meilleure pratique 112, 121, 123
Metaplan® 38, 102, 129

N

norme 103
notation 12, 35, 57, 59, 63, 85, 90,
102, 106, 114, 116, 131
note globale pondérée 63

O

objectif global 17, 22, 26, 100
objectif implicite 27, 31, 103, 106
objectif spécifique 21
opérateur 27, 28, 31, 41, 54

P

panel d'experts 60, 64, 66
performance 54
pertinence 103
pondération 39, 64, 77, 110, 114,
116, 117, 118, 119, 123
priorité politique 106

programme 17
programme intégré 77
projet 18
protagonistesfinanceur 12
public concerné 54

Q

quantification 100

R

réalisation 18
résultat 18
réussite 35, 119
robustesse 23

S

sélection des projets 65, 69
situation contrefactuelle 80
succès 25, 123
suivi 21, 54
surclassement 110
SWOT 31, 102
synergie 57, 77, 78, 83, 87, 91, 92
synergie interne 80
synergie externe 80
synergie négative 81

T

tableau synoptique des conclusions
56, 59, 62, 68

V

vote coloré 102

Remerciements

La préparation de ce Volume s'est étalée sur plusieurs années, de 1993 à 1998. Les travaux préparatoires ont comporté les étapes suivantes :

- Plusieurs ateliers et réunions ont permis de rapprocher les perspectives des experts des différents thèmes et la perspective transversale (ou multi-thématique) des responsables des Fonds structurels.
- Une série d'évaluations pilotes a permis de tester et de stabiliser les solutions techniques proposées dans cet ouvrage.

La coordination des travaux de ce volume a été assurée par Miguel Benito Alonso et Andrea Mairate (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 et Anastassios Bougas.

Ce Volume a été rédigé par le Centre Européen d'Expertise en Evaluation, notamment avec la participation de C.Díaz-Gomez et Céline Fontaine.

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 : J.Aldas (IVIE – Valencia), J. Anselmo (Université Bocconi – Milan), C. Bana e Costa (CESUR, Lisbonne), J.Bigne (IVIE – Valencia), J.Cuadrado (QUASAR – Madrid), C.Gillo (Ministère de l'Équipement, Paris), S.Green (Department of Environment, Londres), C. Knox (University of Ulster), J.Malone (Ernst & Young, Londres), S Martin (Warwick Business School), C. Mouilleron (RIDER – Université de Louvain), M.Rey (CEAT, Lausanne), P. Tyler (Cambridge Policy Consultants,UK), Ph.Vincke (Université Libre de Bruxelles), P.Viveret (Cour des Comptes, Paris).

Les auteurs souhaitent remercier également les responsables des institutions régionales qui ont coopéré de façon particulièrement ouverte à ces évaluations pilotes et notamment : Luc Vandendorpe (Région Wallonne, Belgique), Javier Quesada-Ibañez (Communauté autonome de Valencia, Espagne), Manuel Vaquero et Delphine Vincent (Région Poitou-Charentes et Préfecture de Région, France).

Commission européenne

Évaluer les programmes socio-économiques
Solutions techniques pour évaluer dans un cadre de partenariat

Luxembourg: Office des publications officielles des Communautés européennes

1999 — 141 p. — 14,8 x 21 cm

ISBN 92-828-6627-0

BELGIQUE/BELGIÉ

Jean De Lannoy
Avenue du Roi 202/Koningslaan 202
B-1100 Bruxelles/Brussel
Tel. (32-2) 538 43 08
Fax (32-2) 538 08 50
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/Wetstraat 244
B-1040 Bruxelles/Brussel
Tel. (32-2) 295 25 30
Fax (32-2) 735 08 50
E-mail: mail@libeuro.be
URL: <http://www.libeuro.be>

Moniteur belge/Belgisch Staatsblad
Rue de Louvain 40-42/Louvenseweg 40-42
B-1000 Bruxelles/Brussel
Tel. (32-2) 552 22 11
Fax (32-2) 51 01 84

DANMARK

J. H. Schultz Information A/S
Harslevvej 10-12
DK-2620 Albertslund
Tel. (45) 43 63 23 00
Fax (45) 43 63 19 69
E-mail: schultz@schultz.dk
URL: <http://www.schultz.dk>

DEUTSCHLAND

Bundesanzeiger Verlag GmbH
Verlagsabteilung
Amsterdamer Straße 192
D-50735 Köln
Tel. (49-221) 97 66 80
Fax (49-221) 97 66 82 78
E-mail: vertnet@bundesanzeiger.de
URL: <http://www.bundesanzeiger.de>

ΕΛΛΑΔΑ/GREECE

O. C. Eleftheroudakis SA
International Bookstore
Panepistimou 17
GR-10564 Athens
Tel. (30-1) 331 41 80/132/4/5
Fax (30-1) 353 98 21
E-mail: elebooks@net.gr

ESPAÑA

Boletín Oficial del Estado
Tráfila, 27
E-2827 Madrid
Tel. (34) 915 38 21 11 (Libros).
913 84 17 15 (Suscrip.)
Fax (34) 915 38 21 21 (Libros).
913 84 17 14 (Suscrip.)
E-mail: clientes@com.boe.es
URL: <http://www.boe.es>

Mundi Press Libros, SA
Castell, 37
E-28001 Madrid
Tel. (34) 914 36 37 00
Fax (34) 915 75 39 98
E-mail: libreria@mundiprensa.es
URL: <http://www.mundiprensa.com>

FRANCE

Journal officiel
Service des publications des CE
26, rue Desaix
F-75727 Paris Cedex 15
Tel. (33) 140 58 77 31
Fax (33) 140 58 77 00
URL: <http://www.journal-officiel.gouv.fr>

IRELAND

Government Publications Agency
Publications Section
4-5 Harcourt Road
Dublin 2
Tel. (353-1) 661 31 11
Fax (353-1) 475 27 60

ITALIA

Ilcosse SpA
Via Duca di Calabria, 1/1
Casella postale 552
I-50125 Firenze
Tel. (39) 055 94 83 1
Fax (39) 055 84 12 57
E-mail: ilcosse@ibcc.it
URL: <http://www.ibcc.it/icosse>

LUXEMBOURG

Messageries du Livre SARL
5, rue Rafflissen
L-2411 Luxembourg
Tel. (352) 40 10 20
Fax (352) 49 06 61
E-mail: mail@mdl.lu
URL: <http://www.mdl.lu>

NETERLAND

SDU Servicecentrum Uitgevers
Christoffel Plantinstraat 2
Postbus 20014
2500 EA Den Haag
Tel. (31-70) 378 96 80
Fax (31-70) 378 96 83
E-mail: sdu@sdn.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

Distribuidores de Livros Bertrand Ltd.
Grupo Bertrand, SA
Rua das Terras dos Vales, 4-A
Aparado 80037
P-2700 Anadia
Tel. (351-1) 495 80 50
Fax (351-1) 496 02 55
Imprensa Nacional-Casa de Moeda, EP
Rua Marquês Sá da Bandeira, 16-A
P-1050 Lisboa Codex
Tel. (351-1) 353 03 99
Fax (351-1) 353 02 94
E-mail: delincom@mail.telepac.pt
URL: <http://www.incm.pt>

SUOMI/FINLAND

Academica Kirjakauppa/
Keskuskatu 1/Centralgatan 1
PL/PB 128
FIN-00101 Helsinki/Helsingfors
P. nro (358-9) 121 44 18
F. fax (358-9) 121 44 35
Sähköposti: akka@akateeminen.com
URL: <http://www.akateeminen.com>

SVERIGE

BTJ AB
Traktorvägen 11
S-221 82 Lund
Tfn (46-46) 18 00 00
Fax (46-46) 30 79 47
E-post: btju-pub@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 80
Fax (44-171) 873 84 63
E-mail: gas.enquiries@theso.co.uk
URL: <http://www.the-stationery-office.co.uk>

ISLAND

Bókabúð Larusar Blöndal
Skólavörðslög, 2
IS-101 Reykjavík
Tel. (354) 551 56 50
Fax (354) 552 55 60

NORGE

Swets Norge AS
Østergate 18
Boks 6512 Esterstad
N-0506 Oslo
Tel. (47-22) 97 45 00
Fax (47-22) 97 45 45

SCHWEIZ/SUISSE/SVIZZERA

Euro Info Center Schweiz
c/o OSEC
Stämpfenbachstrasse 85
PO 492
CH-8035 Zürich
Tel. (41-1) 365 53 15
Fax (41-1) 365 54 11
E-mail: eics@osec.ch
URL: <http://www.osec.ch/eics>

BÁLGAIRIA

Europress Euromedia Ltd
59, Blvd Vitosha
BG-1000 Sofia
Tel. (359-2) 960 37 66
Fax (359-2) 960 42 30
E-mail: Milena@mbx.cit.bg

ČESKÁ REPUBLIKA

USIS
NIS-prodovna
Hayekova 22
CZ-130 00 Praha 3
Tel. (420-2) 24 23 14 66
Fax (420-2) 24 23 11 18
E-mail: nikpso@oc.nic.cz
URL: <http://usis.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@ccci.org.cy

EESTI

Eesti Kaubandus-Tööstuskoda (Estonian Chamber of Commerce and Industry)
Toom-Kooli 17
EE-00011 Tallinn
Tel. (372) 546 92 44
Fax (372) 646 02 45
E-mail: einfo@koda.ee
URL: <http://www.koda.ee>

HRVATSKA

Mediator Ltd
Pavla Hatza 1
HR-10000 Zagreb
Tel. (385-1) 481 94 11
Fax (385-1) 481 94 11
MAGYARORSZÁG
Euro Info Service
Europa Ház
Margitáiget
PO Box 475
H-1396 Budapest 62
Tel. (36-1) 350 80 25
Fax (36-1) 350 90 32
E-mail: euroinfo@mail.matazv.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: gwrth@usa.net

POLSKA

Ans Polona
Krakowskie Przedmiescie 7
PL-00-950 Warszawa 1001
Tel. (48-22) 826 12 01
Fax (48-22) 826 62 40
E-mail: ans_pol@berv.hsn.com.pl

ROMÂNIA

Euromedie
Str. G-ral Berthelot Nr 41
RO-70749 Bucuresti
Tel. (40-1) 315 44 03
Fax (40-1) 314 22 86

ROSSIYA

CEEC
60-Iyevya Otkrybiya Av. 9
117312 Moscow
Tel. (7-095) 135 52 27
Fax (7-095) 135 52 27

SLOVAKIA

Centrum TV SR
Nám. Slobody, 19
SK-81223 Bratislava
Tel. (421-7) 54 41 83 64
Fax (421-7) 54 41 83 64
E-mail: europ@ttb1.slt.stuba.sk
URL: <http://www.slt.stuba.sk>

SLOVENIJA

Gospodarski Vestnik
Dunajska cesta 5
SLO-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 Infotel AS
100. Yıl Anıtkabresi 34440
TR-80050 Bagcilar-Istanbul
Tel. (90-212) 629 46 89
Fax (90-212) 629 46 89
E-mail: infotel@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: gdayies@ozemail.com.au

CANADA

Les éditions La Liberté Inc.
3020, chemin Sainte-Foy
G1X 3P1 Sainte-Foy, Québec
Tel. (1-418) 658 67 83
Fax (1-800) 567 54 49
E-mail: liberte@mediom.qc.ca
Renouf Publishing Co. Ltd
5369 Chénier-Capitol Road Unit 1
K1J 5J3 Ottawa, Ontario
Tel. (1-613) 745 26 65
Fax (1-613) 745 76 80
E-mail: orderdept@renoufbooks.com
URL: <http://www.renoufbooks.com>

EGYPT

The Middle East Observer
41 Sherif Street
Cairo
Tel. (20-2) 392 69 39
Fax (20-2) 393 97 32
E-mail: melfoude@meobserver.com.eg
URL: <http://www.meobserver.com.eg>

INDIA

EBIC India
3rd Floor, V. B. Chavan Centre
Gen. J. Bhosale Marg
400 021 Mumbai
Tel. (91-22) 282 60 64
Fax (91-22) 285 45 64
E-mail: ebic@psbnet01.vsnl.net.in
URL: <http://www.ebicindia.com>

ISRAËL

ROY International
41, Mahmar Hayarden Street
PO Box 13066
61310 Tel Aviv
Tel. (972-3) 649 94 39
Fax (972-3) 648 80 69
E-mail: roy@netvision.net.il
URL: <http://www.royint.co.il>

Sub-agent for the Palestinian Authority:

Index Information Services
PO Box 19502
Jerusalem
Tel. (972-2) 827 16 34
Fax (972-2) 827 12 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 62 88
Fax (60-3) 262 61 98
E-mail: ebic-id@mol.net.my

MEXICO

Mundi Prensa Mexico, SA de CV
Rio Pánuco No 141
Colonia Cuauhtémoc
MX-06500 Mexico, DF
Tel. (52-5) 533 56 58
Fax (52-5) 514 67 99
E-mail: 1515452361@compuserve.com

PHILIPPINES

EBIC Philippines
19th Floor, PS Bank Tower
Sen. Gil J. Puyat Ave. cor. Tindaio St.
Makati City
Metro Manila
Tel. (63-2) 759 69 80
Fax (63-2) 759 66 90
E-mail: eicpcom@globe.com.ph
URL: <http://www.eicp.com>

SRI LANKA

EBIC Sri Lanka
Trans Asia Hotel
115 Sir chittampalam
A. Gardiner Mawatha
Colombo 2
Tel. (94-11) 074 71 50 78
Fax (94-11) 44 87 79
E-mail: ebicsl@itmin.com

THAILAND

EBIC Thailand
29 Vanichsa Building, 8th Floor
Sri Chulalongkorn
Ploenchit
10330 Bangkok
Tel. (66-2) 555 06 27
Fax (66-2) 555 06 28
E-mail: ebicth@kcsit.th.com
URL: <http://www.ebicthk.org>

UNITED STATES OF AMERICA

Bernan Associates
4611 F Assembly Drive
Lanham MD20706
Tel. (1-800) 274 44 47 (toll free telephone)
Fax (1-800) 865 34 50 (toll free fax)
E-mail: query@bernan.com
URL: <http://www.bernan.com>

ANDERE LANDER/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 55
Fax (352) 29 29 42 58
E-mail: info@opco.cec.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