

COMPLET

Suivi & Evaluation Axés sur les Résultats en Afrique Francophone



Liste des auteurs

Edoé Djimitri Agbodjan, Eva Anstett, Boubacar Aw,
Abdoul Aziz Kane, Désiré Kanga, Amos Menard,
Hadiza Moussa-Salley, Miché Ouédraogo,
Clément Sekongo Ouollo, Moussa Thiaw,
Donipoho Tuo.

Relecture orthographique : Doudou Ndiaye, Edem Anselme Afoutou

Infographie : Doudou Ndiaye, Samir Said C. M. ZOURE

Impression : Centre Africain d'études supérieures en gestion (CESAG)

Toute reproduction de ce document à des fins commerciales est
strictement interdite et passible de poursuite.

© CLEAR FA

Boulevard Charles de Gaulles, Dakar, Sénégal,
Mars 2023.

Table Des Matières

REMERCIEMENTS	VIII
INTRODUCTION GENERALE	IX
MODULE1 : CONCEPTS DE SUIVI ET ÉVALUATION EN CONTEXTE DE DÉVELOPPEMENT	1
I. LE DEVELOPPEMENT ET SES INSTRUMENTS DE MISE EN ŒUVRE	4
1.1. Qu'est-ce que le développement ?	4
1.2. Instruments de gestion du développement	7
1.2.1. <i>Le plan de développement</i>	7
1.2.2. <i>Le programme</i>	8
1.2.3. <i>Le projet</i>	8
II. CONCEPTS DE SUIVI ET EVALUATION EN CONTEXTE DE DEVELOPPEMENT	9
2.1. Définition du suivi	9
2.1.1. <i>Suivi des réalisations (Délais- Coûts- Avancement)</i>	10
2.1.2. <i>Suivi du déroulement (processus)</i>	11
2.2. Définition de l'évaluation en contexte de développement	12
III. DIFFERENCES ET COMPLEMENTARITES ENTRE LE SUIVI ET L'EVALUATION ET LEUR LIEN AVEC LES FONCTIONS CONNEXES	18
IV. ETHIQUE ET ÉVALUATION	21
MODULE 2 : ENJEUX ET TENDANCES EMERGENTES DE L'EVALUATION DANS LES PAYS EN DEVELOPPEMENT	26
I. CULTURE DE L'EVALUATION	29
1.1. L'évaluation dans les pays de l'OCDE	29
1.2. Evaluation dans les pays en développement	30
1.3. L'Agenda global des Nations Unies et son rôle dans la culture d'évaluation	33
1.4. L'Agenda de l'Initiative Globale d'Evaluation – The Global Evaluation Initiative (GEI)	34

II. THEMES A LA UNE DU DEVELOPPEMENT INTERNATIONAL ET LEUR LIEN AVEC L'EVALUATION	35
2.1. Des initiatives qui encouragent l'utilisation de l'évaluation	35
2.2. Emergence de nouveaux acteurs dans l'aide au développement	36
2.3. Initiative PPTE	37
2.4. Prévention des conflits, reconstruction post-conflit	37
2.5. Démocratie (exigence du contrôle de l'action de l'exécutif par les parlementaires et les élus locaux) et gouvernance	38
2.6. Financement de l'action anti-blanchiment d'argent et anti-terroriste	39
2.7. Approche genre : de « la place des femmes dans le développement » à « la parité dans le développement » puis à « l'égalité des sexes »	39
2.8. Développement du Secteur Privé (DSP) et climat des investissements	41
2.9. La durabilité environnementale et sociale	42
2.10. Biens publics, biens communs et contrôle citoyen	44
III. ETAT DES LIEUX DE L'EVALUATION EN AFRIQUE	45
3.1. Evaluation des actions de développement : Où en sommes-nous en Afrique ?	45
3.2. Vers quel type d'évaluation pour les pays en développement ?	46
MODULE 3 : PROCESSUS DE PLANIFICATION, PROGRAMMATION, SUIVI ET ÉVALUATION EN CONTEXTE DE DÉVELOPPEMENT	49
I. LIEN ENTRE LA PLANIFICATION LE SUIVI ET L'ÉVALUATION EN CONTEXTE DE DÉVELOPPEMENT	51
1.1. Planification stratégique	51
1.1.1. Définition	51
1.1.2. Démarche	53
1.2. Lien entre la conception des projets, la programmation annuelle et le suivi évaluation	54
II. OUTIL DE PLANIFICATION DE PROJET AXEE SUR LES RESULTATS	55
2.1. Le cadre logique (Principes et méthodologie)	55
2.1.1 Définitions et usage du cadre logique	55
2.1.2. Structure du cadre logique	57
2.2. Le cadre d'analyse et de suivi des risques	60
2.3. Le cadre de mesure de rendements	61

III. PLANIFICATION OPERATIONNELLE ET LE SUIVI DES ACTIVITES DE DEVELOPPEMENT	63
3.1. Objet de la planification opérationnelle	63
3.2. Elaboration de la Structure de Fractionnement des Tâches du projet (SFT)	65
3.3. La charte ou matrice de responsabilités	67
3.4. La programmation des activités	68
3.4.1. Ordonnancement : modélisation	68
3.4.2. La méthode PERT (<i>Program Evaluation and Review Technic</i>)	70
3.4.3. La méthode du chemin critique (<i>Le CPM</i>)	72
3.4.4. Le diagramme de GANTT	72
3.5. La planification des coûts	73
3.5.1. Evaluation des coûts directs	74
3.5.2. Evaluation des coûts indirects	75
IV. LIEN ENTRE LE SUIVI, EVALUATION ET LA PLANIFICATION	76
MODULE 4 : CONCEPTION D'UN SYSTEME DE SUIVI ÉVALUATION AXÉ SUR LES RÉSULTATS	80
I. DEFINITION D'UN SYSTEME DE SUIVI EVALUATION AXE SUR LES RESULTATS (SSEAR) ET DE SES ELEMENTS CONSTITUTIFS	82
1.1. Définition du SSEAR	82
1.2. Éléments constitutifs d'un SSEAR	83
1.3. La valeur ajoutée d'un SSEAR	84
II. LES DIX (10) ETAPES DE CONCEPTION D'UN SYSTEME DE S&E AXE SUR LES RESULTATS	86
2.1. Etape 1 : Evaluation de l'état de préparation	88
2.2. Etape 2 : Accord (s'entendre) sur les résultats devant faire l'objet de suivi et d'évaluation	89
2.3. Etape 3 : Choisir les indicateurs-clés pour le suivi des résultats	91
2.4. Etape 4 : Recueillir (Rassembler) des données de référence (de base) sur les indicateurs	98
2.5. Etape 5 : Planifier pour une amélioration : choisir (sélectionner) des objectifs réalistes	99
2.6. Etape 6 : Suivi axé sur les Résultats	102
2.7. Etape 7 : Production d'informations grâce à l'évaluation	103
2.8. Etape 8 : Rappporter les conclusions (présenter les informations)	104
2.9. Etape 9: Utilisation des conclusions	106
2.10. Etape 10 : Maintenir le système de S&E dans l'organisation	108

III. AUTRES APPROCHES DE DEVELOPPEMENT D'UN SYSTEME DE SUIVI EVALUATION	110
MODULE 5 : CONTEXTE DE L'ÉVALUATION ET THÉORIE DU CHANGEMENT	113
I. LA LOGIQUE D'INTERVENTION	115
II. LA THEORIE DU CHANGEMENT	118
MODULE 6 : ÉLABORATION DES QUESTIONS D'ÉVALUATION ET DE LA MATRICE D'ÉVALUATION	123
I. ELABORATION DE QUESTIONS PERTINENTES D'EVALUATION	125
1.1. Questions descriptives	126
1.2. Questions évaluatives	127
1.3. Questions sur les rapports de causalité	128
II. LA MATRICE D'EVALUATION	129
2.1. Elaboration de la matrice d'évaluation	129
2.2. Variables et indicateurs	131
2.3. Sources et méthodes de collecte des données	133
2.4. Les méthodes d'évaluation	133
MODULE 7 : CHOIX DU DESIGN DE L'ÉVALUATION SELON LE TYPE DE QUESTIONS ET LE CONTEXTE	136
I. CHOIX D'UNE METHODE D'EVALUATION	138
II. TYPES DE METHODES D'EVALUATION	140
2.1. Méthode expérimentale ou aléatoire	140
2.2. Méthodes quasi-expérimentales	143
2.3. Les menaces à la validité des méthodes d'évaluation des effets	146
2.4. Autres méthodes d'évaluation quantitatives, non-expérimentales	148
2.5. Méthodes qualitatives	149
MODULE 8 : SÉLECTION DE LA STRATÉGIE D'ÉCHANTILLONNAGE	153
I. PRINCIPE, OBJECTIF ET OBJET DE L'ECHANTILLONNAGE	155
II. LES TECHNIQUES PROBABILISTES	157
2.1. Sondage aléatoire simple (SAS)	158
2.2. Sondage à probabilités inégales	162
2.3. Sondage stratifié	166

2.4. Sondage à plusieurs degrés	167
2.5. Sondage en grappes	168
III. LES TECHNIQUES DE SELECTION NON PROBABILISTES	169
3.1. Technique des quotas	169
3.2. Technique des itinéraires	170
3.3. Technique des unités-types	171
3.4. Technique du volontariat	171
3.5. Technique boule de neige	171
3.6. Technique accidentelle	172
IV. DEFINITION ET DETERMINATION DE LA PRECISION D'UNE ENQUETE PAR SONDAGE	174
4.1. Notion d'erreurs	174
4.2. Erreurs de types I et II, et intervalles de confiance	175
V. DETERMINATION DE LA TAILLE DE L'ECHANTILLON	177
MODULE 9 : STRATEGIE DE COLLECTE DE DONNEES	181
I. PRINCIPAUX CONCEPTS	183
1.1. Conseils pour la collecte de données	183
1.2. Caractéristiques d'une bonne mesure	184
1.3. Type de données	185
II. OUTILS DE COLLECTE DE DONNEES QUANTITATIVES	188
2.1. Les registres secondaires	188
2.2. Questionnaires	188
III. OUTILS DE COLLECTE DE DONNEES QUALITATIVES	190
3.1. Les procédés participatifs	190
3.2. L'observation	191
3.3. Groupes de discussion	195
3.4. Les journaux et listes de contrôle des données auto-déclarées	197
3.5. Le jugement d'experts	198
MODULE 10 : ANALYSE ET INTERPRÉTATION DES DONNÉES	201
I. ANALYSE DES DONNEES QUANTITATIVES	203
1.1. Compilation et traitement des données quantitatives	203
1.2. Différents niveaux d'analyse	204
1.2.1. <i>Analyse descriptive</i>	205

1.2.2. Analyse explicative	211
II. ANALYSE ET INTERPRETATION DES DONNEES QUALITATIVES	214
2.1. Compilation et traitement des données qualitatives	214
2.2. L'analyse inductive vs l'analyse déductive	215
2.2.1. L'analyse inductive - Processus général	215
2.2.2. L'analyse déductive - Processus général	216
2.3. Analyse de contenu	216
2.3.1. Codage du Contenu	216
2.3.2. Matrice de codage	218
2.4. Exemple d'un logiciel d'analyse qualitative	218
III. REDACTION DES FICHES D'ANALYSE	219
MODULE 11 : COMMUNICATION DANS LE PROCESSUS DE SUIVI ET D'EVALUATION	222
I. COMMUNICATION ET RAPPORTAGE PENDANT LE CYCLE DE VIE D'UNE INTERVENTION	224
II. COMMUNICATION AVANT ET PENDANT LES TRAVAUX D'EVALUATION	226
2.1. Finalités de la communication dans le processus d'évaluation	226
2.2. Communication en amont de l'évaluation	227
2.3. Communication pendant l'évaluation	228
III. COMMUNIQUER ET RENDRE COMPTE DES RESULTATS DE L'EVALUATION	229
IV. STRATEGIE DE COMMUNICATION	234
4.1. Objectifs de la communication	234
4.2. Communication informelle et formelle	235
V. PRESENTATION DES RAPPORTS D'EVALUATION	238
5.1. Moment, pertinence et utilisation de l'évaluation	238
5.2. Conseils pour la rédaction de rapports	238
5.3. Prise en compte des commentaires des parties prenantes	241
5.4. Suivi des recommandations	241
5.5. Texte et visuel	242
5.6. Qualité du rapport final et de l'évaluation	243
VI. EXPOSES ORAUX	244

MODULE 12 : GESTION DU PROCESSUS D’EVALUATION	248
I. PLANIFICATION DU PROCESSUS DE L’EVALUATION	250
1.1. Étape 1 : Préparation du dossier de l’évaluation	250
1.2. Étape 2 : Contractualisation de l’évaluation	254
1.3. Étape 3 : Identification des parties prenantes	257
1.4. Étape 4 : Élaboration d’un plan des ressources, des tâches et du budget	259
1.4.1. <i>Planification des activités</i>	259
1.4.2. <i>Comment établir le budget d’une évaluation ?</i>	261
1.5. Étape 5 : Note méthodologique	264
II. COORDINATION DES ACTIVITES DE L’EVALUATION	265
2.1. Gestionnaire d’évaluation	265
2.2. Logistique de l’évaluation	266
2.3. Evaluation la qualité d’une évaluation	267
MODULE 13 : EVALUATION DES INTERVENTIONS COMPLEXES	271
I. DEFINITION ET CARACTERISTIQUES D’UNE INTERVENTION COMPLEXE	273
II. QUELQUES PISTES POUR L’EVALUATION D’UNE INTERVENTION COMPLEXE	275

Remerciements

Ce manuel est le support pédagogique du séminaire de formation en suivi et évaluation axé sur les résultats en Afrique francophone (SERAF) développé et délivré par le Centre de renforcement des capacités en suivi et évaluation et gestion de la performance pour l'Afrique francophone (CLEAR FA). Sa rédaction a bénéficié des contributions et soutiens d'une diversité d'acteurs de la grande famille du CLEAR FA. Le Centre les remercie vivement pour leur contribution.

C'est le lieu d'exprimer notre profonde gratitude au Dr El Hadji Gueye et M. Boubacar Aw, tous deux anciens Coordonnateurs du CLEAR FA, pour l'initiative de développer un tel séminaire. Nous remercions tout particulièrement M. Boubacar Aw, qui a pris l'initiative de lancer l'édition de ce manuel.

Ces remerciements s'adressent ensuite à feu Marie Gervais pour la coordination scientifique du processus de développement du SERAF, entre 2012 et 2013.

Nos remerciements vont aussi à l'endroit du collège des formateurs du SERAF qui avait pris part à l'atelier pilote dédié à la contextualisation et l'enrichissement du matériel pédagogique du SERAF, en 2013, et tous ceux qui ont poursuivi la mise à jour du matériel, sous la coordination pédagogique de Dr Hadiza Moussa-Salley.

Le contenu de ce manuel a en outre été amélioré grâce aux commentaires de Jean-Michel Servet et aux travaux de relecture de Edem Kwami Afoutou, Doudou Ndiaye, Jérôme Babatoundé Agué, Nassibou Bassongui et Fatiya Bague. Qu'ils en soient remerciés !

Nous exprimons enfin notre plus profonde gratitude aux participants des séminaires du SERAF dont les commentaires et recommandations ont guidé et enrichi le fonds et la forme du présent manuel.

Introduction générale

Le programme de formation de base en suivi et évaluation du CLEAR FA dénommé « Suivi et Évaluation axés sur les Résultats en Afrique francophone » (SERAF) vise à présenter les fondamentaux en matière de suivi évaluation. Il est de ce fait destiné à un très large public composé de praticiens de l'évaluation dans les structures gouvernementales et les autres organismes, de consultants en suivi évaluation, d'universitaires et des décideurs.

Initialement développé à partir de supports des programmes de l'International Program for Development Evaluation Training (IPDET) et des documents de référence comme "The Road to Results : Designing and Conducting Effective Development Evaluations" de Morra et Rist (2009), le SERAF a été mis à jour en fonction des enjeux et tendances émergentes de l'évaluation en Afrique francophone et des pratiques sur le terrain. Il a par ailleurs bénéficié d'une refonte intégrale, tant au niveau de son contenu pédagogique, que des méthodes d'enseignement et d'évaluation des apprenants afin d'être conforme aux standards internationaux.

Les treize chapitres de ce manuel présentent les enjeux, approches, processus, méthodes et pratiques du S&E dans les pays d'Afrique francophone.

Le chapitre 1 clarifie les concepts de suivi et évaluation en contexte de développement. Il présente leur complémentarité, leur spécificité ainsi que les normes encadrant la pratique.

Le chapitre 2 décrit les déterminants culturels de l'évaluation à travers les systèmes de suivi et d'évaluation (SSE), afin de dégager les tendances programmatiques émergentes. Il aborde également les enjeux liés à la promotion du suivi évaluation de projets/programmes, la culture de l'évaluation et l'état de la pratique.

Le chapitre 3 du manuel est consacré aux processus de planification, de programmation, de suivi et évaluation. Il examine le lien entre ces trois fonctions dans un concept de gestion axée sur les résultats.

Le chapitre 4 du manuel s'attelle ainsi d'une part à rappeler quelques concepts de base de la notion de suivi-évaluation dans une perspective systémique, puis à articuler différentes approches de conception et d'édification d'un tel système de S&E axé sur

les résultats d'autre part. Le S&E est ici abordé dans une perspective à la fois organisationnelle qu'à l'échelle sectorielle et nationale.

Le chapitre 5 est consacré à un outil important dans la pratique de l'évaluation, à savoir les différentes représentations de la logique d'intervention. Il outillera l'évaluateur à mieux comprendre la démarche d'élaboration d'une théorie du changement.

Le chapitre 6 porte sur les premières étapes de l'élaboration d'une évaluation. Ces étapes comprennent la définition du type de questions auxquelles l'évaluation vise à répondre et les différentes composantes d'une matrice d'évaluation. Il fournit aux lecteurs et lectrices les connaissances essentielles pour l'élaboration des questions pertinentes d'évaluation et la construction subséquente d'une matrice d'évaluation essentielle à la conduite d'un mandat d'évaluation.

Le chapitre 7, qui porte sur le choix du design de l'évaluation selon le type de questions et le contexte, offre au lecteur l'opportunité d'apprendre les techniques d'identification des différentes méthodes d'évaluation, pour ainsi les relier aux types de design, mais aussi identifier les éléments du contexte susceptibles d'influencer le choix de la méthode d'évaluation. Il sera ainsi outillé pour choisir la méthode en fonction des questions et du contexte d'évaluation.

Le chapitre 8 expose le lecteur aux différentes méthodes de sélection de cas et d'échantillonnage, les conditions du choix de l'une ou l'autre des méthodes et le calcul de la taille de l'échantillon. En effet, le choix de la méthode d'échantillonnage et la taille de l'échantillon peuvent varier en fonction de la méthode et du contexte de l'évaluation. Par ailleurs, le choix de la méthode d'échantillonnage a des implications sur le traitement et l'exploitation des données collectées. Les différentes implications seront également présentées dans ce chapitre.

Le chapitre 9 met l'accent sur les différentes stratégies de collecte des données, qualitative et quantitative. Il présente aux lecteurs les modes opératoires pour la collecte des données, et fait un tour d'horizon des principaux outils de collectes les plus utilisés en évaluation. Le sujet de la collecte de données est primordial, car il constitue une des tâches les plus importantes en temps du travail de l'évaluateur. Le bon déroulé et la rigueur de la mise en œuvre de cette étape de travail vont aussi

fortement déterminer la crédibilité de l'argumentaire évaluatif ainsi que le rapport d'évaluation.

Le chapitre 10 porte sur les différentes stratégies d'analyse des données et consiste à donner une signification aux données collectées. À l'aide d'une diversité de techniques, elle permet à l'évaluateur de résumer, interpréter, comparer et classer les données collectées, dans le but d'en tirer des conclusions.

Le chapitre 11 porte sur la communication dans le processus de suivi et d'évaluation. Il présente dans un premier temps au lecteur l'intérêt de la communication avant, pendant et après la conduite du suivi et de l'évaluation, puis décrit les enjeux de la communication dans la présentation des résultats de l'évaluation. Il présente ensuite les stratégies et outils de communication adaptés au public ciblé, le processus d'élaboration de rapports d'évaluation de qualité et pour la présentation des résultats.

Le chapitre 12 initie le lecteur à la gestion du processus d'évaluation. Il est structuré autour de trois thèmes. Le premier porte sur la planification du processus de l'évaluation et couvre la préparation du dossier de l'évaluation, les formes de contractualisation, l'identification des parties prenantes, et l'élaboration du budget et de la note méthodologique. Le deuxième est dédié à la coordination des activités de l'évaluation pour examiner les questions organisationnelles et logistiques. Enfin, le dernier thème présente quelques directives pour conduire une évaluation de qualité.

Le dernier chapitre du manuel porte sur l'évaluation des interventions complexes. À travers ce chapitre, le lecteur sera initié à l'identification des situations de complexité dans les interventions de développement, à la prise en compte de la complexité dans le dispositif d'évaluation et à la compréhension du déficit de capacités des acteurs impliqués dans l'évaluation des interventions complexes.

Comment lire le manuel ?

Le manuel est développé suivant une approche graduelle qui permet aux lecteurs de découvrir les notions de base en suivi et évaluation, puis de progresser vers le développement de la méthode d'évaluation et l'analyse de la complexité des phénomènes autour du S&E. Une première manière de lire ce manuel pour les non-initiés est de suivre cette démarche graduelle.

Pour les initiés, le manuel peut aussi être aisément lu en faisant le focus sur des chapitres, ou sections spécifiques en vue de combler des besoins précis d'apprentissage. A l'entame de chaque chapitre, le(la) lecteur(trice) pourra retrouver le plan de présentation, les mots-clés et un résumé du contenu.

Enfin, selon les besoins de spécialisation, les chapitres peuvent être aussi catégorisés en deux blocs majeurs : un premier bloc couvrant les modules 1 à 5 ainsi que le module 12 sont orientés sur les besoins des commanditaires, gestionnaires des évaluations et un deuxième bloc composé des modules 6 à 11 et module 13 réponds à des besoins spécifiques des évaluateurs.

Le Coordonnateur du CLEAR pour l'Afrique francophone

Edoé Djimitri Agbodjan

MODULE 1

CONCEPTS DE SUIVI ET ÉVALUATION EN CONTEXTE DE DÉVELOPPEMENT

Objectifs d'apprentissage :

- Définir les concepts de suivi et de l'évaluation axés sur les résultats
- Etablir les différences et complémentarités entre le suivi et l'évaluation
- Distinguer certaines typologies d'évaluation
- Décrire les principales étapes de suivi et de l'évaluation axés sur les résultats
- Discuter des normes de l'évaluation et des principes éthiques de l'évaluateur

Thèmes couverts

- I. Notions de plan, Programme et Projet
- II. Définitions des concepts du suivi et de l'évaluation axés sur les résultats
- III. Différences et complémentarités entre le suivi, l'évaluation et l'audit
- IV. Typologie des évaluations
- V. Principales étapes du suivi et de l'évaluation axés sur les résultats
- VI. Principes éthiques en évaluation
- VII. Normes de pratique en évaluation

Ce chapitre correspond au premier des treize modules du SERAF. Il vise à développer une compréhension commune des concepts de suivi et d'évaluation et dans un contexte d'interventions publiques visant le développement.

Dans le cadre de ces interventions, ces deux notions de suivi et d'évaluation sont intimement liées. En raison de cette proximité, les frontières qui les séparent sont parfois floues. De plus, elles sont souvent confondues à des notions voisines telles que celles de l'audit, la vérification, le contrôle, l'inspection, etc.

Le présent chapitre amène donc le lecteur à mieux cerner les différences entre ces notions, leur complémentarité et leur spécificité.

Mots-clés : contexte du développement, suivi, évaluation, critères

I. Le développement et ses instruments de mise en œuvre

1.1. Qu'est-ce que le développement ?

Au regard de la diversité des débats, parfois contradictoires, dont elle fait l'objet, la notion de développement n'est pas aisée à définir. Les caractérisations de l'état de « développement », de « sous-développement » ou de « pays en développement » ont été différemment appréhendées au cours de l'histoire. Les postures paradigmatiques varient en fonction de l'accent mis sur la croissance économique comme pièce centrale de la trajectoire de développement, ou d'une remise en cause de celle-ci en faveur des postures favorables à la « décroissance ».

Il est important de pouvoir faire la différence entre la croissance et le développement. La croissance tient en effet compte uniquement de la dimension économique alors que le développement intègre les dimensions sociales, politiques, culturelles, etc. La notion de développement est apparue à la fin des années 40 dans le programme du candidat aux élections présidentielles aux Etats Unis, Harry Truman.

Dans les années 1960, l'américain Walt W. Rostow dans son ouvrage « Stages of economic growth : a non-comunist manifesto » en propose une explication linéaire. Le développement selon lui se traduit par une transformation de l'économie vers la satisfaction des besoins à travers cinq grandes étapes :

- la société traditionnelle;
- les conditions préalables au décollage: à cette étape, épargne et investissement se développent;
- le décollage ou le take-off: les sociétés entrent dans un rythme de croissance continue;
- la marche vers la maturité: les premières industries sont relayées par de nouveaux secteurs industriels;
- l'ère de la consommation de masse: les productions de biens et de services deviennent les principaux secteurs de l'économie.

Avec cette explication, l'accent est mis sur la croissance économique comme déterminant principal du développement. La croissance est obtenue lorsque les individus investissent leurs capitaux et travaillent avec l'aide de la technologie. Les pays qui n'ont pas beaucoup de moyens (capitaux) pour investir parce qu'ils sont pauvres (déficit d'épargne interne) devront alors avoir recours à des prêts ou aides extérieurs pour croître, d'où l'utilité de l'aide au développement et dans une acception plus large de la coopération au développement.

Confrontée à la réalité des pays bénéficiaires de l'aide au développement, cette explication est apparue trop linéaire et simpliste. La croissance économique, qui est l'augmentation de la production des biens et services, n'est pas totalement expliquée par le niveau de capitaux ni de travail qu'un pays utilise. Cette croissance peut être limitée par la structure des échanges entre les pays, favorisant certains par rapport à d'autres, en raison de leurs puissance et positionnement dans le système mondial. Par ailleurs, la croissance ne se traduit pas automatiquement en développement, c'est-à-dire l'amélioration des conditions de vie des populations. Au contraire, elle peut entretenir des structures de dépendances des pays sous-développés par rapport à ceux dits « développés » et entraîner l'augmentation de la pauvreté et des inégalités.

Sous l'impulsion des travaux d'économistes institutionnalistes comme Douglass North, Ronald Coase et Olivier Williamson, il est important de mettre en place des mesures visant à rendre les systèmes d'échange et de redistribution des richesses plus équitables. L'État est également un acteur-clé qui peut mener des politiques sociales pour améliorer les conditions de vie de ses citoyens. Le développement vient alors de la combinaison cohérente et efficace de toutes les politiques économiques (pour la croissance) et sociales.

C'est ainsi que les analyses sur le thème du développement vont montrer qu'afin de croître et de se développer, les pays ont besoin de se doter d'un ensemble de règles formelles et informelles, c'est-à-dire d'institutions qui permettent de réduire les coûts de transaction et de stabiliser les anticipations des agents. Par ricochet, les pays ont aussi comme devoir de promouvoir la "bonne gouvernance" dans ces institutions pour que leurs actions soient efficaces.

Cette vision du développement a été menée par les travaux des dépendantistes ou structuralistes dont les précurseurs sont : André Gunder Frank (1972), Celso Furtado (1972), Samir Amin (1973) et Raúl Prebisch (1977).

Le développement étant le résultat de l'action humaine, il est également apparu important de promouvoir les capacités, c'est-à-dire les capacités des personnes à avoir accès et faire usage de leurs droits. (Amartya Sen, PNUD).

Enfin, une des dernières évolutions majeures de la notion de développement est la prise en compte de la question environnementale, avec l'émergence du concept de développement durable. Selon le Rapport Brundtland (1987) qui énonça ce concept, « le développement durable est un développement qui répond aux besoins du présent sans compromettre la capacité des générations futures à répondre aux leurs ». Deux concepts sont inhérents à cette notion : le concept de "besoin", et plus particulièrement des besoins essentiels des plus démunis, à qui il convient d'accorder la plus grande priorité, et l'idée des "limitations" que l'état de nos techniques et de notre organisation sociale impose sur la capacité de l'environnement à répondre aux besoins actuels et à venir. "

En somme, le développement est un concept très vaste et multidimensionnel, et il n'y a pas de consensus sur sa définition exacte, notamment parce qu'on le conçoit souvent dans une logique téléologique de rattrapage (les pays en développement doivent rattraper les pays développés) et parce que les dimensions de pouvoirs entre pays donateurs et pays bénéficiaires et l'histoire difficile, souvent liée à la colonisation, de l'aide au développement ne sont pas simples. On peut toutefois, synthétiser qu'il fait référence à une variété d'interventions qui ont pour finalité d'améliorer les conditions de vie sur le long terme en favorisant :

- la promotion des investissements et de la croissance ;
- le développement des institutions fiables et efficaces ;
- la promotion de la liberté, des droits, l'inclusion et l'équité ;
- la protection de l'environnement.

La notion de développement traduit le niveau qualitatif des conditions de vie d'une population donnée.

1.2. Instruments de gestion du développement

Plusieurs instruments de gestion des interventions visant le développement sont répandus dans les pays dits « en développement ». Il est évident que le vocabulaire et les pratiques ne sont pas uniformes dans tous les pays, mais on peut retrouver des similitudes, notamment dans les pays d'Afrique francophone.

1.2.1. *Le plan de développement*

Un plan est un modèle systématique qui est conçu avant de mettre en œuvre une action de telle sorte qu'elle puisse être ciblée aux fins attendues. Le plan fixe les intentions et les directives d'une vision de développement.

Un plan est donc un instrument de gestion qui aide à promouvoir le développement en fixant une vision claire sur laquelle toutes les actions devront se greffer. Les plans de cette nature visent donc à faire l'inventaire des attentes des populations en matière de développement, pour concevoir ensuite des actions visant à améliorer leur qualité de vie.

Même si des organisations non gouvernementales ou internationales peuvent travailler pour le développement d'une société, le concept de plan de développement s'apparente généralement à l'action de l'État et à ses politiques et stratégies.

Le plan de développement comprend une approche stratégique d'avenir dans la mesure où il cherche à offrir une vision sur le temps long. Ainsi, les plans doivent être durables, en apportant des améliorations visibles dans la société même une fois le plan terminé.

En ce sens, le plan de développement implique un engagement politique pour être mené à bon port et requiert un investissement conséquent, permettant de le concrétiser en actions concrètes.

1.2.2. *Le programme*

La notion de programme renvoie généralement à un ensemble (portefeuille) de projets visant une même finalité ou but. Le programme est aussi la déclinaison et l'opérationnalisation d'un plan en axes stratégiques d'intervention, en objectifs opérationnels. Il est ensuite constitué d'actions orientées vers l'atteinte d'un objectif global commun de développement.

Un programme est plus orienté vers l'atteinte d'un objectif à moyen ou long terme, plusieurs programmes peuvent donc se succéder sur la durée d'un plan de développement.

Même si on parle ici de programme, il est possible que d'autres termes puissent être utilisés pour faire référence à ce type de document (stratégie, politique sectorielle, plan d'action, etc.)

1.2.3. *Le projet*

Selon le Project Management Institut (PMI), un projet est une action réalisée une seule fois, dotée d'un début et d'une fin déterminée et qui vise à créer un produit ou un service unique. Le projet est l'un des instruments principaux de l'action publique (Hirschman, 1967). Pour Pierre-Marie Aubert, un projet de développement peut être défini comme une organisation temporaire, dotée de moyens dédiés – humains, financiers, techniques – fonctionnant comme un dispositif à double détente : sa mise en œuvre vise des objectifs propres et est censée servir de levier de transformation de l'action publique à plus grande échelle (Aubert, 2014 : 220). Le projet est un processus de mise en œuvre d'activités spécifiques et temporaires ayant pour but la production de biens livrables uniques permettant l'atteinte de résultats précis.

Ainsi, loin d'être une fin en soi, le projet est un moyen pour réaliser efficacement un changement permettant à un demandeur), d'atteindre durablement des objectifs prédéterminés. Il est un ensemble complexe d'activités visant un objectif précis et connu au départ, dont l'atteinte est vérifiable objectivement parce qu'elle correspond à la réalisation d'un produit. Le projet se caractérise par son unicité, sa durée limitée, la multiplicité et la variété des intervenants dans son exécution et par l'exigence de respect des contraintes de qualité, de coûts et de temps.

Un programme de développement va donc regrouper un certain nombre de projets qui vont tous concourir à la réalisation des objectifs généraux du programme.

II. Concepts de suivi et évaluation en contexte de développement

La gestion des interventions de développement est un sujet d'étude et de débats, avec en toile de fond la nécessaire attention à la performance de ces interventions pour les populations bénéficiaires, à la fois pour les Etats des pays en développement dont les ressources sont limitées, mais aussi pour les organisations et pays donateurs de l'aide au développement.

A l'heure actuelle, un consensus a été atteint depuis la Déclaration de Paris sur l'efficacité de l'aide au développement, pour adopter un mode de gestion axé sur les résultats, qui a d'ailleurs été défini comme un des piliers de la déclaration. Parallèlement, de nombreuses administrations adoptent un mode de gestion axée sur les résultats pour leur propre action, ce que l'on retrouve notamment dans les réformes liées plus largement au « budget-programme ». On retrouve également la gestion axée sur les résultats au cœur de la formulation des Objectifs de Développement Durable (ODD) qui structurent nombre d'actions de développement pour la période 2015-2030, et sur lesquels les prochains chapitres reviendront. Programmation, suivi et évaluation des interventions sont trois disciplines interreliées qui font partie intégrante de la gestion axée sur les résultats. La section suivante est consacrée aux définitions du suivi et de l'évaluation et la programmation sera abordée dans le chapitre 3.

2.1. Définition du suivi

Dans tous les domaines de l'activité humaine, le suivi s'impose au quotidien plus particulièrement au sein des organisations. Pour mieux comprendre la notion de suivi dans le domaine spécifique du management des projets de développement, on peut retenir la définition suivante : pour le PNUD (2008), « le suivi est un processus itératif de collecte et d'analyse d'informations pour mesurer les progrès d'un projet au regard des résultats attendus. Il fournit donc aux gestionnaires, un retour régulier d'informations pouvant aider à déterminer si l'avancement du projet est conforme à la programmation ».

Cette définition insiste sur la **collecte, le traitement et l'analyse de l'information** par le système de suivi. En effet, le rôle du suivi est d'assurer un contrôle continu et systématique des activités et des résultats d'un projet par la surveillance, la vérification et la maîtrise du processus de mise en œuvre tout au long de son exécution. Il s'agit de vérifier si les ressources humaines, matérielles et financières mises en place sont bien définies, administrées et judicieusement utilisées. Dans ce contexte, l'objectif principal du suivi est de constater les anomalies, tirer la sonnette d'alarme et attirer l'attention des décideurs du projet, lorsqu'il y a déviation ou dérapage par rapport aux buts initiaux et aux incidences désirées, afin qu'ils puissent y apporter des solutions pour un nécessaire réajustement.

Il existe plusieurs types de suivi dans le management des projets, mais leur regroupement permet de faire ressortir deux grandes catégories à savoir, le **suivi des réalisations** et le **suivi du déroulement**

2.1.1. Suivi des réalisations (Délais- Coûts- Avancement)

Le suivi des réalisations consiste à vérifier la manière dont les ressources du projet sont employées en se référant au budget alloué et au calendrier prévu. Il vise également à savoir si les résultats sont obtenus dans les délais et s'ils tiennent compte de l'efficacité et de l'efficience dans la gestion. A travers le suivi des réalisations, il s'agit d'identifier et de corriger rapidement les problèmes rencontrés dans la mise en œuvre du projet. Le suivi des réalisations est lui-même subdivisé en trois catégories de suivi à savoir : le suivi des délais, des coûts et le suivi des aspects physiques.

Le suivi des délais consiste à veiller à ce que l'ensemble des activités du projet soient réalisées à temps, dans les délais impartis et selon le calendrier prévu. Son objectif principal est de mesurer les écarts (lorsqu'ils se produisent) entre les délais réels et ceux planifiés, et d'essayer d'y apporter des explications aussi bien sur les causes que sur les conséquences que cela pourrait avoir sur le début et l'achèvement des autres activités.

Le suivi des coûts sert à vérifier si l'exécution des activités du projet est réalisée en suivant la ligne budgétaire qui avait été prévue. Il permet de comparer tout au long de l'exécution du projet, les coûts effectifs avec les coûts prévus et relever les écarts à

l'attention des décideurs, expliquer les causes et proposer des mesures correctives pour éviter ainsi tout dérapage de coûts.

Pour ce qui est du suivi des aspects physiques, il s'occupe de l'état d'avancement des réalisations physiques du projet. Il couvre trois aspects essentiels à savoir : le suivi des ressources, le suivi des approvisionnements et le suivi de la conformité physique des réalisations.

- ▶ **Le suivi des ressources** : Le suivi des ressources concerne l'élaboration et le suivi des procédures de mise en œuvre qui utilisent de façon optimale les ressources disponibles. Son rôle principal est de veiller au planning d'utilisation des ressources, de dénoncer le cas échéant et à temps, toutes les déviations et les erreurs de gestion constatées et ceci par comparaison au planning initial.
- ▶ **Le suivi de la conformité physique des réalisations** : Le suivi de la conformité physique des réalisations cherche à vérifier et s'assurer que l'exécution du projet s'effectue suivant les règles de l'art, tout en respectant scrupuleusement ce qui avait été prévu dans le cahier de charge, en conformité avec les normes de qualité.
- ▶ **Le suivi des approvisionnements** : Le suivi des approvisionnements permet de vérifier que tous les intrants nécessaires à la réalisation sont fournis et mis à la disposition du projet en quantité suffisante, en qualité et suivant le niveau de coûts prévu tout au long de l'exécution du projet.

2.1.2. Suivi du déroulement (processus)

Le suivi du déroulement permet d'examiner et de vérifier le degré d'efficacité des méthodes et outils utilisés dans le suivi de la réalisation du projet. Il mesure les moyens par lesquels les objectifs sont atteints ; ceci inclut l'utilisation des données saisies, l'information sur le progrès des activités, le **suivi des risques** et la façon dont les activités sont menées.

En effet, pour assurer un bon suivi, il faut que les méthodes et outils utilisés permettent de collecter et de générer une information pertinente, capable d'éclairer les décideurs dans leur processus décisionnel.

Le suivi du déroulement étudie également l'attitude des bénéficiaires tout au long du projet, ainsi que la qualité du produit et /ou du service fourni(s), il est aussi question de voir comment l'environnement externe affecte la mise en œuvre normale du projet.

2.2. Définition de l'évaluation en contexte de développement

Dans le domaine spécifique du développement et suivant les plus récentes définitions approuvées internationalement par les spécialistes du domaine, l'évaluation est perçue comme « **une appréciation aussi systématique et objective possible de la conception, la mise en œuvre et des résultats d'un projet en cours ou achevé. Elle sert à déterminer la pertinence des objectifs du projet ainsi que leur degré de réalisation, son efficacité, son impact et sa viabilité au regard du développement** » (Comité d'aide au développement - CAD/OCDE, 1999). Une évaluation devrait donc fournir des informations crédibles et utiles, permettant d'intégrer les renseignements dans le processus de prise de décision tant par les bénéficiaires que par les promoteurs d'un projet.

Plus précisément, l'évaluation est un jugement argumenté et rigoureux porté à intervalles réguliers sur un certain nombre de questions de performance (aussi appelés critères), notamment la pertinence, l'efficacité, l'efficience, ou la viabilité d'une intervention au regard des objectifs énoncés lors de la planification. Le mot « intervention » est ici utilisé pour parler des différents objets d'évaluation, qui au sens de la classification présentée à la section précédente peut être un plan, un programme ou un projet.

Plus qu'un simple jugement de valeur, la réalisation d'une mission d'évaluation obéit donc au respect d'une démarche rationnelle basée sur des critères et des indicateurs sur le terrain.

La démarche d'évaluation demande donc un temps préalable et un travail déterminant pour la suite, de définition précise de l'objet à évaluer, des critères à utiliser, de choix des questions à poser, et de définition d'une stratégie, c'est-à-dire de la marche à suivre, des données à rassembler et des personnes à associer ou à consulter.

L'évaluation ne cherche pas forcément à débusquer des écarts à la norme, mais souhaite plutôt émettre un jugement sur la performance d'une intervention, ce qui

nécessite souvent de mettre en évidence tous les effets attribuables à l'objet évalué, attendus ou inattendus, qualifiés de positifs ou de négatifs. Si cette analyse met en évidence des faiblesses ou des dysfonctionnements, l'évaluation doit être accompagnée de recommandations pour les faire disparaître. Les réussites notables et les pratiques particulièrement adaptées doivent aussi être mises en valeur.

L'évaluation possède quatre grandes finalités qui intègrent les spécificités de l'objet à évaluer et les besoins en information des parties prenantes. Dans une même mission, ces quatre finalités peuvent se combiner (Rapport Banque mondiale, 2008) :

- ▶ **Décider** : l'évaluation permet au responsable de l'intervention de prendre une décision dans la mesure où elle apporte les éléments qui l'éclairent et l'orientent dans sa prise de décision (évaluation sommative ou décisionnelle).
- ▶ **Gérer** : l'évaluation apporte aux gestionnaires et opérateurs de l'intervention les informations nécessaires pour une bonne orientation de la gestion ainsi qu'un bon contrôle des opérations sur le terrain (évaluation formative ou managériale).
- ▶ **Informier et rendre compte** : L'évaluation permet d'informer les bénéficiaires sur l'évolution de l'intervention, de faire connaître aux contributeurs et financeurs les résultats atteints en adéquation avec les objectifs de départ.
- ▶ **Savoir, comprendre et tirer des enseignements** : en comparant les observations effectuées au cours de l'évaluation d'une intervention avec les constats réalisés dans d'autres, l'évaluateur peut explorer les raisons de cette variation en essayant d'expliquer les écarts, de les comprendre et enfin, d'en tirer des enseignements dans le sens de l'application dans des expériences futures (évaluation éducationnelle ou recherche évaluative).

Selon ces différents buts des missions d'évaluation, ainsi que selon un certain nombre d'autres paramètres (les acteurs et leurs rôles, l'objet d'évaluation et les moyens techniques de l'équipe d'évaluation), de multiples formes d'évaluations sont donc réalisées. C'est souvent aux « équipes d'évaluation » de diagnostiquer la situation qui se présente pour choisir le type d'évaluation le plus approprié.

On distingue par exemple des formes différentes d'évaluation selon le moment du cycle de vie de l'intervention auquel on se trouve, notamment :

- ▶ **l'évaluation prospective / ex ante / a priori** : elle est réalisée avant la mise en œuvre de l'intervention et peut se prononcer sur les chances de réussite d'une intervention, sur sa pertinence, et sur les meilleures pistes d'action.
- ▶ **L'évaluation à mi-parcours ou in itinere** : elle est réalisée en cours de mise en œuvre de l'intervention et est généralement centrée sur l'amélioration continue de celle-ci, puisque des recommandations d'évaluation peuvent être directement mises en place pour augmenter la performance d'intervention.
- ▶ **L'évaluation finale / ex post / a posteriori** : elle est réalisée à la fin de la période de mise en œuvre de l'intervention ou après celle-ci et est généralement centrée sur l'étude des effets de l'intervention, sur la reddition de compte et sur la prise de décision pour la poursuite de l'action sur la thématique.
- ▶ **L'évaluation en continu** : cette forme d'évaluation, plus minoritaire, mais utilisée de plus en plus fréquemment, notamment lorsque des équipes incluent un chargé de suivi-évaluation, consiste à réaliser des travaux évaluatifs dès que le besoin s'en fait sentir et non de manière périodique. Cela conduit à ne pas rédiger d'importants rapports d'évaluation, mais à informer en temps réel la prise de décision.

Enfin, l'évaluation diffère grandement selon les questions qu'elle se pose, c'est-à-dire les critères de performance choisis. En effet, il n'y a pas de liste de questions « obligatoires » à se poser dans une évaluation, mais un angle est à chaque fois développé sur chacune des interventions. De nombreuses variations des définitions des critères existent, et il s'agit surtout de personnaliser les questions d'évaluation à chaque situation, et de se référer aux cadres d'évaluation de chaque mission. Les critères de performance et leurs définitions révisés en 2019 par le CAD/OCDE sont par exemple très utilisés :

- ▶ **Critère de la Pertinence : l'intervention répond-elle au problème ?**
Mesure dans laquelle les objectifs et la conception de l'intervention correspondent aux besoins, aux politiques des personnes affectées, du pays, de la communauté internationale et des partenaires. Le but recherché c'est non seulement d'examiner les différences et les arbitrages entre des priorités ou des besoins différents, mais également d'analyser toute évolution du contexte afin de déterminer dans quelle mesure l'intervention peut être adaptée.

- ▶ **Critère de la Cohérence : l'intervention s'accorde-t-elle avec les autres interventions menées ?**
Mesure dans laquelle l'intervention est compatible avec les autres interventions menées au sein d'un pays, d'un secteur ou d'une institution. Il a pour but d'examiner comment d'autres interventions renforcent ou affaiblissent l'intervention évaluée, et inversement.

- ▶ **Critère de l'Efficacité : l'intervention atteint-elle ses objectifs ?**
Mesure dans laquelle les objectifs et les résultats de l'intervention ont été atteints, ou sont en train de l'être, y compris les résultats différenciés entre populations. Il doit permettre l'analyse des objectifs tout au long de la chaîne des résultats.

- ▶ **Critère de l'Efficience : les ressources sont-elles utilisées de manière optimale ?**
Mesure dans laquelle l'intervention produit, ou est susceptible de produire, des résultats de façon économique et dans les temps. Il vérifie si les réalisations et résultats ont été atteints de la façon la plus économiquement avantageuse possible et dans le respect des délais.

- ▶ **Critère de l'Impact : quelle différence l'intervention fait-elle ?**
Mesure dans laquelle l'intervention a produit, ou devrait produire, des effets importants et de vaste portée, positifs ou négatifs, intentionnels ou non. Le

critère d'impact porte sur l'importance ultime et les effets potentiellement transformateurs de l'intervention. Le but visé consiste à déterminer quels sont ses effets sociaux, environnementaux et économiques à plus long terme ou à plus vaste échelle que ceux déjà évalués au titre du critère d'efficacité. Au-delà des résultats immédiats, le critère d'impact vise à évaluer les conséquences indirectes, secondaires et potentielles de l'intervention, à travers l'examen des évolutions globales et durables des systèmes ou des normes, ainsi que des effets potentiels sur le bien-être des populations, les droits humains, l'égalité des sexes et l'environnement.

▶ **Critère de la Viabilité : les bénéfices seront-ils durables ?**

Mesure selon laquelle les bénéfices nets de l'intervention perdureront ou sont susceptibles de perdurer. Il vise à examiner les capacités financières, économiques, sociales, environnementales et institutionnelles des systèmes nécessaires à la continuité des bénéfices nets dans le temps à travers l'analyse de la résilience, des risques et des arbitrages potentiels entre priorités.

D'autres organisations actives en évaluation, notamment l'UNICEF, proposent d'autres critères qui peuvent être ajoutés dans la boîte à outils des évaluateurs :

- ▶ **Équité** : principe fondamental de justice dans les processus décisionnels et leurs résultats. Pour l'UNICEF, cela implique que tous les enfants aient le droit de survivre, de se développer et de s'épanouir pleinement, sans faire l'objet de discriminations, de préjugés ou de favoritisme ;
- ▶ **Égalité entre les sexes** : pour l'UNICEF, cela signifie de promouvoir l'égalité des droits des femmes et des filles, et de soutenir leur pleine participation au développement politique, social et économique des communautés dans lesquelles elles vivent ;
- ▶ **Approche de la programmation fondée sur les droits de l'homme** : les cinq principes fondamentaux sous-tendant cette approche peuvent être utilisés comme critères d'évaluation : la normativité, la non-discrimination, la participation, la transparence et la responsabilité.

- ▶ **Couverture** : la nécessité d'atteindre des catégories importantes de la population victimes de souffrances qui mettent leur vie en danger, quel que soit le lieu où elles se trouvent, pour leur apporter une aide et une protection qui soient à la mesure de leurs besoins et exemptes de toute considération d'ordre politique inutile ;
- ▶ **Coordination** : la nécessité pour les différentes parties prenantes d'une intervention d'urgence de coordonner leur action ; l'intervention d'une seule organisation ne peut être évaluée indépendamment, notamment parce que les mesures qui peuvent paraître appropriées du point de vue d'un acteur ne le seront pas forcément du point de vue du système dans son ensemble ;
- ▶ **Protection** : la fourniture de services de sécurité et de protection aux membres de la population cible (p. ex. protection contre des éléments armés opérant dans la zone d'intervention, ou fourniture de services de sécurité à l'intérieur d'un camp de réfugiés ou de personnes déplacées) est essentielle pour garantir l'efficacité de l'action humanitaire et des secours ;
- ▶ **Cohérence** (des politiques) : la nécessité de s'assurer que les politiques concernant l'aide humanitaire et les politiques en matière de sécurité, de développement, de commerce et de défense militaire sont compatibles.

III. Différences et complémentarités entre le suivi et l'évaluation et leur lien avec les fonctions connexes

Le suivi et l'évaluation sont deux fonctions différentes, mais complémentaires. L'évaluation ne peut être facilement réalisée que s'il y'a eu au préalable un suivi rigoureux. En effet, l'évaluation s'effectue en s'appuyant sur l'exploitation des informations générées par les instruments mis en place par le système de suivi. Sans ces informations, l'évaluateur se retrouve souvent à devoir faire un travail « d'archiviste » pour reconstituer les données de suivi et perd du temps précieux. Dans les interventions où le système de suivi est faible ou inexistant, l'évaluation reste possible, mais elle sera plus longue et peut-être plus limitée. L'évaluation dans ce cas sera certainement l'occasion de recommander la mise en place d'un dispositif de suivi, basé sur des critères et indicateurs pertinents et qui pourra alimenter de futures évaluations. Dans ce sens, et conformément à la vision de la gestion axée sur les résultats comme un cycle, l'évaluation peut donc également aider à mettre en place un bon suivi.

Toutefois, effectuer le suivi d'une intervention n'équivaut pas à en faire l'évaluation et inversement, et les deux disciplines restent distinctes en ce qui concerne leurs objectifs, les principales activités qui les composent et la fréquence de réalisation. Elles coexistent aussi avec d'autres disciplines qui ont à cœur la performance des interventions publiques, notamment l'audit, le contrôle ou la vérification.

Tableau 1.1: Tableau comparatif des fonctions de suivi, évaluation, audit et contrôle

	Suivi	Evaluation	Audit / contrôle
Objectifs	<p>Améliorer l'efficacité, modifier le plan ou l'affectation des ressources</p> <p>Clarifier les objectifs et leur transformation en indicateurs de performance</p> <p>Comparer régulièrement les réalisations</p> <p>Communiquer les progrès aux responsables et les alerter sur les difficultés</p>	<p>Examiner différents critères de performance opérationnelle et stratégique :</p> <p>efficacité, efficacité, efficacité, etc.</p> <p>Examiner les relations causales conduisant des activités aux effets, expliquer pourquoi certains résultats attendus n'ont pas été atteints</p>	<p>Garantir le respect des normes de gestion (notamment financières)</p> <p>Mettre fin aux pratiques non conformes</p> <p>Audit de performance : s'assurer d'une mise en œuvre efficace (niveau opérationnel)</p>
Principales activités	<p>Définition des indicateurs, recueil régulier d'informations, comparaison avec plan, comptes rendus</p>	<p>Appréciation basée sur la collecte et l'analyse de données probantes, mesure des effets et recherche des causalités</p>	<p>Contrôle de la conformité aux normes, identification des irrégularités et recommandations de changements</p>

	Suivi	Evaluation	Audit / contrôle
Fréquence	Périodique : journalier, mensuel, trimestriel ...selon les variables et les interventions	Généralement épisodique, en début de programme, à mi-parcours ou à la fin	Episodique selon des normes de gestion ou de façon inopinée

Sources : Banque mondiale.

On peut aussi établir des similitudes entre l'évaluation et la recherche en sciences sociales, notamment parce que l'évaluation se sert des méthodes de recherche pour développer un argumentaire rigoureux. Malgré ce point commun, l'évaluation se doit cependant de fonctionner dans le cycle des politiques publiques, et donc être faite plus rapidement que la recherche, ce qui nécessite de trouver des méthodologies rigoureuses de collecte et analyse de données, tout en restant dans un échéancier et un budget raisonnable. Selon les besoins, on peut parfois, face à des situations complexes, prendre plus de temps pour l'évaluation, dès lors la distinction avec la recherche devient floue, et on peut parler de recherche évaluative.

IV. Ethique et évaluation

L'éthique est un ensemble de valeurs et de croyances qui déterminent des choix de comportements. Puisque le suivi et l'évaluation sont des disciplines qui ont le pouvoir d'influencer l'action publique, plusieurs défis éthiques peuvent se poser. Un sondage mené par l'American Evaluation Association a par exemple, identifié plusieurs problèmes liés à l'éthique (Morris & Cohn, 1993). En voici quelques-uns :

- une pression exercée par les parties prenantes pour altérer les conclusions ;
- avant l'évaluation, les parties prenantes ont déjà décidé de ce qui « devrait être » ;
- les parties prenantes suppriment ou font fi des conclusions ;
- l'évaluateur est complètement réticent aux conclusions ;
- l'évaluateur découvre un comportement illégal, non conforme à l'éthique ;
- L'évaluateur ne présente pas toutes les garanties de son objectivité.

Pour répondre à ces défis, plusieurs organisations ont donc développé des normes et des principes directeurs pour encadrer ces dérives, particulièrement en évaluation. Leur but est de promouvoir la redevabilité, faciliter la comparaison, augmenter la fiabilité et la qualité des services. Somme toute, les normes et principes directeurs dépendent des organisations dans lesquelles les professionnels agissent, ou des réseaux auxquels ils adhèrent. Quelques principes de base sont partagés sur le rôle de l'évaluation dans la gestion de l'aide comme l'impartialité et l'indépendance des professionnels, la crédibilité, la rigueur et l'utilité de leurs travaux, la participation des parties prenantes et leur coopération.

Le Groupe des Nations Unies pour l'évaluation (GNUE), qui se compose de praticiens, a notamment entrepris de définir des normes, afin de contribuer à la professionnalisation de l'évaluation et de fournir une orientation aux bureaux d'évaluation chargés d'élaborer des politiques d'évaluation ou d'autres aspects de leurs opérations. Ils sont souvent utilisés par des organisations internationales et des gouvernements qui les adaptent également à leur réalité.

Ces normes au nombre de treize (13) régissent le mode de fonctionnement de l'évaluation applicable dans le système des Nations Unies.

- ▶ **La définition de l'évaluation** : l'évaluation a pour objet d'apprécier de manière aussi systématique et impartiale que possible, une activité, un projet/programme, une stratégie, une politique. Elle est axée sur le principe du rapport des résultats escomptés sur les résultats effectivement obtenus.
- ▶ **La responsabilité de l'évaluation** : elle incombe aux organes directeurs et/ou des chefs d'entités des Nations Unies qui s'assurent que la fonction de l'évaluation est clairement prise en compte.
- ▶ **La politique** : il s'agit de définir le cadre institutionnel, les rôles et les responsabilités d'évaluation sur tout projet /programme.
- ▶ **L'intentionnalité de l'évaluation** : la planification de l'évaluation doit être partie intégrante de la planification globale du projet/programme.
- ▶ **L'impartialité de l'évaluation** : l'évaluation doit être réalisée avec l'ensemble des parties prenantes au projet/programme et sans parti pris de sorte à renforcer la crédibilité de l'évaluation, lui conférant ainsi une plus forte légitimité.
- ▶ **L'indépendance** : la fonction d'évaluation doit être à ISO positionnement des autres fonctions de la gestion d'un projet/programme de sorte à la mettre à l'abri de toute influence.
- ▶ **L'évaluabilité** : il faut que le sujet à évaluer reflète des intentions claires avec des indicateurs mesurables et des sources d'informations fiables.
- ▶ **La qualité de l'évaluation** : les rapports d'évaluation doivent présenter de façon complète et équilibrée les éléments de preuve, les constatations certifiées, les conclusions et recommandations concrètes et faciles à comprendre.
- ▶ **Les qualités requises pour réaliser une évaluation** : les évaluateurs doivent posséder les qualités de base requises pour réaliser des études d'évaluation tout en gérant les évaluateurs contractuels externes engagés sur le projet/programme.
- ▶ **Transparence et consultation** : les mandats d'évaluation et les rapports qui en découlent doivent être accessibles aux parties prenantes et au public de sorte à garantir une transparence du processus évaluatif.

- ▶ **L'éthique de l'évaluation** : les évaluateurs doivent mener leurs actions en se référant à la Déclaration universelle des droits de l'homme, garantissant la prise en compte des réalités contextuelles ainsi que la non-discrimination.
- ▶ **Le suivi évaluation** : l'application des recommandations de l'évaluation devrait faire l'objet d'un rapport périodique adressé aux organes directeurs de l'organisation (secrétariat général, présidence)
- ▶ **La contribution à l'acquisition des connaissances** : la façon de réaliser les évaluations tout comme la présentation des constatations et recommandations qui en découlent doivent être aisément compréhensibles par le public cible.

Un autre ensemble de normes est celui développé par l'Association Africaine d'Evaluation (AfrEA), fondée en 1999, en réponse à l'appel croissant de l'Afrique pour le plaidoyer, le partage d'informations et le renforcement des capacités de pointe en matière d'évaluation. Il faut noter que les 20 dernières années ont d'ailleurs vu l'essor des associations d'évaluateurs, souvent nationales, pour ces mêmes raisons. L'AfrEA est donc un réseau qui regroupe les associations nationales africaines. On peut également noter l'importance du Réseau francophone de l'évaluation (RFE) qui est lui aussi actif auprès des évaluateurs de langue française. Les normes de l'AfrEA sont construites à partir de 4 principes directeurs ; l'utilité, la faisabilité, la protection et la précision.

Les principes directeurs de **l'utilité** sont destinés à assurer que l'évaluation réponde aux besoins en informations des futurs utilisateurs et que les parties prenantes se les approprient.

Les principes directeurs de **faisabilité** sont destinés à assurer que l'évaluation est réaliste, prudente, diplomatique et économique.

Les principes directeurs de **propriété** sont destinés à assurer que les évaluations sont conduites légalement avec éthique et avec une attention particulière pour le bien-être de ceux qui sont impliqués et touchés par les résultats.

Les principes de **précision** sont destinés à assurer qu'une évaluation révélera et apportera une information techniquement adéquate autour des caractéristiques qui déterminent la valeur du programme évalué.

Messages clés :

- ✓ La gestion axée sur les résultats a été mise en place dans les organisations publiques nationales et internationales, pour augmenter la performance de leurs interventions. Le suivi et l'évaluation sont deux disciplines distinctes, mais complémentaires, qui sont intégrées dans ce système plus large de gestion des interventions publiques qu'est la gestion axée sur les résultats.
- ✓ D'autres disciplines telles que l'audit et la recherche peuvent s'approcher du suivi et l'évaluation, et peuvent-elles aussi s'intéresser à la performance des interventions publiques. Toutefois, chaque discipline a ses finalités propres et utilise des méthodologies particulières.
- ✓ Du fait de l'influence du suivi-évaluation sur la prise de décision publique, des normes de pratiques et des principes éthiques sont développés par les organisations qui réalisent des évaluations, et des associations d'évaluateurs se créent également pour encadrer et protéger la pratique.

Bibliographie :

- Comité d'aide au développement - CAD/OCDE. (1999). Évaluer la coopération pour le développement récapitulatif des normes et standards de référence. <https://www.oecd.org/fr/developpement/evaluation/49898837.pdf>
- Donaldson et al. (2013) *Emerging Practices in International Development Evaluation*, IAP, 269p.
- Hirschman A.O., (1967), *Development projects observed*, Washington: Brookings Institution Press
- Mackay, K. (2007). *Comment mettre en place des systèmes de S&E pour améliorer les performances du secteur public*. Groupe Indépendant d'Evaluation du Groupe de la Banque Mondiale. Washington, DC, 190p.
- Morra Imas, L.G. et Rist, R.C. (2009). *Road to Results : Designing and Conducting Effective Development Evaluations*. Washington, The World Bank, 592p. Chapitres: 1 et 2
- Morris, M., & Cohn, R. (1993). Program evaluators and ethical challenges: A national survey. *Evaluation Review*, 17(6), 621-642.
- OCDE. (2011). *Rapport Final de l'Évaluation de la mise en œuvre de la Déclaration de Paris*, Phase 2, 18p.
- OCDE. (2013). *Evaluating Development Activities, 12 Lessons from the OECD DAC*, 55p.
- ONU. (1987). *Notre avenir à tous*, Rapport de la Commission mondiale sur l'environnement et le développement présidé Gro Harlem Brundtland, 811p.
- Page Web des ODD :
<http://www.undp.org/content/undp/fr/home/sdgooverview/post-2015-development-agenda.html>
- Pierre-Marie A., *Projets de développement et changements dans l'action publique*, *Revue Tiers Monde* 2014/4 (n° 220), p. 221-237.
- PNUD (2009). *Guide de la planification, du suivi et de l'évaluation du développement*, 232 p.

MODULE 2

ENJEUX ET TENDANCES EMERGENTES DE L'ÉVALUATION DANS LES PAYS EN DÉVELOPPEMENT

Objectifs d'apprentissage :

A la fin de la session, les participants seront capables de :

- Décrire les déterminants d'une culture de l'évaluation à travers les systèmes de suivi et d'évaluation (SSE) des pays du Nord et du Sud ;
- Cerner les tendances émergentes programmatiques dans les pays en développement et au niveau international et les implications pour l'évaluation.

Thèmes couverts

- I. Enjeux liés à la promotion du suivi évaluation de projets/programmes
- II. Culture de l'évaluation : Aperçu dans les pays de l'OCDE et dans les pays en développement.
- III. Tendances émergentes : Les implications pour l'évaluation.
- IV. L'évaluation des actions de développement : où en sommes-nous en Afrique ?

Le suivi et l'évaluation sont devenus de nos jours, une des préoccupations majeures des organisations tant publiques que privées. Leur importance s'est accrue au fil du temps, au regard de leur pertinence, en tant qu'outils de mesure du progrès et de capitalisation pour les interventions en cours de réalisation et ceux en conception au regard des leçons apprises des interventions exécutées. C'est ainsi que, dans la majorité des organisations, les chercheurs et les praticiens se sont attelés à identifier les méthodes et outils qui peuvent permettre au suivi et à l'évaluation de contribuer efficacement à l'atteinte des résultats. La valorisation de la fonction de suivi et évaluation répond notamment à des exigences de **bonne gouvernance**, de **redevabilité** et d'**efficacité** dans la **production de résultats** de développement.

L'importance croissante du suivi et de l'évaluation des politiques publiques trouve son fondement dans les interrogations sur la pertinence, l'efficacité et la légitimité des interventions publiques. Dans ce contexte, l'évaluation prend sa place dans la conduite de l'action, en tant qu'instrument de promotion de la **bonne gouvernance**.

Le concept de bonne gouvernance, intimement lié au processus d'évaluation de performances en matière de gestion de l'action publique a émergé à la fin des années 80 et dans les années 90 en réponse au constat de l'échec des politiques mises en œuvre dans un grand nombre de pays en développement. Cet échec s'est traduit le plus souvent par un gaspillage de ressources illustré par ce qu'on a appelé les projets « éléphants blancs » et par l'inefficacité des individus, soit des règles non implicites et explicites qui gouvernent la mise en œuvre des interventions.

La bonne gouvernance est ainsi devenue un enjeu central dans les débats sur le financement du développement, avec une attente forte pour le suivi et l'évaluation de l'action publique. C'est ainsi que la question de bonne gouvernance et de suivi et évaluation s'est retrouvée au cœur de l'initiative sur les pays pauvres très endettés (PPTE) et dans les deux générations de Documents Stratégiques de Réduction de la Pauvreté (DSRP) entre 2003 et 2010. Elle se retrouve ensuite dans les différentes Stratégies Nationales de Développement Economique et Social (SNDES) et les Plans Nationaux de Développement (PND), qui constituent le cadre général consensuel de coordination des interventions publiques en Afrique.

Pour répondre à ce défi de performance des interventions de développement, les pays francophones d’Afrique ont engagé des réformes ambitieuses pour adopter une gestion axée sur les résultats au sein de l’administration publique, en accord avec les principes de la Déclaration de Paris sur l’efficacité de l’aide et des recommandations des hauts forums sur le financement du développement (Monterrey, 2002, Busan, 2011, Addis Abeba).

Ces réformes entre autres cherchent à orienter la gestion des finances publiques vers la recherche de la performance.

A partir de 2020, les pays membres de l’UEMOA ont adopté un nouveau système budgétaire appelé budget-programme. Cette nouvelle loi de finances communautaire obéit aux instructions inhérentes au nouveau cadre harmonisé issu des directives de l’UEMOA en matière de réformes et d’harmonisation du cadrage des huit (8) pays de l’Union. En entreprenant cette réforme profonde du budget général qui fait passer le système budgétaire du budget moyen au budget programme, ces pays entendent viser plus d’efficacité dans l’action publique. Ces pays se sont donc lancés sur les chantiers de la transformation structurelle de leur économie, car cette nouvelle forme de budget s’appuie sur des résultats clairs et précis à obtenir en privilégiant la performance, compte tenu des actions et objectifs intermédiaires ou spécifiques intégrant l’ensemble des contraintes.

Mots-clés : culture d’évaluation, développement, tendances dans les pays du Sud

I. Culture de l'évaluation

Qu'elle soit obligatoire ou volontaire, l'évaluation doit être comprise comme un ensemble de pratiques permettant de juger de la valeur des interventions. L'évaluation s'inscrit dans une logique d'aide à la réflexion et à la bonne gouvernance. Elle fait désormais partie des mécanismes de fonctionnement de la gestion axée sur les résultats, en portant une double nécessité d'apprentissage et de redevabilité.

De nature largement participative, l'évaluation stimule le dialogue et l'échange des savoirs entre des acteurs des politiques de développement, de sorte que les représentations et les valeurs de ceux-ci évoluent à mesure que plus d'évaluations sont réalisées. On appelle ce mouvement la culture de l'évaluation.

1.1. L'évaluation dans les pays de l'OCDE

Selon un sondage de Furubo, Rist Sandhal (2002), la plupart ont adopté la gestion axée sur les résultats, des formes différentes, mais l'usage de suivi-évaluation y est souvent systématisé. Au sein de ces pays, l'Australie, le Canada, les Pays-Bas, la Suède et les Etats-Unis ont les niveaux les plus élevés dans le classement en culture de l'évaluation selon ces auteurs.

La culture de l'évaluation repose sur 9 critères essentiels qui sont entre autres :

1. Une évaluation dans plusieurs domaines ;
2. Des évaluateurs multidisciplinaires utilisant des méthodes différentes et variées ;
3. Un discours national sur l'évaluation adapté au contexte national ;
4. Existence des associations professionnelles de l'évaluation participant régulièrement aux rencontres des sociétés internationales pour discuter sur les normes et l'éthique
5. Existence de dispositions institutionnelles au niveau du Gouvernement pour conduire les évaluations et disséminer les résultats ;
6. Existence de dispositions institutionnelles au niveau du Parlement ou autres organes législatifs pour conduire les évaluations et disséminer les résultats;
7. Existence dans chaque secteur des personnes ou des agences chargées de conduire les évaluations ;

8. La mise en œuvre d'activités d'évaluation au sein de l'Institution suprême chargée des audits et contrôles ;
9. L'emphase sur les effets des politiques ou des programmes aussi bien que sur les intrants, les activités et les extrants.

1.2. Evaluation dans les pays en développement

Dans les pays en développement, on assiste depuis près de deux décennies à la manifestation d'exigences minimales pour l'élaboration de système de suivi et d'évaluation, notamment à travers :

- des demandes grandissantes pour la réalisation de travaux de suivi et d'évaluation ;
- la constitution d'un groupe minimum d'acteurs intéressés ;
- un engagement d'acteurs notamment citoyens pour la transparence et la bonne gouvernance ;
- un renforcement de la capacité institutionnelle avec la formation d'évaluateurs ;
- des preuves de volonté politique avec l'inscription du suivi-évaluation dans les missions des parlements, institutions supérieures de contrôle ou d'autres organisations ;
- une coopération renforcée sur les thèmes du suivi-évaluation.

Plusieurs pays ont mis en place des initiatives intéressantes et font de nombreux efforts pour institutionnaliser le suivi et l'évaluation, on peut notamment citer le Maroc ou encore le Bénin¹.

La mise en place de nouveaux systèmes d'évaluation solides dans les pays en développement nécessite cependant encore un accroissement de la volonté de redevabilité de l'État, doté d'institutions plus fortes et plus stables, avec des défenseurs de la gestion axée sur les résultats à de hauts niveaux de responsabilité. Des diagnostics des capacités nationales d'évaluation réalisés par les centres

¹ Site web du Bureau de l'évaluation des politiques publique :
<https://evaluation.gouv.bj/a-propos>

CLEAR dans les pays francophones aussi bien qu'anglophones révèlent que le suivi-évaluation y est encore fortement associé aux interventions des bailleurs de fonds internationaux comme le système des Nations-Unies. Les Etats ne se l'approprient encore que faiblement.

Encadré 2.1 : Institutionnalisation de l'évaluation au Bénin

Le Bénin est l'un des rares pays d'Afrique francophone disposant d'un système national d'évaluation assez explicite et bien structuré. Au nombre des avancées de l'institutionnalisation de l'évaluation au Bénin, il y a la décision N°16/PR/SGG/REL/ORD du conseil des ministres du 2 mai 2018 qui recommande l'élaboration d'une théorie du changement pour tout projet ou programme sous financement public. Aussi, les ministères doivent planifier et mettre en œuvre les évaluations dans le cadre de leurs plans d'action annuels. Ils doivent également prévoir des fonds pour l'évaluation des nouveaux projets et l'évaluation systématique des projets et programmes existants depuis plus de 5 ans sous financement public.

Ce résultat a été le fruit de plusieurs actions initiées depuis 2007. En effet, en juin 2007 il y a eu la mise en place par la Présidence, d'un Bureau chargé de l'Evaluation des Politiques Publiques et rattaché au Cabinet du ministre d'Etat chargé du Développement, du Plan et de la Coordination de l'Action Gouvernementale. Le Bureau de l'Evaluation des Politiques Publiques et de l'Analyse de l'Action Gouvernementale (BEPPAAG) sera par la suite placé sous tutelle du Secrétariat Général de la Présidence de la République.

En 2010, il y a eu une évaluation des capacités nationales en matière d'évaluation qui a conduit à l'élaboration d'une Politique Nationale de l'Evaluation et à la mise en place d'un Conseil National de l'Evaluation. Ce dernier dont le BEPPAAG assure le secrétariat, est composé d'acteurs de l'évaluation provenant du gouvernement, des administrations communales, des Universités et de la société civile.

De plus, de nombreux outils et guides ont été développés pour renforcer la pratique de l'évaluation, ainsi que la diffusion et l'utilisation des résultats. On peut notamment citer :

- Le plan stratégique d'Evaluation ;
- Le guide méthodologique national des évaluations ;
- Le guide méthodologique pour le développement ou la reconstitution de la théorie du changement ;
- Le guide méthodologique des évaluations sensibles au genre ;
- La base de données des évaluations ;
- La plateforme de gestion des connaissances en évaluation ;
- La stratégie de communication sur l'évaluation ;
- Les Journées béninoises de l'évaluation organisées tous les 2 ans depuis 2010.

De plus, une série de collaborations ont été initiées avec des acteurs de l'évaluation issus d'autres pays afin d'apprendre des expériences extérieures pour améliorer le processus national. On peut citer :

- Le programme TWENDE MBELE qui est un partenariat impliquant l'Afrique du Sud, L'Ouganda et le Bénin ;
- Le Programme de renforcement des capacités et d'évaluation d'impact en Afrique de l'Ouest / West Africa Capacity Building and Impact Evaluation (WACIE) visant à promouvoir l'institutionnalisation de l'évaluation dans les systèmes gouvernementaux des pays de l'UEMOA (WACIE) en collaboration avec l'Initiative Internationale pour l'Evaluation d'Impact (3ie).

Enfin, des programmes de renforcement de capacités ont été initiés que soit de façon singulière ou en collaboration avec d'autres parties prenantes afin de garantir la qualité des ressources humaines pour des évaluations mieux élaborées et plus utiles. On peut notamment citer :

- Le camp des évaluateurs émergents ;
- Le programme régional de formation et d'appui à la conduite des évaluations d'impact dénommé (PRéFA-EI) conjointement organisé par le Center For Learning on Evaluation and Results pour l'Afrique francophone (CLEAR - FA) et le WACIE.

1.3. L'Agenda global des Nations Unies et son rôle dans la culture d'évaluation

Au cours de la 55e session de l'Assemblée Générale des Nations Unies, appelée Sommet du Millénaire, un plan d'action pour le XXIe siècle a été élaboré, afin d'inciter les pays membres de faire en sorte que la mondialisation profite à tous. Ce plan d'action dénonce le fait que les progrès économiques des 50 dernières années n'aient pas permis d'éradiquer l'extrême pauvreté, et que de profondes inégalités subsistent et se creusent au sein de certains pays, ainsi qu'entre les pays. Ce rapport donna naissance aux « Objectifs du Millénaire (OMD) », un plan d'action qui fixe huit objectifs à atteindre à l'horizon 2010, puis 2015.

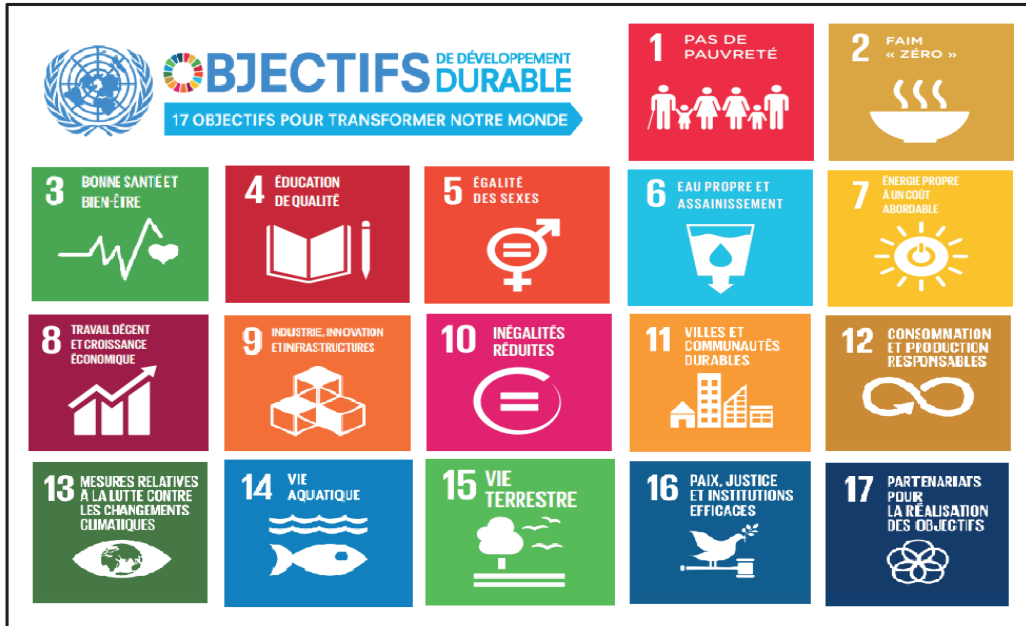
À la suite de l'adoption des OMD et leur mise en œuvre, les gouvernements du monde entier ont constaté au cours des deux dernières décennies, des pressions de plus en plus fortes, dans la mesure où ils se doivent de démontrer qu'ils utilisent efficacement les ressources publiques pour l'atteinte des objectifs de développement. Il en résulte, une prise de conscience que la réalisation d'activités n'est pas nécessairement synonyme de changement et d'amélioration pour les populations ou groupes cibles des pays aidés. Par conséquent, la performance des interventions doit se faire désormais sur la base de résultats escomptés et non sur celle des réalisations.

Les Objectifs de Développement Durable (ODD) à la fois prolongent, remplacent et complètent les OMD, adoptés et mis en œuvre sur la période 2000 - 2015. Les 17 ODD qui sont entrés en vigueur en 2016 pour les 15 prochaines années sont chacun composés de plusieurs cibles, au nombre total de 169 et d'indicateurs associés.

Les ODD offrent une opportunité précieuse pour repenser la pratique de l'évaluation afin de l'adapter aux défis de développement qui se présentent aujourd'hui.

Alors que l'évaluation des progrès accomplis par les pays dans la réalisation des ODD en est encore à ses débuts, ce processus est l'occasion d'améliorer les politiques et les programmes nationaux, mais aussi de fournir des informations aux pairs sur le plan international. Les évaluateurs doivent se montrer innovants afin de valoriser au mieux ce potentiel. C'est en ce sens que les pays en développement sont appelés à renforcer leurs capacités nationales sur les systèmes de suivi et évaluation. Ils bénéficient également d'un accompagnement pour la mise en place de systèmes performants de suivi-évaluation de leurs interventions pour mesurer l'atteinte des ODD.

Figure 2. 1 : Les objectifs du développement durable



Source : Site officiel des ODD : <https://www.un.org/sustainabledevelopment/fr/>

1.4. L'Agenda de l'Initiative Globale d'Évaluation – The Global Evaluation Initiative (GEI)

L'Initiative mondiale pour l'évaluation (GEI) soutient le renforcement des capacités institutionnelles en collaborant avec les pays en développement afin d'améliorer la performance et la durabilité de leurs systèmes de suivi et d'évaluation (S&E). En intervenant à la croisée des institutions, des politiques et des personnes, la GEI aide les gouvernements à intégrer plus efficacement les données probantes dans les processus de planification et d'élaboration des politiques publiques.

Le soutien apporté par la GEI est adapté aux besoins spécifiques de chaque pays et repose sur des diagnostics approfondis, notamment à l'aide d'outils tels que l'Analyse des systèmes de suivi et d'évaluation (MESA). L'initiative vise à réduire la fragmentation, à favoriser la collaboration sur des enjeux transversaux majeurs tels que le genre, le climat et la fragilité, et à fournir des services de conseil qui contribuent à des systèmes nationaux plus cohérents et adaptés aux objectifs.

La GEI encourage également l'innovation en inspirant, en incubant de nouvelles solutions et en accompagnant le déploiement à grande échelle des approches efficaces. À travers la production de connaissances, l'apprentissage global et les partenariats, la GEI promeut l'utilisation concrète des données probantes dans les politiques publiques. Son insistance sur le leadership national et l'adaptation garantit que les systèmes d'évaluation soient inclusifs, réactifs et durables.

II. Thèmes à la une du développement international et leur lien avec l'évaluation

2.1. Des initiatives qui encouragent l'utilisation de l'évaluation

Parmi de nombreuses initiatives entourant l'organisation des interventions de développement qui mettent de l'avant l'évaluation, notamment en lien avec l'Agenda des Nations Unies, on peut citer une initiative africaine née d'un Nouveau Partenariat Mondial pour le Développement (NEPAD) : le Mécanisme Africain d'Évaluation par les Pairs (MAEP). Le MAEP a été créé en 2003 par le Comité des chefs d'État et de gouvernement (HSGIC) chargé de la mise en œuvre du NEPAD comme un instrument d'autoévaluation volontaire de la performance des États membres en matière de gouvernance.

Le MAEP est un accord volontaire conclu entre les États africains pour évaluer et réviser systématiquement la gouvernance à l'échelon des Chefs d'État, afin de promouvoir la stabilité politique, l'intégration économique accélérée, la croissance économique et le développement durable dans la sous-région que partout sur le continent. En adhérant au MAEP, les États membres conviennent d'évaluer indépendamment, la conformité à leurs engagements africains et internationaux en matière de gouvernance.

La performance et les progrès sont mesurés dans quatre domaines thématiques :

1. la démocratie et la gouvernance politique ;
2. la gouvernance et la gestion économiques ;
3. la gouvernance d'entreprise ;
4. le développement socio-économique.

2.2. Emergence de nouveaux acteurs dans l'aide au développement

Au-delà de la mise en place de dispositifs comme le MAEP, on note l'émergence de nouveaux acteurs tels que les organisations de la société civile, les ONG, les organisations communautaires de base qui constituent de nouveaux interlocuteurs des pouvoirs publics dans la formulation des politiques publiques, et qui bénéficient de la reconnaissance des partenaires du développement. Il en est ainsi des acteurs non étatiques (ANE) auprès de l'Union européenne, qui jouissent de pouvoirs sociaux, économiques et financiers qui prennent des proportions importantes depuis plusieurs décennies à l'échelle du monde.

Ce sont notamment :

- Les collectivités territoriales,
- la société civile et leurs regroupements ou associations,
- les parlementaires

Il faut ajouter à ces acteurs les grandes fondations² qui sont également de plus en plus visibles et impliquées dans le dialogue sur les politiques, programmes ou projets globaux, nationaux et sectoriels. Le rapport de l'OCDE « *La Philanthropie privée au service du développement* » indique que les fondations privées ont consacré 23.9 milliards USD au développement au cours de la période 2013-2015, soit 5 % du volume de l'APD. Cependant, les apports de source philanthropique jouent un rôle majeur dans certains secteurs, comme celui de la santé. Elles sont par ailleurs parties prenantes du plaidoyer en faveur de la gestion axée sur les résultats, du fait de leur forte inspiration du secteur privé. Pour elles, le suivi et l'évaluation des résultats sont une évidence et les interventions subventionnées par ces fondations font d'ailleurs généralement l'objet d'un suivi-évaluation rigoureux, ce qui crée des compétences et des habitudes sur les terrains qu'elles occupent.

² Par exemple la fondation Bill et Melinda Gates ou encore la fondation Soros / open Society

2.3. Initiative PPTE

L'initiative en faveur des pays pauvres et très endettés (PPTE), lancée conjointement par la Banque Mondiale et le Fonds monétaire international vise à réduire la dette extérieure des pays les plus pauvres, pour la plupart situés en Afrique. Un certain nombre de critères doivent être respectés par les pays pour faire partie de l'Initiative PPTE, dont plusieurs sont liés à la mise en place de réformes économiques dans la lignée idéologique des ajustements structurels. Ces critères diminuent la marge d'appréciation des créanciers quant aux montants d'allègement souhaitables et permettent de vérifier le respect de leurs engagements. Toutefois, cette initiative a permis de dégager des moyens financiers importants au service des politiques sociales dans les pays bénéficiaires. Elle a favorisé la création d'une certaine capacité d'évaluation puisque les pays bénéficiaires doivent s'engager en faveur de la responsabilité et de la transparence à travers : le suivi, l'évaluation et les réalisations.

2.4. Prévention des conflits, reconstruction post-conflit

Les années 1990 ont été marquées par la multiplication de foyer de tension et de conflit en Afrique sous l'effet conjugué de plusieurs facteurs : chute du mur de Berlin, déclenchement de processus de démocratisation fragilisant les régimes autoritaires, conjonctures économiques et sociales liées aux effets pervers des politiques d'ajustement structurel. La recherche de solutions idoines pour prévenir, gérer les conflits et reconstruire les pays et régions post conflit a fait émerger la préoccupation de gérer efficacement et consciencieusement les politiques, programmes et projets humanitaires et de développement. Ces préoccupations ont fait émerger des notions comme celle du « do no harm » dont l'application nécessite un suivi-évaluation rigoureux des interventions.

Plusieurs préoccupations vont favoriser ce recours au suivi – évaluation :

- le besoin d'analyser l'impact et la coordination de l'action des bailleurs et des programmes multisectoriels financés par la multitude de bailleurs bilatéraux et multilatéraux ;
- la gestion efficace des actions de déminage, de démobilisation et de réintégration d'ex-combattants ;
- la nécessité de trouver les stratégies pour empêcher les conflits d'éclater.

Encadré 2.2 : Le concept du Do no Harm

Le principe du Do no Harm a été développé par l'économiste américaine Mary B. Anderson. Elle s'inspire de l'expérience d'un vaste programme du même nom qui a débuté en 1994 avec la participation de plusieurs institutions d'aide humanitaire ainsi que d'organisations de développement local et/ou international intervenant en situation de crise, de conflit ou de catastrophe naturelle. L'objectif est d'analyser les interactions possibles entre l'aide humanitaire et le contexte de sa mise en œuvre. Partant du constat qu'une intervention humanitaire peut avoir des effets non prévus, parfois néfastes en fonction du contexte, les objectifs des interventions humanitaires en situation de crise ont été revus comme suit : faire si possible un peu de bien ou, au moins, « ne pas nuire » d'où l'appellation (Do no Harm).

De nombreux organismes humanitaires et d'aide au développement comme Oxfam ou la Coopération Suisse s'inspirent aujourd'hui de ce concept pour intégrer davantage la prise en compte du contexte dans leurs interventions de par le monde.

Pour en savoir plus, lire: Mary B. Anderson, Do No Harm: How Aid Can Support Peace-Or War, Cambridge 1999. Pour de plus amples informations sur le « Do no Harm ».

2.5. Démocratie (exigence du contrôle de l'action de l'exécutif par les parlementaires et les élus locaux) et gouvernance

Les années 1990 ont été aussi marquées par le développement du courant institutionnaliste et néo institutionnaliste qui vont concourir à la prise de conscience du rôle des institutions et de l'État dans le développement des programmes de gouvernance visant à promouvoir la lutte contre la corruption, la gestion saine des dépenses publiques, la réforme du service social, la réforme judiciaire, la décentralisation de l'administration.

Elles vont donner lieu à des revendications en faveur de :

- la lutte contre la corruption : Transparence Internationale (TI) « Indice de Perception de la Corruption » (voir site : www.transparency.org)
- la participation politique plus forte des femmes, des jeunes et des groupes vulnérables.

Les bailleurs et les évaluateurs utilisent ces éléments pour mesurer l'efficacité de l'aide. Pour le PNUD, la valeur d'une évaluation de la gouvernance parrainée par les pays réside dans le fait que celle-ci sert de mécanisme fondamental d'obligation de rendre compte et de transparence, qui permet de mesurer les performances en matière de gouvernance. La gouvernance démocratique est d'autant plus susceptible d'être soutenue lorsque les cibles des réalisations et des attentes concernant les choix adoptés par le gouvernement sur la manière de s'y prendre pour atteindre les objectifs fixés sont formulées par ceux-là mêmes qui animent le système politique. Lorsque les acteurs locaux prennent part au processus d'évaluation, une telle participation les engage à réclamer davantage d'informations sur la gouvernance et renforce leurs exigences pour une plus grande obligation de rendre compte de la part du gouvernement.

2.6. Financement de l'action anti-blanchiment d'argent et anti-terroriste

En référence à la problématique de conflits, de la sécurité, de la gouvernance s'est imposée la nécessité de lutter contre les actions tendant à convertir ou transférer les produits de l'activité criminelle. Le Groupe d'Action Financière (GAF) de l'OCDE sur le blanchiment de capitaux est un exemple de mécanisme mis en place.

Le suivi et l'évaluation font partie du mandat du GAF mené au niveau multilatéral, par la revue des pairs et l'évaluation mutuelle.

2.7. Approche genre : de « la place des femmes dans le développement » à « la parité dans le développement » puis à « l'égalité des sexes »

L'évaluation de nos jours s'intéresse à l'analyse de la question du genre : l'accès et le contrôle que les hommes et les femmes ont sur les ressources et la façon systématique de déterminer les besoins, les préférences et l'impact.

Les dispositifs de suivi et d'évaluation qui intègrent les rapports de genre reprennent des critères "classiques" que sont la pertinence, l'efficacité, l'efficience, la viabilité et l'impact. Un autre critère qui est plus récent, mais de plus en plus répandu, et qui est le degré de participation de la population concernée.

L'intégration de l'aspect genre demande une réinterprétation de ces critères, fondée sur une analyse moins centrée sur l'action en soi, mais plus sur les personnes et les processus en marche. En effet, il est intéressant de considérer critère par critère les implications d'une telle réinterprétation (dans une perspective genre). (Suivre et évaluer selon le genre – Genre en action)

Encadré 2.3 : Quelques thématiques émergentes

Outre la question des évaluations sensibles au genre, plusieurs thématiques nouvelles émergent et/ou sont de plus en plus récurrentes.

Par exemple, la question de l'inclusion est étroitement liée à la notion de genre et implique que toutes les couches sociales soient touchées par les projets, programmes et politiques de développement.

Il y a également la notion d'impact Investment visant à créer un changement social positif parallèlement aux rendements financiers, créant ainsi une valeur mixte. La principale différence entre l'impact Investment et les autres formes d'investissements socialement responsables réside dans le fait qu'il vise à obtenir un résultat social positif contrairement aux autres qui ont généralement pour objectif d'amoindrir les impacts sociaux négatifs.

Par ailleurs, le contexte socio politique actuel dans certaines régions notamment africaines fait ressortir la nécessité d'intégrer les questions de terrorisme dans l'élaboration et la conduite des évaluations.

Enfin, la crise sanitaire de la Covid 19 a fait émerger la question de la prise de décision dans un contexte sanitaire aussi délicat que celui imposé par cette pandémie mondiale. Dès lors, plusieurs conférences et autres événements internationaux ont été consacrés à la prise de décisions dans le contexte de la crise Covid-19.

2.8. Développement du Secteur Privé (DSP) et climat des investissements

Compte tenu du rôle fondamental des entreprises, la nécessité de trouver les moyens de mobiliser le savoir-faire, les capacités productives et les ressources financières du secteur privé en faveur du développement figure au premier plan des enjeux de développement. Pour encourager un secteur privé dynamique, les organisations internationales d'aide au développement ont apporté un soutien important aux pays partenaires afin de mettre en place les conditions préalables à son développement, parmi lesquelles un climat des affaires favorable, des investissements dans l'infrastructure et une facilitation de l'entrepreneuriat grâce au financement du développement et à l'assistance technique. Pour promouvoir une croissance inclusive et durable tirée par le secteur privé, les pays en développement doivent privilégier la réduction des coûts du commerce et de l'investissement. Même si la création d'un environnement propice à la croissance du secteur privé incombe aux gouvernements des pays en développement, la coopération pour le développement peut aider ces gouvernements à orienter les activités du secteur privé de façon qu'elles contribuent à une croissance plus inclusive et plus durable pour l'environnement.

Pour faciliter la croissance du secteur privé — et faire en sorte que les personnes les plus défavorisées puissent aussi en bénéficier — il faut mettre en place des réglementations où les nouveaux venus qui sont dynamiques et animés de bonnes idées, quels que soient leur sexe ou leur origine ethnique, peuvent créer leur propre entreprise et où les entreprises peuvent investir, se développer et créer davantage d'emplois.

Le programme Doing Business part du principe fondamental que l'activité économique doit reposer sur de bonnes règles — des règles qui établissent et clarifient les droits de propriété et réduisent le coût de règlement des différends, des règles qui améliorent le caractère prévisible des interactions économiques et qui offrent aux partenaires un contrat la certitude et la protection contre les abus. L'objectif est de mettre en place des réglementations efficaces, accessibles à tous et faciles à mettre en place.

Doing Business accorde ainsi de meilleures notes pour une réglementation qui assure une protection plus solide des droits des investisseurs, à l'instar des normes plus rigoureuses en matière de publication de l'information dans les transactions entre parties apparentées.

Le projet Doing Business a une portée limitée. Il ne cherche pas à évaluer tous les coûts et avantages d'une loi ou réglementation particulière pour la société dans son ensemble. Il ne mesure pas non plus tous les aspects du climat des affaires qui revêtent de l'intérêt pour les entreprises et les investisseurs ou qui ont une incidence sur la compétitivité d'une économie. Son but consiste simplement à fournir aux dirigeants d'entreprise et aux pouvoirs publics une base factuelle pour les aider à prendre des décisions éclairées et offrir aux chercheurs, des données librement accessibles concernant les répercussions de la réglementation des affaires et des institutions sur des résultats économiques tels que la productivité, l'investissement, le secteur informel, la corruption, le chômage et la pauvreté.

2.9. La durabilité environnementale et sociale

De plus en plus, la durabilité environnementale et sociale fait l'objet d'appréciations dans le cadre des évaluations de projets, programmes et politiques, qu'ils soient directement liés au domaine de l'environnement ou non. En effet, certains défis comme les inégalités sociales, les changements climatiques ou encore la déforestation, auxquels l'humanité fait face depuis quelques décennies, ont fait naître la nécessité de rendre les interventions de développement plus sensibles à ces questions afin de garantir un développement durable et inclusif.

- Dans cette perspective, de nombreux concepts comme la RSE ainsi que divers principes et normes de réglementation ont fait leur apparition. Responsabilité Sociale des Entreprises (RSE) : tenir activement compte des conséquences et des impacts économiques, environnementaux et sociaux des activités d'affaires ;
- Les principes d'Equateur (2009) :
 - base de référence et cadre commun, normes de financement des activités ;
 - définition des voies et moyens de déterminer, d'évaluer et de gérer les risques (environnementaux et sociaux) ;
 - engagement à un financement socialement responsable.

→ Les normes couvrent les éléments relatifs :

- à l'environnement ;
- à la santé et la sécurité ;
- aux populations indigènes ;
- aux habitats naturels ;
- à la réinstallation.

En 2020, le Groupe d'Évaluation des Nations Unies et la Société Canadienne de l'Évaluation ont réalisé des bilans sur la mesure dans laquelle les questions environnementales sont prises en compte dans les évaluations. Il ressort que les considérations environnementales sont très peu prises en compte dans l'élaboration des interventions et par conséquent, elles étaient moins susceptibles d'être prises en compte dans les évaluations.

En effet, bien que les critères de l'OCDE incluent d'une manière ou d'une autre les impacts environnementaux, la plupart des évaluations d'impact peinent à prendre en compte cette dimension environnementale et sociale. La raison est que la durabilité environnementale n'est pas souvent un objectif explicite des projets, programmes et politiques (à moins qu'ils soient spécifiquement dans le domaine de l'environnement). Ainsi, les commanditaires des évaluations ont du mal à les intégrer dans les TDR et par la suite, les évaluateurs ont du mal à prendre en compte la dimension environnementale si elle n'est pas incluse dans les TDR. Face à cette situation, les pistes de solutions qui s'offrent aux évaluateurs en dehors des critères d'évaluation de l'OCDE sont :

- Développer des questions d'évaluation clés qui tiennent compte des enjeux environnementaux.
- Faire référence aux engagements environnementaux existants le cas échéant (accords internationaux, politiques ou stratégies existantes, etc.)

2.10. Biens publics, biens communs et contrôle citoyen

Les biens publics sont des biens accessibles à tous dont la quantité disponible pour les autres ne se trouve pas réduite à la suite de la consommation par une personne. En revanche, un bien commun est un bien accessible à tous dont l'utilisation par une personne empêche l'utilisation par une autre personne. La différence entre un bien public et un bien commun tient à leur mode de gestion. Un bien public est administré alors qu'un bien commun fait l'objet d'une gestion soumise à une délibération entre des parties prenantes. L'évaluation des biens publics mondiaux est largement absente en raison du manque d'objectifs clairement définis et d'indicateurs de performance vérifiables.

Encadré 2.4 : Les biens publics et les biens communs

La théorie des biens publics a été développée pour la première fois par l'économiste Paul Samuelson en 1954 dans son article « The Pure Theory of Public Expenditure ». Il identifie deux (02) principes fondamentaux qui caractérisent les biens au sens large. Il s'agit de la notion de rivalité et de celle d'exclusion. La première fait appel à la disponibilité du bien pour tous et la seconde fait référence à l'accessibilité du bien pour tous. Ainsi, un bien public est caractérisé par la non-rivalité et la non-exclusion. C'est-à-dire que la consommation du bien par une personne n'entrave pas la quantité disponible pour toute autre personne, et personne ne peut être exclu de la consommation de ce bien. De ce fait, il est impossible de faire payer pour un bien public.

En maintenant, le principe de non-exclusion et en annulant la non-rivalité, on obtient la catégorie des « biens communs » plus simplement appelés « communs ». Les travaux de Samuelson (1954) ont inspiré de nombreux économistes des temps modernes comme le prix Nobel d'économie 2014 Jean Tirole, grand précurseur du principe des biens communs et notamment auteur de l'ouvrage « Économie du bien commun ». Il y a également le professeur Jean-Michel Servet qui a notamment travaillé sur la monnaie comme un « commun ».

III. Etat des lieux de l'évaluation en Afrique

3.1. Evaluation des actions de développement : Où en sommes-nous en Afrique ?

Les analyses ont identifié plusieurs causes courantes pour expliquer les causes des échecs des politiques publiques dans les pays africains. Il s'agit entre autres de l'insuffisance de liens clairs entre les interventions publiques et les priorités stratégiques énoncées dans les plans de développement, de l'insuffisance d'un management affirmé, de l'insuffisance d'engagement effectif des parties prenantes, du manque d'une approche efficace de gestion des risques et de la faible attention accordée à certains événements affectant les différentes phases de gestion de l'intervention et la livraison des résultats.

Face à ce constat, la grande question qui se pose est la suivante : comment l'Afrique peut-elle accroître la réussite des actions de développement ?

La pratique du suivi et l'évaluation apparaissent comme une des réponses à cette question. Trois raisons militent pour une installation durable de l'évaluation dans les actions de développement.

La première est que l'évaluation participe d'un souci de transparence dans l'usage des fonds publics, qui est l'une des composantes de la démocratie. Cette exigence figure à l'article 15 de la Déclaration des droits de l'homme et du citoyen qui dispose que « la société a le droit de demander compte à tout agent public de son administration ». Cela impose de s'intéresser aux résultats de toute action publique.

La deuxième raison est que les citoyens ont une exigence accrue quant à la qualité de leurs services publics. Ils en attendent davantage d'accessibilité, de simplicité, d'efficacité, de fiabilité, de célérité et de capacité de s'adapter aux situations individuelles. Cela impose une interrogation constante sur l'organisation et les moyens des services publics pour que leur performance puisse être en permanence améliorée. Il s'agit de faire en sorte que l'action publique rime avec efficacité, au sens large du terme.

Enfin, les politiques publiques sont de plus en plus complexes. Elles font intervenir un nombre toujours croissant d'acteurs. Les interventions économiques et sociales génèrent, pour la plupart, des externalités dont la mesure est essentielle pour en apprécier l'efficacité et l'efficience. L'existence de ces externalités nécessite que soient identifiés tous les acteurs impliqués dans la politique et de mesurer ses résultats sur la société et l'économie.

Ces trois exigences, conjuguées avec la nécessité de mieux justifier chaque denier public, rendent durablement nécessaire un recours beaucoup plus généralisé aux évaluations de politiques publiques avec un double mouvement caractérisé par :

- le passage de modèles traditionnels d'évaluation centrés sur la mise en œuvre aux modèles d'évaluation basés sur les résultats ;
- un processus mettant l'accent sur les projets individuels, ou l'approche partielle à une approche plus globale.

3.2. Vers quel type d'évaluation pour les pays en développement ?

La nouvelle vision du développement exige une compréhension plus large des secteurs, des pays, des stratégies de développement et des politiques.

- Nombre croissant de partenaires ;
- La performance des partenaires individuels nécessite maintenant d'être évaluée en fonction de leurs contributions et obligations respectives ;
- L'évaluation est devenue plus difficile à concevoir avec une approche au développement plus contraignante, plus fragmentée et plus participative ;
- Nécessite d'institutionnaliser l'évaluation.

Bibliographie

- Furubo J.E, Rist R.C., Sandahl R. (2002), *International Atlas of Evaluation*, Edition Transaction Publishers, New Brunswick (U.S.A.) and London (U.K.)
- OCDE. (2018). La philanthropie privée pour le développement.

MODULE 3

PROCESSUS DE PLANIFICATION, PROGRAMMATION, SUIVI ET ÉVALUATION EN CONTEXTE DE DÉVELOPPEMENT

Objectifs d'apprentissage :

- Familiariser les participants avec les méthodes et les outils de planification stratégique et opérationnelle en contexte de développement
- Présenter la démarche de planification axée sur les résultats
- Présenter la démarche globale de suivi et d'évaluation

Thèmes couverts

- I. Liens entre la planification et le suivi évaluation en contexte de développement
- II. Planification stratégique axée sur les résultats : Approche cadre de résultat
- III. Méthodes et outils de planification opérationnelle
- IV. Démarche de suivi et d'évaluation

Les chapitres précédents nous ont rappelé que les acteurs au développement sont de plus en plus préoccupés par l'atteinte des résultats. Ils sont soucieux de savoir si les actions, pour lesquelles des ressources ont été dégagées ont apportés les changements projetés. Dans cette dynamique, le suivi et l'évaluation peuvent aider une organisation à tirer des informations pertinentes sur des activités passées et en cours pour en faire la base d'une réorientation des initiatives et ressources pour un meilleur impact. Le suivi et l'évaluation jouent des rôles fondamentaux dans la réalisation des résultats du développement. Mais ils sont précédés dans un cycle de gestion axée par les résultats d'une phase de planification et de programmation.

La planification peut être définie comme un « processus permettant de définir les objectifs d'un projet, d'élaborer les stratégies, de tracer les grandes lignes des dispositions de mise en œuvre et d'attribuer les ressources nécessaires à la réalisation de ces objectifs » (PNUD 2009).

Le présent chapitre décrit le processus de planification des actions de développement ainsi que son lien avec le suivi et l'évaluation. Il fournit entre autres des définitions claires des termes et concepts clés liées à la gestion axée sur les résultats.

Mots-clés : programmation ; planification, GAR.

I. Lien entre la planification le suivi et l'évaluation en contexte de développement

Le suivi et l'évaluation sont au cœur des actions de développement. Les acteurs sont de plus en plus préoccupés pour savoir si les actions pour lesquelles ils fournissent des ressources apportent des changements souhaités. Le suivi et l'évaluation sont aujourd'hui une composante importante de la plupart des projets. Toutefois, sans une réelle planification et une articulation bien définie des résultats espérés, ce qui doit être suivi n'est pas clairement défini. Par conséquent, le suivi ne peut pas être fait correctement. Sans une planification efficace (des cadres de résultats clairs), la base de l'évaluation n'est pas non plus solide et l'évaluation ne peut donc pas être faite correctement.

Dans ce chapitre nous allons aborder la démarche de la planification stratégique pour établir ensuite son lien avec le suivi et l'évaluation.

1.1. Planification stratégique

1.1.1. Définition

Planifier, c'est réfléchir sur le **pourquoi** de l'action ; c'est aussi décider de **ce que** l'on veut faire, **comment** on procédera, **quand** on agira, **quelles** ressources seront nécessaires, à **quel** prix et **qui** accomplira **quoi**.

La planification stratégique est un outil de gestion qui aide à :

- ❖ définir les orientations et les objectifs stratégiques à atteindre dans un horizon temporel donné ;
- ❖ préciser les priorités et, pour chacune d'elles, formuler les interventions à mettre en œuvre et formaliser le cadre de référence pour la gestion axée sur les résultats ;
- ❖ déterminer les ressources nécessaires ;
- ❖ établir un calendrier.

La planification est un processus par lequel une vision de développement est transformée en action pour produire un résultat. Ce processus comprend deux principaux niveaux que sont le niveau stratégique et le niveau opérationnel).

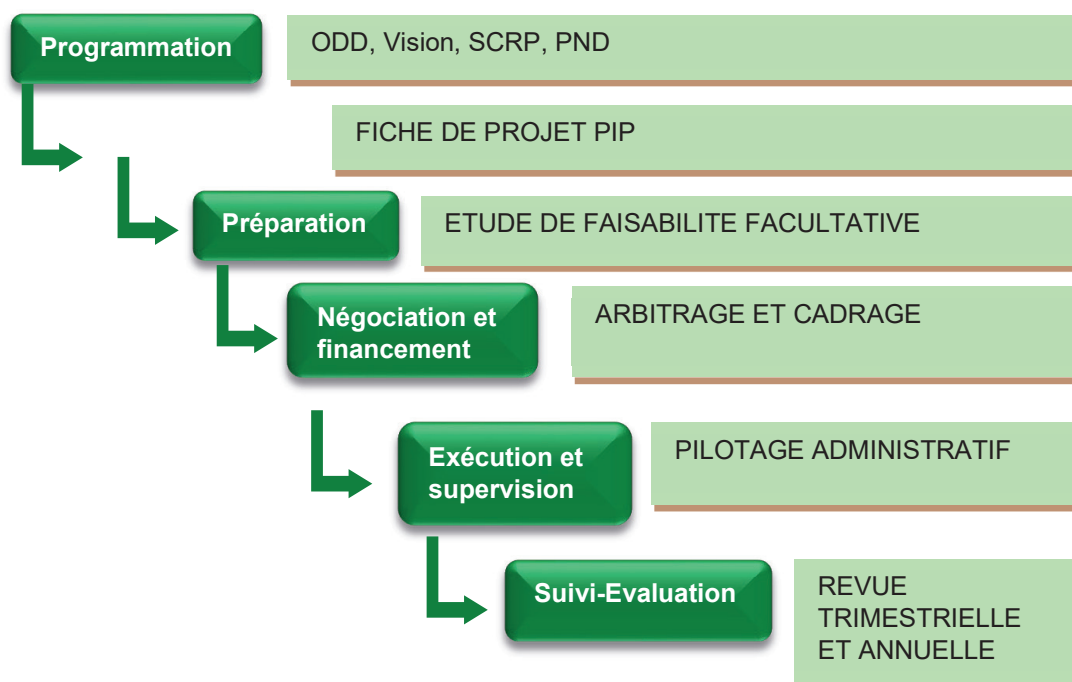
La planification stratégique est orientée sur le long terme et inclut une vision dynamique et intégrée du développement. Celle, dite opérationnelle est axée sur la réalisation des programmes et projets. Les différents niveaux de planification, aussi appelés cadre programmatique se présentent comme suit :

Graphique 3.1 : Les différents niveaux de planification



Ce graphique décrit les différents niveaux de planification selon le critère de l'horizon temporel en distinguant les niveaux politique ou stratégique (long terme avec les documents d'orientation stratégique ou plans nationaux de développement économique et social), sectoriel (moyen terme avec les programmes sectoriels) et opérationnel (court terme avec les programmes et les projets).

Graphique 3.2 : Présentation du cycle programmatique



1.1.2. Démarche

La planification stratégique suit un processus systématique d'analyse au cours duquel on examine :

- l'environnement interne de l'organisation pour déterminer ses forces et ses faiblesses ;
- l'environnement externe pour identifier ses opportunités et menaces ;
- l'élaboration des interventions à travers les objectifs et les moyens pour les atteindre.

Il s'agit de prendre la bonne décision au regard d'une vision, d'une orientation à moyen ou long terme et/ou d'une problématique particulière.

Pour y parvenir, on doit donc tenir compte des aspects suivants :

- ❖ du **contexte** (événement à l'origine de la décision). Il s'agit d'analyser :
 - l'opportunité du problème ou de la crise ;
 - l'envergure de la problématique et
 - la nécessité et/ou l'urgence de réagir ;
- ❖ de la **conjoncture** (événement/facteurs susceptibles d'influencer la décision) ;
- ❖ de l'**environnement social, culturel, économique, concurrentiel**, etc.

Cette exploration sera fonction de (des) :

- la pression interne ou du milieu ;
 - compétences distinctives de l'entreprise ou de l'organisme ;
 - tendances, faits porteurs d'avenir, etc.
- ❖ du **futur** (prospectif) :
 - futur plausible, possible, souhaitable ;
 - types et modalités d'interventions requises ;
 - attentes buts et objectifs à satisfaire.

1.2. Lien entre la conception des projets, la programmation annuelle et le suivi évaluation

Les fondements du lien entre la conception des projets, la programmation annuelle et le suivi-évaluation résultent du cadre programmatique et de la chaîne de résultats. Le cadre programmatique permet de traduire les objectifs politiques et stratégiques en objectifs opérationnels. L'articulation se fait à travers les documents de planification à moyen et court terme notamment dans de nombreuses administrations (on peut retrouver d'autres dénominations) :

- ❖ Le Budget-programme ;
- ❖ Le Plan de Travail Annuel (PTA) ;
- ❖ Le Plan de passation de marchés (PPM) ;
- ❖ Le Plan de consommation des crédits (PCC).

II. Outil de planification de projet axée sur les résultats

2.1. Le cadre logique (Principes et méthodologie)

2.1.1 Définitions et usage du cadre logique

La Méthode du Cadre Logique (MCL) ou l'approche Cadre Logique (ACL), a été mise au point en 1969 aux Etats-Unis par la firme Pratical Concepts Inc. (PIC), pour le compte de l'Agence Américaine pour le Développement International (USAID). L'idée de départ était d'élaborer une grille d'analyse simple et systémique, permettant de mieux comprendre les implications et la complexité des projets américains d'aide pour le développement international. Le succès de cet outil a amené plusieurs organisations à adopter la MCL et à y apporter des adaptations particulières. Le Cadre Logique est un instrument qui vise à renforcer les procédures de conception, d'exécution et d'évaluation des projets. Aussi, est-il conçu pour être utilisé pendant tout le cycle du projet.

L'approche du cadre logique sert à plusieurs usages dont :

- **La formulation en une seule phrase de toutes les composantes essentielles d'un projet.** Formuler toutes les composantes clés d'un projet ou d'un programme de manière systématique, concise et cohérente permet de clarifier et de démontrer la logique qui sous-tend la mise en œuvre de tout projet ou programme. Cette formulation permet également d'établir les différents niveaux de hiérarchisation des objectifs et partant, d'assurer qu'il n'existe aucun amalgame entre les entrées et les sorties d'une part, et entre celles-ci et les objectifs d'autre part.
- **La satisfaction des exigences d'une bonne conception de projet :** Cette formulation permet de poser les questions fondamentales et d'analyser les différentes faiblesses en vue de fournir aux décideurs des informations plus judicieuses et plus pertinentes. Elle sert également de canevas à une analyse systématique et logique des interactions entre les différents éléments essentiels d'un projet bien conçu. Cette approche permet en outre, l'amélioration de la

planification dans la mesure où elle met l'accent sur les rapports qui existent entre les éléments du projet et les facteurs externes déterminants.

- **L'apprentissage et l'utilisation aisés.** Les notions élémentaires de l'approche du Cadre Logique peuvent être enseignées de manière effective en quelques jours.
- **La réduction du temps ou des efforts nécessaires dans la gestion du projet.** Une fois assimilé, le Cadre Logique permet de sauver beaucoup de temps dans la réalisation des interventions.
- **L'utilité interne et externe en matière de conception et d'évaluation des processus**
L'approche du Cadre Logique peut être utile pour la conception et ressort au moment du suivi et de l'évaluation. En plus, le Cadre Logique appelle une approche pluridisciplinaire en matière de conception et de supervision des projets.
- **L'anticipation de la mise en œuvre**
L'approche du Cadre Logique permet de concevoir les activités de l'intervention avec en esprit, un objectif bien précis. Elle facilite la compréhension mutuelle ainsi qu'une meilleure communication entre les décideurs, les gestionnaires et les autres intervenants. De même, l'utilisation des Cadres Logiques assortis d'un mécanisme de suivi systématique garantit la continuité de l'approche même lorsqu'un membre initial de l'équipe de gestion est muté ou remplacé.
- **L'établissement d'un cadre de suivi & évaluation permettant la comparaison des résultats envisagés et concrets.**
Avec des objectifs et des indicateurs de succès clairement définis bien avant le lancement du projet, l'approche présente déjà un certain cadre d'évaluation. Il est généralement très difficile de procéder à une évaluation rétrospective des projets surtout lorsque les objectifs initiaux n'ont pas été clairement définis. L'approche du Cadre Logique permet alors d'élucider les rapports qui sous-tendent les jugements de valeur en matière d'efficacité et d'efficience probables autant qu'elle permet d'identifier les principaux facteurs de succès des interventions. Même si les évaluations peuvent aller plus loin que ces éléments, ils constituent une très bonne base.

- **La présentation sous forme de Cadre Logique permet d'apprécier les facteurs de qualité et de réussite notamment :** la pertinence, la faisabilité, la durabilité et l'efficacité. On les apprécie en tâchant de répondre aux questions suivantes :
 - ▶ **La pertinence:** est-ce que l'intervention répond à des objectifs préalablement définis, à un besoin émanant du contexte d'intervention ?
 - ▶ **La faisabilité :** est-ce que les activités envisagées peuvent être exécutées avec les moyens proposés ? Est-ce que ces activités garantissent effectivement les résultats escomptés ? Est-ce que ces résultats vont mener à la réalisation de l'objectif spécifique ? Est-ce que les suppositions sont réalistes ?
 - ▶ **Une analyse de la durabilité :** les bénéficiaires du projet ont-ils des chances de continuer une fois que le support extérieur sera arrêté ?
 - ▶ **Une appréciation de l'efficacité :** le rapport entre les intrants ou les coûts engagés et les bénéficiaires ou résultats (output), est-il raisonnable ?

Une fois que l'intervention aura démarré, les facteurs de qualité et de réussite énoncés plus haut devront être constamment contrôlés à travers le suivi. Celui-ci doit être continuellement réadapté aux nouvelles constatations en fonction des changements dans l'environnement.

2.1.2. Structure du cadre logique

Le cadre logique est concrètement une matrice à 4 colonnes et 4 lignes. La première colonne présente la chaîne des résultats de l'intervention (qui sera présentée en plus de détails au chapitre 5) Du bas vers le haut, elle présente le processus qui part des activités jusqu'à l'effet ultime.

La seconde et la troisième colonne du Cadre Logique (les indicateurs et les sources de vérification) contiennent les informations qui constituent la base du suivi.

La dernière colonne (Hypothèses) présente les conditions externes aux projets qui doivent être remplies pour garantir le succès des différentes actions à mener. En soi, il y a donc des aides à la gestion incluses dans le Cadre Logique.

Tableau 3.1 : Canevas de logique

Résultats escomptés	Indicateurs de performance	Sources de vérification	Hypothèses / Risques
<p>Impact La finalité du projet : Pourquoi a-t-on réalisé ce projet ?</p>	<p>Comment mesurer l'atteinte de ce(s) résultat(s) ? (%, taux, niveau, nombre, volume, coût, etc.)</p>	<p>Sources à consulter pour s'assurer que les indicateurs sont atteints</p>	<p>Hypothèses : Quelles sont les conditions nécessaires à l'obtention des résultats ? Risques : Qu'est-ce qui pourrait nuire à la réalisation des résultats ?</p>
<p>Effets (produits) Les conséquences des extrants : qu'ont permis d'obtenir les extrants ?</p>	<p>Comment mesurer l'atteinte de ce(s) résultat(s) ? (%, taux, niveau, nombre, volume, coût, etc.)</p>	<p>Sources à consulter pour s'assurer que les indicateurs sont atteints</p>	<p>Hypothèses : Quelles sont les conditions nécessaires à l'obtention des résultats ? Risques : Qu'est-ce qui pourrait nuire à la réalisation des résultats ?</p>

Résultats escomptés	Indicateurs de performance	Sources de vérification	Hypothèses / Risques
<p>Extrants (réalisations) Produits immédiats des activités : Qu'a-t-on produit ?</p>	<p>Comment mesurer l'atteinte de ce(s) résultat(s) ? (%, taux, niveau, nombre, volume, coût, etc.)</p>	<p>Sources à consulter pour s'assurer que les indicateurs sont atteints</p>	<p>Hypothèses : Quelles sont les conditions nécessaires à l'obtention des résultats ?</p> <p>Risques : Qu'est-ce qui pourrait nuire à la réalisation des résultats ?</p>
<p>Activités</p>	<p>Intrants</p>		<p>Hypothèses : Quelles sont les conditions nécessaires à l'obtention des intrants ?</p> <p>Risques : Qu'est-ce qui pourrait nuire à l'obtention des intrants ?</p>

2.2. Le cadre d'analyse et de suivi des risques

Dans le domaine de la gestion de projet, les risques représentent les événements dont la survenance est susceptible de constituer une menace pour le succès d'un projet. Ils sont représentés par les hypothèses dans le Cadre Logique et on les retrouvera dans la théorie du changement utilisée par les évaluateurs.

En réalité, des risques existent tout au long du cycle de vie d'un projet, dans la mesure où les choses ne se déroulent toujours pas comme elles ont été planifiées. Il existe des facteurs indépendants de la volonté de l'équipe projet qui peuvent influencer le cours normal du projet. Gérer les risques d'un projet, consiste à les identifier, à les évaluer, à les mitiger et à les contrôler.

Il est d'usage d'élaborer lors de la planification, un cadre d'analyse des risques présentant les risques recensés, les résultats de leur évaluation, ainsi que les stratégies définies pour les gérer.

Tableau 3.2 : Cadre d'analyse des risques

Résultats	Hypothèses	Risques	Niveau		Stratégies d'atténuation du risque
			Impact	Probabilité	
Impacts					
Effets					
Extrants					

La colonne "Résultats" présente la chaîne des résultats du projet. Les hypothèses sont des conditions externes au projet qui doivent être réunies pour garantir sa réussite et qui sont reprises du cadre logique.

La colonne des risques recense les principaux risques du projet identifiés. L'impact que peut produire le risque s'il se réalise et la probabilité de survenance de celui-ci

permettent de déterminer la criticité du risque. Les acteurs du projet doivent déterminer avec précision ces éléments qui permettent de définir la valeur du risque ainsi que son importance dans le projet. Pour chaque risque identifié, une stratégie de gestion doit être envisagée, soit pour l'éviter, pour l'endiguer ou pour l'atténuer.

2.3. Le cadre de mesure de rendements

La mesure du rendement est aussi un élément très important de la gestion axée sur les résultats. Elle permet d'évaluer le niveau de performance et d'atteinte des résultats des actions réalisées. Pour évaluer le rendement, il est nécessaire de mettre en place un système efficace de collecte et d'analyse des données. La mesure du rendement permet de mettre à la disposition des responsables de projet et des parties prenantes des informations pertinentes pour analyser les succès et les points faibles du projet et de favoriser, le cas échéant, la prise de mesure corrective pour améliorer les résultats.

Le cadre de mesure du rendement est « un plan permettant de recueillir systématiquement des données pertinentes pendant toute la durée d'investissement afin d'évaluer et de mettre en évidence les progrès vers les résultats escomptés. » (ACDI, 2009 :12) Il est réalisé au moment de la planification et devient ensuite un outil de suivi et d'évaluation.

Tableau 3.3 : Canevas de cadre de mesure du rendement

Résultats	Indicateurs	Valeur de l'indicateur		Sources d'information	Méthode et fréquence de collecte de données	Responsable
		Référence	Cible			
Impact						
Effets						
Extrants						

Le cadre de mesure du rendement comprend sept colonnes dont les deux premières sont basées sur le cadre logique. La colonne « Résultats » reprend les résultats

définis dans la chaîne de résultat du cadre logique et les « Indicateurs » sont ceux définis pour évaluer les résultats.

Les indicateurs de référence (colonne référence) représentent les conditions qui existaient avant le début de la réalisation du projet. Ils constituent le point de repère pour l'évaluation des performances.

La cible est une valeur particulière visée par le projet dans un intervalle de temps donné par rapport à un résultat escompté.

Les sources d'information sont les personnes, les organisations, les documents, etc. qui fournissent des données pour évaluer les indicateurs. En fonction de la nature des indicateurs, les responsables du projet peuvent opter pour des sources primaires (collecte de données brutes) ou pour des données secondaires (rapports, documents, études déjà réalisées).

La méthode de collecte de donnée indique comment les informations nécessaires seront recueillies. Le choix de la méthode est fonction de la nature de l'indicateur et de la fréquence de production des indicateurs. La fréquence de collecte représente l'intervalle de temps à laquelle les données sont collectées. Le responsable est la personne, qui est chargée de collecter les données.

III. Planification opérationnelle et le suivi des activités de développement

3.1. Objet de la planification opérationnelle

La planification est un processus par lequel une vision est déclinée en activités pour atteindre un résultat. La planification opérationnelle permet d'avoir le planning détaillé de réalisation d'une intervention. Elle décline les activités liées à chaque résultat en séquences logiques.

Le domaine principal de la planification opérationnelle au niveau du projet est caractérisé par une articulation de ressources humaines et matérielles dans le but d'atteindre un objectif à travers des résultats dans un contexte de contraintes de coût, de délai et de performance.

En outre, elle donne une vision claire sur les :

- ❖ objectifs de performance : quoi et par qui ?
- ❖ Objectifs de délai : quand. ?
- ❖ Objectifs de coûts : à quel prix ?
- ❖ Budgets et procédures de suivi ;
- ❖ dossier de réalisation ; mémo de mise en œuvre du projet ;
- ❖ Personnes / structures impliquées.

La planification opérationnelle est une activité fondamentale dont aucun gestionnaire de projet ne peut faire l'économie. Ainsi, une fois que les objectifs et activités du projet ont été clairement déterminés et que les différentes phases d'analyse de sa faisabilité ont été bouclées, quatre questions essentielles doivent constamment occuper l'esprit du chef de projet jusqu'à ce qu'il parvienne à répondre précisément à chacune d'elle:

- ▶ Que dois-je faire pour réaliser le projet ?
- ▶ Qui fera quoi ?
- ▶ Quand faire quoi ?
- ▶ A quel coût ?

Ces questions répondent respectivement aux exigences de planification de la structure, de l'organisation, des délais et du coût du projet.

Planifier au niveau opérationnel c'est donc :

- ▶ cerner les activités à réaliser ;
- ▶ définir les actions (tâches) à accomplir pour la mise en œuvre ;
- ▶ en déduire les moyens nécessaires ;
- ▶ comparer les moyens disponibles aux moyens nécessaires ;
- ▶ en déduire des délais prévisionnels.

Tout le but de la planification opérationnelle est de suivre le projet en le comparant sans cesse à la planification de référence. Le principe consiste à :

- ❖ mesurer l'avancement ;
- ❖ comparer à la planification de référence ;
- ❖ évaluer les impacts des écarts :
 - sur les jalons ;
 - sur la date de fin du projet ;
- ❖ estimer les actions correctrices nécessaires pour :
 - les ressources ;
 - le coût ;
- ❖ mettre en œuvre les actions correctrices pour respecter les jalons ;
- ❖ respecter les échéances, en particulier celles pénalisables (contrat) ;
- ❖ respecter le délai final prévu.

Toutes ces bonnes pratiques peuvent être également appliquées à la gestion d'une mission d'évaluation comme on le verra au chapitre 12.

3.2. Elaboration de la Structure de Fractionnement des Tâches du projet (SFT)

Un des éléments clés de toute planification opérationnelle d'un projet est sans aucun doute la **structure de fractionnement de travail**. Cette charpente consiste à découper le projet en lots de tâches d'une façon logique, selon une démarche descendante. La planification structurelle ou l'élaboration de la structure de découpage d'un projet vise à produire essentiellement les trois types de résultats suivants :

- ❖ La liste **exhaustive** des tâches à accomplir pour réaliser le produit constituant la finalité du projet ;
- ❖ La **description** du contenu de chaque tâche ou activité du projet ;
- ❖ La **structure** logique regroupant les tâches selon certains paramètres utiles pour la gestion de la réalisation du projet.

La structure de découpage du projet est un processus de planification au terme duquel l'on doit aboutir au fractionnement du travail à effectuer. Dans la terminologie du management de projet, cette opération de fractionnement est appelée **lotissement structuré du travail** (Work Breakdown Structure, d'où le sigle **WBS**). On la désigne aussi par les expressions : **Organigramme des Tâches (OT)**, **Structure de Fractionnement des Tâches (SFT)**, **Structure de Découpage du Projet (SDP)**. Nous retiendrons ici l'expression SDP.

La SDP est le résultat d'un travail de décomposition progressive et exhaustive de l'ensemble d'un projet en ses éléments constitutifs détaillés.

Pour élaborer la SDP, le chef de projet doit décomposer le projet en forme pyramidale comprenant :

1. Le projet en entier découpé en un sous-ensemble de lots ;
2. Les lots sont décomposés en des tâches simples ou activités ;
3. Les tâches simples structurées en un bloc de travail.

Tout projet, du plus simple au plus complexe, se traduit par une liste de travaux à faire. Pour gérer le projet de façon adéquate, il est nécessaire de découper chaque lot de travaux en éléments simples.

Quelles sont les exigences relatives à la description des tâches ?

- Elle doit être écrite ;
- Elle doit être assez claire et sans ambiguïté (pour que l'on puisse en confier la responsabilité à une personne ou une entité unique) ;
- Elle doit être suffisamment détaillée (pour servir de document contractuel entre le chef du projet et le responsable de la tâche) ;
- Elle doit être suffisamment précise (pour servir de base d'évaluation du temps requis pour son exécution et le calcul de son coût) ;
- Elle doit comporter une description du produit dont la livraison marquera la fin de la tâche.

Encadré 3.1: Résumé du processus d'élaboration de la SDP

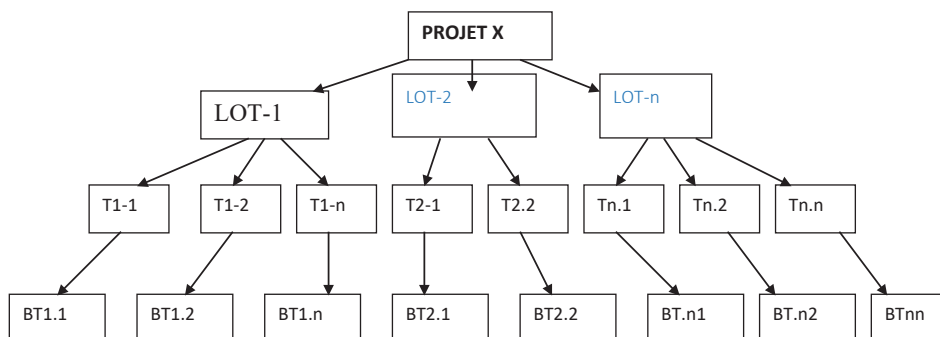
Étape 1 : Il s'agit ici de dresser la liste de toutes les tâches que l'on doit effectuer dans le projet.

Étape 2 : Une fois que l'on a déterminé toutes les tâches à effectuer pour mener à bien le projet, il s'agit maintenant de regrouper celles qui ont des points communs. Ces regroupements, qui s'appellent des « *lots* », forment le niveau 2 de la structure de fractionnement.

Étape 3 : Cette étape consiste à mettre dans un ordre logique et séquentiel, les lots que nous venons d'établir. Elle permet ainsi de bâtir la SDP

Étape 4 : La dernière étape consiste à faire la codification des lots et des activités du projet. Cela permet de rendre la structure de découpage du projet compatible avec le logiciel de gestion de projet.

Graphique 3.3 : Exemple du schéma de la SDP d'un projet



Abréviations : T : Tâches B.T. : Bloc de travail

3.3. La charte ou matrice de responsabilités

Lors de la planification organisationnelle, **chaque intervenant est naturellement informé des responsabilités d'exécution qui le concernent directement** : les tâches qui lui sont assignées ont été discutées avec lui.

Mais chaque intervenant doit aussi être informé de la répartition des responsabilités relatives à d'autres tâches si elles comportent des interfaces avec les tâches dont il est lui-même responsable.

La charte ou la matrice des responsabilités permet de présenter de façon simple et commode une synthèse de cette opération.

Tableau 3.4 : Charte ou matrice de responsabilités

INTERVENANT	LOTS				
	LOT1	LOT2	LOT3	LOT4	LOT5
A	RX		RX		
B		RA		RA	RX
C		RX	I		
D	I		RA	RX	I

Code de responsabilité

RA : Responsabilité d’approbation

RX : Responsabilité d’exécution

I : Doit être informé

3.4. La programmation des activités

La programmation procède à l’ordonnancement (mise en séquence) des activités du projet. Cette opération consiste à identifier les liaisons logiques ou les relations d’ordre qui existent entre les différentes activités du projet. Les liaisons logiques entre les activités peuvent être de plusieurs natures.

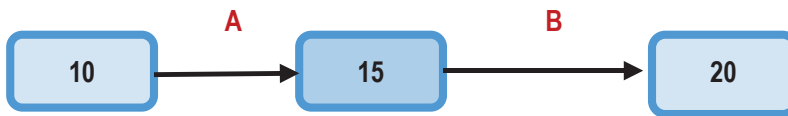
3.4.1. Ordonnancement : modélisation

Il s’agit de représenter les éléments du projet (les activités, leur durée et leurs dépendances) au moyen d’un modèle. Ce modèle dit **réseau est un ensemble cohérent de nœuds connectés par des arcs** (branches orientées).

- **Un nœud est l’élément de base du réseau** : les nœuds sont représentés soit par des cercles, des rectangles ou des carrés ;
- **Un arc est une paire orientée de nœuds** : l’arc AB formé par les nœuds A et B diffère de l’arc BA formé des deux mêmes nœuds, mais connecté en sens inverse.

Deux types :

Activités sur les arcs : ADM (Arrow Diagramming method)



On dit que l'activité A précède l'activité B

L'évènement n° 15 représente deux situations :

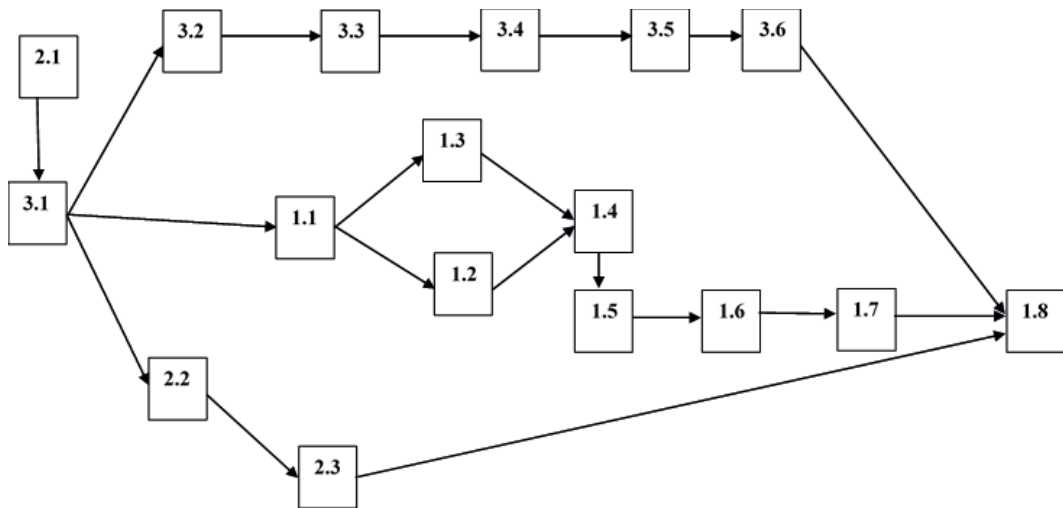
1. La fin de A
2. Le début de B

Activités sur les nœuds: PDM (Precedence Diagramming Method)



On dit que l'activité A précède l'activité B

Graphique 3.4 : Exemple de réseau PDM de projet



L'ordonnancement nécessite l'utilisation d'outils en fonction de la nature et de la complexité du projet. Parmi les nombreuses techniques d'ordonnancement des tâches qui existent, le plus utilisé est :

- La Méthode PERT ;
- La Méthode du Chemin Critique

3.4.2. La méthode PERT (Program Evaluation and Review Technic)

Le PERT (Programme Evaluation and Review Technic) a été développé aux USA pour contrôler l'adjudication du programme de la Fusée Polaris. C'est une variante des techniques de réseaux utilisées en planification de projet. Les autres variantes sont le PDM et le CPM (Critical Path Method). La principale différence entre le PDM et le PERT est que le PERT utilise des estimations probabilistes alors que la PDM se limite à des estimations simples. Comme tel, le PERT est une technique de planification des activités qui permet essentiellement de traiter l'incertitude sur la durée d'exécution des activités. Les estimations probabilistes permettent aussi au chef de projet de calculer les marges de manœuvre dont il dispose dans l'exécution de certaines activités.

La particularité du PERT est qu'il permet de présenter l'enchaînement des tâches qui n'autorisent pas de retard au risque de décaler la date de fin du projet. Il favorise le calcul des marges. Le PERT est un graphique orienté du début du projet vers la fin de celui-ci.

Un réseau PERT doit être sans boucle ni circuit. C'est-à-dire qu'on ne peut plus repasser par le même point une fois qu'il a été franchi. Dans la gestion d'un projet, cela signifie qu'on ne peut reprendre une tâche déjà réalisée.

La gestion des projets est pleine d'incertitude et les choses ne se déroulent presque pas comme elles ont été planifiées. Pour donner plus de précision aux prévisions, la méthode PERT prend en compte l'incertitude dans les prévisions à cet effet, elle utilise des estimations probabilistes. La durée de chaque tâche est considérée comme une variable aléatoire qui peut varier entre une valeur dite optimiste et une autre appelée durée pessimiste. Ainsi, la durée du chemin critique (et aussi celle du projet), est également considérée comme une variable aléatoire, vu qu'elle est la somme de variable aléatoire. La durée probable pour une tâche correspond donc à une distribution proche de la loi normale.

Dans la pratique, les gestionnaires de projet déterminent trois estimations de la durée d'une tâche.

- ▶ *topt* : durée optimiste qui correspond au temps minimum si les choses se déroulent dans les meilleures conditions ;
- ▶ *tpess* : durée pessimiste, qui correspond au temps maximum de la tâche si tout se déroule dans les mauvaises conditions ;
- ▶ *tpro* : durée la plus probable, égale au temps le plus probable si les acteurs devaient faire une seule prévision.

La durée estimée de la tâche est déterminée par la relation suivante (Bouchaouir et al, 2017 :152) :

$$\text{Durée estimée} = \frac{top + 4 * tpro + tpess}{6}$$

3.4.3. La méthode du chemin critique (Le CPM)

Le chemin critique d'un projet (CPM) se détermine à partir de l'élaboration de son réseau PERT. Il désigne l'enchaînement logique sur lequel toutes les activités ont une marge nulle. C'est aussi le chemin le plus long du réseau des activités du projet. Il est fondamental pour le chef de projet de toujours déterminer le chemin critique de son projet, car par ce moyen il arrive à repérer les activités critiques du projet. Les activités critiques du projet sont celles sur lesquelles le chef de projet doit porter toute son attention s'il veut terminer le projet dans la limite du délai prévu. Les activités sur le chemin critique sont donc les activités sur lesquelles il faut orienter les efforts d'accélération des travaux, ainsi que les efforts de suivi durant la phase d'exécution du projet. Le chef de projet pourrait par exemple décider d'affecter plus de ressources à ces activités pour se donner plus de chances de respecter la contrainte du délai de livraison des produits et services du projet. Par ailleurs, il faut garder à l'esprit que les activités non critiques sont aussi importantes pour le succès du projet. Une activité non critique qui prend assez de retard peut devenir critique et compromettre les échéances du projet.

3.4.4. Le diagramme de GANTT

Le diagramme de Gantt est l'outil de planification et de suivi des délais de projet le plus simple et le plus répandu. Il sert à la visualisation des lots de travail ou des activités ainsi que de celle de leurs dates de début et de fin.

Le diagramme de GANTT est un graphique à deux dimensions qui représente chaque tâche d'un projet par un trait ou un rectangle horizontal. Les dates sont définies sur l'abscisse du graphique (axe horizontal) et les tâches sont organisées sur l'ordonnée (axe vertical). L'échelle de temps doit être adaptée à la durée du projet. La longueur des traits ou des rectangles représente la durée affectée à la tâche (Bouchaoui et al, 2017 :147). La charte de GANTT est un outil efficace de communication qui permet d'avoir une vision globale du projet et de donner les détails sur le déroulement des activités dans le temps.

Tableau 3.5: Allure générale d'un diagramme de GANTT

Activités	Durée	Responsables	Calendrier / Chronogramme							
			T1	T2	T3	T4	T5	T6	T7	T8
1.1			■							
1.2			■							
1.3				■						
2.1					■					
2.2						■				
2.3							■			
3.2								■		
4.1									■	
4.2									■	

3.5. La planification des coûts

La réalisation des activités d'un projet nécessite l'utilisation de ressources. Les ressources peuvent être des ressources humaines (internes ou externes) ou des ressources matérielles. L'évaluation du coût des différentes ressources permet d'établir le budget des activités et par ricochet celui du projet entier. De même que les objectifs et les délais, le budget peut faire l'objet d'arbitrage. C'est pourquoi la budgétisation doit faire partie intégrante du processus de planification des programmes et projets.

La planification financière d'un projet vise à :

- déterminer les ressources nécessaires pour chaque activité (ou groupe d'activité) ;
- **vérifier si le budget** octroyé par le promoteur pour la réalisation du projet est **suffisant**.

- **Répartir ce budget** entre les tâches à exécuter et les intervenants de ces tâches.

Pour ce faire, il s'agit de déterminer au mieux le **budget de réalisation de chaque tâche (coûts directs)**, les **coûts indirects** du projet et le **budget octroyé à chaque intervenant**.

***N. B.** Si le Chef de projet s'engage à réaliser le projet avec un budget démontré insuffisant, il devient responsable du dépassement éventuel de ce budget.*

3.5.1. Evaluation des coûts directs

Les coûts directs concernent toutes les dépenses qui seront occasionnées par le projet. Il peut s'agir de main-d'œuvre, de matériels (ou matières), de coûts de transport, de déplacement, de packaging, etc.

Coût direct = coût encouru du fait de l'exécution d'une tâche

L'évaluation des coûts directs d'une tâche procède le plus souvent par l'évaluation de la quantité de ressources humaines et matérielles requises pour l'exécution de cette tâche.

L'évaluation du coût des ressources humaines se fait en personnes - jours. A cet effet, il est donc important de connaître :

- la liste exhaustive des tâches;
- la nature et le nombre des intervenants / tâche;
- la durée d'exécution de chaque tâche ;
- la productivité des intervenants;
- le tarif d'honoraires des intervenants.

Exemple : Soit des travaux de crépissage d'une maison qui doivent être réalisés par un maçon et un ouvrier.

Données : maçon: 3 500 F / jour de travail

Ouvrier : 1 500 F / jour de travail

Durée de travail : 4 jours

Calcul du coût direct (Main d'œuvre)

Maçons	Ouvriers	Nb Jours	Personnes-jours	Coût/ personne-jour	Montant
1	1	4	8	$2500 = (3500+1500) / 2$	20 000

3.5.2. Evaluation des coûts indirects

Les coûts indirects sont constitués de l'ensemble des frais qui ne sont pas directement induits par le projet. En d'autres termes, c'est un coût qui est supporté parce que le projet est réalisé et qui à l'inverse ne le serait pas, si le projet n'était pas réalisé, mais non imputable à une tâche spécifique.

Coût indirect = coût qui est imputable au projet, mais pas à une tâche ou à une activité précise.

Exemples : Salaire du chef de projet, frais de location d'un local pour l'équipe de projet, salaires du personnel administratif affecté à plein temps au projet pour sa durée.

Ne pas confondre coûts indirects et frais généraux, qui sont liés à la réalisation du projet.

- **Coûts indirects dans le cas d'un projet réalisé en régie**

Si le projet est réalisé en régie, c'est-à-dire par l'entreprise promotrice, les frais indirects ne sont pas considérés explicitement, mais regroupés avec les frais généraux de l'entreprise.

- **Coûts indirects dans le cas d'un projet réalisé à contrat**

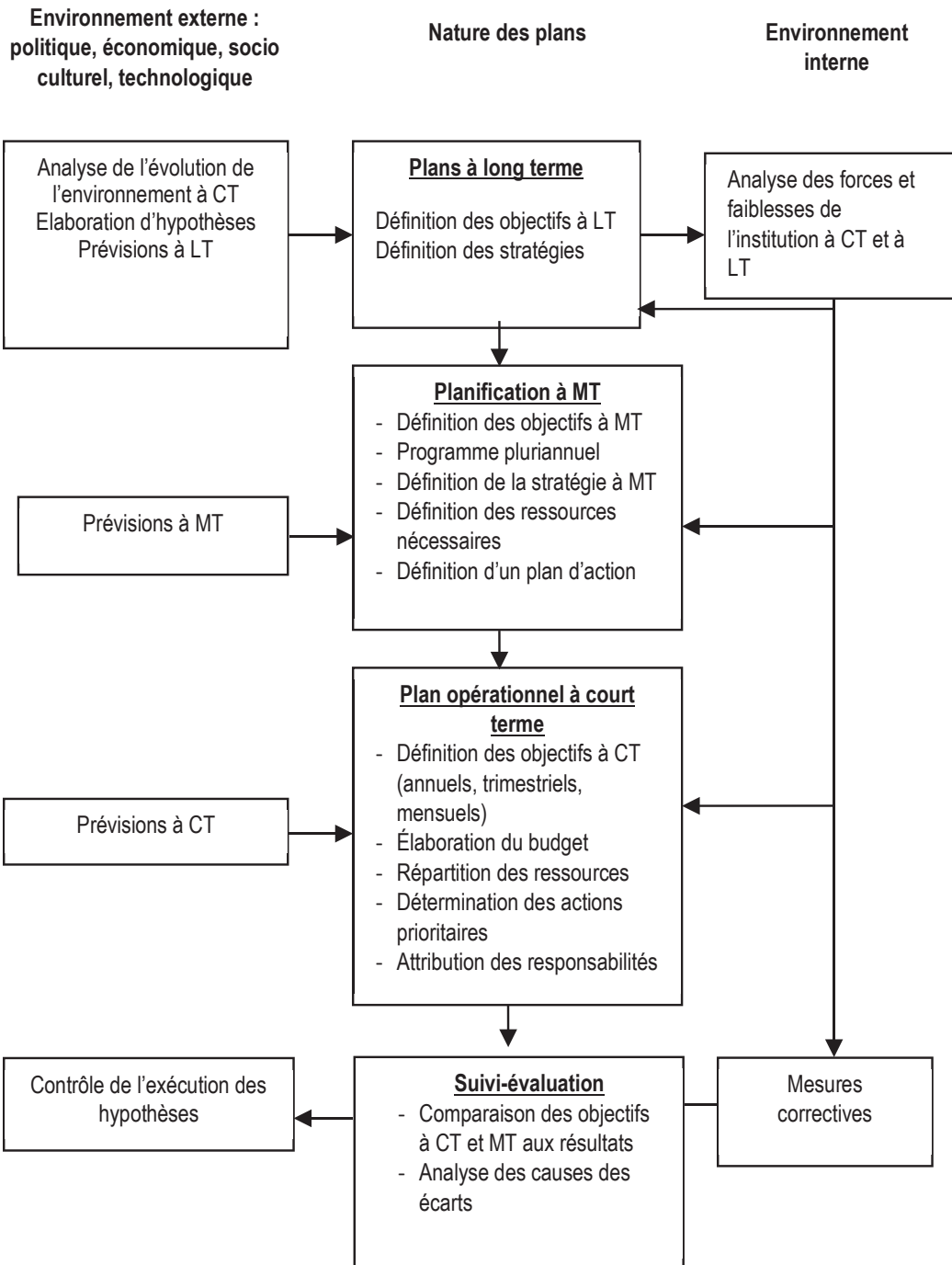
Dans le cas où un mandataire réalise un projet à contrat pour un promoteur d'une organisation différente, il est important pour la rentabilité de l'opération du mandataire que celui-ci identifie tous les frais indirects (résultant de la réalisation du projet), les évalue correctement et les impute au budget du projet.

IV. Lien entre le suivi, évaluation et la planification

Le suivi et l'évaluation permettent d'apprécier et d'améliorer l'exécution au niveau opérationnel (suivi-évaluation des PTA, PPM, PCC, suivi annuel de performance) et en remontant le cadre programmatique, d'apprécier et d'améliorer l'exécution au point de vue stratégique. Le suivi et l'évaluation ont lieu tout au long du processus de mise en œuvre des interventions. Ils permettent d'évaluer le niveau et les conditions de réalisation des activités vis-à-vis de la planification. Sans planification, il ne peut y avoir de suivi, car les acteurs n'auront aucune base de référence et d'évaluation des changements. Aussi, sans suivi, il serait difficile de savoir si les changements planifiés ont été réalisés. Le suivi, l'évaluation et la planification sont donc à la fois complémentaires et indispensables l'un à l'autre.

De la planification à long terme découle celle à moyen terme et de celle à moyen terme découle celle à court terme et enfin la planification opérationnelle. Les différents niveaux du processus de planification sont représentés par le graphique suivant.

Graphique 3.5 : Processus de planification



Messages clés :

- ✓ Une bonne connaissance du processus de gestion axée sur les résultats, inclut le suivi et l'évaluation, mais aussi la planification. Les trois disciplines distinctes s'entrecroisent pour mener les interventions à leurs résultats.
- ✓ On distingue la planification stratégique, sur le long terme souvent à l'échelle d'une administration en entier, ou pour un secteur, et la planification opérationnelle qui vise la gestion de chaque projet.
- ✓ La planification axée sur les résultats amène à élaborer des outils particuliers, dont le cadre logique, le cadre de mesure de rendement, la matrice des risques. Ces outils nourrissent par la suite le suivi et l'évaluation des interventions et leurs propres outils, d'où l'intérêt pour les évaluateurs de les maîtriser.

Bibliographie

- ACDI (1999), La gestion axée sur les résultats à l'ACDI : guide d'introduction aux concepts et aux principes ;
- ACDI (2009), Les outils de la GAR à l'ACDI, Québec, Canada
- Bouchaour F. ; Dentinger Y., Englender O. (2017), Gestion de projet ; 50 outils pour agir, 4^e édition, Vuibert, Paris
- Morley C. (2008), Management d'un projet système d'information, Principes, techniques, mise en œuvre et outils, 6^e édition, Dunod, Paris
- PMI (2013), A Guide to the Project Management Body of Knowledge, (PMBOK® Guide)- Fifth Edition,
- PNUD, Bureau de l'évaluation (2002), Guide du suivi et de l'évaluation axés sur les résultats, New York, USA
- PNUD, (2009), Guide de la planification, du suivi et de l'évaluation axés sur les résultats du développement, New York, USA.
<http://web.undp.org/evaluation/handbook/french/ch1-1.html>
- Secrétariat du Conseil du Trésor (2003), Guide sur les indicateurs, Québec.

MODULE 4

CONCEPTION D'UN SYSTÈME DE SUIVI ÉVALUATION AXÉ SUR LES RÉSULTATS

Objectifs d'apprentissage :

A la fin de la session, les participants seront capables de :

- Définir un système de S&E axé sur les résultats (SSEAR)
- Déterminer les éléments constitutifs d'un SSEAR
- Concevoir un SSEAR :
Étapes clés

Thèmes couverts

- I. Définition d'un système de suivi évaluation et de ses éléments constitutifs
- II. Étapes (10) de conception d'un système de S&E axé sur les résultats

Les concepts de "suivi" d'une part, et d'"évaluation" d'autre part, au regard de leur définition respective telle qu'exposée dans les chapitres précédents, apparaissent comme des notions distinctes, mais aussi, et surtout complémentaires. Ainsi, ce chapitre focalise l'attention sur le suivi-évaluation en tant qu'un système intégré de collecte de données, d'analyse et de communication d'informations, à l'instar d'un tableau de bord de pilotage de projet. Le module 4 s'attelle ainsi dans un premier temps à rappeler quelques concepts de base de la notion de suivi-évaluation, vue comme un système, puis dans un deuxième temps à exposer différentes approches de conception et de mise en place d'un tel système de S&E axé sur les résultats. Le S&E est ici présenté dans une perspective à la fois organisationnelle, mais aussi systémique, c'est-à-dire à l'échelle sectorielle ou nationale.

Mots-clés : système de suivi-évaluation, SSEAR

I. Définition d'un Système de Suivi Evaluation axé sur les résultats (SSEAR) et de ses éléments constitutifs

1.1. Définition du SSEAR

Le système de suivi-évaluation axé sur les résultats peut être défini comme l'ensemble des dispositions (objectifs, organisation, ressources, processus) définies et mises en œuvre tout au long du déroulement d'un projet/programme pour :

- suivre l'état d'avancement et les performances du projet/programme ;
- obtenir à des jalons prédéfinis une analyse du projet/programme suivant des critères précis ;
- faciliter la prise de décisions sur la base de données d'appréciation construites, disponibles et fiables.

Il s'agit d'un ensemble d'activités de planification et de gestion, de collecte, d'analyse, de rapportage, et d'utilisation de l'information sur un projet, un programme ou une politique, ses résultats (gestion, produits/services, effets, impacts) pour en mesurer la performance (bonne exécution, efficacité, efficacité, pérennité, etc.), afin d'améliorer la prise de décisions (sur la stratégie et les opérations) en temps réel et la capitalisation des bonnes pratiques.

Son dispositif doit être :

- ▶ **léger** : il ne doit pas exiger beaucoup de temps, de moyens matériels et financiers.
- ▶ **ciblé** : cela signifie que les informations à connaître et à suivre doivent être définies d'avance; les indicateurs à utiliser doivent être précis, peu nombreux et simples.
- ▶ **participatif** : les informations à rechercher seront choisies en concertation avec tous les acteurs du projet ; ces derniers doivent bien comprendre le processus et participer à toutes les étapes.

Il doit par ailleurs :

- ▶ **permettre la diffusion des informations collectées** auprès des équipes et des destinataires-bénéficiaires/clients.
- ▶ **permettre le feed-back et les échanges autour des données** - le SSEAR est un outil d'autoévaluation et de capitalisation systématique des procédures, des savoir-faire et des pratiques.
- ▶ **intégrer à la fois le quantitatif et qualitatif** - les informations à caractère quantitatif et qualitatif doivent être prises en compte pour permettre d'affiner l'analyse et de mieux comprendre les effets du projet/programme/politique.

1.2. Éléments constitutifs d'un SSEAR

Le SSEAR offre un cadre conceptuel et surtout pratique pour la planification, le suivi-évaluation des interventions y compris la communication. Les fonctions et éléments constitutifs sont résumés dans le tableau ci-dessous.

Tableau 4.1 : Utilité et outils d'un SSEAR

Le SSEAR permet : ...	et mobilise à cet effet les outils/concepts tels que
<ul style="list-style-type: none"> ● d'utiliser des données de base pour décrire un phénomène et/ou analyser la robustesse et l'adéquation de la théorie de changement avant le démarrage même des interventions 	<ul style="list-style-type: none"> ○ Cadre logique/chaîne de résultats ○ Indicateurs ○ Données de références (base-line) ○ Cibles/Objectifs (de résultat)

- | | |
|---|---|
| <ul style="list-style-type: none"> ● de suivre des indicateurs de résultat en collectant des données sur les intrants, les activités, les extrants (produits/services), de même que leurs contributions à l'atteinte des résultats | <ul style="list-style-type: none"> ○ Cadre de mesure des résultats ○ Outils et techniques de collecte et d'analyse de données ○ Techniques de gestion et d'archivage des données ○ Manuel/Plan de S&E |
| <ul style="list-style-type: none"> ● de rendre compte aux parties prenantes à travers un mécanisme qui s'efforce de fournir à tout moment des informations crédibles et utiles à la prise de décision. | <ul style="list-style-type: none"> ○ Techniques de gestion des parties prenantes (Rôles et responsabilités / Intérêts / Contraintes) ○ Plan de communication (Dissémination / Diffusions des résultats) |

1.3. La valeur ajoutée d'un SSEAR

L'approche traditionnelle de gestion focalise l'attention sur les questions de conformité liées à la mise en œuvre des interventions : les intrants nécessaires ont-ils été mobilisés et convenablement utilisés ? Les activités prévues ont-elles été initiées et complétées ? Les extrants attendus ont-ils été produits ? Les systèmes axés sur les résultats, au contraire, permettent de mieux comprendre la nature des succès ou des échecs des interventions.

Un système de suivi-évaluation axé sur les résultats fournit des informations cruciales sur les performances du secteur public ou d'une organisation. Il peut aider les décideurs et les autres parties prenantes à répondre aux questions fondamentales à savoir : les engagements ont-ils été tenus ? Les résultats visés sont-ils atteints ? Etc. Le S&E est le moyen par lequel les améliorations - ou l'absence d'améliorations - peuvent être démontrées.

Selon Osborn and Gaebler (1992), mesurer les résultats est essentiel pour les raisons suivantes :

- Si vous n'évaluez pas les résultats, vous ne pourrez pas distinguer la réussite de l'échec.
- Si vous ne voyez pas la réussite, vous ne pourrez la sanctionner positivement.
- Si vous ne sanctionnez pas positivement la réussite, alors vous êtes certainement en train de récompenser l'échec.
- Si vous ne voyez pas la réussite, vous ne saurez en tirer des enseignements.
- Si vous ne reconnaissez pas l'échec, vous ne pourrez y remédier.
- Si vous pouvez démontrer les résultats, vous pourrez gagner la confiance du public.

II. Les dix (10) étapes de conception d'un système de S&E axé sur les résultats

La démarche de mise en place d'un système de S&E axé sur les résultats n'est pas uniforme et universelle. La séquence des étapes varie selon les experts et l'on retrouve des modèles en quatre (4) ou en sept (7) étapes.

Indépendamment du nombre d'étapes, mettre en place un système de S&E axés sur les résultats consiste à concevoir un dispositif capable de réaliser une série d'actions incontournables à savoir :

- (i) formuler les réalisations et choisir des indicateurs en lien avec ces dernières ;
- (ii) recueillir des données sur la situation de référence ;
- (iii) établir des cibles spécifiques à atteindre et leurs échéances ;
- (iv) collecter et analyser des données régulièrement pour évaluer si les cibles sont en voie d'être atteintes ;
- (v) présenter les résultats.

De plus, la mise en place du système de suivi et évaluation requiert une excellente compréhension des parties prenantes et de leurs intérêts, ainsi qu'une bonne maîtrise des facteurs organisationnels, politiques et culturels liés à l'action.

La démarche ici décrite est inspirée des travaux de Jody et Ray (2004). Il s'agit d'un modèle complet visant à guider les praticiens du développement à travers la conception, la gestion et, surtout, la durabilité des systèmes de S&E axés sur les résultats.

Le modèle est articulé en dix (10) étapes à savoir :

- i. Procéder à une évaluation de l'état de préparation** - Appréciation de la capacité organisationnelle et de la volonté politique/managériale à disposer d'un cadre/système de gestion orienté vers la performance.
- ii. S'entendre sur les résultats (réalisations) devant faire l'objet de S et E** - Arbitrage à effectuer pour identifier et retenir les réalisations à suivre et à évaluer, sous la contrainte de l'optimisation des ressources disponibles.
- iii. Déterminer les principaux indicateurs pour suivre les réalisations (résultats)** - Identification et choix des meilleurs indicateurs de performance nécessaires au suivi des progrès à tous les niveaux de la chaîne de résultats.
- iv. Rassembler les données de référence (de base) sur les indicateurs** - Identification, au début de la période de suivi, des points de comparaison nécessaires pour mieux apprécier ultérieurement la performance, tant d'un point de vue qualitatifs que quantitatifs.
- v. Planifier les améliorations : choisir des objectifs réalistes (établir des points de comparaison)** - Identification des valeurs cibles pour les résultats, c'est-à-dire des points d'arrêt intermédiaires vers le résultat à plus long terme.
- vi. Suivre les résultats** - Recueil continu des données de qualité sur la performance.
- vii. Utilisation des informations pour l'évaluation** - Evaluation, ses applications et sa synchronisation avec le dispositif tout entier.
- viii. Transmission des conclusions (présenter les informations)** - Présentation des résultats obtenus de façon à ce qu'ils participent au processus de prise de décision.
- ix. Utiliser les résultats (conclusions/informations)** - Techniques et outils pour générer et partager les connaissances et les enseignements produits par le dispositif au sein des gouvernements et des organisations.
- x. Soutenir et maintenir le SSEAR au sein de l'organisation** - Maintien des systèmes de S&E axés sur les résultats, notamment les questions de fiabilité et de crédibilité de l'information, de responsabilité, de capacités et des incitations appropriées.

Ce modèle en dix (10) étapes peut être utilisé pour l'ensemble des interventions publiques. Il ne s'agit pas d'un processus linéaire, malgré son aspect rectiligne. Dans la pratique, il appelle un va-et-vient entre les étapes, et parfois l'on s'attaque à plusieurs étapes à la fois.

2.1. Etape 1 : Evaluation de l'état de préparation

Il s'agit d'une approche systématique pour déterminer la capacité et la volonté d'un gouvernement ou d'une organisation à élaborer un système de Suivi-Evaluation basé sur les résultats. Elle met l'accent sur un certain nombre de facteurs critiques tels que : la présence ou l'absence de défenseurs et d'incitatifs, les rôles et responsabilités, la capacité organisationnelle, les obstacles au démarrage, etc. Elle répond à la question centrale suivante : ***pourquoi se lancer dans la construction de systèmes de Suivi-Evaluation axé sur les résultats ?***

Il est important de déterminer si les motivations appropriées, qu'elles soient politiques, institutionnelles ou personnelles et qu'elles existent avant de se lancer dans la conception et la construction du système de Suivi-Evaluation axé sur les résultats. Il convient de tenter de trouver des réponses à ces questions :

- *Qu'est-ce qui rend nécessaire la création d'un système de S&E ? Est-ce une prescription légale ? Une demande de la part des usagers/citoyens ? Une exigence des bailleurs ?*
- *Qui sont les défenseurs (promoteurs) de la création et de l'utilisation du système de S&E ? Est-ce la haute direction ? Le gouvernement ? Le parlement ? La société civile ? etc.*
- *Qu'est-ce qui motive ceux qui défendent la création d'un système de S&E ? Des ambitions personnelles ? Politiques ? Un impératif de réformer ? La quête de performance ? etc.*
- *A qui profitera le système une fois qu'il sera créé ?*
- *Qui n'en profitera pas ?*

Il est également important de repérer les rôles et les responsabilités, ainsi que les structures existantes qui permettent déjà de suivre et d'évaluer les objectifs de développement.

- *Quels sont les rôles des ministères centraux et de tutelles associées à l'évaluation des performances ?*
- *Quel est le rôle du Parlement, de la cour des comptes, de la société civile ?*
- *Y a-t-il échange d'informations entre les ministères et les agences ?*
- *Qui produit les données au niveau du pays ?*
- *Quels sont les différents niveaux du gouvernement où les données sont utilisées ?*

Il est par ailleurs nécessaire de sonder les capacités actuelles de S&E de l'organisation ou du pays en ce qui concerne :

- les compétences techniques ;
- les compétences managériales ;
- les systèmes de données existants et leur qualité ;
- la technologie disponible ;
- les ressources disponibles ;
- l'expérience institutionnelle.

L'examen de l'état de préparation permet par ailleurs d'attirer l'attention sur les barrières existantes ou potentielles à la construction d'un système de S&E. Ce sont :

- l'absence de ressources fiscales ;
- l'absence de volonté politique ;
- l'absence de défenseur du système ;
- l'absence de stratégie ou d'expérience en congruence avec les résultats (pouvant aider à atteindre les résultats).

2.2. Etape 2 : Accord (s'entendre) sur les résultats devant faire l'objet de suivi et d'évaluation

Il est important de mettre l'emphase sur les résultats pour diverses raisons, notamment :

- rendre explicites les objectifs déclarés de l'action gouvernementale ;
- obtenir une meilleure idée de ce à quoi le succès ressemblera. Les résultats prouveront que les efforts ont été couronnés de succès. Ils montrent le chemin à suivre ;
- les résultats sont tout ce qui génère des bénéfices ;
- le fait de déterminer des résultats de manière claire est essentiel dans la conception et la création d'un système de S-E basé sur les résultats

On dit souvent qu'il faut *budgetiser jusqu'aux extrants, mais gérer jusqu'aux résultats !*

Les questions à se poser pour le choix des résultats à suivre et évaluer

La conception et le maintien d'un système de S&E axé sur les résultats requièrent des ressources (techniques, matérielles et financières). Or les pays disposent de ressources limitées et à juste titre, doivent établir des priorités pour l'allocation qu'ils en font. Il en est de même pour les organisations. Ainsi, durant le processus de conception d'un système de S&E axé sur les résultats, il faudrait s'assurer de faire un arbitrage participatif et très cohérent, de manière à consacrer les rares ressources et l'effort aux tâches et activités les plus productives. Plusieurs questions doivent être examinées lorsqu'on choisit les résultats qu'il convient de suivre et évaluer :

- *y a-t-il des buts nationaux/régionaux/sectoriels définis ou annoncés ?*
- *y a-t-il des promesses politiques ou engagements qui font cas d'une amélioration de performances dans un domaine particulier ?*
- *est-ce que des évaluations et les sondages des citoyens montrent des préoccupations particulières ?*
- *les prêts des bailleurs sont-ils liés à des buts spécifiques ?*
- *le gouvernement s'est-il sérieusement engagé à réaliser les ODD ?*

Illustration : Développement des résultats dans un domaine politique : l'Education

Le tableau ci-après présente deux exemples de résultats possibles dans le domaine de l'éducation. Ce cadre de performance servira d'illustration dans la suite. Les indicateurs, les valeurs de référence et les cibles découleront de cette étape initiale.

Tableau 4.2 : Cadre de performance – développer des résultats

Résultats :	Indicateurs :	Points de référence (Situation de référence):	Objectifs (cible):
1. Amélioration de la couverture des programmes préscolaires			
2. Amélioration des résultats des enseignements/apprentissages dans l'élémentaire			

2.3. Etape 3 : Choisir les indicateurs-clés pour le suivi des résultats

Comment saurons-nous que nous avons atteint le but ou le succès visé ? Où en sommes-nous dans la progression vers l'atteinte des réalisations souhaitées ? Tels sont deux principales questions auxquelles répondent les indicateurs.

Les indicateurs se définissent comme des mesures quantitatives ou qualitatives qui fournissent un moyen simple et fiable de mesurer les faits et donc de savoir ce qu'on a atteint, de mettre en lumière les changements liés à une intervention, ou de mesurer la performance d'une organisation en regard des réalisations prescrites. La formulation des indicateurs dans un SSEAR inclut souvent une cible à atteindre (contrairement à la conception plus stricte des indicateurs en évaluation développée au module 6). Ces cibles doivent notamment être précises en termes de

Quantité	→	Combien ?
Qualité	→	Comment ?
Temps	→	Quand ?
Lieu	→	Où ?
Coût	→	Unités monétaires

Développer des indicateurs adaptés (ou plus précisément des indicateurs avec des cibles) et leur renseignement permet aux gestionnaires d'estimer dans quelle mesure les résultats escomptés ou promis sont en voie d'être atteints. Ainsi, le développement des indicateurs est une activité centrale dans la construction d'un SSEAR. Elle détermine toute la collecte de données, l'analyse et la production de rapports qui viendront par la suite.

- **Pour la formulation des indicateurs dans un SSEAR, on retiendra notamment qu'ils doivent être** : objectivement vérifiables, quantifiables et observables. Quelle que soit la personne qui utilise la méthode, les résultats doivent être les mêmes ;
- **pertinents** : permet de mesurer les réels résultats d'une intervention;
- **explicites** : chaque niveau de réalisation d'un objectif doit être spécifiquement défini en termes de quantité, qualité et temps lieu et coût;
- **indépendants** : chaque indicateur doit être rapporté à un objectif/résultat donné.

La construction d'indicateurs et particulièrement des cibles associées nécessite beaucoup de travail et requiert particulièrement la prise en compte de plusieurs perspectives et dimensions, notamment les questions de fond, celles techniques et des politiques, la temporalité, etc. Les exemples ci-après décrivent les étapes successives à respecter.

Exemple 1/2 :

- commencer par l'objectif de base : les rendements de riz des petits exploitants ont augmenté
- ajouter la dimension quantité : les rendements de riz, en tonne par hectare. Les 30 000 tonnes, des petits exploitants, ont augmenté de 30%
- ajouter la dimension qualité : les rendements de riz (même qualité qu'en 2007) de 30 000 tonnes des petits exploitants (possédant 3 hectares ou moins) ont augmenté de 30%
- ajouter la dimension durée : les rendements de riz (même qualité qu'en 2007) de 30 000 tonnes des petits exploitants (possédant 3 hectares ou moins) ont augmenté de 30% à la fin de la campagne de 2008.

Exemple 2/2 :

Objectif : vacciner les enfants contre les maladies du PEV

1. Préciser l'objectif de base : enfants de 0 à 11 mois vaccinés
2. Fixer la quantité et/ou la qualité : 15 000 enfants vaccinés / et qui n'ont pas contracté la maladie
3. Préciser le cadre temporel : entre le 1er Janv. 2005 et le 31 Déc. 2006
4. Préciser le lieu : District de DOSSO, etc.

Enoncé de l'Indicateur :

- 15 000 enfants de 0 à 11 mois correctement vaccinés contre les maladies du PEV entre le 1er janv. 2005 et le 31 déc. 2006 dans le district de DOSSO (Niger) ; ou
- 15 000 enfants de 0 à 11 mois correctement vaccinés contre les maladies du PEV entre le 1er janv. 2005 et le 31 déc. 2006 dans le district de DOSSO, moyennant 100f/enfant vacciné.

Critères de validation d'un indicateur

Il existe essentiellement deux critères de validation des indicateurs et de leurs cibles. Il s'agit du critère CREAM dont les définitions sont présentées dans le tableau ci-dessous.

« CREAM »

- **C:** Clear (clair): précis et non ambigu, suffisamment précis pour assurer une mesure objective.
- **R:** Relevant (pertinent): approprié pour le sujet en question; reflétant aussi directement que possible l'objectif ou le processus qu'il est supposé représenter.
- **E:** Economic (économique): disponible à un coût raisonnable; permettant une collecte de données de manière pratique à un coût raisonnable.
- **A:** Adequate (adéquat) : fournissant une base suffisante pour évaluer la performance; sensible aux changements dans l'objectif ou le processus qu'il est supposé mesurer, mais relativement non affecté par d'autres changements; pouvant être désagrégé selon les besoins de l'analyse.
- **M:** Monitorable (susceptible d'être suivi): pouvant faire l'objet d'une mesure fréquente et objective (indépendance de l'évaluateur).

Matrice pour l'élaboration et l'utilisation des indicateurs

La sélection de bons indicateurs est donc un processus compliqué. La qualité de l'indicateur s'apprécie par ailleurs par la disponibilité de données pour le renseigner, le coût d'accès et de traitement de ces données, la pertinence ou la contribution attendue de l'information à produire dans le processus de décision, etc. La matrice pour l'élaboration et l'utilisation des indicateurs ou encore "la matrice de suivi" (tableau ci-dessous) est un bon outil qui permet de caractériser chaque indicateur en précisant les moyens de collecte et de traitement, la fréquence, et le niveau d'utilisation des données. Il renseigne par ailleurs les rôles et responsabilités, y compris les obstacles anticipés, et peut indiquer si le tout forme un ensemble réaliste, qu'il sera possible de renseigner sur le long terme.

Tableau 4.3: Matrice de suivi

Indicateurs	Source de données	Méthode de recueil de données	Qui recueille les données	Fréquence du recueil de données	Coûts liés au recueil de données	Difficultés liées au recueil de données	Qui analyse avec quelle méthode ?
1.							
2.							
3.							
4.							

Illustration :

Le tableau 4.3 illustre le choix d'indicateurs dans le cas de l'exemple dans le domaine de l'éducation, introduit à l'étape précédente.

Tableau 4.4 : Cadre de performance – choisir les indicateurs

Résultats :	Indicateurs :	Points de référence (Situation de référence)
1. Amélioration de la couverture des programmes préscolaires	<ol style="list-style-type: none"> 1. Taux d'enfants éligibles inscrits au préscolaire en zones urbaines 2. Taux d'enfants éligibles inscrits au préscolaire en zones rurales 	
2. Amélioration des résultats des enseignements/apprentissages dans l'élémentaire	<ol style="list-style-type: none"> 1. Pourcentage d'élèves du CM2 ayant une note d'au moins 70 sur 100 aux évaluations standardisées en maths et sciences 2. Pourcentage d'élèves du CM2 ayant, en évaluations standardisées en maths et sciences, une note supérieure à celles fixées dans les points de référence 	

2.4. Etape 4 : Recueillir (Rassembler) des données de référence (de base) sur les indicateurs

Une fois les bons indicateurs choisis (étape 3), il convient à présent de définir des points de comparaison, c'est-à-dire les conditions qui prévalent actuellement et par rapport auxquelles les changements attendus du résultat planifié pourront être comparés. On ne peut pas projeter un niveau de performance dans le futur ou fixer des cibles sans au préalable avoir établi une situation pour servir de référence. Le point de comparaison est la première mesure de l'indicateur. Il sert à en savoir plus sur les niveaux actuels ou récents de performance et les tendances qui s'en dégagent. Les données peuvent soit faire l'objet d'une collecte particulière, on parlera alors de données primaires, où être récupérées d'une autre collecte, ce sont alors des données secondaires. Alors qu'on peut contrôler la qualité des données primaires en effectuant sa propre collecte, cela voudra dire que pour tenir à jour le SSEAR, une collecte sur cet indicateur sera toujours nécessaire. En utilisant des données secondaires, la fiabilité est à vérifier, mais leur utilisation est moins coûteuse, et pourra certainement être plus régulière. Les chapitres 9 et 10 abordent en profondeur le sujet de la collecte et de l'analyse des données.

Illustration :

Le cadre de performance ci-dessous (tableau 4.4) présente un exemple de données témoins pour illustrer le cas de l'éducation introduit dans les chapitres précédents

Tableau 4.5 : Exemple de données témoins

Résultats :	Indicateurs :	Niveau de référence :	Objectifs :
1. Amélioration de la couverture des programmes préscolaires	1- Taux d'enfants éligibles inscrits au préscolaire en zones urbaines 2- Taux d'enfants éligibles inscrits au préscolaire en zones rurales	1. 75 % en zones urbaines en 1999 2. 40 % en zones rurales en 2000	
2. Amélioration des résultats des enseignements/a pprentissages dans l'élémentaire	1- Pourcentage d'élèves du CM2 ayant une note d'au moins 70 sur 100 aux évaluations standardisées en maths et sciences 2- Pourcentage d'élèves du CM2 ayant, en évaluations standardisées en maths et sciences, une note supérieure à celles fixées dans les points de référence	1. En 2002, 47% des élèves ont eu 70% ou plus en math et 50% ou plus en sciences. 2. En 2002 la plus faible note pour le CM2 était de 68% en math et 53% en sciences.	

2.5. Etape 5 : Planifier pour une amélioration : choisir (sélectionner) des objectifs réalistes

Dépendamment de la formulation des indicateurs (avec ou sans cible associée), l'étape 5 peut être superflue. Si des indicateurs sans cibles ont été retenus, c'est à cette étape que l'on fixe ces dernières.

La « cible » est « ...un but spécifique qui indique le nombre, l'horizon temporel et l'endroit de ce qui doit être réalisé » (FIDA, 2002, p. A-11). Fondamentalement, les cibles sont les niveaux quantifiables de progrès/rendement/performance/changement qu'un pays, une société ou une organisation souhaite atteindre à un moment spécifique.

Fixer des cibles constitue une étape décisive dans la mise en place d'un cadre de performance. Elle dépend des réalisations, des indicateurs et des points de comparaison. Le processus de raisonnement est déductif, en ce sens que tout découle des réalisations souhaitées. L'une des méthodes utilisées pour établir des cibles est de déterminer le point de départ des indicateurs et d'inclure le niveau souhaité d'amélioration (en tenant compte des ressources disponibles durant une période précise) pour arriver à la cible de performance. Il est crucial que le point de départ soit connu, tout comme les ressources disponibles pour progresser vers la cible pendant une période spécifique. On obtient ainsi la performance visée (la cible).

Exemple :

Les exportations agricoles augmenteront de 20% au-dessus du point de référence au cours des trois prochaines années

Attention toutefois !

- Fixer une cible de performance peut se révéler un exercice complexe. Il y a un certain nombre de facteurs à considérer, dont les points de comparaison qui doivent être clairement établis. Pour cela, on peut tenir compte de l'historique des données existantes pour projeter de nouvelles cibles de performance (par exemple, une moyenne de la performance des trois dernières années, la performance de l'an dernier, la tendance moyenne, les données au cours des six derniers mois, etc.)
- Les niveaux de ressources. La capacité existante, les budgets, le personnel, les ressources financières, les installations, etc. disponibles/mobilisables sur la période concernée ; et la capacité organisationnelle existante pour soutenir les activités et de fournir les extrants.
- l'échelle temporelle et le risque que des imprévus liés aux ressources et aux intrants affectent les cibles trop lointaines. Les cibles pour des périodes qui dépassent trois ou quatre ans ne sont pas recommandées. Dans le cas de politiques de développement qui s'inscrivent en général sur le long terme (10 à 20 ans, voire plus), il convient de fixer des cibles intérimaires sur des périodes plus courtes, qui correspondent à la performance escomptée à différents intervalles.
- Les préoccupations politiques, surtout lorsqu'il s'agit de programmes et politiques publiques. Il convient d'intégrer le fait que telles promesses faites par le

gouvernement ou l'administration pourraient influencer positivement ou négativement l'action visée.

Les effets de l'amélioration ne se constatent pas du jour au lendemain. Il est dès lors important d'être réaliste en fixant les cibles (résultats ciblés) et en évitant de trop promettre. Ce serait une façon de programmer à l'avance son échec. Cette fixation peut aussi être instrumentalisée pour montrer que l'on a réussi avec des cibles peu ambitieuses, ou inversement que l'on a encore besoin de soutien avec des cibles qui le sont trop.

Illustration : Fixer des cibles de performance pour un domaine politique : l'éducation

Résultats :	Indicateurs :	Niveau de référence :	Objectifs :
1. Amélioration de la couverture des programmes préscolaires	1- Taux d'enfants éligibles inscrits au préscolaire en zones urbaines	1. 75 % en zones urbaines en 1999	1. 85 % en zones urbaines à l'horizon 2006
	2- Taux d'enfants éligibles inscrits au préscolaire en zones rurales	2. 40 % en zones rurales en 2000	2. 60 % en zones rurales à l'horizon 2006
2. Amélioration des résultats des enseignements / apprentissages dans l'élémentaire	1- Pourcentage d'élèves du CM2 ayant une note d'au moins 70 sur 100 aux évaluations standardisées en maths et sciences	1. En 2002, 47% des élèves ont eu 70% ou plus en math et 50% ou plus en sciences.	1. A l'horizon 2006, 80% des élèves auront 70% ou plus en math et 67 % auront 70% ou plus en science
	2- Pourcentage d'élèves du CM2 ayant, en évaluations standardisées en maths et sciences, une note supérieure à celles fixées dans les points de référence	2. En 2002 la plus faible note pour le CM2 était de 68% en math et 53% en sciences.	2. En 2006 la plus faible note sera de 78% en math et 65% en sciences.

2.6. Etape 6 : Suivi axé sur les Résultats

Une fois les cibles de performance sélectionnées, vient ensuite l'étape de la collecte de données pour renseigner de manière continue les indicateurs de performance préétablis et fournir une base d'évidence pour la prise de décision de gestion qui pourrait induire des changements d'orientation ou de cap pour une intervention.

La mise en place d'un système pour suivre continuellement la performance est absolument essentielle en ce sens qu'il fournit de l'information en continu (par l'entremise des indicateurs) sur la direction, la vitesse et la magnitude des changements. Ceci est d'une importance critique pour savoir si les interventions évoluent dans la direction escomptée.

On distingue essentiellement deux (2) principaux types de suivi : le suivi axé sur la mise en œuvre et le suivi axé sur les résultats.

Des systèmes de suivi réussis

Pour réussir, tout système de suivi a besoin de s'appuyer sur ces quatre (4) éléments:

- l'appropriation par tous les acteurs – au niveau national, régional et local. Elle est soutenue par la demande d'information sur la performance à tous les niveaux par ceux qui utilisent le système. Avoir un champion politique assure l'appropriation du système, en ce sens que ce dernier soulignera l'importance de produire, partager et diffuser correctement de bonnes données sur la performance.
- gestion des données (Dispositif de collecte et de gestion de données) – qui est critique pour la durabilité du système. Qui s'en charge ? Comment et où ?
- entretien du système. Il est important de savoir qui récoltera quelles sortes de données et quand, et de s'assurer que l'information circule horizontalement et verticalement.
- crédibilité du système. Qu'il tient de sa capacité à produire aussi bien les bonnes que les mauvaises données.

2.7. Etape 7 : Production d'informations grâce à l'évaluation

Les données du suivi, telles qu'étudiées à la section précédente, ne permettent pas d'attribuer les changements à des causes particulières ni d'expliquer pourquoi ils se produisent. Elles permettent de noter qu'ils se produisent ou non. L'information générée par l'évaluation quant à elle permet d'aborder ces questions et d'autres questions importantes qui touchent l'atteinte de résultats.

Evaluation signifie des informations sûres :

Stratégie	<ul style="list-style-type: none"> • Si nous faisons ce qu'il faut faire <ul style="list-style-type: none"> – Justification – Une théorie du changement claire
Opération	<ul style="list-style-type: none"> • Si nous faisons les choses comme il le faut <ul style="list-style-type: none"> – Efficacité dans l'atteinte des résultats attendus – Efficience dans l'optimisation des ressources – Satisfaction des clients
Apprentissage	<ul style="list-style-type: none"> • Si nous faisons les choses comme il le faut <ul style="list-style-type: none"> – Efficacité dans l'atteinte des résultats attendus – Efficience dans l'optimisation des ressources – Satisfaction des clients

Source : adapté des travaux de Kusek, J. Z. (2006). Vers une culture du résultat. Dix étapes pour mettre en place un système de suivi et d'évaluation axé sur les résultats : Un Guide pour les praticiens du développement.

On déclenche des travaux d'évaluation :

- Chaque fois qu'il y a un résultat inattendu ou une performance aberrante qui nécessite une étude ou enquête complémentaire,
- Au moment où il est procédé à la répartition de l'allocation de budget ou de ressources au niveau des projets, programmes ou politiques,
- Au moment de prendre la décision d'étendre ou non une expérience pilote,
- Quand il y a une longue période sans amélioration, et ce pour des raisons non clarifiées,
- Quand des programmes similaires fournissent des résultats divergents.

Le suivi et l'évaluation sont ainsi complémentaires, soit que l'un ou l'autre pris isolément ne suffira pas. Leur complémentarité est avant tout séquentielle, ce qui signifie que l'information issue du suivi peut générer des questions auxquelles l'évaluation répondra plus tard – ou inversement, de l'information issue de l'évaluation peut mener à un suivi dans de nouveaux champs ou domaines. Leur complémentarité est aussi interactionnelle, c'est-à-dire que le suivi et l'évaluation peuvent utiliser les mêmes données, mais posent des questions différentes et structurent l'analyse différemment. Les analystes se servent alors en même temps de l'un et l'autre pour mener leurs initiatives.

2.8. Etape 8 : Rapporter les conclusions (présenter les informations)



Source : Kusek, J. Z. (2006). *Vers une culture du résultat. Dix étapes pour mettre en place un système de suivi et d'évaluation axé sur les résultats : Un Guide pour les praticiens du développement.*

Les données sur les résultats de performance découlent aussi bien du suivi que de l'évaluation, comme nous l'avons vu précédemment. Lorsqu'elles sont collectées, archivées et analysées, il est important ensuite de définir ce qui sera présenté, par qui, sous quel format, quand et à qui.

La raison d'être principale du suivi tout comme de l'évaluation est d'informer des parties prenantes ciblées, des résultats et des conclusions issus de la collecte, de l'analyse et de l'interprétation de l'information produite. Il est alors important de tenir compte de leurs intérêts et attentes, mais aussi de leurs moyens de communication préférés.

Quelques règles à observer

- Pour que le système d'information puisse fournir une rétroaction continue sur la performance, il faut instituer une communication permanente. Les résultats du suivi et de l'évaluation devraient être continuellement diffusés pour fournir une rétroaction aux décideurs.
- Les informations devraient être présentées d'une façon succincte et précise et refléter les intérêts de l'auditoire ciblé ; prendre garde à les formater différemment selon les intérêts et les préférences de chacun.
- Seules les données les plus importantes devraient être présentées ; prendre garde à ne pas noyer l'information dans un torrent de données.
- La comparaison des réalisations concrètes avec les cibles est un élément central de la présentation des résultats, de même il est recommandé de présenter les données en les comparant aux données précédentes et aux données de départ.
- La présentation continue de données peut et doit s'étendre aux conseils pour guider les décideurs dans la mise en œuvre des recommandations.
- Comparer les données relatives aux indicateurs, aux points de référence, aux objectifs et fournir ces informations dans des écrans d'affichage (supports) faciles à comprendre.
- Comparer les informations actuelles avec les données passées et chercher les caractéristiques et les tendances.

- Attention à ne pas tirer des conclusions générales basées sur peu d'informations : plus vous avez de données (d'informations), plus vous serez certain que les tendances sont réelles.
- Protéger le messager : les gens qui livrent de fausses nouvelles/informations ne devraient pas être punis. Les conclusions controversées peuvent révéler de nouvelles tendances ou prévenir très tôt les chefs de projet sur des problèmes éventuels, leur permettant ainsi d'avoir le temps nécessaire pour trouver des solutions.

Pourquoi communiquer les données obtenues ?

- Pour livrer des renseignements sur le statut des projets, programmes ou politiques ;
- Pour fournir des indices aux problèmes ;
- Pour créer des occasions pour prendre en charge les changements ;
- Pour fournir des informations dans le temps sur les tendances et les orientations ;
- Pour aider à confirmer ou à contester la validité de la théorie du changement.

2.9. Etape 9: Utilisation des conclusions

L'objectif principal d'un système de suivi-évaluation n'est pas simplement de générer continuellement de l'information axée sur les résultats, qui soit formatée et présentée adéquatement. Il est aussi important que cette information soit transmise aux utilisateurs concernés au bon moment, de façon qu'elle alimente la prise de décision basée sur l'évidence - en vue de l'amélioration de la performance. C'est la raison d'être principale d'un SSEAR. Si l'on prend le temps de surveiller, mesurer, évaluer et intégrer les données produites, c'est pour mieux fonder la prise de décision. S'il n'existe aucune donnée pour appuyer les décisions, celles-ci seraient arbitraires et inefficaces.

Il existe plusieurs usages pour les données obtenues :

1. Satisfaire l'exigence de responsabilité de la part des autorités élues et du public
2. Aider à formuler et à justifier des requêtes budgétaires.
3. Aider à prendre des décisions opérationnelles relatives à l'allocation de ressources.
4. Déclencher des examens approfondis sur les problèmes de performance soulevés et les solutions y afférentes.
5. Aider à motiver le personnel à poursuivre les améliorations des programmes.
6. Faire le suivi des performances du projet ou programme par rapport aux objectifs de résultats.
7. Fournir des données pour des évaluations spéciales et en profondeur de programme.
8. Faire le suivi du rendement de la prestation de services par rapport aux objectifs précis de résultats.
9. Fournir un soutien stratégique et d'autres efforts de planification à long terme.
10. Communiquer avec le public pour instaurer une confiance publique.

Les stratégies de communication doivent être adaptées pour se conformer aux besoins des parties prenantes spécifiques qu'on vise à atteindre – le Parlement, les ministres, les médias, le secteur privé, les ONG, les organisations de la société civile, le public, etc. En ce sens, il peut s'avérer un exercice complexe. Ci-après quelques orientations :

- Renforcer les médias et les associer à la diffusion de l'information – Dans leur rôle de gardien de l'information et tireur d'alerte, ils peuvent s'avérer d'excellents partenaires pour diffuser les résultats obtenus par les systèmes de suivi-évaluation axés sur les résultats ;

- Appliquer la loi sur la “liberté d’information” – autorisant quiconque à accéder aux informations sur la performance du gouvernement, lorsqu’elles ne compromettent pas la sécurité du pays ;
- Instituer le e-gouvernement
- Ajouter les informations aux sites internet internes et externes – créer des bases de données en ligne et ouvertes d’accès, comprenant les résultats du suivi-évaluation ;
- Publier les rapports des budgets annuels
- Impliquer la société civile et les associations citoyennes
- Renforcer et impliquer le contrôle législatif ou les institutions supérieures de contrôle tels que la Cour des Comptes
- Partager et comparer les conclusions sur les résultats avec les partenaires au développement

2.10. Etape 10 : Maintenir le système de S&E dans l’organisation

La conception d’un système de suivi-évaluation est un processus complexe qui devrait s’inscrire dans la durée. Six (6) éléments favorisent le maintien des systèmes de suivi-évaluation axés sur les résultats.

- La demande - si celle-ci est épisodique et varie au gré des circonstances, les systèmes de suivi-évaluation axés sur les résultats ne seront pas utilisés ni entretenus. Stimuler une demande soutenue requiert des exigences formelles en termes de redevabilité et de reddition des comptes, incluant des lois et des réglementations assorties de mesures incitatives et coercitives.
- Les rôles et responsabilités clairement définis - une hiérarchie organisationnelle et politique formelle doit être établie, affichant les rôles et des responsabilités clairs en termes de l’organisation et les personnes clés responsables de la collecte, de l’analyse, de la diffusion des données et de l’utilisation de l’information.
- L’information sur la performance crédible et digne de foi – en ce sens qu’elle doit être impartiale et transparente, générer à la fois de bonnes et mauvaises

nouvelles en toute objectivité et bonne sincérité, et surtout sujette à vérification indépendante.

- La responsabilité – nul (parmi les parties prenantes) n'est exonéré ou exempté face à ses responsabilités.
- Les capacités et compétences techniques appropriées pour assurer la collecte et l'analyse des données nécessaires pour le bon fonctionnement et le maintien du système.
- Les incitations – en ce sens qu'il faut encourager l'utilisation de l'information sur la performance, soit que le succès est reconnu et récompensé ; qu'il faille s'attaquer aux problèmes ; que l'apprentissage organisationnel est valorisé ; etc.

Mais il existe aussi des facteurs de risque pour le maintien d'un système de suivi-évaluation :

- Les coûts et la faisabilité du système dans son ensemble ;
- La peur et la résistance des gestionnaires des programmes ;
- Les besoins en formation du personnel ;
- Les changements dans les priorités législatives et celles du Gouvernement ;
- La stabilité des indicateurs dans le temps ;
- L'instabilité politique ;
- Etc

Voici donc quelques recommandations pour maintenir un SSEAR dans la durée :

- Maintenir un plaidoyer (ministère des finances, parlement, etc.) sur le besoin de données régulières du SSEAR pour satisfaire aux exigences de la gestion axée sur les résultats ;
- Chercher toute opportunité pour établir une relation entre les informations sur les résultats, le budget et les décisions relatives à l'allocation de ressources pour prouver l'utilité du SSEAR ;
- Commencer avec des expériences pilotes pour montrer l'efficacité du SSEAR pour intéresser les décideurs et les habituer à utiliser les données (ils en demanderont plus ensuite !) ;
- Commencer par une stratégie isolée (ex: des îlots d'innovations) au lieu d'une approche pangouvernementale ;

III. Autres approches de développement d'un système de suivi évaluation

En dehors de la démarche en dix (10) étapes proposées par la Banque Mondiale pour la mise en place d'un SSEAR, d'autres approches sont développées par différents acteurs en fonction des types d'actions à suivre et à évaluer :

Approche globale

Approche isolée

Approche mixte

Approche globale du SSEAR

Le S&E est étendu à tous les niveaux du gouvernement

- Adoptée dans la plupart des pays pionniers en matière de S&E
- Difficile à adopter dans les pays où différents ministères sont à différents stades du processus de S&E

Les Objectifs du Millénaire pour le Développement ont créé une impulsion de cette approche dans beaucoup de pays en développement qui est poursuivie avec les ODD et leur suivi.

Approche isolée du SSE

Le S&E est centré sur une partie ou un secteur du gouvernement

- S&E piloté dans quelques ministères ou agences – clés
- Menée au niveau local, national ou régional

Approche mixte

Elle combine l'approche globale et l'approche isolée. Certains secteurs ont une approche globale, d'autres ont une attention plus sporadique.

Messages clés :

- ✓ La mise en place d'un SSEAR est un travail de longue haleine, qui est l'opérationnalisation d'une gestion axée sur les résultats sur l'ensemble du cycle de vie d'une intervention.
- ✓ 10 étapes de travail complémentaires sont constitutives de la mise en place et de l'utilisation d'un SSEAR, du choix des indicateurs, à leur renseignement par des collectes de données régulières et des travaux évaluatifs lorsque nécessaire, à la communication et l'utilisation des résultats par les décideurs.
- ✓ La mise en place et le maintien d'un SSEAR est l'occasion de comprendre les complémentarités entre le suivi et l'évaluation dans la gestion d'une intervention, et ainsi de savoir quand déclencher une évaluation en fonction des données de suivi, pour informer la prise de décision.

Bibliographie

- FIDA (2002), « Guide pratique de suivi et d'évaluation des projets : pour une gestion orientée vers l'impact ».
- Jody et Ray (2004). Vers une culture du résultat vers une culture du résultat : Dix étapes pour mettre en place un système de suivi et d'évaluation axé sur les résultats, Nouveaux Horizons, Editions Saint-Martin, Montréal (Québec), 234p.
- Osborne, D et T. Gaebler (1992). Reinventing government: How the entrepreneurial spirit is transforming government. New York : Penguin Books
- World Bank, "Ten Steps to a Results-Based Monitoring and Evaluation System: A Handbook for Development Practitioners." Washington, DC.

MODULE 5

CONTEXTE DE L'ÉVALUATION ET THÉORIE DU CHANGEMENT

Objectifs d'apprentissage :

A la fin de la session, les participants seront capables de :

- définir la théorie du changement ;
- comprendre la démarche d'élaboration d'une théorie du changement ;
- situer le suivi et l'évaluation dans la théorie de changement.

Thèmes couverts

- I. La logique des interventions
- II. La chaîne de résultat
- III. Les hypothèses dans les interventions
- IV. Le suivi et l'évaluation dans la théorie de changement

Ce module est consacré à un outil important dans la pratique des évaluateurs, à savoir les différentes représentations de la logique d'intervention. Une des premières tâches de l'évaluateur consiste en effet à bien comprendre l'objet d'évaluation, ce qui s'effectue souvent, grâce à de l'analyse documentaire (les documents de présentation de l'intervention, et les principaux documents de suivi), ou à l'aide d'entretiens exploratoires, par exemple avec l'organisation responsable de l'intervention.

Cette représentation graphique permet de présenter l'intervention sous la forme d'un changement entre une situation diagnostiquée comme problématique, et une direction souhaitée : les effets de l'intervention, qui doivent en principe résoudre le problème à la base de l'intervention. C'est pourquoi on parle de la logique d'intervention.

Puisque les interventions publiques, notamment celles qui visent le développement ne sont pas mises en œuvre dans des contextes simples, et qu'elles touchent des situations problématiques complexes, il apparaît pertinent de se demander comment inclure ces difficultés dans les représentations des logiques d'intervention, pour en faire des cartes d'identité réalistes des interventions évaluées.

Mots-clés: théorie de programme, logique d'intervention, chaîne des résultats, théorie de changement

I. La logique d'intervention

La logique d'intervention, également appelée théorie d'intervention, est une représentation de la manière dont une intervention entend mener à des résultats, c'est-à-dire à un changement par rapport à une situation problématique de départ, que l'on appelle souvent la raison d'être de l'intervention. Cette conception des interventions publiques constitue la base de la gestion axée sur les résultats. Les analyses de la logique d'intervention sont très importantes en évaluation et se fondent sur une analyse des liens de cause à effet (causalité) reliant une intervention publique à ses effets. Plusieurs représentations graphiques sont utilisées pour traduire ces effets. Il s'agit de la chaîne des résultats, le modèle logique ou la théorie du changement. Ces outils visuels contribuent à clarifier les relations de cause à effet entre des ressources investies dans une intervention, les activités que cela permet de mener et les résultats engendrés par ces activités, qui consistent, si l'intervention est performante, en une solution adéquate au problème identifié au démarrage.

La première étape de l'élaboration de la théorie d'intervention consiste habituellement à identifier un modèle simple que l'on appelle la chaîne des résultats, qui peut être enrichie par la suite.

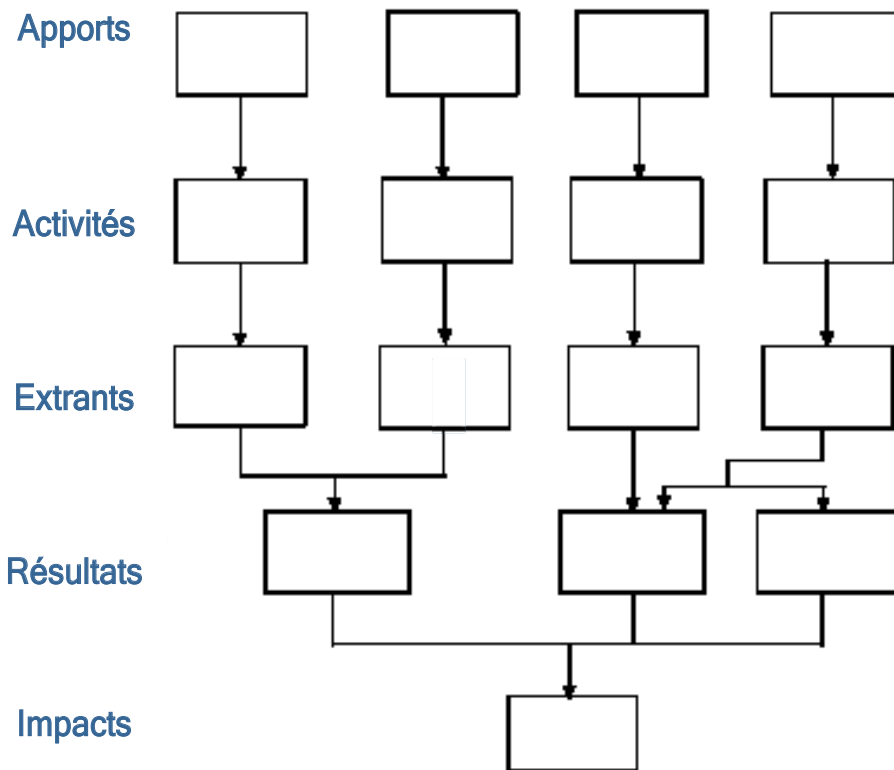
Les principales composantes de la chaîne des résultats sont les suivants :

- **Les effets:** ce sont les conséquences de la livraison des extrants sur les populations cibles de l'intervention. Ils peuvent s'échelonner à plus ou moins long terme et être donc caractérisés de directs ou immédiats (arriver en même temps que les extrants), d'intermédiaires ou de finaux (on peut dans ce cas également les appeler impacts).
- **Les extrants:** ce sont les livrables directs des activités, sous la responsabilité du maître d'œuvre de ces dernières. Ils sont généralement concrets et dénombrables.
- **Les activités:** c'est l'ensemble des opérations de terrain menées grâce à l'investissement des intrants. Elles sont souvent formulées à l'aide d'un verbe d'action.

→ **Les intrants:** il s'agit des ressources financières, humaines, matérielles, ou d'autres natures qui sont investies dans l'intervention pour réaliser les activités.

Les chaînes de résultats peuvent être tracées de manière verticale des intrants aux effets finaux comme ici de façon descendante, en mettant en valeur les liens entre les cases grâce à des flèches :

Graphique 5.1 : Représentation verticale d'une chaîne de résultats



D'autres représentations préfèrent la forme horizontale, plus compacte pour résumer une intervention

Graphique 5.2 : Représentation horizontale d'une chaîne de résultats



La diversité des formes de représentation ne doit pas trop inquiéter les évaluateurs novices, puisqu'il s'agit avant tout de synthétiser la logique d'intervention en une page, de manière graphique, pour comprendre les liens de cause à effets que l'intervention souhaite provoquer. La forme, voire le nom du graphique (le vocable modèle logique est également souvent utilisé pour ces deux types de représentations) dépendra donc souvent des manuels d'évaluation de chaque organisation publique, ou à défaut de cadre ou d'un manuel de pratiques, des préférences des évaluateurs et des parties prenantes de l'évaluation.

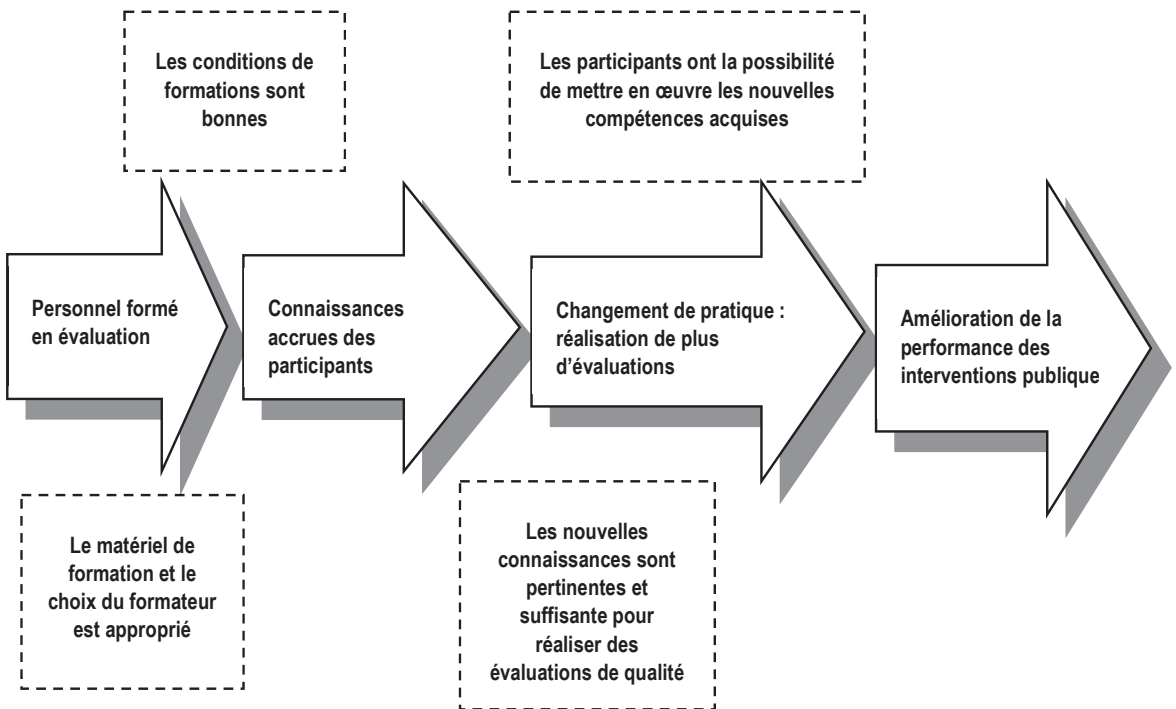
La théorie de programme représentée par une chaîne des résultats examine nos idées ou plutôt la situation idéale de changement, qui est supposée se réaliser à la suite d'une intervention. Il n'est pas évident de pouvoir organiser une évaluation, ou de comprendre une intervention sans aucune théorie de programme. Celle-ci permet donc de :

- montrer comment l'intervention est censée fonctionner et peut ainsi mettre en valeur des défauts sur le plan de la logique ;
- démarrer l'élaboration des questions d'évaluation ;
- élaborer une compréhension commune des interventions et des attentes des parties prenantes en tant qu'outil participatif pour une interprétation partagée du changement souhaité ;
- de fournir une carte d'identité simple et graphique d'une intervention pour les besoins de gestion (la chaîne des résultats est d'ailleurs souvent utilisée dans les cadres logiques pour la mise en œuvre de l'intervention).

II. La théorie du changement

La chaîne des résultats n'est pas pour autant la seule représentation utilisée par les évaluateurs et elle tend plutôt à être remplacée par une représentation plus complète : la théorie de changement. Comme dans les modèles précédents, la théorie de changement est construite sur la logique d'intervention, allant des intrants aux effets, mais elle y ajoute des hypothèses ou conditions de succès. L'ajout de ces hypothèses rend la carte d'identité de l'intervention beaucoup plus réaliste et complexe, car elle permet d'entourer le scénario idéal de changement représenté par la chaîne des résultats, de son contexte d'implantation et des difficultés de terrain qui pourraient entraver la mise en œuvre de l'intervention. Une théorie de changement simple d'une formation en évaluation pourrait être représentée comme suit :

Graphique 5.3 : Exemple de théorie du changement simplifiée d'une formation en évaluation

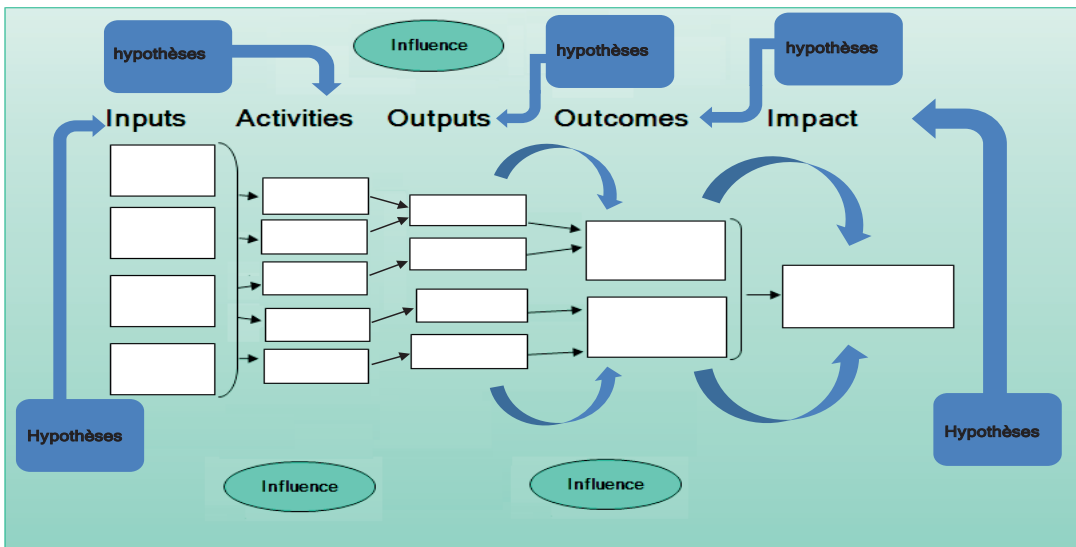


On comprend ici que les hypothèses viennent entourer la chaîne des résultats, pour présenter les conditions qui doivent être présentes pour que l'intervention soit un succès et amène les effets escomptés. Sans la présence de ces hypothèses, il est

probable que l'intervention soit nettement moins performante. La carte d'identité ainsi formée s'avère donc plus réaliste et complète et évite de donner l'impression de liens de cause à effet automatiques. C'est pourquoi de nombreuses organisations ont adopté la théorie de changement.

En plus des hypothèses à proprement parler, on peut également choisir de mettre en valeur des facteurs externes à l'intervention qui ont aussi exercé une influence sur la réalisation des effets, s'ils éclairent la compréhension de l'intervention. Il revient à l'évaluateur de garder l'esprit synthétique de la représentation, tout en donnant une vue la plus réaliste possible du contexte d'implantation de l'intervention.

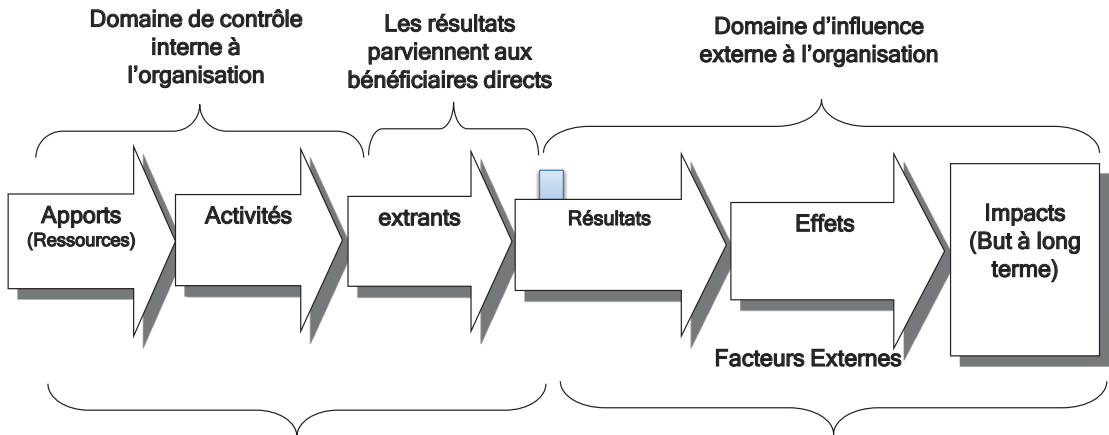
Graphique 5.4 : Modèle de théorie du changement



Pour expliquer la théorie de changement, on utilise souvent le raisonnement "si-alors" : Si **X** se produit, alors **Y** devrait suivre (X étant ici une hypothèse, Y le passage d'une case à une autre de la chaîne de résultats).

On se pose aussi les questions suivantes : Est-ce que ces activités ou extrants ont des chances de mener aux résultats souhaités ? Est-ce que toutes les conditions sont réunies pour que les résultats émergent? (*est-ce plausible ?*). Est-ce que les intrants (capacités et ressources) suffisent pour les activités ? « Est-ce faisable ? ».

Il convient de noter que plus on avance dans la chaîne des résultats, plus le contexte et les facteurs externes peuvent entraver la réalisation d'effets. C'est pourquoi on distingue souvent entre le domaine de contrôle de l'organisation maître d'œuvre ou la chaîne de production (intrants — activités — extrants) et les effets qui sont plutôt une zone d'influence de l'intervention sur une population cible, les bénéficiaires..



Généralement, ce sont les évaluateurs ou les planificateurs qui élaborent la théorie. Mais de plus en plus, notamment dans le cadre d'une évaluation participative, les parties prenantes peuvent être impliquées dans son élaboration. Ceci permet notamment aux parties prenantes, notamment les gestionnaires de l'intervention évaluée de s'approprier les travaux d'évaluation, et de se resservir ensuite de la théorie de changement comme outil de gestion axée sur les résultats. D'autre part, l'évaluateur ne part jamais de rien pour élaborer la théorie de changement, d'autant plus que la gestion axée sur les résultats a été adoptée dans de nombreuses administrations. Il est donc possible de trouver dans la documentation de programme certains morceaux de la théorie de changement, par exemple dans un cadre logique ou dans le système de suivi-évaluation de l'intervention ou des idées d'hypothèses dans un registre de risque. La tâche de l'évaluateur revient alors à remettre toutes les pièces ensemble, en s'assurant qu'il n'en manque aucune (certaines étapes dans la chaîne des effets peuvent par exemple manquer), en montrant la dynamique générale grâce aux flèches et en systématisant la réflexion sur les hypothèses à chaque niveau de la chaîne des résultats, tout en gardant le vocabulaire utilisé par l'organisation responsable de l'intervention.

Messages clés :

- ✓ L'élaboration de la théorie d'intervention est une des étapes-clés du travail d'évaluateur pour comprendre les objets d'évaluation, et le changement qu'ils souhaitent apporter à une situation jugée comme problématique.
- ✓ Il existe différentes représentations graphiques des théories d'intervention, selon les auteurs, les usages dans différentes organisations et les préférences des évaluateurs et de leurs clients. Une représentation n'est pas en soi meilleure qu'une autre et il s'agit surtout de comprendre leur diversité et leurs usages.
- ✓ La théorie du changement est une représentation graphique de plus en plus utilisée, qui combine la représentation de la chaîne des résultats et du contexte d'implantation de l'intervention représenté par l'ajout d'hypothèses. Ces hypothèses permettent d'obtenir une vue plus réaliste d'interventions qui s'attaquent à des problèmes complexes pour lesquels de nombreuses conditions de succès doivent être réunies.

Bibliographie :

- Marceau R. et Sylvain F. (2014) *Dictionnaire terminologique de l'évaluation : politiques, programmes, interventions*, Québec, Les Éditions GID, 105p.
- Marceau, R., Otis, D. et Simard, P. (1992). « La planification d'une évaluation de programme », dans Parenteau, R. (éd.) *Management public : comprendre et gérer les institutions de l'État*, Sillery, Presses de l'Université du Québec, p. 445-480.
- Mayne, J. (2008). « Contribution analysis: An approach to exploring cause and effect », *ILAC Brief*, 16.
- Mayne, J. (2012). « Making causal claims », *ILAC Brief*, 26.
- Mayne J. (2015) «Useful Theory of change models», *Canadian Journal of Program Evaluation*, 30(2), pp.119-142.
- Morra Imas, L.G. et Rist, R.C. (2009). *Road to Results: Designing and Conducting Effective Development Evaluations*. Washington, The World Bank, 592p. Chapitre : 4
- Rogers, P. J. et Weiss, C. H. (2007), «Theory-based evaluation: Reflections ten years on: Theory-based evaluation: Past, present, and future», *New Directions for Evaluation*, pp. 63–81.

MODULE 6

ÉLABORATION DES QUESTIONS D'ÉVALUATION ET DE LA MATRICE D'ÉVALUATION

Objectifs d'apprentissage :

A la fin de la session, les participants seront capables de :

- Élaborer des questions pertinentes d'évaluation
- Construire une matrice d'évaluation

Thèmes couverts

- I. Élaboration de questions pertinentes d'évaluation
- II. Conception d'une matrice d'évaluation

Le présent module s'attèle aux premières étapes de l'élaboration d'une évaluation. Ces étapes comprennent la définition du type de questions auxquelles l'évaluation répond et les différentes composantes d'une matrice d'évaluation.

La définition des questions d'évaluation est un moment clé de l'évaluation puisqu'elle permet de formuler le fil conducteur que l'évaluation doit suivre pour répondre aux besoins d'informations. La formulation de questions d'évaluation porte d'une manière générale sur les critères d'évaluation, que sont par exemple la pertinence, l'efficacité, l'efficience, la durabilité ou viabilité, la cohérence (et plusieurs autres selon les approches d'évaluation et les référentiels des organisations publiques).

Outil d'organisation pour aider à planifier une évaluation (CDRT, 2015), la matrice d'évaluation est quant à elle décrite comme un tableau présentant la stratégie d'évaluation qui sera déroulée pour chaque question d'évaluation, avec les besoins en matière de collecte d'informations (y compris les variables concernées et les indicateurs choisis) et éventuellement les méthodes d'évaluation proposées. La matrice d'évaluation aide donc à dresser systématiquement le plan de l'évaluation et décrit avec précision, rigueur et transparence la méthodologie qui va être suivie par l'équipe, afin qu'elle puisse asseoir la crédibilité du jugement évaluatif qu'elle va dérouler.

Mots clés : questions d'évaluation ; matrice d'évaluation

I. Elaboration de questions pertinentes d'évaluation

Les questions de l'évaluation sont les fils conducteurs du travail d'évaluation. Ce sont ces questions qui déterminent ce qu'il est important d'aborder et d'évaluer, c'est l'angle d'attaque de l'équipe sur l'intervention, compte tenu de l'impossibilité de « tout évaluer » dans une même mission. De leur choix découle ensuite un grand nombre de décisions : approche d'évaluation, données à considérer, méthodes à employer, collectes, analyses... Un de ces éléments peut aussi guider la formulation des questions pour aboutir à une stratégie d'évaluation cohérente. C'est en trouvant les bonnes questions, celles qui sont pertinentes au moment de l'évaluation, qui vont apporter de nouvelles informations et qui sont utiles pour la prise de décision sur l'intervention, que l'évaluateur maximise l'utilité de sa mission. Le travail d'évaluation consiste ensuite à générer des arguments crédibles pour répondre à cette question (Gertler et al., 2011, p 40).

Les questions d'évaluation jouent donc un rôle essentiel par conséquent, il est nécessaire de prévoir assez de temps et de discussions avec vos principaux intervenants (équipe d'évaluation et commanditaires voire comité d'évaluation) afin de les mettre au point et de les sélectionner.

Dans le processus d'évaluation, les questions évaluatives permettent de :

- établir une orientation pour l'évaluation ;
- refléter les préoccupations majeures des décideurs de l'intervention ou du citoyen selon la mission donnée à l'équipe d'évaluation ;
- donner un fil conducteur aux travaux qui vont être menés et au rapport ;
- communiquer simplement l'angle d'attaque que les évaluateurs adoptent.

La **provenance des questions** d'évaluation est multiple : i) conclusions des évaluations précédentes en lien avec l'intervention ; ii) conclusions des évaluations d'interventions similaires ; iii) hypothèses et facteurs contextuels identifiés dans le développement de la théorie du changement ; iv) préoccupations et questions du client et des intervenants ; v) normes professionnelles ; vi) points de vue et connaissances des experts consultés ou questionnements de la population sur l'intervention ; vii) jugement professionnel de l'évaluateur. Elles diffèrent aussi grandement en fonction de l'approche d'évaluation (le paradigme dans lequel l'équipe

d'évaluation se place) et les buts de l'évaluation (formatif ou sommatif) et le moment de l'évaluation (ex-ante, ex-post, à mi-parcours ou en continu).

A cet effet, on distingue généralement plusieurs types de questions d'évaluation : i) questions descriptives ; ii) questions évaluatives ; iii) questions sur les rapports de causalité

1.1. Questions descriptives

Les questions descriptives :

- cherchent à comprendre ou à décrire un programme, un processus ou des attitudes envers ce programme ;
- s'intéressent aux questions fondamentales qui sont (Qui ? Quoi ? Où ? Quand ? Comment ? Combien ? Pourquoi ?) ;
- peuvent être utilisées pour décrire des rapports, des activités ou des résultats ;

Les questions descriptives sont multiples et peuvent se présenter sous différentes formes. On peut citer en guise d'exemple :

- *Quels sont les objectifs du programme du point de vue des acteurs ?*
- *Quelles sont les premières réalisations du programme ?*
- *Où se déroule le programme ?*
- *Qui a bénéficié de quels services ?*
- *Quels sont les obstacles auxquels le programme fait face ?*
- *Quelle était la moyenne des gains en rapport avec les résultats des participants ?*
- *À quel point la conception du programme reflète-t-elle les leçons tirées de programmes similaires précédents ?*
- *Quel est le niveau des différences systématiques entre les sites dans la manière dont le programme s'exécute ?*
- *Quelles sont les qualifications des fournisseurs ou prestataires de services ?*
- *De quand date la mise en œuvre complète du programme ?*
- *Quelle est la proportion des femmes ayant participé au programme ?*
- *Comment le coût du programme se compare-t-il aux coûts des programmes similaires ?*
- *Quels sont les moyens de communication informels dont dispose l'organisation ?*

- À quel point les participants ont-ils trouvé utile le programme ?

Bien que la description de la réalité de mise en œuvre d'une intervention sur le terrain soit à la base du travail de l'évaluateur, les questions descriptives ne font pas forcément les meilleurs rapports. Elles encouragent en effet l'évaluation à s'en tenir uniquement à une description sans ensuite interpréter et « trancher » sur ce que cette réalité veut dire pour la performance de l'intervention évaluée. C'est pourquoi elles constituent parfois des préalables à des questions plus intéressantes : les questions évaluatives.

1.2. Questions évaluatives

Les questions évaluatives :

- identifient une comparaison entre “ce qui est” et “ce qui devrait être” ou ce qui est performant;
- font un lien entre l'existant et un critère d'évaluation ;
- obligent l'évaluateur à conclure sur la performance de l'intervention évaluée.

Les questions évaluatives incluent donc souvent de se référer à une norme (ce qui est performant). C'est pourquoi on parle également de questions normatives. On peut retrouver ces normes :

- dans les documents autorisant le programme ;
- à l'aide des indicateurs avec des cibles spécifiques dans les systèmes de suivi ;
- dans des systèmes d'accréditation, des référentiels du métier, des comités d'experts et d'organisations professionnelles ;
- à travers l'évaluation, c'est-à-dire discuter de ce qui est performant ou non au moment des collectes d'information.

Des questions normatives peuvent être reformulées de la façon suivante :

- *le budget a-t-il été dépensé de manière efficiente?*
- *Dans quelle mesure peut-on dire que l'intervention a répondu au besoin des populations ?*
- *Avons-nous atteint les objectifs fixés par le système de suivi-évaluation du programme ?*

- *En quoi le processus de sélection des participants est-il juste et équitable ?*
- *A quel point le programme reflète-t-il les principes de la parité ?*

Ce qui est intéressant ici, c'est qu'en plus d'une simple description, on fait le lien avec un critère d'évaluation et on encourage l'évaluateur, en répondant à la question, à réellement émettre un jugement évaluatif (pas personnel, mais étayé avec des faits).

1.3. Questions sur les rapports de causalité

Globalement, les questions sur les rapports de causalité :

- visent à déterminer la différence faite par l'intervention;
- cherchent à savoir si les résultats souhaités ont été atteints et si ces résultats ont été le fruit de l'intervention ;
- impliquent la situation avant et après l'intervention et la notion de changement;

Ces questions sont souvent abordées grâce aux méthodes d'évaluation des effets présentés dans le chapitre suivant.

Des illustrations de questions de cause à effets peuvent se présenter de la forme suivante :

- *la stratégie de partenariat entre les trois pays a-t-elle préservé la biodiversité de la zone affectée tout en assurant les moyens de subsistance durables ?*
- *À la suite du programme de formation à l'emploi, les participants ont-ils obtenu des emplois plus rémunérateurs que ceux auxquels ils pouvaient prétendre avant la formation ?*
- *Le programme relatif aux microentreprises a-t-il contribué à réduire la pauvreté dans les milieux où il s'est déroulé ?*
- *Quels sont les domaines d'effet ou d'impact attribuables à l'intervention ?*

II. La matrice d'évaluation

La matrice d'évaluation est un tableau présentant les questions relatives à la performance de l'intervention, opérationnalisées en variables et indicateurs ainsi que les besoins en matière de collecte d'informations. Cette matrice indique comment, quand, par qui et où les données seront recueillies (FIDA, 2000). Toutefois, plusieurs modèles de matrice coexistent et chaque organisation ou équipe d'évaluation a tendance à préférer un modèle plutôt qu'un autre. Il est surtout important de bien comprendre que la matrice d'évaluation est un outil d'organisation pour aider à planifier une évaluation (CDRT, 2015), en d'autres termes, elle est la clé pour réussir en évaluation. Ainsi peu importe sa mise en page exacte, elle doit lier les questions d'évaluation à la stratégie de collecte et d'analyse de données qui sera mise en œuvre par l'équipe.

2.1. Elaboration de la matrice d'évaluation

Globalement l'on peut distinguer quatre étapes pour la définition de la matrice d'évaluation :

Etape 01 : Formulation des questions d'évaluation selon les différents critères ;

Etape 02 : Identification des indicateurs/mesures associés à chacune des questions d'évaluation ;

Etape 03 : Identification des sources, des méthodes/outils de collecte et méthodes/outils d'analyse des données ;

Etape 04 : Analyse de la cohérence entre la matrice et les termes de préférences.

La matrice d'évaluation organise les questions et les stratégies de collecte de l'information afin de répondre aux questions (CDRT, 2015). De façon spécifique, la matrice d'évaluation permet de :

- présenter une synthèse des questions d'évaluation ;
- présenter pour chacune des questions-clés, les variables et indicateurs retenus ;
- présenter les sources, méthodes de collecte et d'analyse des données ;

- présenter l'orientation méthodologique que l'équipe d'évaluation a choisie.

Lorsque les questions ont été précisées et organisées selon les différents critères, il s'agit de préciser quelles variables de la théorie de changement de l'intervention sont concernées, les indicateurs retenus pour comprendre la variable, les sources de données et les stratégies de collecte qui seront employées, et les méthodes d'évaluation à utiliser, particulièrement pour les questions d'évaluation qui touchent les effets. Encore une fois, les modèles de matrice peuvent varier d'une organisation à une autre, le tableau ci-dessous en propose une forme possible.

Tableau 6.1 : Ebauche d'une matrice d'évaluation

	Variable	Indicateurs	Source et méthode de collecte de données	Méthode d'évaluation (si nécessaire)
Question 1				
Question 2				
Question 3				
...				

2.2. Variables et indicateurs

La première étape à réaliser une fois les questions d'évaluation fixées est de faire le recensement des variables concernées par la question. La colonne des variables dans la matrice d'évaluation fait le lien entre la théorie de changement et la stratégie d'évaluation, car les variables ne sont rien d'autre que les cases de la théorie du changement qui sont concernées par la question d'évaluation.

Les indicateurs sont des facteurs ou preuves mesurables qui montrent l'étendue des progrès, de la réussite ou des accomplissements réalisés dans l'intervention (Zarinpoush, 2006). Ils sont l'opérationnalisation des variables en mesure, leur formulation est importante pour synthétiser quelles données exactement seront utiles, et pour guider l'évaluateur dans la formulation précise de sa stratégie d'évaluation. Un indicateur est donc un instrument de mesure, de nature quantitative ou qualitative, qui constitue un moyen simple et fiable de mesurer et d'informer des changements liés à l'intervention ou d'aider à apprécier la performance d'un acteur du développement (OCDE, 2002).

Les indicateurs peuvent être de nature quantitative, comme le nombre de participants, le nombre de visites d'un site Web, le taux et le classement d'opinions. Ils peuvent aussi être de nature qualitative, comme des réactions positives ou négatives, des problèmes, des perceptions.

Généralement, les indicateurs peuvent être classés sous plusieurs catégories : indicateurs quantitatifs, indicateurs qualitatifs, indicateurs binaires, indicateurs de substitution.

► Indicateurs quantitatifs

Les indicateurs quantitatifs servent à mesurer des éléments foncièrement objectifs ou fondés sur des faits et non subjectifs. Généralement, ces indicateurs peuvent être exprimés sous forme numérique et répondre à des questions comme : *combien ? A quelle fréquence ? Quelle quantité ? Pendant combien de temps ?*

► Indicateurs qualitatifs

Les indicateurs qualitatifs servent à démontrer, décrire ou mesurer des éléments de nature subjective, comme des expériences, des opinions ou des perceptions. Ils peuvent signaler des changements dans la sensibilité, la satisfaction, l'influence, la prise de conscience, la compréhension, les attitudes, la perception, le dialogue ou le sentiment de bien-être.

► Indicateurs binaires

Les indicateurs binaires peuvent être qualitatifs ou quantitatifs et servent à vérifier l'existence d'une variable tangible et précise qui n'existait pas avant la mise en œuvre du projet. Une existence qui peut être confirmée par un « oui » ou par un « non ».

► Indicateurs de substitution

Les indicateurs de substitution désignent une mesure qui est indirecte ou, plus précisément, qui se substitue à la mesure directe. Par exemple, le nombre de femmes occupant des postes à responsabilité dans des gouvernements nationaux peut servir d'indicateur indirect de la perception de l'égalité des genres dans ces pays.

À travers les « Lignes directrices Gestion du cycle des projets » de l'UE (2004), d'autres classifications d'indicateurs peuvent être identifiées et peuvent aider les évaluateurs dans leur formulation : les indicateurs d'inputs, les indicateurs de résultat (output), les indicateurs d'objectifs (outcomes), les indicateurs d'impact.

Les indicateurs d'inputs permettent de mesurer les ressources financières, administratives et réglementaires fournies par l'État et les bailleurs de fonds. Il est nécessaire d'établir un lien entre les ressources utilisées et les résultats obtenus afin d'évaluer l'efficacité des activités menées : par exemple la part du budget consacré à l'éducation, la suppression des uniformes scolaires obligatoires.

Les indicateurs de résultat (output) permettent de mesurer les conséquences immédiates et concrètes des décisions prises et des ressources utilisées : par exemple le nombre d'écoles construites, le nombre de professeurs formés.

Les indicateurs d'objectifs (outcomes) permettent de mesurer les résultats en termes d'avantages pour les groupes cibles : par exemple la scolarisation, le pourcentage de filles parmi les enfants entrant en première année d'école primaire.

Les indicateurs d'impact permettent de mesurer les conséquences à long terme des résultats. Ils mesurent les objectifs généraux en termes de développement national et de réduction de la pauvreté : par exemple le taux d'alphabétisation.

Certaines matrices incluent à côté des indicateurs, une colonne avec des données de référence aussi appelées baseline. Les bases de référence fournissent un fondement pour mesurer le changement au fil du temps.

La base de référence établit la valeur de l'indicateur au début de la période de planification du projet. Les données de référence sont généralement ventilées par caractéristiques clés (par ex. âge, sexe), selon ce qu'il convient.

2.3. Sources et méthodes de collecte des données

La source et la méthode de collecte des données indiquent la provenance des informations et comment elles seront assemblées en vue de mesurer l'indicateur. L'identification de l'information dont on a besoin est en effet un préalable au choix et à l'élaboration des outils de collecte d'information qui sera effectuée après l'élaboration de la matrice. L'analyse de l'information disponible (information secondaire) à travers les documents du projet, la revue de la littérature existante dans le domaine étudié ou à travers le système de suivi est d'une grande aide pour savoir quelles sont les sources importantes pour l'évaluation. On essaye en effet, de constituer en équilibre entre une diversité de source reflétant les différents points de vue sur la performance de l'intervention et l'optimisation des travaux de collecte et d'analyse dans le temps et le budget imparti à la mission d'évaluation. La collecte et l'analyse de données sont traitées plus en profondeur dans les chapitres 8 9 et 10 du présent ouvrage.

2.4. Les méthodes d'évaluation

Les méthodes d'évaluation renvoient aux designs d'évaluation ou aux façons dont on va construire et répéter des mesures sur le terrain, généralement pour répondre à des questions relatives aux effets. Encore une fois, le chapitre 7 du présent ouvrage y est consacré.

Messages clés :

- ✓ Les questions d'évaluation constituent la colonne vertébrale de l'évaluation et de leur choix découle le déploiement d'une stratégie de collecte d'information que l'on résume dans un tableau appelé matrice d'évaluation ;
- ✓ Le choix des questions d'évaluation est basé sur un certain nombre de critères de performance que sont la pertinence, l'efficacité ou l'efficience pour ne citer qu'eux. Ce choix n'est pas anodin, et renvoie à l'approche d'évaluation et aux paradigmes adoptés par l'équipe d'évaluation. On ne pose pas les mêmes questions dans une évaluation d'impact que dans une évaluation évolutive ;
- ✓ Les matrices d'évaluation se retrouvent sous plusieurs formes selon les habitudes des organisations et les préférences des évaluateurs. Au minimum, elles doivent rassembler les questions d'évaluation, les variables et indicateurs de mesure, ainsi que la méthode de collecte et la source des données.

Bibliographie :

- Centre de Développement de la Région de Tensift (CDRT) (2015), « Guide de l'évaluation », International Program for Development Evaluation Training.
- Fataneh Z. (2006), Guide d'évaluation de projet à l'intention des organismes sans but lucratif : Méthodes et étapes fondamentales pour procéder à l'évaluation de projets, Imagine Canada.
- FIDA (2000) guide pratique de suivi évaluation,
- Lignes directrices Gestion du cycle des projets - 2004
- Michèle R., Jean-Louis D., André-Pierre C., Michel R., Henriette B., Geneviève S. M., Valérie L. (2011), « Évaluation de l'implantation et de l'impact du programme PRÉVICAP », Institut de recherche Robert-Sauvé en santé et en sécurité du travail (IRSST), Études et recherches, RAPPORT R-716.
- OCDE (2002) - Glossaire des principaux termes relatifs à l'évaluation et la gestion axée sur les résultats.
- Paul J. G., Sebastian M., Patrick P., Laura B. R., Christel M. J. V. (2011). « L'évaluation d'impact en pratique », Banque mondiale, ISBN 978-0-8213-8541-8 -- ISBN 978-0-8213-8593-7 (électronique).
- UE - Lignes directrices Gestion du Cycle de Projet, 2004.

MODULE 7

CHOIX DU DESIGN DE L'ÉVALUATION SELON LE TYPE DE QUESTIONS ET LE CONTEXTE

Objectifs d'apprentissage :

A la fin de la session, les participants seront capables de :

- Identifier les différentes méthodes d'évaluation
- Relier les différents types de design aux questions de l'évaluation
- Identifier les éléments du contexte de l'évaluation susceptibles d'influencer le choix de la méthode d'évaluation
- Choisir la méthode d'évaluation en fonction des questions et du contexte

Thèmes couverts

- I. Différentes méthodes d'évaluation
- II. Relations entre questions et méthode d'évaluation

Après avoir élaboré les questions d'évaluation, l'évaluateur doit choisir la manière dont il va construire son argumentation à l'aide de données qu'il va collecter (données primaires) ou récupérer pour en faire sa propre analyse (données secondaires). Selon le type de questions, cette tâche s'avère plus ou moins complexe : pour des questions descriptives, il s'agit surtout de trouver les outils de collecte qui permettront de mieux comprendre la réalité de l'intervention ; pour des questions de cause à effet une méthodologie particulière est généralement nécessaire pour isoler les effets propres à l'intervention évaluée. On distingue notamment **l'effet brut** (la différence entre la situation avant l'intervention et la situation après), de **l'effet net** (le changement qui est attribuable à l'intervention).

Ainsi, selon les questions d'évaluation, selon l'approche sélectionnée, l'évaluateur devra choisir une méthode d'évaluation appropriée, parfois également appelée *design* (ou traduit par le terme devis en français). De nombreuses méthodes sont utilisées en évaluation, certaines plus quantitatives, issues des statistiques, d'autres plus qualitatives, permettent de comprendre des dynamiques de changement plus complexes.

Mots-clés : méthode d'évaluation ; design expérimental, quasi-expérimental ou pré-expérimental ; méthodes qualitatives d'évaluation des effets

I. Choix d'une méthode d'évaluation

Baker (2000) rappelle que « *malgré les milliards de dollars injectés dans l'aide au développement chaque année, nous savons encore peu de chose de l'impact réel des projets sur les pauvres. [...] Pour un programme ou projet spécifique d'un pays donné, les questions suivantes se posent toujours : l'intervention produit-elle les bénéfices escomptés et quel a été l'impact global sur la population ? Le programme ou projet pourrait-il être mieux conçu pour atteindre les résultats escomptés ?* ».

L'évaluation d'impacts est en effet un exercice difficile techniquement et institutionnellement. Techniquement, parce que de nombreux pièges attendent l'évaluateur. Tout d'abord, il peut exister une corrélation entre une politique et ses résultats. Cette corrélation ne traduit pas un effet causal ; c'est-à-dire que les effets observés ne sont pas nécessairement dus à l'intervention. L'évaluateur doit, en l'occurrence, tenir compte des causalités inverses et des interactions de la politique considérée avec de multiples autres facteurs. En outre, il doit aussi être conscient du fait que le bénéficiaire observé d'une intervention n'est pas forcément celui qui est visé, ou que celle-ci peut avoir de multiples effets, attendus ou inattendus, désirables ou indésirables. Elle est aussi difficile institutionnellement car elle requiert des compétences pointues (pour les méthodes quantitatives et qualitatives), et une planification des activités de suivi-évaluation dès le démarrage de l'intervention si l'on veut effectuer un travail efficient. Mais l'évaluation des effets est pourtant au cœur du métier d'évaluateur, puisque l'évaluation est une des seules disciplines à s'y intéresser (contrairement au suivi ou à l'audit). C'est donc pour s'attaquer à ces difficultés que les évaluateurs ont développé une grande diversité de méthodes d'évaluation quantitatives et qualitatives, inspirées de plusieurs disciplines, notamment les statistiques.

Si l'on revient aux différents types de questions d'évaluation présentés par Imas et Rist (2009), au chapitre précédent, (questions descriptives, normatives et de cause à effet) on comprend dès lors que les méthodes ou *design* d'évaluation seront particulièrement utiles pour répondre aux questions de cause à effet, afin pour d'isoler l'effet net de l'intervention évaluée. Les questions descriptives ou normatives vont en revanche, requérir des méthodes plus simples se concentrant plutôt sur une analyse

approfondie d'éléments observables, souvent appréhendées grâce à des études de cas et des collectes de nature qualitatives.

D'autres facteurs peuvent également guider le choix d'une méthode plutôt qu'une autre, comme la disponibilité des données dans le temps imparti pour l'évaluation, l'approche d'évaluation choisie, les ressources financières, humaines, matérielles et l'échéancier disponibles etc. Il n'y a pas de méthode parfaite en toute occasion, mais chacune a ses forces et faiblesses et s'applique dans des situations précises. Il s'agit donc de faire un compromis entre la faisabilité des méthodes et leur capacité à offrir une réponse aux questions d'évaluation.

II. Types de méthodes d'évaluation

La boîte à outils des méthodes d'évaluation s'est remplie au fil du développement de la discipline et continue de s'étoffer au fur et à mesure que les évaluateurs partagent de nouvelles techniques élaborées sur le terrain.

Plusieurs méthodes permettent tout d'abord de contourner le problème de l'isolation de l'effet net (et non l'effet brut) en reconstruisant ce qui se serait passé en l'absence de l'intervention évaluée, c'est ce qu'on appelle utiliser un **contrefactuel**. C'est le cas de la méthode expérimentale.

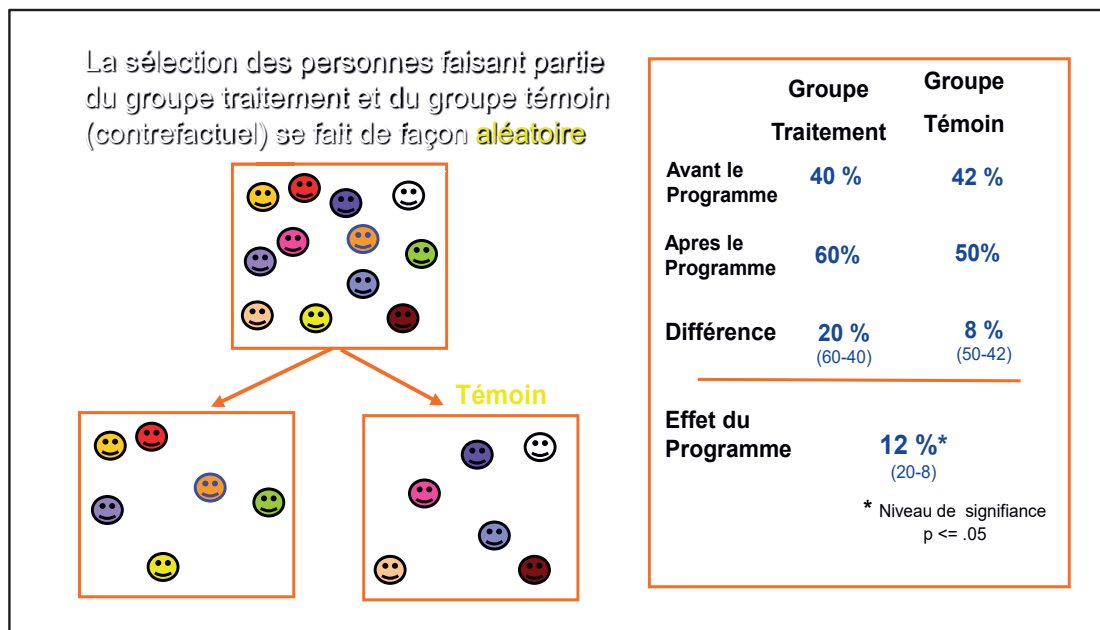
2.1. Méthode expérimentale ou aléatoire

Les méthodes expérimentales consistent à répartir les individus de façon aléatoire dans deux groupes : un groupe qui bénéficiera du programme (groupe traité ou groupe expérimental) et un groupe témoin ou de contrôle (cf. Figure 7.1). Il est impératif de s'assurer que les deux groupes ne diffèrent en moyenne que par le fait que l'un reçoit le traitement (l'intervention évaluée) et que l'autre ne le reçoit pas. Cette qualité assure que les deux groupes sont comparables et que leurs caractéristiques ne sont pas statistiquement significatives en absence de l'intervention. L'hypothèse sous-jacente est que s'il y a une différence entre les deux groupes à la fin de l'intervention, cette différence est attribuable à l'intervention et non à d'autres facteurs.

Cette assurance est difficile à obtenir en pratique et il est nécessaire de contrôler ces erreurs et de les réduire au strict minimum en agissant par exemple sur la taille de l'échantillon.

Il faut noter que l'approche expérimentale exige que l'évaluateur et l'organisme chargé du projet collabore avant l'exécution du projet. En effet, le choix des groupes témoin et traitement se fait avant l'intervention et en accord avec toutes les parties prenantes. Cette approche ne peut donc pas se faire si le programme a déjà débuté.

Figure 7.1 : Répartition de la population en traité et contrôle



La mise en œuvre de la méthode expérimentale peut se résumer en six étapes décrites ci-dessus :

- 1 **Formuler une hypothèse** : On formule une *hypothèse par défaut* (ou *hypothèse nulle*) que *l'intervention n'a pas d'impact*. La variable d'intérêt de l'intervention doit être clairement identifiée et mesurable.
- 2 Obtenir **une base de référence** sur la variable à mesurer : Typiquement, on conduit une *enquête de référence* (si les données adéquates n'existent pas) avant l'intervention sur la population susceptible de bénéficier de l'intervention.
- 3 **Assigner** les individus au hasard à la fois au groupe qui reçoit l'intervention et au groupe qui n'en reçoit pas (groupe de contrôle) : Cette assignation se fait en utilisant les données de référence et l'on s'assure que les deux groupes sont *équilibrés* (ou *comparables*) en moyenne.
- 4 **Mettre en œuvre l'intervention** uniquement au niveau des bénéficiaires désignés ou dans le groupe traité. Le groupe de contrôle ne doit pas être exposé

à l'intervention. Il faut éviter toute contamination dans la gestion des deux groupes. Il faut limiter la communication entre les deux groupes de sorte que des personnes du groupe témoin ne bénéficient pas de l'intervention. Si un tel cas se produit, il faut en tenir compte dans l'évaluation d'impact.

5 Mesurer la variable d'intérêt dans les deux groupes : c'est une seconde collecte de données portant sur les deux groupes qui permet de d'obtenir les valeurs des variables d'intérêt à la fin de l'intervention. Les informations collectées sur le groupe témoin permettent de construire le *contrefactuel* ; c'est-à-dire une estimation de ce qu'aurait été le *résultat* pour un participant au programme en l'absence du programme. Par exemple, quel aurait été l'incidence du paludisme dans un village donné si l'on n'avait pas distribué gratuitement les moustiquaires imprégnées. Cette quantité est par définition inobservable puisque le village a bénéficié du projet de distribution gratuite de moustiquaire imprégné. Il faut donc l'estimer. C'est à cela que sert les observations faites sur le groupe de contrôle.

6 Calculer les différences entre les groupes et évaluer la significativité de la différence éventuelle observée. Si l'on soupçonne que d'autres facteurs ont pu affecter les résultats, il faut les éliminer avant de tirer les conclusions appropriées.

L'approche expérimentale est la méthode d'évaluation d'impact permettant d'isoler avec le plus de certitude l'effet net d'une intervention, puisque lorsqu'elle est rigoureusement mise en œuvre, elle permet d'annuler de nombreuses menaces à la validité. Elle exerce donc un grand attrait dans la mesure où son application est finalement assez simple lorsque les groupes sont sélectionnés à l'avance. Cette approche exige par conséquent, une étroite collaboration entre l'évaluateur et le gestionnaire du projet depuis la conception de l'intervention et durant la mise en œuvre de l'évaluation. C'est une approche qui permet d'intégrer l'évaluation à l'implémentation de l'intervention. Néanmoins, cette approche comporte également de lourdes contraintes. Elle ne peut pas s'appliquer à toutes les interventions, par exemple un programme d'ampleur nationale (dans ce cas de figure, on ne peut pas trouver de groupe témoin puisque tous les habitants sont bénéficiaires). Cette approche peut aussi poser des questions d'ordre pratique et éthique suivants : est-ce acceptable de faire une sélection de bénéficiaires (et donc de laisser des potentiels

bénéficiaires sans intervention uniquement pour les besoins de l'évaluation) Est-il réaliste de sélectionner des bénéficiaires au hasard et avant l'intervention ? Peut-on retrouver les mêmes non-bénéficiaires après plusieurs années d'intervention ?

Pour ces raisons, la méthode expérimentale est peu utilisée en dehors des travaux de recherche, et des projets-pilote, qui sont créés avec un protocole d'évaluation rigoureux dès le départ. Elle inspire toutefois plusieurs autres méthodes, dites « quasi-expérimentales » qui vont tenter de recréer les conditions de la méthode expérimentale, en présence de différentes contraintes.

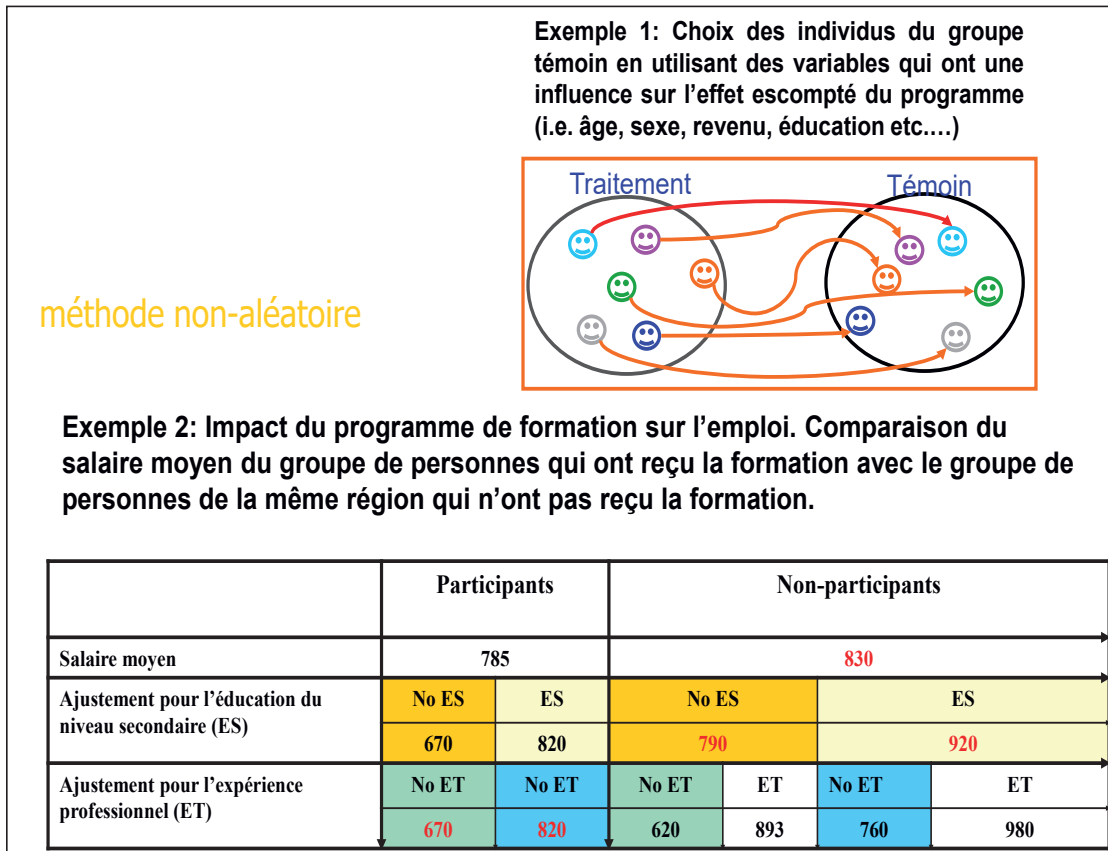
2.2. Méthodes quasi-expérimentales

Les méthodes quasi-expérimentales (MQE) sont des alternatives aux méthodes d'assignation aléatoires qui permettent, en se basant sur le modèle contrefactuel, d'évaluer les effets d'interventions (Benmarhnia et Fuller, 2019). Selon Cook et al (1979), le terme « quasi-expérimental » fait référence à une expérimentation avec une intervention donnée, au moins deux unités expérimentales, à savoir les bénéficiaires et les non-bénéficiaires, mais sans recours à la randomisation pour créer les conditions d'inférence causale. En d'autres termes, on n'a pas recours à une assignation aléatoire pour constituer les groupes (témoin et traité). Une illustration est donnée par la Figure 7.2. On peut constater dans cet exemple que les deux groupes n'ont pas la même taille et les salaires moyens sont différents entre les deux groupes.

Tout comme la méthode expérimentale, l'approche quasi-expérimentale utilise également la notion de contrefactuel pour estimer des effets de nature causale.

Néanmoins, chaque méthode quasi-expérimentale nécessite de respecter des hypothèses de travail supplémentaires. Parfois ces hypothèses sont impossibles à vérifier empiriquement. L'usage de ces méthodes requiert ainsi généralement un échantillon plus large pour minimiser les menaces à la validité de la méthode.

Figure 7.2 : Illustration des méthodes quasi-expérimentales



Le Tableau 7.1 fait une synthèse des méthodes quasi-expérimentales les plus utilisées. **L'appariement** consiste à utiliser des techniques statistiques pour construire le groupe de comparaison alors que pour appliquer la discontinuité de la régression, on suppose qu'il existe un critère d'éligibilité qui permet de séparer les deux groupes. Quant à **la double différence**, elle présente l'avantage de travailler sur deux groupes (traitement et témoin) qui sont a priori hétérogènes. La contrainte fondamentale est qu'on suppose que l'évolution des deux groupes est similaire dans le temps en absence de l'intervention (*constance des tendances*).

Tableau 7.1 : Quelques méthodes quasi-expérimentales

	Appariement	Discontinuité de la Régression	Double-différence
Principe	L'appariement des scores de propension (PSM) construit un groupe de comparaison statistique qui est basé sur un modèle de la probabilité de participer au traitement, en utilisant les caractéristiques observées. Les participants sont ensuite appariés sur la base de cette probabilité, ou score de propension, aux non-participants. L'effet de traitement moyen du programme est ensuite calculé comme la différence moyenne des résultats entre ces deux groupes.	Dans un cadre non expérimental, les règles d'éligibilité du programme peuvent parfois être utilisées comme des instruments pour identifier de manière exogène les participants et les non-participants. Pour établir la comparabilité, on peut utiliser des participants et des non-participants dans un certain voisinage du seuil d'admissibilité comme échantillon pertinent pour estimer l'impact du traitement.	La méthode de la double différence (DD) suppose qu'il existe une hétérogénéité non observée de la participation, mais que ces facteurs sont invariants dans le temps. Par conséquent, avec les données sur les observations de projet et de contrôle avant et après l'intervention du programme, cette composante fixe peut être différenciée.
Conditions d'utilisation	On peut utiliser le PSM même lorsqu'on dispose d'une seule donnée (enquête d'évaluation)	L'affectation au groupe de traitement ou de contrôle repose sur une règle bien définie (indicateur et seuil).	La méthode exige de disposer des données avant l'intervention et après l'intervention sur le groupe traité et le groupe de contrôle.
Contraintes	La validité du PSM dépend de deux conditions : (a) l'indépendance conditionnelle (à savoir, que les facteurs non observés n'affectent pas la participation) et (b) un support commun entre les échantillons des participants et des non-participants. Elle repose uniquement sur les caractéristiques observables uniquement, les inobservables ne sont pas considérées.	Elle nécessite une grande quantité de données pour autour du seuil d'éligibilité	Elle est valide à condition que les tendances soient constantes (hypothèse pas toujours aisée à vérifier).

Ces différentes méthodes recréent ainsi les conditions expérimentales en l'absence de sélection aléatoire qui est souvent manquante lorsque l'on effectue une évaluation ex-post. Toutefois, puisque l'on n'a souvent pas la preuve d'une équivalence des groupes, elles restent plus sujettes à différentes menaces à leur validité.

2.3. Les menaces à la validité des méthodes d'évaluation des effets

Dans une évaluation d'impact, on peut faire face à plusieurs menaces pouvant invalider les résultats d'une évaluation : en statistique, on parle également de *biais* ou de *résultats biaisés*. Illustrons le biais par un exemple relativement simple.

Nous savons qu'un *groupe de comparaison valide* doit vérifier des propriétés spécifiques, notamment être *similaire au groupe de traitement*. Intuitivement, un groupe de comparaison non valide est un groupe qui diffère du groupe de traitement autrement que par la seule absence du traitement à l'étude. Ces autres différences peuvent rendre l'estimation d'impact invalide ou, en termes statistiques, *biaisée*. En effet, en présence d'autres différences entre les groupes de traitement et de comparaison, l'estimation ne permettra pas de déterminer l'impact réel du programme, car elle confondra l'effet du programme avec les effets des autres différences. Trois types de biais peuvent menacer la validité des résultats, et sont regroupés dans le Tableau 7.2.

Tableau 7.2 : Définition et sources de biais

Biais	Définition
Biais de sélection	<p>Un biais de sélection apparaît lorsque les raisons pour lesquelles une personne participe à un programme sont corrélées aux résultats. Par exemple, des jeunes peuvent décider de participer à un programme de formation professionnelle parce qu'ils anticipent que le programme leur permettra d'améliorer leurs conditions de vie et espèrent beaucoup bénéficier de la formation. Ce biais se produit généralement lorsque le groupe de comparaison n'est pas éligible au programme ou décide de ne pas y participer. Le groupe ayant décidé de ne pas participer au programme ne permet pas d'obtenir un contrefactuel convaincant. Ce problème se traduit par une différence systématique entre les caractéristiques des deux groupes et peut provenir d'une auto-sélection des participants ou que les gestionnaires du programme ont sélectionné les participants les plus susceptibles de réussir.</p>
Attrition	<p>Une attrition se produit lorsqu'il y a une déperdition de certaines unités de l'échantillon d'une ronde à l'autre de la collecte des données. Certains individus quittent le programme. On peut tenir compte de ce facteur en calculant la taille de l'échantillon.</p>
Spillover	<p>Ce sont des effets appelés encore <i>effets de diffusion</i> ou <i>effets de débordements</i>. On parle aussi de <i>contamination</i> si les effets affectent le groupe de comparaison. L'effet de diffusion se produit lorsque le groupe de comparaison est affecté par le traitement administré au groupe de traitement, même si le traitement n'est pas directement administré au groupe de comparaison. Si l'effet de diffusion sur le groupe de comparaison est négatif, c'est-à-dire si le programme nuit à ce groupe, la différence directe entre les résultats du groupe de traitement et ceux du groupe de comparaison produit une surestimation de l'impact du programme. Par contre, si l'effet de diffusion sur le groupe de comparaison est positif, c'est-à-dire si le programme est bénéfique pour ce groupe, le résultat est alors une sous-estimation de l'impact du programme.</p>

2.4. Autres méthodes d'évaluation quantitatives, non-expérimentales

D'autres méthodes d'évaluation inspirés des statistiques sont utilisées par les évaluateurs, notamment lorsqu'il n'est pas possible d'obtenir des données de non-bénéficiaires (donc de groupes témoins).

Série chronologique ininterrompue ou design longitudinal

Dans cette approche, l'évaluateur ne dispose pas d'un groupe de contrôle. Néanmoins, il dispose d'une suite d'observations des variables évaluées dans le temps, il est donc possible de comprendre l'évolution de ces dernières sur le long terme et de faire différentes projections statistiques. On s'intéresse notamment aux changements dans le temps, soit avant l'intervention ou après, voire durant la mise en œuvre de l'intervention. L'usage de cette méthode exige de disposer d'une série temporelle ou chronologique d'au moins trois observations, avant et après l'intervention. Ce type de design s'applique bien lorsque nous disposons même d'une série chronologique longue, comme le taux de croissance du Sénégal entre 1960 et 2019. On peut alors analyser les effets de différentes mesures, comme la dévaluation du franc CFA. On parle alors de *design longitudinal*. Le principe de cette méthode est donc de comprendre, à partir des observations passées, comment aurait évolué la variable étudiée, sans intervention. Ici, le contrefactuel n'est donc pas mesuré auprès de non-bénéficiaire, mais calculé grâce à la connaissance acquise sur la variable. De cette manière, on peut donc isoler l'effet net d'une intervention, si elle se distingue bien dans le temps (il n'y a pas d'autres explications concurrentes pour les changements observés pendant cette période).

Méthodes à contrôle statistique ou méthodes avec variable indépendante

Designs pré-expérimentaux ou implicites

Les méthodes pré-expérimentales, sont des designs simplifiés, c'est-à-dire que l'on fait des mesures sur les indicateurs clés de l'intervention avant et/ou après l'intervention, sans ajouter de groupe témoin séparé. Bien que très sujets aux menaces à la validité qui ont été présentées plus haut, ils sont particulièrement utiles dans le cas de projets simples, limités dans le temps, où il n'y a que peu d'explications concurrentes aux changements sur la variable cible (l'effet net est égal à l'effet brut). Considérons par exemple, une intervention qui consiste à dispenser une formation

professionnelle aux jeunes pendant un an. On peut se poser la question de savoir si les participants ont accru leurs connaissances des techniques présentées dans la formation. Dans la conception de l'intervention, on ne suit que les jeunes formés et on ne dispose donc pas de groupe témoin. Toutefois, si l'on soumet les jeunes à un test de connaissances avant et après leur formation, il est tout à fait vraisemblable de considérer que les gains de connaissances sont attribuables à la formation. Dans ce cas, former un groupe témoin de jeunes, non formés serait une perte de temps. Il revient donc aux évaluateurs de bien comprendre la situation qui leur est présentée pour ne pas compliquer la mesure des effets sans que cela soit nécessaire.

2.5. Méthodes qualitatives

Les méthodes quantitatives d'évaluation d'impact supposent que l'on puisse mesurer simplement des variables d'effet, et « démêler » différents facteurs explicatifs pour isoler les effets attribuables à une seule intervention. Dans la pratique, notamment pour des interventions dans des contextes complexes comme celui du développement, il est souvent difficile d'appliquer ces méthodes. C'est pourquoi, de nombreux auteurs (IE4D Group, 2011) ont développé des approches complètement différentes de celles présentées plus haut, préférant trouver les moyens d'évaluer la contribution d'une intervention à un changement observable sur le terrain, ce qui serait finalement plus utile pour la prise de décision, que de se perdre à vouloir chiffrer un impact précis. Trois méthodes reconnues sont présentées ici, mais il en existe une grande diversité, s'adaptant à la multitude d'interventions qu'un évaluateur peut rencontrer.

La méthode du changement le plus significatif (MSC)

La méthode du « changement le plus significatif » (MSC) est une forme participative de suivi-évaluation basée sur le recueil de récits personnels de changements significatifs, développée par Davies et Dart (2005). Technique que les auteurs ont détaillée en 10 étapes distinctes tout au long du cycle de vie du programme, la méthode du MSC fait du suivi-évaluation « sans indicateur ». Elle consiste plutôt à recueillir des histoires de changements liés à l'intervention évaluée auprès des bénéficiaires et de gérer l'intervention, en sélectionnant de manière participative les histoires de changement qui sont souhaitées (par exemple dans la théorie du changement) et en orientant ainsi l'intervention. Des mesures sur l'ampleur de

l'existence de chaque type de changement (en classant les différentes histoires) auprès de l'ensemble des bénéficiaires permettent également de comprendre les impacts d'une intervention, sans passer par les designs quantitatifs, dans le cas d'interventions complexes aux théories de changement incertaines.

La cartographie des incidences (*Outcome mapping*)

Cette deuxième méthode qualitative peut également être utilisée comme un système de suivi-évaluation tout au long du cycle de vie d'une intervention. La cartographie des incidences, méthode développée au Centre de recherches pour le développement international (Earl, Carden et Smutylo, 2002), se concentre sur les « incidences », plutôt que sur les effets, soit des preuves de changements qui peuvent être des modifications de comportement ou d'activité de groupes ou d'organisations limitrophes à l'intervention. Cette méthode, développée en 6 étapes distinctes par les auteurs, vise à documenter, par un système de liens plausibles de cause à effet, dans une vision complexe et non-linéaire du changement, les contributions de l'intervention évaluée aux incidences observées. Une réflexion plus large est menée sur les réelles possibilités de changement apportées par interventions en contexte de développement.

La récolte des incidences (*Outcome Harvesting*)

La troisième méthode choisie est la récolte des incidences ou *Outcome Harvesting* en anglais. Déclinaison de la cartographie des incidences, puisqu'elle se fonde sur la même observation des incidences de changement, cette méthode a été développée par Ricardo Wilson Grau et ses collègues (Wilson-Grau R., 2019). Cette méthode fonctionne comme une enquête policière, visant à établir des liens entre des incidences documentées et une intervention qui vise ce type de changements. On ne cherche donc pas à mesurer de possibles effets en fonction d'une théorie établie, mais on part plutôt de réelles preuves de changements sur le terrain et on étudie si et comment une intervention donnée y a contribué. Cette approche est, elle aussi participative, puisqu'elle nécessite de réaliser de nombreux entretiens avec les groupes ou organisations parties prenantes du contexte pour recenser les incidences et étudier les liens de contribution. Elle repose aussi sur beaucoup d'analyse documentaire.

Messages clés :

- ✓ L'évaluation est une des seules disciplines qui s'intéressent aux effets et aux impacts, disposant d'outils pour appréhender l'effet net d'une intervention, au-delà d'une simple observation avant/après.
- ✓ Il n'existe pas de méthode magique qui s'applique dans tous les cas, et c'est vraiment à l'évaluateur de diagnostiquer le contexte dans lequel l'intervention évaluée agit, pour choisir celle qui s'applique le mieux. Une combinaison de plusieurs méthodes est possible selon les effets que l'on veut évaluer, les données disponibles ou la complexité des facteurs externes jouant également sur les variables cibles.
- ✓ La meilleure méthode d'évaluation est celle qui permet d'émettre un jugement évaluatif crédible, de manière efficiente, et non celle qui se classe comme « la plus rigoureuse » dans votre livre de statistique.

Bibliographie:

- Baker, J. L. (2000). *Evaluating the impact of development projects on poverty: A handbook for practitioners*. The World Bank.
- Benmarhnia, T., & Fuller, D. (2019). 9. Les méthodes quasi-expérimentales. L'effet de l'âge légal minimum sur la consommation d'alcool chez les jeunes aux États-Unis. In *Évaluation des interventions de santé mondiale. Méthodes avancées*. Sous la direction de Valéry Ridde et Christian Dagenais, pp. 241-264. Québec : Éditions science et bien commun et Marseille : IRD Éditions.
- Cook T.D., Campbell D.T. et Day A. (1979). *Quasi-experimentation: Design et analysis issues for field settings*. Boston: Houghton Mifflin.
- Earl, S., F. Carden, et T. Smutylo. *La cartographie des incidences, intégrer l'apprentissage et la réflexion dans les programmes de développement*. Ottawa: Centre de recherches pour le développement international, 2002, 143p.
- IE4D Group. «Impact Evaluation for Development - Principles for Action. » 2011, 8p.
- Imas L. G. M., Rist R. C. (2009). *The road to results: Designing and Conducting Effective Development Evaluations*. The World Bank.
- Wilson-Grau R. *Outcome Harvesting: Principles, Steps, and Evaluation Applications*. Charlotte, NC: Information Age Publishing, 2019.

Pour aller plus loin sur le sujet :

- Gertler P. J. et al. (2010) *L'évaluation d'impact en pratique*, Banque Mondiale, 235p. <https://documents1.worldbank.org/curated/fr/187971468155114516/pdf/599980PUB0FRENCH010evaluation0d0impact.pdf>
- Khandker S. R., G.B. Koolwal et H.A. Samad (2010) *Handbook on impact evaluation : quantitative methods and practices*, The World Bank, 239p. <https://openknowledge.worldbank.org/bitstream/handle/10986/2693/520990PUB0EPI1101Official0Use0Only1.pdf?sequence=1>

MODULE 8

SÉLECTION DE LA STRATÉGIE D'ÉCHANTILLONNAGE

Objectifs d'apprentissage :

A la fin de la session, les participants seront capables de :

- Décrire les différentes techniques d'échantillonnage non probabilistes
- Décrire les différentes techniques d'échantillonnage probabilistes
- Choisir la technique de sélection adéquate.
- Définir la notion de précision d'un échantillonnage
- Utiliser les méthodes de calcul de la taille d'un échantillon

Thèmes couverts

- I. Principe, objectif et objet de l'échantillonnage
- II. Techniques probabilistes
- III. Techniques non probabilistes
- IV. Notion de précision et ses déterminants
- V. Méthodes de calcul de la taille d'échantillon

Les différentes méthodes d'évaluation nécessitent généralement des collectes d'informations auprès de groupes d'individus ou d'organisations (appelée population). Il peut s'agir d'une enquête auprès des bénéficiaires et non-bénéficiaires d'une intervention, d'entretiens avec des personnes-ressources ou de villages dans lesquels l'évaluateur se déplace pour des études de cas. Dans certaines situations, on peut étudier l'ensemble de la population ciblée soit parce qu'elle est peu nombreuse, soit parce que l'on dispose déjà de données sur l'ensemble. Dans d'autres cas, notamment face à de grands groupes, on peut plutôt étudier un petit nombre d'unités dans cette population, on procède dès lors à un échantillonnage. La question fondamentale pour effectuer cette procédure est de savoir comment constituer un échantillon auprès duquel les données seront collectées avec rigueur et la taille de cet échantillon ; c'est-à-dire le nombre d'unités qu'il doit contenir.

Ce module expose les différentes méthodes de sélection de cas et d'échantillonnage, les conditions du choix de l'une ou l'autre des méthodes et le calcul de la taille de l'échantillon. Le choix de la méthode d'échantillonnage et la taille de l'échantillon peuvent varier en fonction de la méthode d'évaluation et de son contexte. Par ailleurs, le choix de la méthode d'échantillonnage a des implications sur le traitement et l'exploitation des données collectées. Les différentes implications seront également présentées dans le module.

Mots-clés : sélection, étude de cas, échantillonnage

I. Principe, objectif et objet de l'échantillonnage

Dans le cadre d'un mandat d'évaluation, la recherche de l'information peut se faire suivant trois directions. On peut procéder à une recherche documentaire – fichiers administratifs – pour obtenir l'information désirée. Cette approche dépend de l'information disponible, et peut constituer un véritable gain de temps. Elle peut cependant fournir des informations partielles, datées et incomplètes au regard des questions d'évaluation.

L'évaluateur peut alors décider de collecter lui-même des données. Deux choix se posent à lui : il peut recourir à une enquête exhaustive auprès de l'ensemble des individus faisant partie du champ de l'étude : c'est ce que l'on appelle le *recensement*. Cependant, le recensement peut devenir très coûteux en temps et en argent si la population étudiée est très grande.

Pour des raisons de rapidité et de coûts, et bien d'autres (qualité), l'évaluateur est donc souvent amené à adopter une technique de collecte qui consiste à sélectionner et interroger seulement une partie de la population, celle-ci étant appelée *échantillon*. L'objectif est de collecter des données permettant d'étudier les paramètres caractéristiques de la population en observant seulement un échantillon de celle-ci.

L'utilisation des échantillons pose cependant les questions suivantes :

- *Quelles procédures et techniques de sélection s'adaptent à chacune des méthodes d'évaluation vues dans les précédents modules et aux différents contextes d'intervention ?*
- *Combien d'individus faut-il sonder pour avoir des résultats satisfaisants? C'est la question de la taille de l'échantillon ;*
- *Quelle procédure de tirage faut-il retenir pour la sélection des individus? Cette question se réfère à l'échantillonnage ou à la technique d'échantillonnage, ou encore comment sélectionner une partie de la population d'intérêt ?*
- *Une fois l'échantillon constitué, comment agréger les données collectées ? Que faire lorsque l'on veut produire des résultats au*

niveau de toute la population ? Il s'agit ici de l'estimation et de l'extrapolation ;

- *Jusqu'à quel point peut-on faire confiance aux résultats de l'inférence statistique ? Cette question renvoie au calcul de la précision des estimateurs (ou estimations effectuées) ;*
- *Qu'est-ce qui influence cette précision en dehors même de la procédure d'agrégation utilisée ? Il s'agit ici de discuter des différentes sources d'erreurs ou de biais.*

Avant d'aller plus loin, il convient de préciser que le choix de la méthode d'échantillonnage (et les autres activités subséquentes) ne peut se faire que si les objectifs de l'étude sont clairement définis et la population cible précisée. De façon spécifique, la conception de la technique d'échantillonnage nécessite que :

- les hypothèses de l'enquête soient précisées ;
- les variables à renseigner soient définies (matrice d'évaluation);
- les critères d'inclusion et d'exclusion soient fixés ;
- les informations plus ou moins détaillées sur la population ont été collectées.

L'échantillonnage renvoie à la (technique de) sélection de l'échantillon. Le tirage ou la sélection d'un échantillon peut se faire de façon aléatoire ou empirique. Dans le premier cas (tirage aléatoire), on parle de **méthodes probabilistes** et dans le second cas (façon empirique), on parle de **méthodes non probabilistes**. Ce choix est à faire selon le type d'information dont on a besoin, l'approche et la méthode d'évaluation employée, ainsi que différents éléments contextuels tels que les données disponibles ou le temps dédié à l'évaluation.

II. Les techniques probabilistes

Les techniques probabilistes consistent à choisir un échantillon de façon aléatoire selon des techniques précises pour ensuite pouvoir généraliser les réponses à l'ensemble de la population étudiée. Elles sont fondées sur le principe que toute unité de l'univers a une probabilité non nulle et connue de figurer dans l'échantillon. On les utilise principalement dans le cadre d'approches quantitatives d'évaluation, par exemple dans les méthodes expérimentales d'évaluation des effets. L'échantillon sera dit **représentatif** s'il permet d'estimer les paramètres d'intérêt avec une précision acceptable, étant donnés les objectifs et les contraintes de l'enquête.

Ces méthodes supposent l'existence d'une base de sondage (cf. Encadré 8.1). A partir de la base de sondage, on tire de façon aléatoire les unités qui vont être observées avant même d'aller sur le terrain. Le plus souvent, on apprend aux agents enquêteurs la procédure de tirage des unités-échantillons c'est-à-dire les individus échantillonnés. Et une fois sur le terrain, ils doivent choisir les individus à enquêter conformément à la procédure qui leur a été apprise.

L'approche probabiliste permet de bénéficier des résultats de l'inférence statistique et permet donc de généraliser les résultats de l'échantillon à l'ensemble de la population, ce qui est très utile pour économiser du temps.

Encadré 8.1 : Définition et caractéristiques d'une base de sondage

On appelle **base de sondage** la liste complète de tous les individus d'une population cible. Pour faciliter l'exploitation, chaque individu doit porter un numéro et une adresse géographique. Vu son importance, une base de sondage doit avoir certaines qualités. Elle doit être complète et à jour, sans omission ni double compte et telle que l'identification de chaque unité se fasse sans ambiguïté. C'est à partir de cette liste que s'effectuera le tirage des unités-échantillons ; c'est-à-dire les individus échantillonnés. Il faut cependant noter que ce n'est pas dans toutes les enquêtes qu'on a une base de sondage fiable.

Cette catégorie des échantillonnages probabilistes comprend :

- ❖ le sondage aléatoire simple ;
- ❖ le sondage stratifié ;
- ❖ le sondage à probabilités inégales ;
- ❖ le sondage à plusieurs degrés ;
- ❖ le sondage par grappes.

2.1. Sondage aléatoire simple (SAS)

On parle de sondage aléatoire simple (SAS) lorsqu'un échantillon de taille n est tirée d'une population de taille N de telle sorte que tous les individus de la population aient la même chance d'être choisis. La mise en œuvre du SAS exige la disponibilité d'une liste de la population cible à jour (recensement, liste des écoles, des élèves, des familles et.). Cette liste appelée *base de sondage* est définie à l'Encadré 8.1.

Avantages

Le sondage aléatoire simple est le modèle de référence dans la mesure où c'est par rapport à ses propriétés que l'on juge les autres plans de sondage. Il est caractérisé par :

- ✓ sa simplicité ;
- ✓ la non-nécessité d'avoir des informations détaillées sur les individus étudiés ;
- ✓ un tirage qui repose sur l'identifiant de la base de sondage sans recourir à une information auxiliaire.

Inconvénients ✖

Cette technique, bien que simple, présente la spécificité qu'elle est particulièrement adaptée lorsque la population est relativement homogène par rapport aux caractéristiques principales d'intérêt à estimer. En d'autres termes, lorsque les individus présentent une forte hétérogénéité vis-à-vis de la variable d'intérêt et qu'il existe des sous-groupes de petite taille à l'intérieur de la population, un échantillonnage par sondage aléatoire simple peut conduire à ne retenir que

des unités statistiques d'un même sous-groupe ou à en oublier un autre et donc à biaiser les résultats de l'analyse.

Procédure de tirage des unités dans un sondage aléatoire simple

Le tirage des unités d'un échantillon aléatoire peut se faire soit à partir d'une *table de nombres aléatoires* ou de façon *systematique* dans la base de sondage. Dans les deux cas, les unités sont tirées avec des probabilités égales avec ou sans remise.

Encadré 8.1: Description du tirage des unités dans un SAS

Table de nombre aléatoire

Une table de nombres aléatoires se lit suivant un sens, en ligne (de gauche à droite ou de droite à gauche), en colonne (de haut en bas ou de bas en haut) ou en diagonale. On peut aussi faire des sauts réguliers. Elle ne doit pas être lue « au hasard » des positions de l'œil sur la table. Une table de nombres aléatoires permet de tirer des échantillons avec remise. Mais on peut s'en servir également pour échantillonner sans remise. Pour cela, il suffit lors de la lecture de la table de rejeter tous les numéros déjà tirés. Pour tirer un échantillon de taille n dans une population de taille N , on procède de la façon suivante :

- Numéroté d'abord les individus de la population de 0.1 à N (les numéros contiennent le même nombre de chiffres que N) ;
- Choisir un repère et un sens pour la lecture ;
- Retenir tous les nombres inférieurs à N à concurrence de n . Les unités dont les numéros ont été obtenus dans la table constituent l'échantillon.

Tirage systématique

Le tirage systématique est le mode de tirage aléatoire le plus célèbre à cause de sa facilité d'exécution. Les étapes sont les suivantes :

- Numéroté les individus de 1 à N selon l'ordre dans lequel ils ont été recensés ;
- Calculer le pas du tirage $PAS=N/n$; c'est-à-dire la taille de la population divisée par la taille de l'échantillon ;
- Générer un nombre aléatoire "ALEA" selon une loi uniforme entre 0 et 1.
- Calculer le départ aléatoire $DA=1+ALEA*PAS$. La partie entière de ce nombre donne le numéro du premier individu de l'échantillon (x_1) ;
- Ajouter au départ aléatoire le PAS et tronquer en prenant la partie entière pour obtenir le numéro du deuxième individu échantillon :
 $x_2= 1+[(ALEA+1)*PAS]$;
- A l'étape i , calculer $1+[(ALEA+i-1)*PAS]$ pour obtenir le numéro du i ème individu-échantillon ;
- Continuer jusqu'à obtenir les n individus-échantillons.

Application : on veut déterminer la prévalence d'une maladie infectieuse parmi les 300 habitants d'un village ($N= 300$). Compte tenu du degré de précision souhaité pour l'estimation, il sera nécessaire de rechercher la présence de la maladie chez 10 habitants ($n = 10$). En d'autres termes, on veut tirer un échantillon de 10 habitants

dans une population de 300 habitants. On suppose qu'il existe (ou on constitue) une dresse une liste de tous les habitants (base de sondage). L'encadré 8.3 présente le tirage de l'échantillon en utilisant les deux techniques précédemment présentées, à savoir la table de nombres aléatoire et le tirage systématique.

Encadré 8.3: Application du tirage du tirage des unités dans un SAS

Principe de la table de nombres aléatoire.

Nous allons tirer les 10 habitants en utilisant la table de nombres aléatoires. On utilise la table de Kendall et Babington Smith (voir Tableau 8.0).

- Numéroté les habitants de 001 à 300. Nous allons choisir dans la table des nombres de 3 chiffres entre 001 et 300 ;
- Point de départ : ligne 3, colonne 4 ;
- Sens= vers le bas, on choisit de lire les 3 derniers chiffres de la colonne 4. On ne fait pas de saut de lignes.
- Résultat : 251, 045, 075, 157, 199, 267, 278, 026, 238, 051, 118, 175, 136, 128.

Principe du tirage systématique

Nous allons tirer les 10 habitants en utilisant le tirage systématique.

- Numéroté les habitants de 1 à 300.
- Le pas du tirage est : $PAS=300/10=30$;
- On génère un nombre compris entre 0 et 1. On peut utiliser la fonction ALEA de Excel. Soit $ALEA=0,25$;
- Le premier individu tiré $x_1=1+[0,25*PAS]=8$. Le crochet [] désigne la partie entière ;
- Pour le second individu, on fait $x_2=1+[[(1+0,25)*25]*PAS]=38$ et pour le troisième $x_3=1+[(2+0,25)*PAS]=68$, etc.

L'échantillon est composé des habitants dont les numéros sont les suivants : 8, 38, 68, 98, 128, 158, 188, 218.

Tableau 8.0 : table de Kendall et Babington Smith

	5	10	15	20	25	30	35	40	45	50
5	13407	62899	78937	90525	25033	56358	78902	47008	72488	57949
	50230	63237	94083	93634	71652	02656	57532	60307	91619	48916
	84980	62458	09703	78397	66179	46982	67619	39254	90763	74056
	22116	33646	17545	31321	65772	86506	09811	82848	92211	51178
	68645	15068	56898	84021	40115	27524	42221	88293	67592	06430
10	26518	39122	96561	56004	50260	68648	85596	83979	09041	62350
	36493	41666	27871	71329	69212	57932	65281	57233	07732	58439
	77402	12994	59892	85581	70823	53338	34405	67080	16568	00854
	83679	97154	40341	84741	08967	73287	94952	59008	95774	44927
	71802	39356	02981	89107	79788	51330	37129	31898	34011	43304

NB : Dans la mise en œuvre du SAS, il est de coutume d'utiliser le tirage systématique. Cette approche donne des résultats relativement satisfaisants. Néanmoins, si les unités sont rangées de façon à présenter pour la variable d'intérêt une certaine périodicité, on risque de sélectionner des unités très particulières, surtout si le pas du tirage est égal à la période ou à un multiple. Dans de tels cas, mieux vaut recourir à la table de nombres aléatoires. Le tirage systématique est donc suggéré lorsqu'on ne soupçonne aucune périodicité dans la base de sondage.

2.2. Sondage à probabilités inégales

Il consiste à sélectionner les individus en leur donnant des chances inégales d'appartenir à l'échantillon. L'objectif est d'améliorer la précision de l'estimation de l'impact programme en exploitant les informations quantitatives disponibles sur les unités à enquêter, notamment la taille des unités.

Par exemple, une « grosse » unité peut apporter plus d'informations qu'une « petite » unité dans le cadre d'une évaluation sur des entreprises. Ainsi il peut être intéressant dans certaines situations d'accorder une plus grande importance à certaines unités, surtout lorsque la plupart des variables de l'enquête sont liées par un effet de taille. On définit alors la dimension taille de chaque unité en fonction d'une variable (appelée variable auxiliaire ou variable de pondération) et on accorde une probabilité de sortie

en fonction de cette taille. Dans notre exemple d'une enquête auprès des entreprises, on peut utiliser une variable comme le chiffre d'affaires, les bénéfices ou encore le nombre d'employés pour estimer la taille.

Avantages

Permet d'améliorer la précision de l'estimation de l'impact du programme dans la mesure où le tirage des unités dépend de leur taille.

Inconvénients ✕

- exige un travail supplémentaire parce qu'il faut trouver une variable auxiliaire (dans notre exemple, la taille) ;
- si la variable auxiliaire n'est pas liée ou est corrélée négativement à la variable étudiée, un échantillonnage à probabilités inégales peut amplifier les variances des estimateurs par rapport à un plan de sondage simple ;
- temps et coût en matière de recherche d'information quantitative.

Procédure de tirage des unités dans un sondage à probabilités inégales

Il existe trois méthodes couramment utilisées pour tirer les unités-échantillons dans le cas d'un sondage à probabilités inégales. Nous avons :

- le tirage systématique à partir des totaux cumulés ;
- le tirage systématique sur fichiers de probabilités cumulées ;
- la méthode de Bernoulli ou tirage de Poisson.

a) Tirage systématique à partir des totaux cumulés

Cette méthode procède comme suit :

- 1 Numérotter les unités de façon séquentielle de 1 à N ;
- 2 Faire le cumul des tailles unités (N_i désigne la taille de l'unité i)

$$C_1 = N_1, C_2 = N_1 + N_2, \dots, C_i = \sum_{k=1}^i N_k, \dots, C_N = \sum_{k=1}^N N_k$$

- 3 Calculer le PAS de tirage $PAS = CN/n$, avec n la taille de l'échantillon ;
- 4 Générer un nombre aléatoire compris entre 1 et le PAS. Soit x ce nombre ;
- 5 On situe ce nombre dans la colonne des cumuls et on retient l'unité i pour laquelle $C_{i-1} \leq x \leq C_i$
- 6 On progresse dans la base en ajoutant à chaque fois le PAS. On situe les nombres ainsi obtenus dans la colonne des cumuls pour retenir les unités-échantillon.

b) Tirage systématique sur fichiers de probabilités cumulées

Cette méthode exige de connaître les probabilités d'inclusion dans l'échantillon de chaque individu de la base de sondage. La probabilité d'inclusion π_i d'une unité de

taille N_i est donnée par $\pi_i = \frac{n \times N_i}{C_N}$. Nous avons besoin de la taille de la population

entière (C_N), de la taille de chaque individu (N_i) et de la taille de l'échantillon (n) pour calculer les probabilités d'inclusion de chaque individu de la base de sondage. Il faut noter qu'en cas de présence d'unités de très grandes tailles dans la base de sondage, la probabilité d'inclusion peut être supérieure à l'unité (voir colonne proba du **Tableau 8.1**). Généralement, les unités avec des probabilités d'inclusion plus grande que l'unité sont systématiquement sélectionnées dans l'échantillon.

Le principe est similaire au tirage systématique à partir des totaux cumulés (méthode précédente), mais en utilisant le cumul des probabilités d'inclusion en lieu et place de la taille cumulée.

c) La méthode de Bernoulli ou tirage de Poisson.

Les étapes sont les suivantes :

- 1 Numérotter les unités de façon séquentielle de 1 à N ;
- 2 Calculer pour chaque unité i de la base de sondage sa probabilité d'inclusion π_i ;
- 3 Pour chaque unité i , générer un nombre aléatoire u_i compris entre 0 et 1 selon une loi de probabilité uniforme ;
- 4 Si $u_i < \pi_i$, on sélectionne l'unité i dans l'échantillon, sinon on passe.

Application de la méthode de Bernoulli

On désire tirer quatre villages parmi 10 suivant un sondage à probabilité inégales. Le tableau 8.1 montre le tirage de l'échantillon selon la méthode de Bernoulli

Tableau 8.1 : Tirage à partir de la méthode de Bernoulli

N°	Taille	Proba	Proba cumulée	Nombre aléatoire	Résultat du test
1	1	0,05714286	0,05714286	0,3764484	on rejette
2	6	0,34285714	0,4	0,21142828	on retient
3	3	0,17142857	0,57142857	0,12107606	on retient
4	2	0,11428571	0,68571429	0,20295364	on rejette
5	9	0,51428571	1,2	0,38336235	on retient
6	10	0,57142857	1,77142857	0,6885306	on rejette
7	6	0,34285714	2,11428571	0,81372377	on rejette
8	5	0,28571429	2,4	0,53827382	on rejette
9	8	0,45714286	2,85714286	0,84610835	on rejette
10	20	1,14285714	4	0,70700543	on retient
Total	70	4			

2.3. Sondage stratifié

Comme expliqué précédemment, si la population est homogène, le sondage aléatoire simple donne de bonnes estimations. Cependant, si elle est très hétérogène, cette méthode peut produire des estimations très imprécises, surtout si le tirage sous représente considérablement certaines catégories de la population.

On peut rendre les estimateurs du sondage aléatoire, simple, beaucoup plus précis en constituant, préalablement au tirage des individus, des groupes, appelés *strates*, à l'intérieur desquels les individus présentent une certaine ressemblance vis-à-vis de la variable d'intérêt ; et procéder ensuite, de façon indépendante, à des sondages partiels dans chaque strate. L'assemblage des sondages partiels précis donnera des résultats beaucoup plus précis qu'un sondage de même taille effectué en vrac. Ce principe est à la base du *sondage stratifié*.

En d'autres termes, dans le sondage aléatoire stratifié, la population est d'abord divisée en groupes distincts, ne se chevauchant pas, appelés *strates* (par ex. âge, sexe, critères géographiques) ; on effectue ensuite dans chaque strate un échantillonnage aléatoire. En d'autres termes, on doit tirer absolument au moins une unité à enquêter dans chaque strate. En pratique, l'échantillonnage est conçu de sorte à tirer au moins deux unités dans chaque strate pour être en mesure de calculer la précision des estimateurs au sein des strates.

Avantages

- possibilité d'obtenir une estimation pour chaque strate de la population et non pas seulement une estimation globale comme dans le cas du sondage aléatoire simple ;
- gain de précision de l'estimation globale de la population, à condition que la méthode de stratification ait été correctement choisie ;
- réduction des coûts de l'enquête lorsque l'on utilise un critère de stratification géographique ;
- avec une bonne stratification, un plus petit échantillon ne peut être sélectionné qu'avec un sondage aléatoire simple.

Inconvénients ✕

- Coût lié à la préparation d'une base de sondage ;
- il faut obtenir certaines informations préliminaires sur la (les) variable(s) à utiliser pour la stratification, ce qui peut être coûteux ou irréalisable ;
- si la stratification pour certaines variables permet d'améliorer la précision de l'estimation concernant un paramètre comme la prévalence d'une maladie, elle peut donner des précisions assez médiocres pour les autres caractères qui doivent être étudiés en même temps (par ex. la fécondité) ;
- lorsque les critères de stratification ne sont pas connus avec précision, les erreurs commises lors de la constitution des strates peuvent, en réduisant l'homogénéité de celles-ci, diminuer l'efficacité de la méthode sans toutefois introduire de biais.

2.4. Sondage à plusieurs degrés

Cette technique est utilisée lorsqu'il n'est pas possible d'avoir une base de sondage des unités à enquêter. On tire d'abord un échantillon aléatoire de groupes (unités collectives ou unités primaires), puis un sous échantillon aléatoire d'éléments dans chacune des unités collectives choisies (unités secondaires). Les unités collectives peuvent être des villages, des entreprises, des écoles, des ménages etc.

Exemple :

On tire des villages, puis des maisons au sein des villages, ensuite des foyers à partir de la liste des maisons, puis des individus au sein des foyers. Il s'agit d'un *échantillonnage à 4 degrés*.

Avantages ✓

- Une base complète de la population n'est pas nécessaire, il faut seulement la liste des unités collectives (ou primaires), et pour celles qui sont tirées, la liste de la population. La seconde étape suppose qu'un recensement est conduit au sein des unités primaires sélectionnées pour constituer la base de sondage au sein de chaque unité primaire ;

- réduction des frais de déplacement, car les sujets à examiner sont groupés dans les unités collectives tirées.

Inconvénients ✕

- L'inclusion dans l'échantillon d'une partie seulement des membres d'un groupe peut parfois donner lieu à des difficultés d'adhésion à l'enquête ;
- le caractère incomplet des informations peut conduire à des résultats biaisés.

2.5. Sondage en grappes

C'est un cas particulier du sondage à plusieurs degrés où au dernier degré on effectue un recensement. Par exemple, ayant tiré un certain nombre d'îlots à l'avant-dernier degré de tirage, on interroge TOUS les individus appartenant à ces îlots.

Avantages ✓

réduction des frais de déplacement, car le regroupement géographique des sujets à examiner facilite les opérations.

Inconvénients ✕

Perte en précision due au fait que les unités d'un même groupe ont tendance à se ressembler (*effet grappe*) et donc l'échantillon fournit une représentativité faible de la population totale.

Exemple : En l'absence d'une base de sondage de la population, dresser une liste des unités d'habitation ; puis tirer un échantillon d'unités d'habitation par sondage aléatoire simple et examiner tous les individus de chaque unité tirée (recensement).

En pratique, il est plus commode de combiner les méthodes. Par exemple, il est possible de faire un échantillonnage par grappe avec stratification des villages et tirage au sein des strates des villages à probabilités inégales en fonction de la taille du village. Dans ce cas, la perte de précision de l'échantillonnage par grappe est compensée par les gains de la stratification et du sondage à probabilité inégale. En outre, cette approche permet de limiter les coûts.

III. Les techniques de sélection non probabilistes

Une méthode de sondage est dite *non probabiliste* (ou *empirique* ou encore *non aléatoire*) lorsqu'elle n'inclut pas de mécanisme de sélection aléatoire des individus à interroger. Les méthodes empiriques font intervenir des facteurs subjectifs dans la sélection des unités à enquêter dans la mesure où le choix est laissé à l'évaluateur, qui devra convaincre le lecteur de la rigueur de sa démarche plutôt que de suivre une technique précise de sélection. Ces méthodes sont souvent plus simples et rapides à mettre en œuvre et particulièrement adaptées lorsque la collecte vise à explorer un phénomène dans toute sa richesse, ou à documenter des cas précis de mise en œuvre plutôt que de disposer d'un échantillon représentatif pour y faire des estimations généralisables à une population plus large. C'est souvent le cas pour des collectes de nature qualitative, des entretiens, des groupes de discussion ou pour réaliser des études de cas. Certaines techniques permettent tout de même d'obtenir des échantillons se rapprochant de ceux qui ont été composés aléatoirement.

Les techniques de sondage empiriques les plus couramment utilisées sont les suivantes :

- la technique des quotas ;
- la technique des itinéraires ;
- la technique des unités-types ;
- la technique du volontariat ;
- la technique boule de neige ;
- la technique accidentelle

3.1. Technique des quotas

La technique des quotas consiste à construire un échantillon qui soit une photo en miniature de la population étudiée. Pour cela, on impose à l'échantillon de respecter des quotas définis par rapport à des critères explicatifs (liés au sujet de l'enquête) dont on connaît la distribution dans la population étudiée.

Elle commence par identifier dans la population d'origine des groupes d'individus selon certains **caractères** en général liés au phénomène investigué (âge, sexe, ethnie, région d'habitation, niveau, etc.).

On définit une composition de la population à étudier qui permet de dire qu'elle est composée de $x\%$ d'individus de sexe masculin, de chaque groupe ethnique, de chaque classe d'âge, de région d'habitation, de niveau socioéconomique élevé, etc.

L'échantillon sera ensuite défini comme devant comporter $x\%$ d'individus de sexe masculin, de chaque groupe ethnique, de chaque classe d'âge, de région d'habitation, de niveau socioéconomique élevé, etc.

Les individus le composant sont identifiés de façon non aléatoire et les enquêteurs ont un « quota » d'individus à trouver dans chaque groupe.

Cette méthode bien utilisée dans les situations où les caractères démographiques et sociaux dans la population d'origine sont bien connus permet d'avoir une bonne estimation de la fréquence du phénomène à étudier.

3.2. Technique des itinéraires

C'est une variante de la technique des quotas : au lieu de laisser la liberté à l'enquêteur de choisir dans la *nature* les individus à interroger, on lui impose de ne réaliser ses interviews que dans des endroits bien prédéterminés (rues, arrêts de bus, gares, immeubles, centres commerciaux, etc.) tout en respectant des quotas. Cette contrainte géographique prend la forme d'un *itinéraire*. La méthode des itinéraires permet surtout de traduire dans l'échantillon les caractéristiques géographiques de la population, et permet de simplifier logistiquement la collecte.

Par exemple, si on veut étudier les caractéristiques sociodémographiques et le chiffre d'affaires des vendeurs ambulants de la capitale, on peut utiliser la méthode des itinéraires : on fixe des quotas suivant un certain nombre de critères sociodémographiques et un itinéraire à chaque enquêteur et des ronds-points où il doit observer ces vendeurs.

3.3. Technique des unités-types

Cette technique consiste à enquêter un individu « type » que l'on considère « représentatif » d'un groupe d'individus possédant les mêmes caractéristiques. Cette méthode est beaucoup utilisée dans les études de cas, où on ne va étudier qu'un tout petit nombre d'unités. On cherche ainsi à observer une diversité de cas, sans s'imposer un tirage aléatoire, mais plutôt en créant la diversité grâce à la connaissance acquise sur la population à l'étude. Cette méthode est donc rapide et exige peu de moyens. Cependant, la difficulté réside dans le choix des unités-types, car il est nécessaire d'avoir une très bonne connaissance de la population étudiée pour effectuer cette sélection raisonnée, sans louper des caractéristiques intéressantes.

3.4. Technique du volontariat

Ici, seuls les individus volontaires sont interrogés. Cette méthode est donc très sensible au biais de sélection et ne donne pas en général une bonne représentation de la réalité, si l'évaluateur ne prête pas une attention particulière à cette menace à la fiabilité des résultats. En effet, ce sont en général ceux qui ont des opinions très favorables ou très défavorables qui accepteront de participer. On aboutira généralement à des surestimations des valeurs réelles des paramètres d'intérêt. En conséquence, les résultats de ce type d'enquête doivent être interprétés avec beaucoup de précautions et ne doivent pas faire l'objet d'une extrapolation hâtive.

3.5. Technique boule de neige

La méthode boule de neige (ou toile d'araignée) consiste à se procurer un échantillon d'individus possédant une caractéristique rare dans la population.

On procède comme suit :

- pour se faire, on repère d'abord un individu ayant la caractéristique en question ;
- après l'avoir interrogé, on lui demande ensuite s'il connaît un autre individu qui a la même caractéristique que lui ;
- ainsi de suite, jusqu'à l'obtention de la taille de l'échantillon.

Cette technique est souvent utilisée pour étudier des phénomènes illégaux, rares ou novateurs, où il est impossible de détenir une liste d'individus. Cette technique est donc souvent utilisée à défaut de pouvoir en utiliser une autre, et s'avère utile pour explorer des effets indésirables d'une intervention ou ses effets auprès de sous-groupes marginaux.

3.6. Technique accidentelle

Cette technique d'échantillonnage est également utilisée lorsqu'on n'a pas d'informations suffisantes sur la structure de la population d'étude. Elle consiste donc à choisir les individus à inclure dans l'échantillon au fur et à mesure qu'ils se présentent à l'observateur, sans aucun tri. Par exemple, sélectionner les 30 premiers votants d'un bureau de vote.

L'échantillonnage accidentel a l'avantage d'être simple, rapide, peu coûteux, même si elle ne donne pas suffisamment de garanties quant au biais de sélection.

Catégorie	Méthode d'échantillonnage	Principe	Avantages	Limites
Probabiliste	SAS	Même chance à tous les individus d'être dans l'échantillon	+ Calcul de précision et d'intervalle de confiance + Évite des choix orientés + Réduit les biais dans le choix des unités interrogés	+ N'est pas utilisable si la liste des unités à enquêter n'est pas disponible + Moins précise si la population est hétérogène
	Stratifié	SAS dans des strates	+ Ciblage de catégories particulières de la population + Amélioration de la précision	+ N'est pas utilisable si la liste des unités à enquêter n'est pas disponible et si on ne connaît pas à priori les catégories
	Probabilités Inégales	Chance inégales aux individus d'appartenir à l'échantillon	+ Améliore la précision des estimations	+ nécessité d'obtenir des informations supplémentaires dans la base de sondage
	Plusieurs degrés	Tirer des groupes d'unité, ensuite tiré des sous-groupes dans les groupes tirés, ... jusqu'au dernier niveau, tirage des unités à enquêter	+ Moindres coûts + Remédier à l'absence de base de sondage	+ Moins précis
	Grappes	Plusieurs degrés avec enquête de toutes les unités des sous-groupes tirés	+ Moindres coûts + Remédier à l'absence de base de sondage	+ Moins précis

Non probabiliste	Quota	Sélection des unités de façon raisonnée en respectant une distribution selon la distribution de critères connus sur la population totale	+ simplicité de mise en oeuvre + Pas nécessaire d'avoir à priori la liste des unités à enquêter	+ Sujet à la subjectivité de l'évaluateur + validité externe limitée
	Itinéraires	Sélection des unités de façon raisonnée en suivant un itinéraire spécifique	+ logistique de collecte simplifiée + Pas nécessaire d'avoir à priori la liste des unités à enquêter	+ peut être source de biais + validité externe limitée
	Unités-types	Sélectionne des unités représentatives d'un groupe	+ Utile pour faire une sélection d'étude de cas	
	Volontariat	On ne prend que les volontaires	+ Pas nécessaire d'avoir à priori la liste des unités à enquêter	+ très sensible au biais de sélection
	Effet boule de neige	Identification des unités à enquêter successivement selon l'orientation des personnes interrogées au fil et à mesure	Utile pour étudier des populations marginales, ou des phénomènes rares ou peu connus	+ validité externe limitée

IV. Définition et détermination de la précision d'une enquête par sondage

Cette section présente les différentes notions utiles pour le calcul de précisions et la détermination de la taille de l'échantillon, elles sont particulièrement utiles lorsque l'on souhaite extrapoler les résultats d'un échantillon à l'ensemble d'une population pour effectuer ensuite un jugement évaluatif sur une intervention.

4.1. Notion d'erreurs

Dans un sondage, on peut rencontrer quatre types d'erreurs : (i) erreur d'échantillonnage, (ii) erreur de couverture (iii) erreur d'observations ou de mesure, et (iv) erreur de non-réponse. Le Tableau 8.2 définit chacune de ces erreurs.

Tableau 8.2: Typologie des erreurs dans un sondage

Type d'erreur	Définition
Erreur d'échantillonnage	Par définition, le sondage se contente d'observer seulement une partie de l'univers. Cette façon de procéder peut introduire des biais dans l'estimation des paramètres d'intérêt en ce sens que les résultats numériques dépendront des individus qui ont été enquêtés. Pour différents choix de l'échantillon, on a donc plusieurs estimations du même paramètre. On parle alors d'erreur d'échantillonnage. L'erreur d'échantillonnage résulte donc du plan d'échantillonnage, c'est-à-dire de la façon de tirer les unités.
Erreur de couverture	Elle intervient lorsque la base de sondage est incomplète. Les unités absentes de la base de sondage ne sont pas enquêtées et cela modifie la structure réelle de la population. Les unités exclues ne seront pas prises en compte dans le calcul des estimateurs ; ce qui peut entraîner une perte de précision au niveau du résultat final. La perte de précision sera moindre si les comportements des unités absentes ne diffèrent pas trop des unités présentes dans la base (mais on ne sait pas juger !).
Erreur d'observations ou de mesure	Ce sont les erreurs qui proviennent de l'inexactitude des réponses des enquêtés. Ces erreurs peuvent être aléatoires et se compenser. Cependant, il existe certaines erreurs systématiques qui introduisent des biais et faussent les résultats de façon irrémédiable. Les erreurs d'observation étant en général difficilement mesurables, on s'efforcera lors de la conception d'une enquête à les maîtriser le plus possible (par exemple par la formation des enquêtés sur la façon de présenter les questions), car leur ordre de grandeur peut être nettement supérieur à celui de l'erreur d'échantillonnage.

Erreur de non-réponse	<p>Elles proviennent de l'absence totale ou partielle d'information concernant certaines unités de l'échantillon. On parle de <i>non-réponse totale</i> lorsque l'enquêté ne répond à <i>aucune</i> des questions posées. Cette situation peut provenir de l'absence de l'individu au moment de l'enquête ; du refus catégorique de l'enquêté de répondre aux questions ; ou de la perte de questionnaire. L'attrition est un cas de non-réponse totale. La <i>non-réponse partielle</i> correspond au cas où l'enquêté ne répond pas à <i>certaines</i> questions (soit parce qu'il les trouve tabous ou sensibles). Mais on peut introduire une non-réponse partielle en considérant une réponse donnée comme aberrante (on la supprime dans ce cas).</p>
------------------------------	---

L'erreur d'échantillonnage est très importante dans le calcul des précisions. Nous en discutons en détail dans la section suivante.

4.2. Erreurs de types I et II, et intervalles de confiance

Lorsque l'on utilise un échantillon, on peut rencontrer deux types d'erreurs d'échantillonnage potentielles, appelées *type I* et *type II*.

Une *erreur de type I* apparaît lorsque l'évaluation amène à conclure quelque chose (par exemple, un programme a eu un impact) alors que ce n'est pas le cas.

Illustration (erreur de type I) : Dans l'évaluation d'un programme de nutrition, une erreur de *type I* serait commise si l'évaluateur conclut que le poids moyen des enfants de l'échantillon traité est supérieur à celui des enfants de l'échantillon de comparaison, alors que le poids moyen dans les deux groupes est en fait équivalent. Dans ce cas, l'impact positif constaté est entièrement attribuable au manque de précision des estimations.

À l'inverse, une *erreur de type II* apparaît lorsque l'évaluation amène à réfuter une hypothèse alors qu'il y a en réalité eu un résultat.

Illustration (erreur de type II) : Dans notre exemple du programme de nutrition, une erreur de *type II* est commise si on arrive à la conclusion que le poids moyen des enfants des deux échantillons est le même, alors que le poids moyen des enfants de la population traitée est en fait différent à celui des enfants du groupe de comparaison. Là encore, l'impact estimé aurait dû être différent de zéro, mais le manque de précision des estimations amène à conclure que le programme n'a pas eu d'impact.

Niveau de confiance : Lorsqu'on teste l'hypothèse qu'un programme a eu un impact, l'évaluateur peut limiter la probabilité d'erreurs de type I. En effet, la probabilité de commettre une erreur de type I est définie par un paramètre appelé le « *niveau de confiance* » (notée α). Le niveau de confiance est généralement fixé à 5%, ce qui indique que vous pouvez être sûr à 95% de votre conclusion selon laquelle le programme a eu un impact. Si vous craignez de commettre une erreur de type I, vous pouvez fixer un niveau de confiance plus faible, de 1% par exemple, pour pouvoir être sûr à 99% de votre conclusion selon laquelle le programme a eu un impact. Le niveau de confiance permet de construire un intervalle de confiance qui peut contenir la vraie valeur du paramètre d'intérêt.

Mise en contexte pour une erreur de type II : Les erreurs de type II sont également une source d'inquiétude pour les décideurs. Plusieurs facteurs influencent la probabilité de commettre une erreur de type II ; toutefois, la taille de l'échantillon est un facteur déterminant. Si le poids moyen de 50 000 enfants traités est le même que le poids moyen de 50 000 enfants de comparaison, vous pouvez probablement conclure avec certitude que le programme n'a pas eu d'impact. En revanche, si les deux enfants de votre échantillon de traitement pèsent le même poids que les deux enfants du groupe de comparaison, il sera plus difficile de formuler une conclusion avec certitude. Vous vous demanderez alors si le poids moyen est similaire parce que l'intervention n'a pas eu d'impact ou parce que les données sont insuffisantes pour tester l'hypothèse à partir d'un échantillon si petit. En prélevant des échantillons suffisamment grands, vous réduisez le risque de n'observer, par hasard que des enfants ayant le même poids. Avec de grands échantillons, la différence de moyennes entre un échantillon de traitement et un échantillon de comparaison fournit une estimation fiable de la véritable différence qui existe entre toutes les unités traitées et toutes les unités du groupe de comparaison.

Puissance statistique : La *puissance* (ou *puissance statistique*) d'une évaluation d'impact correspond à la probabilité qu'elle détecte une différence entre les groupes de traitement et de comparaison, si une telle différence existe. La puissance d'une évaluation d'impact est élevée si le risque de ne pas observer un impact qui existe, c'est-à-dire de commettre une erreur de type II, est faible. La taille de l'échantillon est un facteur déterminant de la puissance d'une évaluation d'impact : la taille de l'échantillon augmente avec le niveau de puissance. Le niveau de puissance est fixé par l'évaluateur. On utilise généralement un niveau de puissance de 0,8.

V. Détermination de la taille de l'échantillon

Dans une collecte en évaluation, deux paramètres viennent déterminer la taille d'un échantillon, à savoir le but recherché (et donc la technique choisie) et le budget disponible.

En augmentant la taille de l'échantillon, on augmente la précision et on diminue la marge d'erreur. Plus la population est réduite, plus le rapport recherché entre la taille de l'échantillon et celle de la population est grand.

L'objectif généralement recherché en évaluation est un niveau de confiance de 95% et une marge d'erreur d'environ 5%.

Plus précisément, la taille de l'échantillon se calcule comme suit :

$$n = \frac{t_{1-\frac{\alpha}{2}} p(1-p)}{\varepsilon^2} c$$

Effectif nécessaire :

p = fréquence relative du phénomène sous forme décimale

n = effectif de l'échantillon

ε = écart réduit pour le risque consenti (1,96 pour le risque consenti de 0,05 à un intervalle de confiance de 95%)

c = facteur correctif de l'effet de grappe (fixé à 2).

Dans le cas particulier de l'évaluation d'impact, la puissance statistique est utilisée pour déterminer la taille de l'échantillon. Le calcul de la puissance indique la plus petite taille d'échantillon requise par un plan d'évaluation pour détecter une différence significative (effet minimum détectable) des résultats entre les groupes de traitement et de comparaison. L'effet minimum détectable (EMD) est donné par

$$EMD = \left(t_{1-\frac{\alpha}{2}} + t_{1-\beta} \right) \times \sqrt{\frac{2\sigma^2}{n}}$$

La quantité $t_{1-\alpha/2}$ est le fractile d'ordre $1 - \alpha/2$ de la loi normale, avec $1 - \alpha$ le niveau de confiance souhaité (α étant l'erreur de type I). La quantité $t_{1-\beta}$ est le fractile d'ordre $1 - \beta$ de la loi normale, avec β l'erreur de type II et $1 - \beta$ la puissance statistique. Le second terme au niveau du membre de droite de l'égalité est l'erreur standard de l'effet estimé : σ^2 est la variance de l'outcome et n est la taille de l'échantillon. On voit que l'erreur de l'effet estimé diminue avec la taille de l'échantillon. Par ailleurs, l'EMD est inversement proportionnelle à la taille de l'échantillon à enquêter. Cela signifie qu'il faudra un échantillon de grande taille pour détecter des phénomènes rares et vice-versa.

La taille minimale de l'échantillon est donnée par :

$$n = 2 \left(\frac{t_{1-\frac{\alpha}{2}} + t_{1-\beta}}{EMD} \right)^2 \sigma^2$$

Application

Dans un pays, un programme est initié pour lutter contre le chômage des jeunes. Le taux de chômage des jeunes de 42% (enquête nationale janv. 2014). Le programme de bons de salaire pour les jeunes ambitionne de réduire le chômage à 20% en 2 ans. Sur un total de 4000 jeunes éligibles, le ministre décide de sélectionner 420 à parts égales dans les groupes de traitement et de contrôle.

Vous dirigez une équipe d'évaluation d'impact et un journaliste demande : ***pensez-vous qu'un échantillon de taille 420 est suffisant pour cette évaluation ?***

Pour répondre à cette question, on suppose une puissance statistique de 80%, soit $\beta = 0,2$ et $\alpha = 5\%$. En appliquant la formule de calcul de la taille minimale, on trouve 79 participants par groupe (traitement et contrôle). Dans le cas de cet exercice, l'EMD est de 22% ; c'est-à-dire la différence entre le taux de chômage visé en fin de programme et le taux de chômage avant l'intervention.

Si on ajoute un taux d'attrition de 5%, on peut enquêter 83 participants ; ce que donnent 166 participants au total. La prise en compte de l'attrition permet de s'assurer que si on perd des individus au cours de la phase de suivi, l'échantillon demeure suffisant pour évaluer l'impact du projet. En conclusion, la taille de 420 est suffisante pour répondre à la question du journaliste.

Messages clés :

- ✓ Il n'existe pas une technique d'échantillonnage s'appliquant à toutes les situations, comme pour beaucoup d'éléments de la stratégie d'évaluation, il s'agit pour l'évaluateur de diagnostiquer la situation pour comprendre la technique qui s'appliquera le mieux.
- ✓ Les techniques non probabilistes ont une validité externe limitée, mais elles sont cependant très utiles dans le cadre d'approches qualitatives, où l'évaluateur devra démontrer la fiabilité de la sélection d'unité effectuée
- ✓ Les techniques probabilistes sont utiles pour des collectes de nature quantitatives, telles que les sondages et pour mettre en œuvre une méthode quantitative d'évaluation des effets. En présence d'un échantillon suffisant et en respectant les techniques de sélection à la lettre, une extrapolation à l'ensemble de la population étudiée est possible.

Bibliographie :

- Ardilly, P. (2006). Les techniques de sondage. TECHNIP.
- Gauthier, B. et Bourgeois, I. (2016) Recherche sociale. De la problématique à la collecte des données, 6^eéd., Québec, Presses de l'Université du Québec, 670p. Chapitres: 7 et 8
- Gertler P. J., Martinez S., Premand P., Rawlings, L. B., & Vermeersch, C. M. (2011). L'évaluation d'impact en pratique. The World Bank.
- Gertler P. J., Martinez S., Premand P., Rawlings, L. B., & Vermeersch, C. M. (2016). Impact evaluation in practice (Second edition). The World Bank.
- Keho, Y (2007). Sondages : Méthodes et calcul d'estimateurs. Note pédagogique. ENSEA.
- Van Campenhoudt L. et Quivy R. (2011). Manuel de recherche en sciences sociales, 4^eéd., Dunod, 262p.

MODULE 9

STRATEGIE DE COLLECTE DE DONNEES

Objectifs d'apprentissage :

A la fin de la session, les participants seront capables de :

- Définir les concepts, stratégies, méthode et outils de collecte de données
- Distinguer les principaux outils de collecte des données
- Choisir et élaborer les outils de collecte de données
- Développer une méthode de collecte de données
- Distinguer les approches quantitatives/qualitatives
- Elaborer un plan de collecte dnnées

Thèmes couverts

- I. Définition des concepts
- II. Méthodes de collecte de données
- III. Outils de collecte de données

Ce module met l'accent sur les différentes stratégies de collecte des données, de nature qualitative et quantitative. Après une définition des concepts, il présente les modes opératoires pour la collecte des données et fait un tour d'horizon des principaux outils de collectes les plus utilisés en évaluation. Le sujet de la collecte de données est primordial, car il constitue une des tâches les plus importantes en temps du travail de l'évaluateur. Le bon déroulé et la rigueur de la mise en œuvre de cette étape de travail vont aussi fortement déterminer la crédibilité de l'argumentaire évaluatif qui sera déroulé dans le rapport d'évaluation.

Le choix d'une stratégie de collecte dépend du type de variable visé dans la matrice d'évaluation, de la précision souhaitée, de la logistique possible pour la collecte, des compétences de l'équipe d'évaluation, etc. Comme pour de nombreux éléments de la méthodologie d'une évaluation, c'est à l'évaluateur de comprendre l'objet d'évaluation et les questions qui lui sont posées et de proposer une stratégie qui permet d'y répondre avec rigueur. La stratégie de collecte de données retrace ainsi tout le processus de la collecte des données. Une stratégie de collecte débute par la définition et la typologie des variables en rapport avec l'objectif de l'évaluation. Elle inclut également la technique de collecte et les outils pour la collecte des données. Les outils de collectes sont les instruments utilisés dans une collecte réalisée de manière rigoureuse (guides d'entretiens, fiches d'observation, questionnaire de sondage, etc.). L'outil de collecte est donc un instrument normalisé dont la structure et le contenu sont construits en fonction des objectifs initiaux de la recherche considérée. Il est fortement conseillé d'inclure également à la réflexion sur la stratégie la façon dont les données seront stockées et gérées, leur traitement et leur exploitation ce qui sera étudié au chapitre suivant.

Mots-clés : collecte, entretiens, questionnaires, observation, outil de collecte

I. Principaux concepts

1.1. Conseils pour la collecte de données

Il n'existe pas une seule meilleure façon ou une seule bonne pratique pour la collecte des données. La décision dépend de ce dont nous avons besoin, à savoir *des chiffres ou des histoires*. La réponse à cette question oriente vers le choix de la stratégie et des outils à utiliser. De plus, les choix de collecte doivent tenir compte des ressources et du temps disponible, de même que la complexité du recueil de données (en fonction des expertises de l'équipe d'évaluation notamment). La fréquence de la collecte des données joue un rôle important dans l'échéancier de l'évaluation et dans les possibilités. Il faut enfin avoir la finalité de la collecte à l'esprit, cela signifie que les formes prévues pour l'analyse des données conditionnent également les données à collecter.

De même qu'il n'existe pas une technique supérieure en matière de collecte de données, aucun outil n'est supérieur à tous les autres. Il convient de ne pas laisser l'outil guider le travail, mais plutôt de choisir l'outil approprié pour aborder la question d'évaluation. Il est intéressant de retenir plus d'une technique de collecte de données et d'utiliser plusieurs outils de collecte de données pour documenter une seule et même variable quand cela est possible pour effectuer une triangulation (c'est-à-dire, confirmer les résultats en disposant de plusieurs sources d'information). Toutefois, cela n'est pas toujours réalisable pour l'ensemble des variables d'une matrice d'évaluation. Dans ce cas, l'on va alors préférer la complémentarité entre différents outils pour documenter chacune des variables, le but étant de retenir au final, une stratégie qui représente un ensemble de points de vue assez importants au regard du budget et de l'échéancier de l'évaluation.

Dans la mesure du possible, il est nécessaire de prioriser l'utilisation des données disponibles appelées secondaires. Néanmoins, il y a nécessité de savoir comment les mesures ont été définies et les données recueillies et traitées pour s'assurer de leur crédibilité et de leur fiabilité avant de les intégrer à l'argumentaire évaluatif. C'est

seulement en l'absence de données déjà disponibles que l'évaluateur procède à une nouvelle collecte de données primaires.

En cas de collecte de données (primaires), il faut également être sensible au fardeau imposé aux répondants, et leur capacité à accorder du temps au processus de collecte. Un certain nombre de principes éthiques doivent également être respectés puisque l'évaluateur est en contact avec des citoyens et peut être amené à détenir des informations personnelles et confidentielles.

1.2. Caractéristiques d'une bonne mesure

Evaluer un projet ou un programme nécessite généralement de collecter des données. La collecte des données répond à un objectif bien précis qui peut être apprécié à travers leur pertinence, leur crédibilité, leur validité et leur fiabilité. Le Tableau 9.1 propose une définition de chacun de ces critères permettant d'évaluer une bonne mesure.

Tableau 9.1 : Critères pour une bonne mesure

Critères	Question	Observations
Pertinence	La mesure saisit-elle ce qui est important ?	Ne pas mesurer ce qui est facile au lieu de ce qui est nécessaire
Crédibilité	Est-ce la mesure fait foi ? Sera-t-elle perçue comme un moyen raisonnable et approprié de recueillir l'information recherchée ?	A-t-on procédé de manière rigoureuse ?
Validité interne	A quel degré de satisfaction la mesure saisit-elle (capture) ce qu'elle est sensée mesurer ?	A-t-on oublié des aspects importants ? Ou des points de vue différents ?
Fiabilité	La précision et la stabilité d'une mesure : A quel point le même résultat serait obtenu avec des essais répétés	Quel degré de fiabilité : Des poids à la naissance des nouveau-nés ? Des vitesses mesurées par un chronomètre ?

1.3. Type de données

On distingue les données qualitatives des données quantitatives (cf. Tableau 9.1) : il s'agit de la *nature* des données. La donnée quantitative est généralement plus structurée et fournit des données sous forme numérique. Elle est utile pour mesurer avec précision des indicateurs comme l'âge, le coût, la longueur, la hauteur, la surface, le volume, etc. Elle est plus difficile à élaborer, mais plus facile à analyser lorsqu'elle est bien conduite.

L'approche quantitative est plus appropriée lorsqu'on voudrait collecter des données auprès d'un large échantillon ou un grand nombre de personnes. Il n'est toutefois pas exclu de collecter des informations qualitatives au cours d'une enquête qui adopte l'approche quantitative.

En outre, l'approche quantitative est adaptée lorsque nous avons une connaissance précise sur ce qui doit être mesuré et les informations recueillies peuvent être standardisées. En effet, sans cette connaissance, l'élaboration de l'outil de collecte sera délicate et celui-ci risque de ne pas être assez rigoureux. L'approche quantitative utilisera principalement des registres ou des bases de données pour la collecte de données, en se basant sur des questionnaires.

L'administration des questionnaires dans l'approche quantitative peut se faire de plusieurs manières, chacune d'elle présentant des avantages et des limites ou des difficultés de mise en œuvre. Typiquement, on utilisera des agents enquêteurs pour administrer le questionnaire en présentiel (ou l'équipe d'évaluation administrera le questionnaire elle-même) aux individus concernés par l'étude. Cette approche nécessite des ressources importantes, c'est pourquoi il ne faut pas sous-estimer des approches alternatives telles que l'enquête par téléphone, sur papier ou internet. Bien que dans les pays en développement, ces approches posent de nombreuses questions telles que le niveau d'alphabétisation des répondants, la langue maternelle utilisée à l'écrit, la réelle diffusion des nouvelles technologies, ainsi que l'accès à internet ; elles restent donc des avenues intéressantes pour différents types de répondants.

En revanche, le but des méthodes qualitatives de recueil de données est de pouvoir appréhender des phénomènes, de chercher à les comprendre et à expliquer leur impact sur les façons de faire et de penser des populations ; phénomènes qui ne peuvent être appréhendés que de cette manière. Toutefois, il s'agit aussi de limiter les inductions, les prénotions ou préjugés que tout être humain possède lorsqu'il va à la rencontre d'autres personnes.

L'approche qualitative est moins structurée et plus facile à élaborer. Là où l'approche quantitative poserait 3 questions à 100 individus, l'approche qualitative posera 100 questions à 3 individus. L'approche qualitative, même si elle porte sur un échantillon faible peut fournir des données riches, détaillées et applicables à large échelle. Néanmoins, la généralisation des résultats n'est pas le but de l'analyse qualitative, il faut donc rester vigilant sur la portée des interprétations que l'on fait. Elle est adaptée lorsque les données peuvent être observées ou auto-déclarées sans avoir besoin d'être forcément quantifiées. Dans les approches qualitatives, on parle de généralisation analytique, c'est-à-dire que les analyses visent plus l'enrichissement théorique.

Cette approche peut toutefois être difficile à analyser et nécessite un travail intensif de codage et d'interprétation et génère habituellement des rapports plus longs, mais très détaillés.

L'évaluateur peut utiliser l'approche qualitative lorsqu'il mène un travail exploratoire ou qu'il cherche à comprendre les thèmes ou les questions spécifiques. Ceci peut être ensuite poursuivi par une enquête plus quantitative pour confirmer ou infirmer des hypothèses. Elle est également utile pour des narrations ou pour récolter des histoires de changement qui peuvent très bien servir d'illustration et d'explication à des résultats quantitatifs. En outre, si l'évaluateur souhaite disposer des informations riches et approfondies ou cherche à comprendre les résultats non attendus, il peut utiliser l'approche qualitative.

Le Tableau 9.2 présente des situations typiques dans lesquelles l'évaluateur peut recourir à une approche qualitative ou à une approche quantitative.

Tableau 9.2: Objectifs et approches de collecte de l'information

Objectifs	Approche appropriée	Exemple
<p>Mener une analyse des statistiques</p> <p>Être plus précis</p> <p>Couvrir un grand groupe d'individus</p> <p>On a une certitude sur ce qu'on veut mesurer</p>	Quantitative	Evaluer l'impact du programme d'emploi des jeunes
<p>On souhaite une information narrative ou approfondie</p> <p>On souhaite comprendre un phénomène en profondeur</p> <p>On n'est pas sûr d'être capable de mesurer ce qu'on est amené à évaluer</p> <p>On n'a pas besoin de quantifier les résultats</p>	Qualitative	<p>Evaluer l'acceptabilité des interventions par les populations</p> <p>Pourquoi le projet a mieux réussi dans certaines zones ?</p>

II. Outils de collecte de données quantitatives

L'outil est un élément d'une activité qui n'est qu'un moyen, un instrument. Dans le cadre de la collecte des données quantitatives, plusieurs outils de collecte de données peuvent être utilisés ; c'est principalement le cas des registres de données secondaires, les bases de données disponibles, les questionnaires. Dans la suite, un accent est mis sur les registres et les questionnaires.

2.1. Les registres secondaires

Un registre est un recueil d'informations concernant des activités particulières. Dans le cas des statistiques maritimes, les registres consisteront en des informations sur les navires de pêche, les sociétés, les engins, les licences ou les pêcheurs eux-mêmes. Il peut servir à effectuer un dénombrement complet dans le cadre d'une obligation légale. On peut aussi y inscrire des renseignements d'ordre fiscal (concernant par exemple la délivrance ou le renouvellement des licences pour une activité dans le pays ou le secteur d'activité). Les registres peuvent être très utiles pour concevoir et mettre en œuvre un système statistique à condition que les données qu'ils contiennent soient fiables, actuelles et complètes.

2.2. Questionnaires

Les questionnaires consistent en des formulaires qui sont remplis par les déclarants seuls ou administrés par un agent enquêteur. Dans le cas de questionnaires auto-enregistrés, ils peuvent être remis en main propre, ou envoyés et collectés ultérieurement, ou renvoyés au moyen d'une enveloppe déjà prête et affranchie ou encore envoyés par voie électronique. Cette méthode peut être adoptée pour la population entière ou pour des secteurs échantillonnés. Cette pratique suppose un niveau d'alphabétisation élevé. Lorsqu'il existe plusieurs langues, les questionnaires doivent être établis dans les principales langues du groupe cible. Il faut, dans ce cas, veiller tout spécialement à la précision des traductions.

Les questionnaires doivent être conçus de manière aussi simple et claire que possible, en ciblant bien les sections et les questions. Il faut assurer une cohérence entre les questions d'une section. Le nombre de sections et la longueur du questionnaire sont

laissés à l'appréciation de l'évaluateur et des objectifs de l'étude. Néanmoins, les questionnaires devraient être aussi brefs que possible. En concevant le questionnaire, il faut penser au répondant qui consacrera du temps, généralement, non rémunéré pour répondre aux questions. En outre, si par exemple, plusieurs organismes demandent les mêmes données, il faut s'efforcer d'en coordonner la collecte afin d'éviter les doubles emplois.

III. Outils de collecte de données qualitatives

Dans le cadre de la collecte des données qualitatives, plusieurs outils de collecte de données peuvent être utilisés, dont principalement :

- les procédés participatifs ;
- l'observation ;
- les entretiens ;
- les entretiens de groupes
- les journaux, revues, listes de contrôle auto-déclarées ;
- le jugement d'experts.

3.1. Les procédés participatifs

Le procédé participatif implique fortement les groupes ou les communautés dans la collecte de données. Il peut s'agir, par exemple, des réunions de la communauté ou de la cartographie

Réunions de communauté.

Une des méthodes participatives les plus courantes

Elles doivent être bien organisées notamment en s'accordant sur les objectifs et en fixant les règles de base comme :

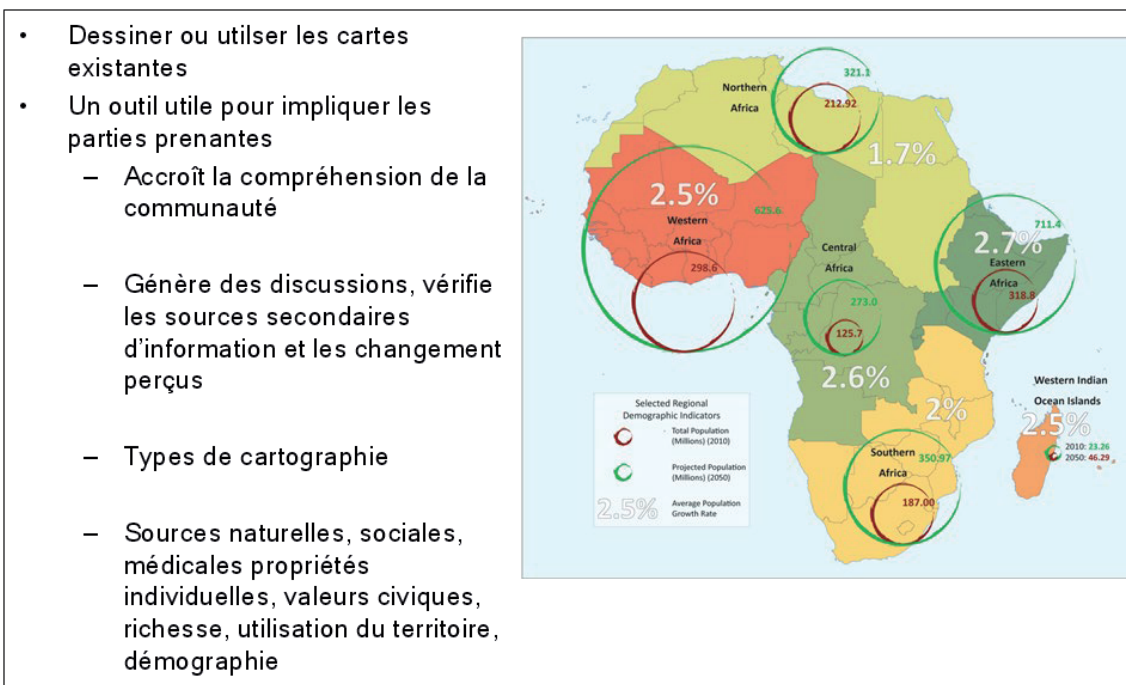
- a) Qui a la parole ?
- b) Temps alloué aux orateurs
- c) Format des questions et des réponses



Cartographie

La cartographie inclut la collecte, l'évaluation et la manipulation des données de base, la création d'une carte, d'un dessin et la production finale du document. La Figure 9.1 fournit quelques informations et un exemple de carte.

Figure 9.1: Information et exemple sur la cartographie



3.2. L'observation

L'observation consiste à observer ou regarder attentivement ce qui se passe au niveau de l'environnement dans lequel la collecte de données s'effectue. La Figure 9.2 donne quelques exemples au niveau desquels l'observation est utilisée pour la collecte des données.

L'observation, souvent oubliée des évaluateurs est utile quand il y a besoin d'informations immédiates et que l'évaluateur peut être témoin du phénomène étudié.

On peut recourir à l'observation si l'on s'intéresse à ce qui est factuel ou lorsqu'il y a des preuves matérielles, des produits, des résultats qui sont observables. Par ailleurs, cet outil se justifie quand il est impossible ou inapproprié de procéder à une nouvelle collecte de données auprès de répondants et que l'on doit fournir une alternative.

Figure 9.1: Collecte de données par observation

Observer ce qui se passe

- Circulation routière
- Modes d'utilisation des terres
- L'aménagement des zones urbaines et rurales
- La qualité des logements
- L'état des routes
- L'état des bâtiments
- Qui fréquente les dispensaires



L'observation peut être structurée, semi-structurée ou non-structurée. Elle est dite structurée lorsqu'avant l'observation, l'on détermine avec précision ce qui doit être observé. Lorsque nous avons une idée générale de ce qui doit être observé, mais sans plan spécifique, on dira que l'observation est semi-structurée. Elle est non-structurée quand le choix de la méthode est fait selon la situation dans un journal d'observations libres, sans idées préconçues ni un plan préalable sur ce qui doit être observé.

Systemes d'informations géographiques (SIG)³

L'observation peut être faite à partir des systèmes d'information géographique ou SIG. Le SIG est un système d'information conçu pour recueillir, stocker, traiter, analyser,

³ Les SIG sont classés dans ce module comme faisant partie des outils qualitatifs. Néanmoins, les SIG peuvent servir à d'autres fins bien au-delà de la collecte qualitative.

gérer et présenter tous les types de données spatiales et géographiques. Le SIG est composé de logiciels, les données, les matériels informatiques, les savoir-faire et les utilisateurs.

En pratique, on utilisera des cartes ou images satellites pour les recherches régionales complexes ou localisées. Un exemple en la matière est Google Earth (<http://earth.google.com/>). Les SIG peuvent avoir une forte influence pour la cartographie, l'évaluation de la distribution des ressources et planifier ou modifier la prestation de services.

Photographie et enregistrement audio / vidéo

La photographie et les enregistrements audio ou vidéo sont également utiles pour l'observation. Les photos historiques peuvent être puissantes et des moyens rapides et utiles pour montrer le degré de progression ou de changement d'une situation.

Avantages et défis de l'observation

Le Tableau 9.3 présente quelques avantages et défis des observations. L'observation bien que permettant de collecter des comportements réels ou factuels, peut présenter de nombreux biais liés à l'observateur et à l'échantillonnage.

Tableau 9.3: Avantages et défis des observations

Avantages	Recueil des données sur les comportements réels au lieu de se fonder sur les informations auto-déclarées ou des perceptions
Défis	Biais des observateurs potentiellement non négligeables Défis liés à l'interprétation et au codage L'échantillonnage peut être problématique et peut demander beaucoup de travail

Les entretiens

Les entretiens sont excellents quand il s'agit de poser aux personnes concernées des questions sur leurs perceptions, leurs opinions et leurs idées. Les entretiens sont utilisés pour explorer en profondeur des problèmes complexes et peuvent donner à l'évaluateur une idée intuitive de la situation. Lors de l'entretien, certaines erreurs peuvent également être tolérées, ce qui en fait un bon outil exploratoire. En effet, les questions non précises peuvent être clarifiées durant l'entretien et changées au cours des entretiens. L'entretien d'évaluation est souvent semi-structuré, c'est-à-dire que lors de son déroulement, la liberté est donnée à la personne interviewée de déborder sur des questions ou des aspects non-inscrits dans le guide, mais que l'évaluateur a tout de même un canevas de quelques questions principales qu'il veut aborder (et qui découle des éléments de sa matrice d'évaluation) Le Tableau 9.4 donne quelques avantages et des défis liés aux entretiens.

Tableau 2.4 : Quelques avantages et défis des entretiens

Avantages	Meilleur outil si vous voulez savoir ce que les personnes pensent, croient et leur perception. Elles sont les seules à pouvoir vous le dire
Défis	<p>Les personnes interrogées peuvent ne pas se rappeler exactement leur comportement ou peuvent être réticents à révéler leur comportement s'il est illégal ou stigmatisé</p> <p>Ce que les personnes interrogées pensent avoir fait ou ce qu'ils disent avoir fait ne correspond pas toujours à ce qu'ils ont effectivement fait.</p> <p>Ils peuvent coûter cher, demander beaucoup de travail et du temps</p> <p>L'écoute sélective de la part de celui qui mène l'entretien peut amener à ne pas capter une information qui n'est pas conforme aux croyances préexistantes</p> <p>La collecte et l'analyse des données des entrevues est sujet à la subjectivité de l'évaluateur</p>

3.3. Groupes de discussion

Les entretiens de groupes ou groupes de discussion sont également très utilisés en évaluation. La forme la plus courante est celle du *focus group* qui est une méthode qualitative utilisée pour connaître les opinions d'un groupe de personnes sur un sujet précis, tel qu'une intervention évaluée. La Figure 9.3 donne une illustration de cet outil. Le processus du groupe de discussion obéit à une logique avec des actions bien précises décrites dans le Tableau 9.5. Les groupes de discussion peuvent être menés assez rapidement et facilement. Ils prennent moins de temps que les entretiens individuels et permettent de confronter les avis et d'observer une dynamique de groupe sur le sujet discuté. Ils autorisent également une flexibilité dans le processus et les questions, mais nécessitent les compétences d'un animateur chevronné pour contrôler la direction prise dans les discussions. En outre, ils peuvent explorer différentes perspectives ; ce qui permet aux individus de se prononcer souvent plus librement et de réfléchir ensemble à ce qu'ils n'auraient pas imaginé en entretien individuel. Les groupes de discussion sont inappropriés quand les obstacles linguistiques sont insurmontables ou quand l'évaluateur a peu de contrôle sur la situation entre les participants (relations de famille, hiérarchiques ou opinions irréconciliables). L'évaluateur qui conduit le groupe de discussion doit veiller à répartir le temps de parole et à éviter les malaises, il est responsable de la dynamique de groupe qui est souvent imprévisible.

Toutefois, l'analyse des données collectées comme en entretien individuel prend beaucoup de temps. En outre, les données sont tributaires de la dynamique qui s'est installée et donc des façons de faire du modérateur ou des membres dominants du groupe.

Figure 9.1: Description du focus-group

- **Type de recherche qualitative** où de petits groupes homogènes de personnes sont mis ensemble pour discuter de manière informelle des sujets spécifiques sous la conduite d'un modérateur
- **Objectifs** : identifier les questions et les thèmes ; pas seulement les informations intéressantes, et pas de "comptes"



Tableau 9.5 : Phases de la conduite de focus-group

Phases	Actions
Ouverture	Briser la glace : expliquer les objectifs, les règles de base, faire les présentations
Réchauffement	Raconter l'expérience ; stimuler l'interaction des groupes ; commencer par des questions les moins intimidantes et les plus simples
Partie principale	Aller vers les questions intimidantes, sensibles ou complexes, provoquer des réponses en profondeur sans les influencer, relier les données émergentes à une participation complexe large
Clôture	Terminer par des questions types de clôture, résumer et préciser ; présenter les théories, etc. ; demander les derniers commentaires ou points de vue ; remercier les participants

3.4. Les journaux et listes de contrôle des données auto-déclarées

Les journaux et listes de contrôle des données sont utilisés quand on veut obtenir des informations sur les événements intervenus dans la vie quotidienne de la population cible. Les participants restituent leurs expériences dans le questionnaire en temps réel, pas plus tard. Ils sont utilisés pour compléter d'autres méthodes de collecte de données.

Cet outil de collecte a l'avantage de recueillir des données détaillées qui, autrement, auraient pu être oubliées. Il peut permettre de recueillir des données sur la manière dont la population cible utilise son temps. Ils complètent les entretiens et fournissent des données plus riches.

Tableau 9.6: Lignes directrices des revues ou journaux

Phases	Processus
1	Recruter des personnes en face à face Encourager la participation, appeler à l'altruisme, assurer la confidentialité et procurer des incitations
2	Donner une brochure à chaque participant La page de garde de la brochure doit contenir des instructions, des définitions et des exemples clairs Fournir de petits aide-mémoires et expliquer les termes. Ces informations sont généralement fournies sur la dernière page de la brochure. Il faut insister également sur le calendrier de remplissage
3	Avoir une bonne appréciation de la période allouée à la collecte de données. Si elle est trop longue, elle peut devenir fastidieuse ou ennuyeuse. Si elle est trop courte, elle peut manquer de contenu.

Les principaux défis sont liés au niveau d’alphabétisation de la population cible et le changement potentiel de comportement des personnes ciblées. Les données collectées peuvent être inexactes ou incomplètes surtout si la personne concernée change de comportement. Cet outil exige un engagement et une autodiscipline des personnes ciblées. En outre, l’écriture peut être illisible et les phrases difficiles à comprendre.

Mis en œuvre pratique des listes de contrôle

En pratique, on effectue un croisement entre un questionnaire et un journal. L’évaluateur spécifie une liste de comportements ou d’évènements et demande aux personnes interrogées de remplir la liste de contrôle. Il spécifie une période au cours de laquelle l’évènement ou le comportement est restitué.

3.5. Le jugement d’experts

Le jugement d’experts fait appel aux avis des experts sur le sujet qui appliqueront le plus souvent la méthode analogique de manière informelle. La méthode analogique est basée sur la comparaison du projet à chiffrer (cas cible) aux projets passés (cas source). La comparaison se fait sur la base d’une série de paramètres descripteurs standardisés. Il s’agit en fait d’utiliser l’expérience passée pour construire le présent.

La mise en œuvre de cet outil nécessite de recruter des experts. Il faut établir des critères pour la sélection d’experts, non seulement sur le fait d’être reconnus comme experts, mais aussi en se fondant sur : (i) les domaines d’expertise; (ii) les points de vue diversifiés ; (iii) les opinions politiques variées et diverses et (iv) une expertise technique diversifiée.

Cette approche peut être très rapide à mettre en œuvre et à moindre coût. La valeur des données dépend du degré de crédibilité accordée aux experts.

Messages clés :

- ✓ L'étape de la collecte et de l'analyse des données est la plus longue dans une mission d'évaluation. Il convient donc de réfléchir à optimiser l'utilisation des données secondaires de bonne qualité, et d'optimiser chacune des collectes réalisées
- ✓ Une réflexion est à mener sur l'ensemble de la stratégie de collecte : quantitative, qualitative ou mixte ? Quel enchaînement entre les collectes ? Quelle collecte est appropriée pour chacune des variables ?
- ✓ Chaque outil de collecte nécessite des façons de faire et des compétences particulières qu'un évaluateur doit maîtriser pour produire des données intéressantes qu'il pourra ensuite analyser et interpréter pour construire son argumentaire évaluatif.

Bibliographie :

- Ardilly, P. (2006). Les techniques de sondage. Editions Technip.
- Gauthier, B. et Bourgeois, I. (2016) Recherche sociale. De la problématique à la collecte des données, 6^eéd., Québec, Presses de l'Université du Québec, 670p. Chapitres: 7 et 8
- Saporta, G. (2011). Probabilités, analyse des données et statistique. Editions Technip.
- Statistique Canada (2003). Méthodes et pratiques d'enquête. No 12-587-X au catalogue. (<https://www150.statcan.gc.ca/n1/fr/pub/12-587-x/12-587-x2003001-fra.pdf?st=JnbvMfWW>)
- Van Campenhoudt, L. et Quivy, R. (2011). Manuel de recherche en sciences sociales, 4^{ème} éd., Dunod, 262p.

MODULE 10

ANALYSE ET INTERPRÉTATION DES DONNÉES

Objectifs d'apprentissage :

A la fin de la session, les participants seront capables de :

- Assurer le contrôle et la compilation des données collectées
- Analyser et interpréter des données quantitatives
- Analyser et interpréter les données qualitatives
- Définir les règles régissant la production du rapport
- Développer une méthode de collecte de données
- Distinguer les approches quantitatives/qualitatives
- Elaborer un plan de collecte de données

Thèmes couverts

- I. Analyse des données quantitatives
- II. Analyse des données qualitatives
- III. Rédaction de fiches d'analyse

Le module précédent a mis l'accent sur les différentes stratégies de collecte de l'information. Il a permis de distinguer les outils quantitatifs des outils qualitatifs. Dans ce module, les différentes stratégies à utiliser pour l'analyse et l'interprétation des données collectées seront présentées. L'analyse des données est l'étape qui consiste à donner une signification aux données collectées. Elle permet à l'aide de différentes techniques de résumer, interpréter, comparer et classer les données collectées, dans le but d'en tirer des conclusions.

L'analyse et l'interprétation des données permettent, en effet à l'évaluateur de dégager les principales informations des données collectées. Pour une bonne exploitation des données, l'évaluateur doit tenir compte du processus de collecte de l'information, à savoir la sélection des répondants ou des documents ainsi que la stratégie de collecte utilisée. L'usage d'une technique d'échantillonnage aléatoire permet d'effectuer des extrapolations à la population entière et le calcul de précision : *on parle d'inférence statistique*. A l'opposé, l'inférence n'est pas possible en cas de recours à un échantillonnage empirique, mais on y trouvera de l'information plus riche et nuancée permettant de mieux interpréter les données quantitatives et de les intégrer dans un argumentaire évaluatif qui fait du sens. Aussi, ne s'étonnera-t-on pas que les outils utilisés pour analyser des données quantitatives soient généralement différents de ceux utilisés pour analyser les données qualitatives.

I. Analyse des données quantitatives

La première étape de l'analyse des données est le résumé de l'information. En général, la collecte des données quantitatives porte sur un nombre important d'individus. L'évaluateur se servira donc des outils statistiques pour résumer l'information, en général la statistique descriptive, en utilisant des indicateurs clés faciles à comprendre par un non-spécialiste (numériques ou graphiques). Il peut par exemple calculer les effectifs de réponse à une certaine question (70% de répondants satisfaits) ou encore une moyenne (les répondants ont en moyenne 5 ans d'expérience dans le poste). Les indicateurs à utiliser pour résumer l'information sont fonction des objectifs visés par l'étude, de la nature de l'information disponible et du message que l'on désire véhiculer. Il n'est jamais possible de présenter toutes les données, et l'art de l'évaluateur est de rassembler les informations qui permettent de répondre à ses questions d'évaluation avec rigueur, mais aussi lisibilité. En outre, les outils à mobiliser sont fonction de la nature des variables disponibles. Par exemple, *on peut* calculer le nombre moyen d'employés par entreprise, mais *on ne peut pas* calculer le sexe moyen du dirigeant de l'entreprise puisque le sexe est une variable qualitative. Enfin, la technique d'échantillonnage affecte l'exploitation des données. Dans le cas d'un sondage probabiliste, on peut par exemple calculer une pondération associée à chaque individu enquêté. La pondération est simplement un poids attribué à chaque individu et est fonction de la sous ou surreprésentation du groupe auquel il appartient.

1.1. Compilation et traitement des données quantitatives

De façon traditionnelle, les informations collectées se faisaient à partir d'un questionnaire papier, qu'il fallait ensuite saisir pour créer une base de données. De plus en plus, on utilise de nos jours des logiciels spécialisés pour programmer le sondage (Google Forms, SurveyMonkey ou Limesurvey pour ne citer qu'eux) qu'il soit auto-administré (par exemple envoi d'un sondage électronique) ou administré par l'équipe d'évaluation (qui saisit les réponses sur un Smartphone ou une tablette plutôt que sur papier). Que ce soit au moment de la saisie de questionnaires papier, ou au moment de la programmation d'un sondage électronique, les réponses font l'objet d'une *codification* ou un *codage*. Le codage des données fait référence à tout processus qui attribue une valeur numérique à une réponse. Par exemple, le sexe de l'individu peut être « masculin » et « féminin » sur le questionnaire papier avec la

possibilité d'entourer ou de cocher la bonne case. La codification consistera, par exemple, à attribuer un code « 1 » à la modalité « masculin » et « 2 » à « féminin ». La codification peut porter aussi sur les questions dites ouvertes qui laissent la possibilité à l'enquêté de se prononcer librement sur une question, mais elle prend beaucoup de temps, on conseille donc de ne pas généraliser leur usage et de s'en tenir à des questions fermées.

Les données saisies constituent la *base de données*. C'est avec cette base de données que les différentes analyses seront faites. Les logiciels couramment utilisés pour l'exploitation des données statistiques sont Excel, Stata, R, Spss, Epi Info, etc.

Toutefois, avant de procéder à l'analyse, il est important de s'assurer que les données sont de bonne qualité. En effet, la qualité des résultats des analyses est fonction de la qualité des données en entrée. Si les données sont douteuses, la crédibilité de l'évaluateur ou de l'analyste est remise en cause. Il ne s'agit donc pas uniquement de la qualité des résultats. Pour s'assurer de la qualité des données, un dispositif d'assurance-qualité doit être mis en place et suivi depuis la conception de l'étude jusqu'à la production du rapport. On vérifiera, par exemple, que les données sont bien renseignées (peu ou pas de valeurs manquantes) et qu'il n'existe pas de valeurs aberrantes. Une valeur aberrante s'écarte en général de la tendance globale (soit trop faible ou trop grande). Par exemple, un chef d'une grande entreprise qui aurait répondu qu'il a 7 ans. Certainement au moment de saisir la réponse, un chiffre a été retiré, et il a plus certainement 47, 57 ou 71 ans. Ceci est un exemple d'erreur de saisie. On vérifiera ainsi toutes les sources potentielles d'erreurs.

1.2. Différents niveaux d'analyse

On distingue globalement deux catégories d'analyse : l'analyse descriptive et l'analyse explicative. Quelle que soit l'approche utilisée, les outils mobilisés sont fortement dépendants de la nature des variables analysées. On distingue deux types de variables, à savoir les variables quantitatives et les variables qualitatives, qui sont subdivisés en plusieurs sous-groupes. Le Tableau 10.1 propose une définition de chaque type de variable.

Tableau 10.1 : Définitions des variables quantitative et qualitative

Variable	Qualité	Définition
Quantitative	Discrète	Variable mesurable mais pouvant prendre un nombre fini de valeurs comme par exemple le nombre d'enfants par femme.
	Continue	Variable mesurable pouvant prendre un nombre infini de valeurs comme par exemple la taille (en cm) ou le poids (en kg)
Qualitative	Binaire	Variable non mesurable comportant deux modalités (sexe du chef d'entreprise)
	Multinomiale	Variable non mesurable comportant plus de deux modalités. Elle est dite <i>ordinaire</i> lorsqu'on peut effectuer un classement ordonné. Par exemple, l'intensité de la douleur est une variable ordinaire. Elle est dite variable nominale dans le cas contraire. Par exemple, la couleur des yeux.

NB : Ne pas confondre les données qualitatives avec les variables qualitatives des données quantitatives (qui sont tout de même exprimées en chiffres dans la base de données).

1.2.1. Analyse descriptive

L'analyse descriptive consiste à dresser un portrait de la situation telle qu'elle nous apparaît suite à la compilation et du classement des variables qualitatives ou quantitatives obtenues. Par exemple, on indiquera les caractéristiques d'un groupe, on établira les liens statistiques ou fonctionnels entre les composantes étudiées, on fera ressortir la valeur des variables significatives, etc.

L'analyse descriptive peut se subdiviser en trois niveaux : l'analyse descriptive univariée, l'analyse descriptive bi-dimensionnelle et l'analyse descriptive multidimensionnelle.

1.2.1.1. Analyse descriptive univariée

Ce type d'analyse s'intéresse aux variables prises de façon isolée. On analysera par exemple, le sexe du manager de l'entreprise ou encore le chiffre d'affaires des entreprises. On ne s'intéresse qu'à une seule variable sans recourir à d'autres. L'analyse descriptive univariée résume l'information en produisant des caractéristiques de tendance centrale, de dispersion et de forme. Le choix des outils dépend de la nature de la variable examinée. Ces indicateurs synthétiques peuvent être numériques ou visuels (graphique). Le Tableau 10.2 propose les outils adaptés en fonction de la nature des variables. En outre, le tableau propose quelques présentations appropriées des résultats en fonction de la nature des variables.

Tableau 10.23: Nature de variables et outils potentiels

Type de variables	Outils	Présentation
Variable qualitative	Fréquence absolue ou relative Mode	Tableau, diagramme en bâtons, camembert
Variable quantitative discrète	<ul style="list-style-type: none"> - Caractéristiques de tendance centrale : moyenne, médiane, Mode (pour les variables discrètes), ... - Caractéristiques de dispersion : variance, écart type, ... - Caractéristiques de forme : Skewness, Kurtosis, ... 	Tableau de synthèse des principaux indicateurs Histogramme ou densité
Variable quantitative continue		

Exemple de statistique descriptive

Sur la base des données des entreprises sénégalaises⁴, nous produisons la répartition des entreprises suivant leur taille et des informations (moyenne, écart-type, minimum et maximum) sur la part du capital détenu par différentes catégories d'investisseurs (Tableau 10.3). Il ressort que 89% des entreprises sont soit de petites ou moyennes entreprises, c'est-à-dire qu'elles emploient entre 5 et 20 employés. Les entreprises sont majoritairement détenues par les investissements privés nationaux. En moyenne, les investisseurs domestiques détiennent 87% du capital des entreprises.

Tableau 10.3 4: Quelques éléments descriptifs

Représentation des entreprises en fonction de leur taille (nombre d'employés)			Part (en %) du capital de l'entreprise détenue par					
	Effectif	Fréquence (%)	Obs.	Moyenne	Écart-type	Min	Max	
Micro (<5)	16	2,66	Actionnaire majoritaire	587	96,79	11,67	15	100
Petite (5 à 19)	406	67,55	Investisseur privé national	599	86,98	31,37	0	100
Moyenne (20 à 99)	131	21,80	Investisseur étranger	599	8,66	26,16	0	100
Grande (100+)	48	7,99	Etat	599	0,63	6,21	0	99
Total	601	100						

1.2.1.2. Analyse descriptive bivariée

Contrairement à l'analyse descriptive univariée, l'analyse descriptive bivariée porte sur deux variables à la fois. Par exemple, on peut s'intéresser à la capacité de création d'emplois en fonction de l'accès au crédit des entreprises. Il s'agit plus précisément de présenter des répartitions croisées et surtout d'analyser les liaisons éventuelles entre les variables considérées. Par exemple, on peut vérifier s'il y a un lien entre le nombre d'emplois créés par les entreprises selon que ces dernières aient accès au crédit ou non. Les outils utilisés pour cette analyse dépendent de la nature des variables en jeu. Le **Tableau 10.4** propose un ensemble d'outils en fonction de la nature des variables mises en jeu. La plupart des statistiques peuvent être facilement calculées à partir d'un logiciel comme Stata.

⁴ Les données sont accessibles sur le site <https://www.enterprisesurveys.org/en/enterprisesurveys>.

Tableau 10.4 : Outils d'analyse en fonction de la nature des variables

Nature des variables	Outils
Deux variables nominales	On peut faire un tableau croisé entre les deux tables et présenter les fréquences ou des profils (fréquence en colonne ou en ligne) On peut aussi se servir du test du khi deux et du V de Cramer pour déterminer l'intensité de la liaison entre les deux variables. Plus la valeur du khi deux est élevée plus le lien entre les deux variables est élevée. Pour les petits échantillons (par exemple, lorsque le nombre d'observations pour une cellule est plus petit que 5), on peut recourir à la statistique exacte de Fisher.
Deux variables ordinales	Calculer la corrélation de rang de Spearman, le τ (tau) de Kendall ou le gamma de Goodman et Kruskal pour analyser l'intensité de la liaison entre les deux variables
Deux variables quantitatives	On effectue généralement un nuage de points pour apprécier s'il se dégage un lien entre ces variables. L'intensité de la liaison est mesurée par le coefficient de corrélation de Pearson ou à l'aide du coefficient de détermination calculée après la régression d'une variable sur une autre.
Une variable quantitative et une variable qualitative	On peut se servir du rapport de corrélation pour évaluer l'intensité de la liaison entre les deux variables Lorsque la variable qualitative est binaire, on peut effectuer des tests de comparaison de moyenne ou de proportion entre les deux groupes. Dans le cas où la variable qualitative est multinomiale, on peut se servir de l'analyse de la variance (ANOVA)

Exemple d'analyse entre deux variables qualitatives

Existe-t-il une relation entre l'accès au crédit et la taille des entreprises au Sénégal ?

Tout d'abord, nous donnons la répartition des entreprises en fonction de la taille et de l'accès au crédit. Dans cette analyse, l'accès au crédit est assimilé à la disponibilité d'une ligne de crédit auprès d'un établissement de crédit. Le Tableau 10.5 donne la répartition en fonction des deux variables. Nous avons exclu les entreprises qui n'ont pas répondu à cette question. On peut constater que la quasi-totalité des micro-entreprises n'ont pas de ligne de crédit tandis que la majorité des grandes entreprises ont accès à au moins une ligne de crédit. Nous examinons en détail la relation entre

les deux variables en calculant la statistique de khi-deux. L'hypothèse nulle de la statistique de khi-deux est *une indépendance (pas de liaison) entre les deux variables*. Le calcul donne une statique de 51,769 et on a calculé la probabilité de rejet de l'hypothèse (p-value) qui est 0,000. De ce fait, on rejette l'hypothèse nulle au seuil de 1%. En d'autres termes, il existe un lien étroit entre les deux. Pour apprécier le degré de la liaison, on a recours au coefficient V de Cramer. Le V de Cramer est compris entre 0 et 1. Lorsqu'il est plus petit que 0,1, la relation est **très faible**. La liaison est dite **faible**, lorsque le V de Cramer est compris entre 0,1 et 0,2. La liaison est dite **modérée** lorsque le V est compris entre 0,2 et 0,4. Entre 0,4 et 0,8, la liaison est **forte** et elle est **très forte (parfaite)** au-delà (égal à 1). Le calcul du V donne 0,34 ; ce qui permet de conclure à l'existence d'une liaison modérée ou moyenne entre la taille de l'entreprise et l'accès au crédit. Cela signifie qu'en général les entreprises de grandes tailles ont une probabilité plus grande d'avoir accès au crédit en comparaison aux plus petites.

Tableau10.5 : Répartition des entreprises suivant leur taille et l'accès au crédit

L'entreprise a-t-elle une ligne de crédit ?			
	Oui	Non	Total
Micro (<5)	1	14	15
Petite (5 à 19)	47	338	385
Moyenne (20 à 99)	25	99	124
Grande (100+)	28	19	47
Total	101	470	571

1.2.1.3. Analyse multidimensionnelle

L'analyse multidimensionnelle est impérative lorsque le résultat du projet qui fait l'objet de l'évaluation est mesuré sur plusieurs dimensions (incidence sur le niveau de vie, création d'activité génératrice de revenus, insertion sociale, etc.). Traditionnellement, on distingue, entre autres, l'Analyse en Composantes Principales (ACP), l'Analyse

des Correspondances Multiples (ACM) ou plus généralement l'analyse factorielle. Le principe de l'analyse factorielle est de réduire un nombre « important » de variables observées en un « petit » nombre de variables inobservées appelées *facteurs*. Les facteurs sont en fait construits comme des combinaisons linéaires des variables observées. On peut se servir des facteurs pour construire des indices synthétiques ou effectuer des classifications, c'est-à-dire un regroupement de la population de départ en classes homogènes. L'analyse factorielle et ces différentes composantes sont utilisées dans la *fouille des données* ou *data mining*. Le Tableau 10.6 propose une description sommaire de quelques techniques utilisées dans l'analyse multidimensionnelle des données.

Tableau 10.6 : Quelques méthodes d'analyse multidimensionnelle

Technique	Description
Analyse factorielle	L'analyse factorielle est une technique statistique utilisée pour expliquer la variabilité parmi les variables aléatoires observées en termes de moins de variables aléatoires non observées appelées facteurs. Les variables observées sont alors des combinaisons linéaires des facteurs plus les termes d'erreur.
Analyse en composantes principales	Elle peut être utilisée pour réduire le nombre de variables ou pour en savoir plus sur la structure sous-jacente des données. Elle porte généralement sur des données quantitatives.
Analyse des correspondances	L'analyse des correspondances donne une représentation géométrique des lignes et des colonnes d'une table de fréquences bidirectionnelle. La représentation géométrique est utile pour comprendre les similitudes entre les catégories de variables et les associations entre variables.
Analyse des correspondances multiples	L'analyse de correspondance multiple (MCA) est une méthode d'analyse des observations sur des variables catégorielles. Elle est généralement considérée comme une extension de l'analyse des correspondances.
Classification	La classification consiste à allouer ou à classer des observations à des groupes dans le cadre d'une analyse discriminante.

1.2.2. Analyse explicative

L'analyse descriptive permet de décrire les informations contenues dans les données. L'analyse explicative, quant à elle, permet d'examiner le lien causal entre les indicateurs de résultats et un ensemble de caractéristiques de la population. Il peut permettre de tester si l'impact du projet ou programme est différencié selon une caractéristique donnée (par exemple le sexe) et selon les spécificités de l'intervention d'un sujet à l'autre. Dans sa mise en œuvre, on distingue deux types de variables à savoir la variable dépendante et la (les) variable(s) explicative(s) ou indépendante(s). Pour mieux comprendre, reprenons notre exemple sur le crédit et la taille de l'entreprise. La question ici sera: quels sont les facteurs qui expliquent l'accès au crédit des entreprises. On dispose d'un ensemble de facteurs, dont la taille de l'entreprise, le sexe du manager, le nombre d'années d'existence de l'entreprise (expérience), le nombre d'années d'expérience du manager dans le domaine d'activité, l'activité principale de l'entreprise.

La variable **dépendante** est celle que l'on veut expliquer. Dans notre cas, c'est l'accès au crédit.

Les variables **indépendantes** sont tous les facteurs qui sont susceptibles d'expliquer l'accès au crédit des entreprises. Dans notre exemple, il s'agit de la taille de l'entreprise, le sexe du manager, le nombre d'années d'existence de l'entreprise (expérience), le nombre d'années d'expérience du manager dans le domaine d'activité, l'activité principale de l'entreprise.

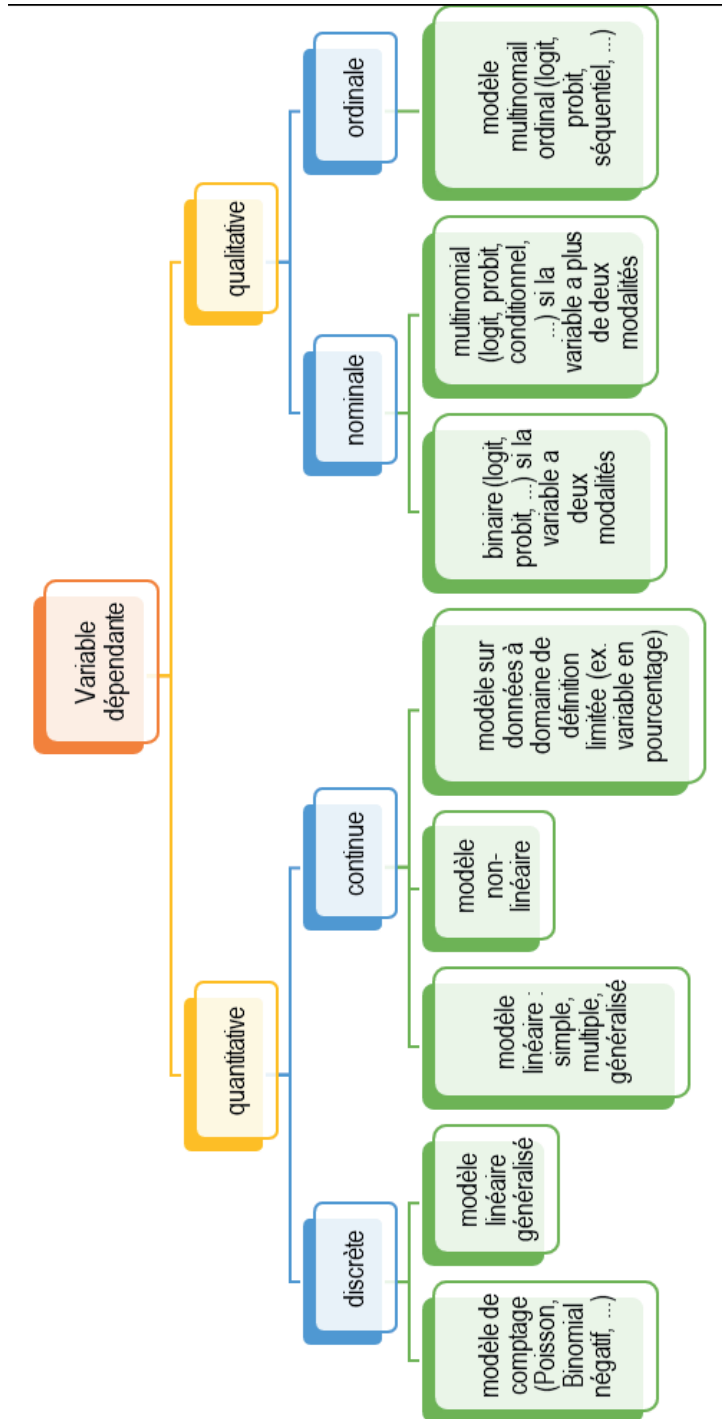
Le principe de l'analyse explicative est d'estimer une équation de la forme :

$$y = f(X, \beta) + \varepsilon$$

où y est la variable dépendante, X est l'ensemble des variables explicatives, f la fonction dite de lien qui permet d'établir la liaison entre la variable dépendante et les variables explicatives. Le choix de la fonction de lien dépend de la nature de y . β est le vecteur des paramètres estimés. Les paramètres indiquent la sensibilité de la variable dépendante aux variations de chacune des variables explicatives. Par exemple de combien augmentent les chances moyennes pour une entreprise d'avoir accès au crédit si l'entreprise a une année d'existence en plus. ε est le terme d'erreurs qui traduit l'imparfaite compréhension de l'analyste du phénomène étudié ou encore la part de la variable dépendante non expliquée par nos variables indépendantes.

L'analyse explicative est généralement basée sur le choix de la fonction de lien. La Figure 10.1 fait une présentation de potentiels modèles en fonction de la nature de la variable dépendante.

Figure 10.1 : Les modèles potentiels en fonction de la nature des variables



Pour apprécier la qualité de la régression, on utilise généralement le coefficient de détermination (R^2). Ce coefficient mesure la part de la variance de y expliquée par les facteurs X . Plus ce facteur tend vers 1, plus la régression est de bonne qualité⁵. Pour d'autres modèles, notamment ceux appliqués aux variables dépendantes qualitatives, le coefficient de détermination n'a pas d'interprétation directe. Par conséquent, d'autres indicateurs sont utilisés pour apprécier la qualité de la régression comme les critères d'information, la log-vraisemblance, ... Il est aussi possible d'utiliser un autre coefficient de détermination appelé le pseudo R^2 de McFadden. Ce dernier correspond au R^2 classique du modèle. La qualité de la régression est bonne si le pseudo R^2 de McFadden est compris en 0,2 et 0,4 McFadden (1977, p.35).

5 Attention, ce n'est pas une panacée. On peut bien avoir des R^2 faibles dans certaines régressions. C'est souvent le cas lorsque la population est très hétérogène. D'ailleurs, lorsque le R^2 est très proche de l'unité, on doute de la véracité des données utilisées dans l'analyse.

II. Analyse et interprétation des données qualitatives

L'analyse qualitative est une recherche par immersion intense ou prolongée dans une situation normale (la vie réelle) en adoptant une perspective contextuelle, holistique ou systémique. La démarche qualitative permet à l'évaluateur de suspendre ses préjugés, l'idée étant d'observer et de tirer des conclusions de l'observation. Elle permet de faire un tri à travers plusieurs interprétations possibles pour trouver celles qui sont convaincantes et elle est basée sur la théorie et la cohérence interne des informations collectées. En outre, l'approche qualitative permet de mettre l'accent sur des modèles, des thèmes et des contrastes qui émergent, au lieu des normes établies à l'avance. Des hypothèses peuvent être a priori formulées ou non dans la démarche qualitative.

Il existe plusieurs façons de conduire l'analyse qualitative qui sont présentées ci-dessous.

2.1. Compilation et traitement des données qualitatives

Le processus de compilation des données qualitatives dépend fortement du mode de collecte adopté. Pour des documents, on peut directement les inclure dans la liste des sources de l'analyse. Pour les entretiens et groupes de discussion, si l'information a été uniquement enregistrée lors de la collecte, il faut la transcrire avant son interprétation, souvent mot à mot, on réalise alors un *verbatim*. Comme pour la saisie des données quantitatives, cette transcription génère des coûts à prévoir dans le budget de la collecte. Si l'information a été enregistrée avec une prise de note simultanée, il faut souvent compléter les notes avec l'enregistrement. Enfin, si l'information n'a été collectée qu'avec une prise de notes, l'analyse portera sur les notes (on s'éloigne alors de la transcription mot à mot pour se baser sur une synthèse, les notes doivent alors être prises au plus près de ce qui a été dit, pour éviter un biais apporté uniquement par la personne qui fait la synthèse). Selon le type d'analyse qualitative que l'on veut effectuer et l'ampleur des données qualitatives dans la stratégie globale de l'évaluation, une synthèse des propos des répondants peut suffire, si toutefois on veut procéder à des analyses avancées et que l'essentiel de l'argumentaire évaluatif repose sur la valorisation de données qualitatives, le *verbatim* s'impose.

2.2. L'analyse inductive vs l'analyse déductive

On distingue généralement entre deux procédures pour réaliser l'analyse qualitative : l'analyse inductive où l'on laisse les sources déterminer les thèmes et l'interprétation, et l'analyse déductive où on analyse les données à partir d'un cadre théorique existant, ce qui est souvent le cas en évaluation avec la théorie du changement de l'intervention.

L'analyse inductive

- Les résultats de recherche **émergent** des thèmes fréquents, dominants ou significatifs trouvés dans les données brutes. Les résultats ne sont pas contraints par des méthodologies structurées, modèles, cadres, etc

L'analyse déductive

- Analyser les données en fonction d'un cadre existant (par exemple, le modèle logique, des recherches antérieures, etc)

26

2.2.1. L'analyse inductive - Processus général

L'analyse inductive consiste à passer en revue les données qualitatives attentivement et entièrement pour identifier les thèmes et catégories qui s'y trouvent, sans rechercher un thème en particulier qui provient de la matrice d'évaluation. Il s'agit donc de la recherche des grandes tendances qui se dégagent, sans prendre en compte les préconçus de l'évaluateur. Pour chaque thème, appelée aussi catégorie, on identifie les déclarations (ou phrases) qui vont l'étayer. Une fois les catégories retenues, on cherche à créer un modèle basé sur ces derniers et les liens entre eux. Cette procédure reste moins utilisée en évaluation, car l'évaluateur démarre souvent avec des hypothèses, et des éléments de sa théorie de changement qu'il souhaite étayer. Il peut toutefois rester ouvert à l'émergence de catégories non prévues, et adopter donc une dose d'analyse inductive dans son travail.

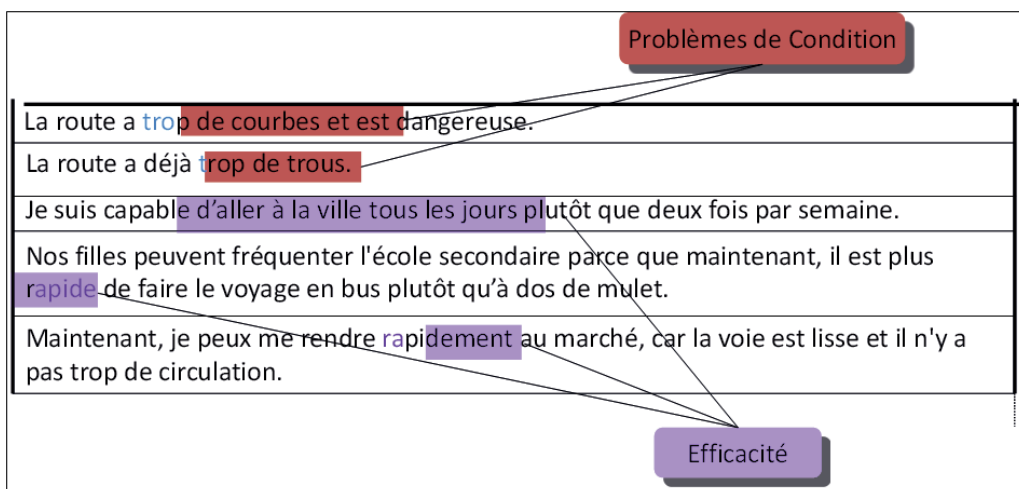
2.2.2. L'analyse déductive - Processus général

Contrairement à l'analyse inductive dont le point de départ est les données collectées, l'approche déductive commence par une révision du cadre du projet. Elle identifie les catégories à partir du cadre théorique de l'évaluation avant de procéder au dépouillement des données collectées. Les catégories sont donc issues du cadre logique de l'intervention. Sur cette base, l'évaluateur passera en revue attentivement et entièrement les données qualitatives en vue de dégager les extraits avec les catégories listées auparavant.

2.3. Analyse de contenu

Quelle que soit l'approche inductive ou déductive, l'analyse qualitative renvoie au contenu des données collectées. Il s'agit de dégager des tendances des données et donc de synthétiser le grand volume de données récoltées. La Figure 10.2 donne une illustration d'une analyse de contenu portant sur la qualité d'une nouvelle route construite.

Figure 10.2 : Analyse de contenu sur la qualité d'une nouvelle route



2.3.1. Codage du Contenu

Pour procéder à une analyse de contenu, on utilise généralement une procédure que l'on appelle codage ou étiquetage. Il s'agit de repérer dans les textes (à la manière d'un surlignage) les extraits qui font référence à chacune des catégories (qu'elles soient créées de manière inductive ou déductive). Si plusieurs personnes dans l'équipe d'évaluation doivent réaliser le codage, il est nécessaire de se doter d'une

liste détaillée de catégories, avec des définitions claires, et de procéder à un contrôle qualité du codage. Une illustration de cette procédure est fournie par la Figure 10.3.

Figure 10.3: Exemple de codification de Contenu

Exemple de codes et catégories Jaune = les revenus ont augmenté Vert = l'accès aux marchés et aux clients est amélioré Rose = amélioration de l'emploi	Bleu = accès à une meilleure éducation Gris = Contournement du village et moins de revenus Pourpre = la pollution de l'air Rouge = les problèmes de circulation et de sécurité
Je vends mes produits dans plus de marchés. Cela me permet de gagner plus d'argent chaque jour.	
Ma fille peut fréquenter l'école professionnelle en ville parce que le service de bus est disponible maintenant.	
Mon revenu a augmenté parce que j'ai pu trouver un meilleur emploi dans la ville.	
J'aurais souhaité que la route n'ait pas été construite. Nous avons plus de trafic se déplaçant à vitesses élevées	
Il y a beaucoup plus de voitures. La pollution de l'air a affecté la respiration de ma grand-mère.	
Une fois la route terminée, moins de voyageurs s'arrêtent à mon magasin. Ils contournent le village et mon revenu mensuel a chuté de 30 %.	
Nous avons plus d'argent. Avec la nouvelle route, j'ai un deuxième emploi.	
Plus de gens de la communauté peuvent fréquenter les formations professionnelles de la ville en raison du service d'autobus.	
Ma femme a réussi à obtenir un emploi à temps et notre revenu familial a augmenté.	
Il y a plus de pollution de l'air, mais l'accès aux marchés a aidé mon entreprise à croître, augmentant nos revenus.	

2.3.2. Matrice de codage

La codification finale des données peut être consignée dans un tableau Excel comme illustré par le Tableau 10.7 où les répondants sont en ligne et les catégories ou code sont en colonne. Pour des petites analyses qualitatives, cette procédure peut parfois être suffisante. Dans d'autre cas, on va souhaiter systématiser cette procédure à l'aide de logiciels spécialisés.

Tableau 5: Exemple de matrice de codifications

Répondant - numero de ID	L'accès au marché amélioré	Augment ation du revenu	L'accès à l'école secondaire amélioré	Amélioration de l'emploi	Plus de trafic	Accident de circulation	Pollution de l'air	Contournant le village	L'accès aux services de santé amélioré
C128	1	1							
K245			1						
M358		1		1					
Etc...
Totale	20	26	15	18	11	5	10	8	16

2.4. Exemple d'un logiciel d'analyse qualitative

Il existe plusieurs logiciels d'analyse de données qualitatives dont NVivo, MAXQDA, ATLAS.ti. En marge de ces logiciels, il existe des logiciels de transcription et d'annotation de la voix comme TRJS, Elan, Transcriber et CLAN. Ces logiciels permettent de systématiser l'opération de codage, en chargeant l'ensemble des sources de données qualitatives (documents, transcriptions d'entrevues et même fichiers audio et vidéo). On rentre ensuite la liste des catégories et on peut ensuite procéder au codage à l'intérieur de l'interface du logiciel. Le logiciel offre ensuite un certain nombre d'analyses manuelles, semi-automatiques voire automatiques, qui permettent de tirer le maximum des données qualitatives.

III. Rédaction des fiches d'analyse

La fiche d'analyse est le produit final de l'étape d'analyse des données qui permet de compiler les informations qui ressortent de chaque collecte ou source de donnée. Elle doit être écrite en ayant en vue le style de rédaction du rapport, en incluant notamment des tableaux et figure et en proposant à chaque fois une interprétation des données pour faire le lien entre l'analyse réalisée et la question d'évaluation à laquelle elle permet de répondre. C'est un document qui a un usage interne et qui permet de rassembler toutes les briques d'informations qui vont constituer l'argumentaire évaluatif dans le rapport. Les fiches d'analyse, bien que non-obligatoires, sont très utiles, notamment lorsque plusieurs personnes doivent travailler en parallèle. Ainsi chaque personne peut s'occuper de l'analyse d'une source, et rassemble dans sa fiche toute l'information qu'elle y trouve.

Messages clés :

- ✓ Le type d'analyse à effectuer dépend des données que l'on a collectées et de l'usage que l'on veut en faire en vue du rapport en fonction de l'approche d'évaluation choisie. Il ne faut pas oublier non plus de considérer le temps de réalisation de chacune des analyses, qui est parfois possible dans le monde de la recherche, mais trop fastidieuse pour un échancier réaliste d'évaluation.
- ✓ Certains gains de temps peuvent être étudiés, comme utiliser des synthèses d'entrevue plutôt que des transcriptions mot à mot ; ou s'en tenir à quelques analyses quantitatives univariées faites sur Excel ou directement dans un logiciel de sondage. Toutefois, la crédibilité du jugement évaluatif repose sur un minimum de rigueur méthodologique que l'équipe d'évaluation doit adopter.
- ✓ Plus une nature de donnée est importante dans l'argumentaire que l'on va construire, plus les analyses vont être poussées. Par exemple si on adopte une approche mixte, mais majoritairement qualitative, alors on pourra s'en tenir à quelques analyses quantitatives succinctes, mais pousser les analyses qualitatives plus loin.

Bibliographie :

- Gauthier, B. et Bourgeois, I. (2016) Recherche sociale. De la problématique à la collecte des données, 6^eéd., Québec, Presses de l'Université du Québec, 670p. Chapitres: 7 et 8
- Saporta, G. (2011). Probabilités, analyse des données et statistique. Editions Technip.
- StataCorp (2019). Stata: Release 16. Statistical Software. College Station, TX: StataCorp LLC.
- Van Campenhoudt, L. et Quivy, R. (2011). Manuel de recherche en sciences sociales, 4^eéd., Dunod, 262p.

MODULE 11

COMMUNICATION DANS LE PROCESSUS DE SUIVI ET D'ÉVALUATION

Objectifs d'apprentissage :

A la fin de la session, les participants seront capables de :

- Décrire l'intérêt de la communication avant, pendant et après la conduite du suivi et de l'évaluation
- Décrire les enjeux de la communication dans la présentation des résultats de l'évaluation
- Présenter les stratégies et outils de communication
- Adapter la communication au public cible
- Elaborer un bon rapport d'évaluation
- Présenter les résultats de l'évaluation
- Distinguer les approches quantitatives/qualitatives
- Elaborer un plan de collecte de données

Thèmes couverts

- I. Communiquer lors du suivi des activités de développement
- II. Communication avant, pendant et après l'évaluation
- III. Quatre étapes pour communiquer et rendre compte dans le processus d'évaluation
- IV. Stratégies de communication
- V. Présentation des rapports d'évaluation
- VI. Exposés oraux

La communication est une composante essentielle pour la gestion d'une intervention de développement. Une étude menée par l'Institut de Management des Projets en mars 2013 confirme d'ailleurs que la communication est un facteur de succès dans une intervention. D'après cette étude, les organisations qui communiquent efficacement obtiennent de meilleures performances avec leurs projets. Par exemple, 80% atteignent les objectifs initiaux (contre 52% sans communication) ; 71% respectent les délais (contre 37% sans communication) et 76% respectent le budget (contre 48% sans communication).

Nombreux sont les évaluateurs qui se plaignent du fait que les rapports qu'ils produisent ne soient pas suffisamment lus ou partagés. En outre, dans certains cas, les recommandations du rapport ne peuvent être utilisées, ce qui est bien dommage, lorsque l'on pense au temps passé à écrire ces rapports. Parmi les raisons qui expliquent ces réalités, les commanditaires indiquent souvent qu'ils ont du mal à comprendre les résultats, présentés dans des rapports touffus et complexes. L'évaluation n'est souvent que résumée dans un rapport et d'autres formes de communication sont délaissées par manque de temps ou par ignorance des possibilités existantes. On comprend dès lors que la communication occupe une place de choix dans le processus d'évaluation et est réellement déterminante pour que l'évaluation remplisse bien son rôle dans la gestion axée sur les résultats. C'est pour cette raison que ce chapitre se concentre sur les stratégies de communication avant et pendant le processus d'évaluation.

Mot-clé : communication, rapport d'évaluation, présentation des résultats

I. Communication et rapportage pendant le cycle de vie d'une intervention

Tout comme pour la gestion d'une intervention, on peut utiliser des outils et des méthodes appropriées, les actions de communication peuvent être structurées selon les phases du cycle de vie d'une intervention.

Les systèmes de suivi sont un véritable outil de communication pour garantir l'efficacité d'une intervention. L'atteinte des objectifs de cette dernière nécessite de répondre à de nombreuses préoccupations dont quelques-unes sont les suivantes :

- *Le projet se déroule-t-il conformément au plan de travail initial ?*
- *Les coûts réels sont-ils inférieurs aux coûts budgétisés ?*
- *Le projet se réalise-t-il correctement ou est-il en retard ?*
- *Le personnel affecté au projet est-il performant ?*

Pour répondre à ces questions, le mécanisme de suivi s'appuie essentiellement sur l'analyse et le rapportage à intervalles réguliers de différents indicateurs de mesure de performances. Ces rapports rendent compte de l'état d'avancement des projets mis en œuvre sur la base des constats et des difficultés rencontrées sur le terrain. Ils servent de base aux parties prenantes pour la proposition d'avenants ou de changements nécessaires pour la meilleure gestion de l'intervention et l'atteinte des résultats escomptés. Les conclusions des appréciations des rapports d'avancement sont également utiles pour les missions d'évaluation à mi-parcours et finales.

Des rapports réguliers de suivi des interventions seront produits par les acteurs du système pour informer les principales parties prenantes des performances réalisées sur le terrain et des résultats intermédiaires obtenus. Ces rapports sont de diverses formes selon le niveau d'exécution de l'intervention, et selon les pratiques des organisations. Le Tableau 11.1 présente selon l'état d'avancement d'un projet (nouveau, en cours, fin) quelques types de rapports qui pourraient être produits par l'équipe responsable du suivi et de l'évaluation, leur périodicité et les destinataires. Par exemple, pour les nouveaux projets, il est recommandé de produire un rapport de démarrage ou plan de mise en œuvre, un rapport d'évaluation ex-ante peut également

être demandé si nécessaire. Dans notre exemple, après le lancement du projet un rapport d'avancement devra être produit à une fréquence trimestrielle. Néanmoins, cette fréquence peut différer en fonction de la durée et des exigences du projet et de son maître d'œuvre. Certains bailleurs peuvent exiger des rapports mensuels ou alors uniquement des rapports semestriels ou annuels. La fixation des délais de production des rapports doit être réaliste en fonction des exigences du projet et des besoins d'information des bailleurs ou des parties prenantes.

Tableau 6: Rapports de suivi et évaluation

Projets	Types de rapports	Périodicité	Destinataires
Nouveau projet	Rapport d'évaluation ex ante	A l'issue de l'analyse de la fiche technique et du document de projet	Initiateur du projet
	Rapport de démarrage	Trois mois après le lancement des activités	Comité de pilotage Maitre d'ouvrage Autres parties prenantes
Projet en cours	Rapports d'avancement du projet	Trimestrielle	Comité de pilotage Maitre d'ouvrage Autres parties prenantes
	Rapport d'évaluation à mi-parcours	A mi-parcours du projet	Comité de pilotage Maitre d'ouvrage Autres parties prenantes
Projet en clôture	Rapport d'évaluation finale	A la fin du projet	Comité de pilotage Maitre d'ouvrage Autres parties prenantes

II. Communication avant et pendant les travaux d'évaluation

La communication joue aussi un rôle crucial dans l'utilisation qui va être faite par le demandeur de l'évaluation de résultats de celle-ci. Le concept dit d'*évaluation axée sur l'utilisation* commence par le principe que les évaluations doivent être jugées en fonction de leur utilité et de leur réel usage. Ainsi, les auteurs qui défendent ce point de vue expliquent que les évaluateurs devraient concevoir et exécuter l'évaluation en tenant soigneusement compte de tout ce qui peut affecter l'utilisation des résultats finaux (Patton, 2003). L'expérience a notamment démontré que l'implication active des parties prenantes à toutes les phases d'une évaluation permet une meilleure appropriation et utilisation de cette dernière (Patton, 1997).

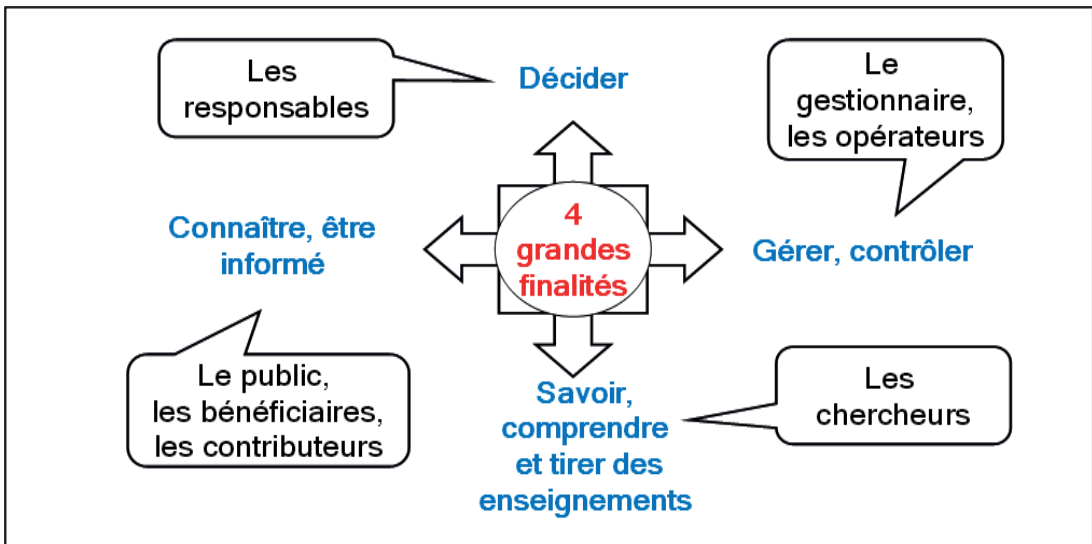
2.1. Finalités de la communication dans le processus d'évaluation

La Figure 11.1 présente les quatre finalités de la communication dans le processus d'évaluation. Communiquer sur un projet permet entre autres (i) de créer un soutien auprès des diverses parties prenantes ; (ii) de recenser les besoins, de recueillir des avis pour adapter l'intervention si nécessaire ; (iii) de donner de la visibilité à l'intervention et contribuer ainsi à sa viabilité ; et (iv) d'accompagner le changement inhérent à toute intervention.

Ces différentes finalités éclairent sur les principales cibles de la communication. En effet, la communication doit être destinée aux responsables ou décideurs pour faciliter ou aider à la prise de décisions (décider). Elle doit aussi apporter des informations utiles aux gestionnaires ou opérateurs de l'intervention dans la gestion quotidienne de celle-ci. De plus, la communication peut également cibler des chercheurs ou groupes d'intérêt sur le sujet qui pourraient se servir de ces résultats dans un contexte plus large. Enfin, la communication doit être destinée au public, aux bénéficiaires et contributeurs du projet, puisqu'elle constitue également un outil de transparence dans la gestion publique. En ce sens, les citoyens ont le droit de connaître les résultats des interventions publiques qui leur sont destinés.

Il faut noter que la stratégie de communication varie en fonction du stade d'exécution de l'intervention ou de celle de l'évaluation. Par exemple, une évaluation ex-ante est avant tout intéressante pour les possibles initiateurs d'un projet. La communication qui entoure cette évaluation ex-ante peut alors être plus limitée. A l'opposé, les résultats d'une évaluation ex-post d'une politique publique d'envergure peuvent être intéressants pour un public plus large, y compris le gestionnaire du projet, les bénéficiaires, les chercheurs, et les décideurs.

Figure 11.1: Finalités de la communication dans le processus d'évaluation



2.2. Communication en amont de l'évaluation

La communication avant l'évaluation est surtout constituée d'un dialogue entre un commanditaire et l'évaluateur. A ce stade, la communication porte sur tout le processus de planification de l'évaluation, y compris les TDR, la contractualisation et les attentes de gestion par rapport aux travaux d'évaluation. Ce dialogue peut-être très codifié dans le cadre d'une passation de marché pour embaucher un évaluateur externe, ou plus simplifié dans le cadre d'une évaluation interne.

Généralement à l'invitation du commanditaire pour lancer les travaux, on effectue une réunion de cadrage avec l'équipe en charge de l'évaluation, et possiblement différentes parties prenantes si l'on met en place un comité de suivi des travaux

multipartite ; pour s'accorder sur la méthodologie à utiliser, l'échantillon, les outils de collecte et d'analyse, le calendrier de l'évaluation.

Outre le commanditaire, l'évaluateur devra très souvent échanger avec différentes parties prenantes, et faire une première analyse documentaire, pour diagnostiquer la situation, son évaluabilité, et trouver les angles d'attaque de l'évaluation. Ces échanges permettront de définir la démarche à suivre pour effectuer l'évaluation, élaborer les outils de l'évaluation et proposer un calendrier indicatif. La méthodologie et les outils seront habituellement discutés et validés par le commanditaire voire le comité de suivi des travaux si tel est le cas avant de procéder à la phase de collecte de données de l'évaluation. Il est important de mener un dialogue ouvert et d'être clair sur les attentes, et sur ce qui est réaliste compte tenu du budget et de l'échéancier.

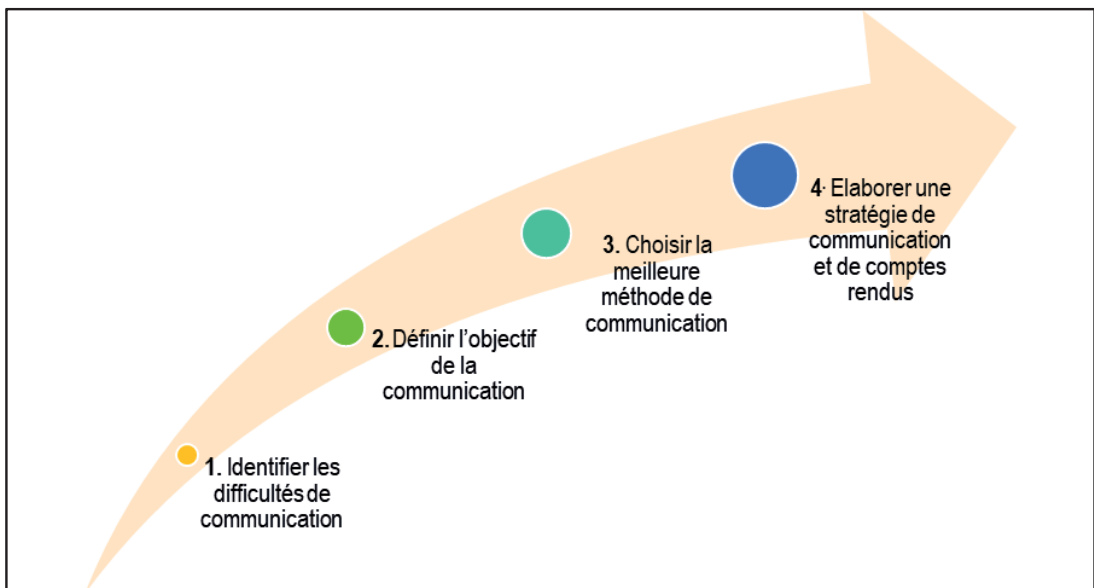
2.3. Communication pendant l'évaluation

Pendant le processus d'évaluation, la communication est tout aussi importante. Elle vise principalement à fournir des informations sur le déroulement global des opérations. Tout en faisant un point sur l'avancement de son travail, l'évaluateur partagera, en particulier, les difficultés rencontrées au cours de l'évaluation avec les parties prenantes afin de trouver des solutions rapides aux éventuels blocages. C'est aussi l'occasion d'impliquer des parties prenantes dans le processus d'évaluation. A ce stade, l'instrument de communication le plus courant est un aide-mémoire envoyé périodiquement, mais elle peut prendre des formes très diverses comme des appels téléphoniques de suivi avec un point focal, des courriels hebdomadaires, ou encore des échanges réguliers sur Whatsapp. Sans communiquer des résultats avant qu'ils ne soient prêts et validés, le fait de garder le lien avec le commanditaire permet d'établir un lien de confiance, et facilitera l'appropriation des résultats. Le commanditaire n'aura pas l'impression d'embaucher l'évaluateur et de ne pas entendre parler de lui avant l'envoi d'un rapport provisoire plusieurs semaines après.

III. Communiquer et rendre compte des résultats de l'évaluation

Le processus pour bien communiquer peut se décliner en quatre étapes résumées par la Figure 11.2, à savoir (i) l'identification des difficultés de communication, (ii) la définition de l'objectif de la communication, (iii) le choix de la meilleure méthode de communication, et (iv) l'élaboration d'une stratégie de communication.

Figure 11.2 : Quatre étapes pour communiquer efficacement



La première étape consiste à identifier les difficultés de communication. Selon Torres et al. (2005) trois obstacles ou difficultés peuvent être rencontrés dans la communication. Tout d'abord, il existe une *anxiété vis-à-vis de l'évaluation*. Très souvent, la confusion entre « audit »⁶ et « évaluation » crée une anxiété qui cause une résistance du fait que les résultats peuvent avoir une influence sur les emplois ou

⁶ L'audit est une procédure consistant à s'assurer du caractère complet, sincère et régulier des comptes d'une entreprise, à s'en porter garant auprès des divers partenaires intéressés de la firme et, plus généralement, à porter un jugement sur la qualité et la rigueur de sa gestion (définition du Larousse).

l'allocation des ressources. Les évaluateurs externes, qui ont besoin de temps pour établir des relations de confiance peuvent augmenter l'anxiété déjà existante. Ensuite, le *manque de planification au début de l'évaluation* peut constituer un frein à la communication. En effet, si on ne communique pas régulièrement avec les parties, cela peut les amener à se désengager, se désintéresser et, en fin de compte, à ne pas utiliser les résultats. C'est la raison pour laquelle, l'évaluateur doit maintenir un contact permanent avec les parties prenantes tout en restant indépendant dans l'élaboration de son jugement évaluatif. Par ailleurs, les équipes d'évaluation découvrent souvent trop tard qu'aucun budget n'est prévu pour la production des rapports, pour les présentations orales ou la vulgarisation des résultats de l'évaluation. Ainsi, un évaluateur expérimenté l'inclut dans son budget et aborde ces questions dès le démarrage. Enfin, la communication *peut buter sur une culture organisationnelle peu appropriée*. Par exemple, la circulation des constatations sur la performance peut être freinée par des directeurs qui se sentent mal à l'aise à partager l'information issue de l'évaluation. Ceci est le cas pour des organisations très rigides avec une hiérarchie relativement forte. De plus, dans d'autres organisations le système de partage de l'information est dysfonctionnel, ce qui empêche une communication continue durant l'évaluation. En outre, dans certaines organisations, une évaluation négative peut être considérée par les employés comme une critique honteuse à leur endroit. Ils peuvent donc résister à discuter ouvertement des résultats de l'évaluation. La communication peut être également inefficace du fait de la perte de mémoire institutionnelle en raison d'un renouvellement rapide du personnel ou pour d'autres raisons connexes.

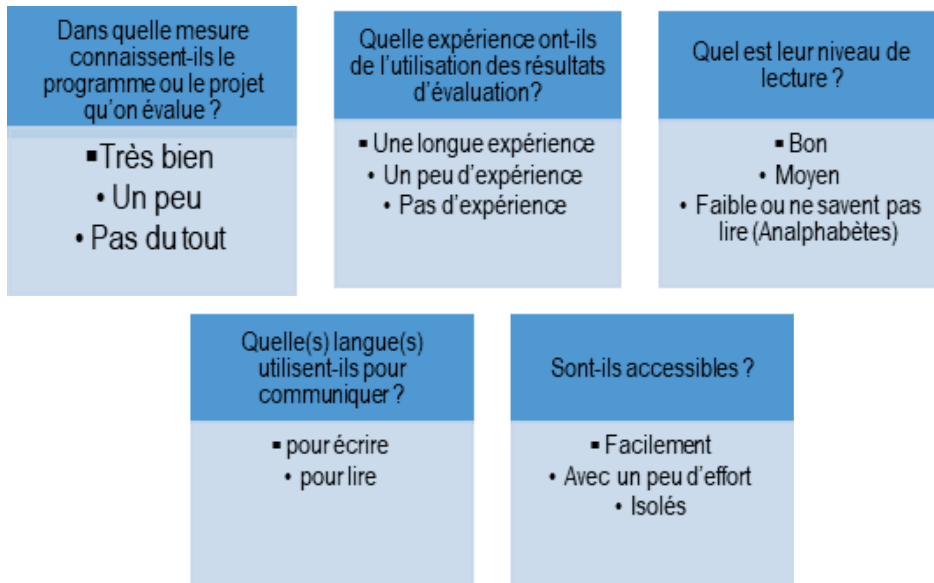
La deuxième étape consiste à définir l'objectif de la communication. Trois catégories de questions sont utiles à ce niveau. Tout d'abord, il faut se demander si chaque individu ou groupe a besoin d'être informé des activités de l'évaluation. Si oui, quand est-ce que cette information est nécessaire et pour quelles raisons ? Cette série de questions permet de cibler les personnes ou groupes à informer ainsi que les raisons pour le faire. En ce qui concerne les raisons, le partage de l'information peut viser une conscientisation des individus ou groupes, une recherche d'appui pour conduire à bien sa mission, ou tout simplement pour montrer du respect aux parties prenantes. Une fois les cibles identifiées, l'on se demande si elles ont besoin de disposer de l'information sur les résultats provisoires ou finaux de l'évaluation. Le partage des résultats provisoires est utile pour voir et suivre la progression de l'évaluation. En

revanche, les résultats finaux sont destinés à apprendre, à s'améliorer ou encore pour promouvoir le dialogue et la compréhension parmi les partenaires. Enfin, l'évaluateur tentera de répondre à la question de savoir si l'individu ou le groupe avec lequel il communique est impliqué dans la prise de décisions. C'est une question essentielle, car elle permet d'évaluer la probabilité d'un appui dans l'avenir, pour aider à élaborer les recommandations ou pour s'assurer que les recommandations seront utilisées. Comme nous l'avions mentionné, les évaluateurs se plaignent que les résultats des évaluations ne sont pas utilisés. Si l'évaluateur a une connaissance des personnes ou groupes d'intérêt à même d'intervenir dans la prise de décision, il peut intensifier la communication avec eux afin de formuler des recommandations utiles pour l'organisation.

La troisième étape consiste à choisir la meilleure méthode de communication pour faciliter les échanges. Les différentes questions pouvant être posées pour un meilleur choix de la méthode de communication sont résumées dans la Figure 11.3. Il faut évaluer les niveaux de connaissance et de langue de son interlocuteur afin d'adapter le message. Lors de l'élaboration du message pour la communication, la connaissance en matière d'utilisation des résultats d'évaluation des interlocuteurs est à considérer avec beaucoup d'intérêt si l'évaluateur veut maximiser l'usage des conclusions et recommandations. Enfin, l'évaluateur doit aussi tenir compte de l'accessibilité des parties prenantes. Afin d'optimiser le processus d'évaluation, il faut être capable de joindre les parties prenantes et avoir un échange avec eux en présentiel si possible, et à distance avec les nouvelles technologies disponibles. La pandémie de Covid-19 a montré l'utilité de nombreux outils de communication à distance, comme Teams, Zoom, WebEx, Skype, Gmeet, etc. La plupart de ces outils ont une version gratuite pouvant être utilisée pour des réunions avec un nombre limité de participants⁷. La seule contrainte liée à leur usage est la qualité de la connexion Internet. Il faut également faire attention aux écarts technologiques entre les utilisateurs surtout si ceux-ci sont de différentes générations.

7 Le nombre maximal de participants varie d'une application à une autre. Par exemple, au plus 100 personnes peuvent assister à une réunion organisée avec la version gratuite de Zoom, tandis que ce nombre atteint 300 personnes avec Teams et même 1 000 personnes avec WebEx.

Figure 11.3 : Questions pour orienter le choix de la méthode de communication



La **quatrième étape** pour une communication efficace est l'élaboration de la stratégie de communication. La stratégie de communication peut être définie comme la manière dont on planifie et coordonne la communication. C'est l'ensemble des décisions déterminant la communication, en l'occurrence la définition des cibles à atteindre, les objectifs des différents types de communication mobilisés, les contraintes budgétaires et temporelles, etc. La Figure 11.4 donne quelques éléments de la stratégie de communication et la section IV donne plus de détails. Pour des questions d'organisation et de cohérence, la section IV intègre les éléments d'explications de la Figure 11.5.

Figure 11.5 : Quelques éléments de la stratégie de communication

Cible
<ul style="list-style-type: none">• Gestionnaire du projet, bénéficiaires, chercheurs• Décideurs, Bailleur de fonds
Que communiquer?
<ul style="list-style-type: none">• Résultats• Recommandations
Quel format?
<ul style="list-style-type: none">• Rapport final d'évaluation avec résumé exécutif• Réunion de bilan avec le bailleur de fonds pour présenter les résultats, les recommandations et les actions prévues
Date
<ul style="list-style-type: none">• Convenir d'une date avec les parties prenantes
Qui préparera et fournira les informations ?
<ul style="list-style-type: none">• L'équipe d'évaluation préparera les rapports écrits.• Désigner une personne (équipe) pour travailler sur l'ordre du jour• Désigner une personne qui travaille sur les présentations
Budget
<ul style="list-style-type: none">• Production et reproduction des rapports• Déplacement (billets d'avion, coût des tickets, ...)• Préparation de la réunion

IV. Stratégie de communication

La Figure 11.6 indique que la première étape de la stratégie de communication est l'identification des parties prenantes ou cibles. L'identification de la cible est importante. Par exemple, le bailleur de fonds a besoin du rapport final d'évaluation pour prendre des décisions sur des financements futurs alors que les chercheurs seraient intéressés par la section « leçons apprises » du projet. Ensuite, il faut s'accorder sur le contenu et le format de la communication. Par exemple, la communication peut porter sur les résultats et les recommandations de l'évaluation. En tout état de cause, un rapport final d'évaluation est requis pour retracer l'ensemble de la procédure d'évaluation et des conclusions. Ce rapport doit comporter un résumé exécutif.

La communication avec les parties identifiées peut se faire sous la forme d'une rencontre physique ou d'une réunion à distance. Dans l'un ou l'autre des cas, il faut convenir d'une date et d'une plage horaire convenables aux différentes parties. Enfin, la stratégie de communication doit inclure le budget pour la conduite de l'ensemble des opérations. Le budget peut comporter les ressources pour la production et la reproduction des rapports, des billets d'avion ou des forfaits internet ...

Cette section revient sur l'ensemble des composantes de la stratégie de communication.

4.1. Objectifs de la communication

La communication dans le processus de l'évaluation vise à partager les résultats de l'évaluation et à orienter ou appuyer la mise en œuvre des recommandations. En effet, l'objectif d'une évaluation est que les résultats puissent éclairer (i) la prise de décisions politiques (ii) les changements de programme et (iii) la reconduite ou la poursuite de programmes. Ainsi, la stratégie de communication doit permettre d'identifier les produits nécessaires pour informer les différentes cibles identifiées. En effet, de nombreux produits peuvent être utilisés pour faciliter la communication. Par exemple, le rapport final d'évaluation incluant le résumé exécutif sera remis au bailleur de fonds alors que des tracts (flyers) sont utiles pour une communication à grande échelle.

Dans la stratégie de communication, il faut effectuer un choix adéquat entre la technique à utiliser, les objectifs visés et les moyens utilisés (voir Tableau 11.2). Par exemple, pour convaincre pour un changement, l'évaluateur doit être persuasif. Il peut utiliser le dialogue tout en montrant les avantages d'une stratégie par rapport à une autre. La technique et les moyens doivent être adaptés s'il s'agit d'informer, de déléguer, de consulter ...

Tableau 11.2 : objectifs de communication dans le processus d'évaluation

Technique	Objectif(s)	Moyens
Persuasion	Convaincre pour le changement	Dialogue, preuve de réussite, démontrer les avantages
Information	Aider à prendre la décision en connaissance de cause	Documentation, visite des sites, mass media
Consultation	Croiser les points de vue	Visite, réunions, sondages
Codécision	Trouver un consensus, la collégialité, anticiper les résistances	Réunions, focus group
Délégation	Partage des responsabilités	

4.2. Communication informelle et formelle

Après la définition des objectifs de la communication, cette section s'intéresse aux formes de communication pouvant être utilisées. Elle se concentre sur les communications informelle et formelle. La communication informelle est une forme d'échange décontractée non officielle entre collaborateurs, sur le lieu de travail. Elle repose sur les relations sociales qui se forment hors de la hiérarchie normale et utilise des canaux comme le téléphone ou autour d'un café. A l'opposé, la communication formelle est encadrée, délimitée ou structurée et renvoie à un ensemble d'informations

programmées relatives au travail et aux obligations organisationnelles. Elle est généralement écrite, voire publiée. La communication formelle peut prendre la forme de briefing, d'exposés ou de communiqués écrits ou oraux.

En pratique, les deux formes de communication se complètent. Les échanges informels peuvent aider à éclaircir, rappeler et confirmer ce qui a été communiqué en utilisant le canal formel. La communication informelle peut être souvent très utile pour la résolution de problèmes complexes ou pour la résolution de conflits. En éliminant la pression, la communication informelle contribue à améliorer l'environnement de travail, car les employés peuvent facilement exprimer leurs plaintes ou points de vue, et les supérieurs hiérarchiques peuvent proposer une solution rapide en conséquence.

Après la définition des formes de communication, nous nous intéressons aux composantes de la stratégie de communication. Ces différentes composantes sont synthétisées par la Figure 11.5. On peut retenir que la stratégie de communication doit identifier les destinataires de l'information, les formats de l'information, la période la plus appropriée pour communiquer l'information, la source de l'information et le canal de délivrance de l'information.

Le destinataire de l'information renvoie aux différentes parties prenantes, comme le bailleur de fonds du programme, des chercheurs, le décideur politique. L'identification du ou des destinataire(s) est fonction de la nature de l'information à véhiculer et de la finalité à atteindre. La Figure 11.1 peut aider au choix des destinataires en reliant ces derniers à la finalité de l'évaluation.

Le contenu et le format renvoient aux résultats et aux recommandations contenus dans le rapport d'évaluation qui est le produit final d'une évaluation. Le rapport final devra contenir un résumé exécutif ; c'est-à-dire un résumé du contenu du rapport destiné à la haute direction sans présenter les détails qui ont peu d'importance. La communication peut se faire sous plusieurs formes, à savoir (i) des réunions virtuelles ou en présentiel (ii) les médias (presse écrite, audiovisuel, radio) (iii) les nouvelles technologies de l'information et de la communication, et (iv) des méthodes orales.

Les réunions ou ateliers de validation constituent le canal traditionnel de la communication des résultats de l'évaluation. Par exemple, une réunion peut être prévue avec le bailleur de fonds du programme à la fin de l'évaluation pour présenter

les résultats, les recommandations et les actions futures. En marge des réunions, les résultats d'une évaluation peuvent être véhiculés par les médias ou encore en utilisant Internet (site web de l'organisation) et les réseaux sociaux (Web, Youtube, Facebook, Instagram ...). Dans certaines communautés rurales, la radio, les artistes chanteurs, griots, leaders d'opinion peuvent être utilisés pour véhiculer l'information.

Il est important de noter qu'à chaque canal correspondent un langage et des mots appropriés. Ainsi, la communication doit être adaptée au choix effectué pour véhiculer l'information.

La stratégie de communication se planifie et intègre un budget comme discuté un peu plus haut. Par ailleurs, le choix de la date ou la période de communication est tout aussi important que le canal de communication. En effet, une évaluation peut servir à éclairer des décideurs. Par exemple, les décideurs discutent des questions de développement au cours des assemblées annuelles du Groupe de la Banque mondiale et du Fonds Monétaire International qui se tiennent chaque année à l'autonome. Publier un rapport d'évaluation juste avant les assemblées annuelles en utilisant de puissants canaux de communication (ex. presse internationale) pourrait contribuer à influencer les décisions.

V. Présentation des rapports d'évaluation

Cette section est dédiée à la rédaction du rapport d'évaluation. En plus de la pertinence du contenu, elle insiste aussi sur les éléments de forme qui contribuent à la qualité du rapport.

5.1. Moment, pertinence et utilisation de l'évaluation

La conception, le processus et le rapport de l'évaluation répondent aux besoins des utilisateurs ciblés. Le rapport doit être organisé de sorte que les conclusions, les recommandations et les enseignements soient suffisamment clairs, pertinents et exploitables afin que l'évaluation atteigne les objectifs souhaités.

Quelle que soit la qualité du rapport d'évaluation, il doit être fourni à un moment opportun pour assurer une utilisation optimale des résultats. L'évaluateur devra assurer un respect du calendrier de l'évaluation, car celle-ci constitue une dimension de la qualité de l'évaluation.

La diffusion, la conservation et la gestion du rapport d'évaluation permettent d'assurer son accès à toutes les parties prenantes et d'optimiser les effets positifs de l'évaluation.

5.2. Conseils pour la rédaction de rapports

Le rapport d'évaluation est un document destiné à un large éventail de parties prenantes. C'est pour cette raison que sa rédaction doit être faite suivant un style simple et compréhensible. Le rapport doit permettre de circonscrire l'évaluation en précisant les conditions et contraintes dans lesquelles l'évaluation a été faite. En outre, il doit fournir des informations détaillées sur la méthodologie pour que les lecteurs puissent juger de la crédibilité du travail effectué.

De façon pratique, le rapport doit être organisé autour des questions d'évaluation ou thèmes majeurs, en commençant par les points les plus importants. La structuration doit être telle que les différents chapitres et sections soient cohérents pour en faciliter sa lecture et son exploitation. Il est conseillé que la première phrase de chaque

paragraphe traduit l'idée principale que l'évaluateur voudrait véhiculer. Cette idée maîtresse peut être mise en gras.

De façon classique, le rapport d'évaluation comporte un résumé analytique, une introduction, une description de la méthodologie de l'évaluation, les résultats, les conclusions, les recommandations et les annexes.

Le **résumé analytique** fournit un aperçu de l'étude en rappelant le contexte, le(s) problème(s) de l'évaluation, les questions et les méthodes utilisées, les constatations ou résultats et les principales recommandations. Idéalement, le résumé doit être de 2 pages, mais il ne faut pas excéder 5 pages. Le résumé doit être lisible pour un non-spécialiste. Il faut donc éviter un style technique.

L'**introduction** du rapport doit comporter une accroche qui attire les lecteurs. Il doit impérativement comporter les questions d'évaluation, l'objectif du rapport, le contexte de l'intervention ainsi que ses objectifs, sans toutefois « raconter l'histoire de l'intervention ».

La **description de la méthodologie de l'évaluation** présente le centre d'intérêt de l'évaluation, la conception de l'évaluation, les questions d'évaluation, la méthodologie de collecte des données, la stratégie de l'analyse des données et les limites de la méthodologie. Ce chapitre peut aussi comporter les groupes ou parties prenantes impliqués. Comme indiqué plus haut, la méthodologie doit être détaillée et reliée aux questions d'évaluation. La qualité de l'évaluation dépend en grande partie de la qualité de la méthodologie. En marge de la méthodologie, le rapport doit **présenter les données de l'évaluation** pour permettre au lecteur de comprendre les sources des données, leur nature, les traitements effectués et d'en juger leur qualité. Les données doivent être présentées de manière sélective, en insistant sur les plus importantes.

L'évaluateur doit veiller à organiser les **résultats** de l'évaluation autour des questions de l'évaluation, des thèmes majeurs ou des composantes de l'intervention. Il peut se servir de graphiques et de tableaux pour illustrer et aider à mettre en exergue les principaux points (voir module 10).

Les résultats sont suivis des principales **conclusions** de l'évaluation. Les conclusions doivent être en adéquation avec les questions d'évaluation ou le centre d'intérêt de l'évaluation. Elles sont basées sur des constatations ou les résultats et n'ajoutent pas

de nouveaux détails. Les **recommandations** peuvent faire partie des conclusions. Elles présentent les actions à entreprendre après la lecture du rapport d'évaluation. Les recommandations devraient être basées sur les constatations ou les résultats. Elles doivent être claires et spécifiques et fortement adaptées au contexte et à la situation (la rédaction des recommandations est très délicate et demande souvent beaucoup d'aller-retour et de diplomatie). Les principales conclusions et les recommandations du rapport d'évaluation doivent être étayées par des faits issus de l'évaluation. Il faut éviter de formuler des recommandations sans s'appuyer sur les résultats de l'analyse. Par exemple, une évaluation qui met en évidence un faible niveau des élèves ne peut conclure systématiquement à un sous-financement du secteur de l'éducation sans avoir pris soin de montrer les raisons de ce faible niveau. La formation insuffisante des enseignants, leur manque de motivation, le contenu des programmes de formation comme la baisse du nombre d'heures d'enseignements clés, la motivation et l'assiduité des élèves sont autant de facteurs pouvant justifier la baisse du niveau des élèves. Ces facteurs appellent à la proposition de réformes qui ne sont pas nécessairement financières.

L'évaluateur doit veiller à alléger le contenu technique du rapport et insister sur l'aspect analytique. Ainsi, les informations techniques peuvent être décrites dans des annexes pour permettre aux lecteurs les plus outillés de les consulter. Il faut également prévoir un temps pour réviser le rapport. Le temps peut varier en fonction du nombre de lecteurs du rapport. Il est conseillé de faire relire le rapport par une personne indépendante non nécessairement spécialiste du domaine sur lequel porte l'évaluation. Une personne qui accorde beaucoup d'importance aux détails est recommandée pour cette activité non moins importante.

5.3. Prise en compte des commentaires des parties prenantes

Les parties prenantes ont la possibilité de formuler des commentaires sur la version provisoire du rapport d'évaluation. Le rapport d'évaluation tient compte de ces commentaires toujours dans le but d'accroître l'appropriation des résultats par le commanditaire. Si le fond du propos s'appuie sur les résultats des collectes et ne peut être changé, la forme peut être adaptée avec diplomatie, sans que l'indépendance de l'évaluateur soit mise en danger. S'il y a des différends au sujet de faits qui peuvent être vérifiés, les évaluateurs doivent effectuer des recherches et modifier le cas échéant la version provisoire du rapport. Lorsque les commentaires des parties prenantes expriment une opinion ou représentent une interprétation, ils doivent être reproduits textuellement, par exemple dans une annexe, à condition que cela ne soit pas incompatible avec le respect des droits et du bien-être des participants.

5.4. Suivi des recommandations

Les membres de l'équipe chargée de l'évaluation peuvent être amenés à suivre la mise en place des recommandations et peuvent alors élaborer un système de suivi en collaboration étroite avec le commanditaire. Ainsi, le système peut se présenter sous la forme d'une matrice avec pour chaque recommandation, la(les) personne(s) chargée(s) de sa mise en œuvre, les dates de début et de fin des activités liées à la recommandation. Le Tableau 11.3 présente une illustration d'un système de suivi des recommandations.

Tableau 7: Suivi des recommandations (modèle de réponse de gestion) du PNUD

Recommandation d'évaluation 1				
Réponse de la direction :				
Mesures principales	Calendrier	Groupe(s) responsable(s)	Suivi	
			Commentaires	Statut
1.1				
1.2				
1.3				
Recommandation d'évaluation 2				
Réponse de la direction :				
Mesures principales	Calendrier	Groupe(s) responsable(s)	Suivi	
			Commentaires	Statut
2.1				
2.2				
3.3				

5.5. Texte et visuel

Le module 10 a présenté différents types d'illustrations en fonction de la nature des données et de l'objectif de l'analyse. Cette section insiste sur la présentation des visuels (tableaux, graphiques) ou d'autres illustrations (schémas, cartes, croquis, encadrés...). Par exemple chaque visuel (tableau ou graphique) ou illustration doit avoir un titre explicite et un numéro unique. L'illustration ou le visuel doit être référencé dans le texte. Par exemple, s'il existe un tableau 2, il faut dans le texte commenter le tableau 2 en le citant explicitement. L'évaluateur doit s'assurer d'établir une relation entre le narratif et les illustrations. De façon spécifique, si l'on utilise une image ou une illustration, il faut être sûr d'utiliser la partie narrative du rapport pour indiquer ce que le lecteur est censé comprendre à partir de cette illustration.

Une bonne pratique est de veiller à ce que les illustrations et visuels soient positionnés le plus près possible de leur mention dans le texte. Une illustration doit communiquer des informations plus clairement et plus efficacement. Elle peut être utilisée pour « alléger » la densité d'un texte continu et procurer un point focal, attirer l'attention du lecteur sur des points-clés.

Les visuels et illustrations doivent communiquer une information sans besoin de texte. Ils doivent présenter une information synthétique, être simples et adaptés au contexte de l'étude. Pour le cas particulier des tableaux, il *n'est pas recommandé* de mettre la base de données à l'état brut sous forme de tableaux dans le corps du document. Il faut synthétiser l'information en présentant des fréquences absolues ou relatives, des profils lignes ou colonnes, des caractéristiques de tendance centrale, de forme et de dispersion pour le cas de variables quantitatives. Les visuels et illustrations doivent facilement s'insérer dans le texte avec des numérotations systématiques et des titres appropriés. Il est indispensable de fournir les sources des données utilisées pour les produire. Les visuels et illustrations doivent être faciles à lire avec des légendes clairement expliquées. Le rédacteur du document ne doit pas hésiter à utiliser des notes placées au bas de l'illustration pour expliquer son contenu.

5.6. Qualité du rapport final et de l'évaluation

Il existe plusieurs dimensions qui permettent d'apprécier la qualité d'un rapport d'évaluation et d'une évaluation de façon générale. On peut néanmoins retenir que le rapport doit être conforme aux termes de référence; c'est-à-dire, répondre aux questions et problèmes au démarrage de l'évaluation. Il doit être clair, bien organisé, facile à lire et utiliser les visuels et illustrations appropriés. Le rapport doit être écrit dans un langage accessible aux futurs utilisateurs ou lecteurs. Il doit être rigoureux et présenter une position neutre et non orientée. Le contenu du rapport doit indiquer l'indépendance de l'évaluateur vis-à-vis du commanditaire. Le rapport doit être fourni à temps conformément au calendrier convenu entre l'évaluateur et le commanditaire.

Pour évaluer une évaluation, les lecteurs peuvent se servir d'une grille dont un exemple est fourni par le Tableau 11.4. La pratique consiste à noter chacune des dimensions ou questions. La note globale sera la note minimale obtenue dans toutes les dimensions.

Tableau 8: Grille d'estimation de la qualité d'une évaluation

	Excellent	Bon	Acceptable	Inacceptable
1. Niveau de réponse à la demande (par rapport aux questions posées à l'origine dans les termes de référence)				
2. Pertinence du champ couvert				
3. Indépendance ou pluralisme de l'équipe d'évaluation				
4. Adéquation méthodologique				
5. Fiabilité des données utilisées				
6. Solidité des analyses				
7. Crédibilité des conclusions				
8. Clarté du rapport (ou de la présentation)				

VI. Exposés oraux

Les exposés oraux peuvent être effectués tout au long de l'évaluation. Tout d'abord au cours de la réunion de démarrage qui permet de présenter la compréhension des termes de référence et la méthodologie proposée par l'évaluateur. Ensuite, au cours des réunions de restitution conformément au calendrier convenu entre l'évaluateur et le commanditaire. Enfin, à la fin de l'évaluation pour rendre compte des résultats obtenus et des recommandations. Certaines organisations organisent des séminaires d'apprentissage à la fin d'une évaluation, que l'évaluateur peut être amené à animer.

Les exposés oraux doivent être préparés et planifiés pour faciliter la communication. Ils doivent être adaptés au public auquel l'on s'attend. Les exposés ne doivent pas être très longs et il faut faire l'effort de respecter le temps imparti. L'évaluateur doit utiliser les ressources disponibles dans la salle de la rencontre (Diapositives, PowerPoint, posters) sans créer d'autres diligences supplémentaires qui peuvent être souvent inutiles pour la circonstance.

La Figure 11.6 présente quelques directives pour mieux présenter ses exposés oraux. Tout comme le rapport écrit nécessite un travail organisé, les exposés oraux doivent être préparés et structurés. L'exposé oral n'est pas une séance de lecture, mais plutôt une séance d'explications et de partage tout en se conformant au contenu du rapport qui est présenté.

Figure 11.6: Quelques directives pour les exposés oraux

Préparation de l'exposé

- Garder toujours votre public à l'esprit
- Annoncer ce que vous allez leur dire
- Expliquer les détails
- Rappeler si possible ce qui a été dit
- Prévoir un plan de secours en cas de coupure d'électricité ou de défaillance de l'équipement (rétroprojecteur, ordinateur, etc.)

Conception des supports

- Utilisez peu de mots
- Utilisez des visuels clairs
- Le public est présent pour écouter, pas besoin de tout écrire dans le Powerpoint

Distribution de documents avant l'exposé

- Aide à suivre le déroulement de la présentation
- Retour en arrière possible pour faire des liens et vérifier la compréhension
- Prise de notes facilitée
- Mais, cela peut distraire

Répétition de l'exposé oral

- Il est indispensable de répéter son exposé seul, et devant une ou d'autres personnes avant le jour de la présentation
- Obtenir de la rétroaction de la part des autres
- Réajuster votre exposé en fonction de ce que vous ressentez ou de ce que les autres disent
- A mesure que vous présentez, tenez compte du temps qui s'écoule
- Parlez aux gens, pas à vos notes, essayez d'établir un contact visuel

Messages clés :

- ✓ La communication dans la gestion d'une intervention s'avère souvent être une condition de succès. Une gestion axée sur les résultats ne peut en effet être menée que dans un climat de production d'informations claires et pertinentes et de partage de cette information.
- ✓ Les évaluations sont parfois sous-utilisées malgré les efforts et les ressources qu'elles représentent, ce qui a encouragé de nombreux auteurs à étudier les bonnes pratiques de communication. Beaucoup d'entre elles peuvent être adoptées par les équipes d'évaluation dans une stratégie intégrée de communication, ce qui améliorera la satisfaction autour de leur travail et la pertinence de leurs rapports.
- ✓ Parmi ces bonnes pratiques, on peut citer des échanges clairs et préparés au démarrage de l'évaluation pour clarifier les attentes et l'angle d'attaque de l'évaluation ; le maintien d'un dialogue sur l'évaluation pendant les travaux ; une attention particulière aux livrables de l'évaluation dont le rapport et les exposés oraux.

Bibliographie :

- Patton, M. Q. (2003). Utilization-focused evaluation. *In International handbook of educational evaluation* (pp. 223-242). Springer, Dordrecht.
- Patton, M. Q. (1997). *Utilization-focused evaluation: The new century text* (3rd ed.). Thousand Oaks, CA: Sage Publications.
- Torres, R. T., Preskill, H., & Piontek, M. E. (2005). *Evaluation strategies for communicating and reporting: Enhancing learning in organizations*. Sage.

MODULE 12

GESTION DU PROCESSUS D'ÉVALUATION

Objectifs d'apprentissage :

A la fin de la session, les participants seront capables de :

- Planifier le processus de l'évaluation
- Préparer la contractualisation de l'évaluation,
- Rédiger une note méthodologique.
- Coordonner les activités de l'évaluation
- Faire le suivi du processus

Thèmes couverts

- I. Planification du processus de l'évaluation
- II. Coordination des activités de l'évaluation
- III. Suivi du processus de l'évaluation

Une mission d'évaluation peut être considérée comme un projet. En effet, elle regroupe souvent une équipe, pluridisciplinaire, qui doit effectuer un certain nombre de tâches, souvent en parallèle et pendant un temps limité. Elle dispose d'un budget, la plupart du temps précis et défini à l'avance dans un contrat dans le cadre d'une évaluation externe. Elle doit également suivre un calendrier serré, car une évaluation déposée trop tard est souvent une évaluation qui n'est plus très utile, puisqu'elle aura certainement loupé les moments de prise de décision. Elle est enfin l'occasion de rassembler de nombreux documents et données, qu'il faut collecter, classer, archiver avec rigueur.

Ainsi, et comme tout projet, l'évaluation doit faire l'objet d'une bonne gestion pour atteindre ses objectifs. Ce module, en complément du précédent sur la communication, porte donc sur la gestion du processus d'évaluation.

Il est structuré autour de trois thèmes : la première porte sur la planification du processus de l'évaluation et couvre la préparation du dossier de l'évaluation, les formes de contractualisation, l'identification des parties prenantes et l'élaboration du budget et de la note méthodologique. Le deuxième est dédié à la coordination des activités de l'évaluation. Pour examiner les questions organisationnelles et logistiques. Enfin, ce chapitre présente quelques directives pour conduire une évaluation de qualité.

I. Planification du processus de l'évaluation

On peut voir la phase de planification de l'évaluation comme un processus en cinq étapes. La première étape consiste à préparer le dossier de l'évaluation, la deuxième porte sur la contractualisation, à la troisième étape sont identifiées les parties prenantes puis est élaboré le budget en relation avec les activités relatives à l'évaluation (étape 4). Enfin, l'étape 5 consiste à produire la note méthodologique de l'évaluation. La suite de cette section décrit chacune de ces étapes.

Le séquençage tel que proposé peut ne pas toujours être suivi à la lettre. En d'autres termes, certaines activités peuvent se faire avant d'autres. Néanmoins, il s'agit tout de même d'une suite d'étapes courantes et qui s'enchaînent de cette façon dans de nombreux cas d'évaluation. On peut toutefois retenir que l'étape 1 (préparation du dossier de l'évaluation) doit être terminée avant de passer aux autres étapes. Il est également recommandé que cette étape déclenche le processus pour s'assurer que l'évaluateur ait bien cerné le travail à accomplir. De même, la note méthodologique ne peut être rédigée qu'après les étapes 1, 2, et 3. Il convient de souligner qu'il existe un lien étroit entre les ressources disponibles, notamment le budget alloué à l'évaluation et la méthodologie finale qui sera retenue. Lorsqu'il s'agit d'une évaluation externe, l'interaction entre les ressources financières et la méthodologie de l'évaluation donne lieu à des négociations entre l'équipe d'évaluation et le gestionnaire du projet.

1.1. Étape 1 : Préparation du dossier de l'évaluation

La première étape de l'évaluation est réalisée par le commanditaire et consiste à préparer le dossier de l'évaluation. Cette préparation se fait généralement en trois phases, à savoir (i) l'élaboration d'un relevé des préoccupations, (ii) la définition des critères de sélection de l'équipe d'évaluation, et (iii) l'élaboration des termes de références (TDR) et la demande de proposition (DP). Ces étapes sont présentées dans le Tableau 12.1.

Tableau 19: Différentes étapes de la préparation du dossier de l'évaluation

No.	Phase	Objet
1	Élaboration d'un relevé des préoccupations	<ul style="list-style-type: none"> • C'est une description de la raison d'être de l'évaluation, son déclencheur. Le commanditaire fait un tour des préoccupations et des interrogations qu'il a, et auxquelles l'évaluation devrait répondre • Ces préoccupations peuvent être rédigées sous la forme de questions d'évaluation en se servant des critères de performance si le commanditaire est à l'aise avec ces concepts
2	Définition des critères de sélection	<ul style="list-style-type: none"> • Ils font appel à la contribution de plusieurs expertises (connaissance des questions de développement, familiarité avec les activités de l'organisation, compétence d'évaluation en conception, en collecte des données, en analyse de données, en préparation de rapports, et à l'utilisation des technologies de l'information et de la communication)
3	Élaboration des termes de référence (TDR) et la demande de proposition (DP)	<ul style="list-style-type: none"> • Souvent utilisés seuls (sans réel document de DP), les TDR décrivent l'évaluation à effectuer et établissent le calendrier des réalisations • Utilisés pour tous types d'évaluation : embauche d'une équipe externe ou interne (forme simplifiée car peu d'éléments contractuels) • Elles permettent à tous les acteurs de comprendre les attentes et le contexte de l'évaluation

Phase 1 : Élaboration d'un relevé des préoccupations

Il est réellement utile pour un commanditaire d'évaluation de réunir les préoccupations de gestion sur l'intervention à évaluer pour guider l'évaluateur dans la proposition qui va être faite. Plus ces préoccupations sont précises, plus l'évaluateur pourra proposer une stratégie d'évaluation pertinente.

Si le commanditaire n'est pas à l'aise avec les concepts d'évaluation ou les critères de performance, ces préoccupations peuvent être exprimées en langage courant le plus naturellement possible. Si toutefois les compétences en évaluation sont présentes, la formulation en question d'évaluation est un plus.

En l'absence de précisions sur ces éléments dans les TdR, l'évaluateur devra poser ces questions et aider le commanditaire à formuler ces préoccupations, pour éviter de produire une évaluation générique, qui risquerait d'être moins utile.

Phase 2 : Définition des critères de sélection

L'évaluation d'une intervention fait souvent appel à plusieurs expertises selon le projet ou le programme concerné. L'équipe d'évaluation est ainsi souvent une équipe pluridisciplinaire. L'évaluateur sénior ou chef de mission doit avoir une excellente connaissance de la discipline de l'évaluation. C'est le critère primordial. Par la suite, on peut notamment demander une connaissance du secteur concerné par l'intervention, mais cela n'est pas forcément nécessaire s'il a une connaissance suffisamment avérée des questions de développement par exemple, ou des dynamiques de travail concernées. Ce dernier travaille habituellement avec plusieurs autres expertises (mais il peut aussi travailler seul), à savoir des spécialistes en collecte et analyse des données (quantitatif ou qualitatif) ou des experts du secteur si cela s'avère nécessaire. L'équipe devra souvent justifier d'une qualité rédactionnelle avérée pour produire des rapports de qualité (par exemple, fournir des exemples de rapports).

Par ailleurs, l'équilibre entre hommes et femmes, ou régional peut être un prérequis selon le projet et/ou le bailleur, la capacité de travailler dans certaines langues locales, l'expertise avec certaines approches particulières, etc. peuvent constituer des critères supplémentaires pour sélectionner une équipe.

Phase 3 : Élaboration des termes de référence (TDR) et la demande de proposition (DP)

La phase de planification culmine avec la rédaction d'un document clair et complet, habituellement appelé « termes de référence (TDR) » présentant différents éléments :

- De fond : comme les préoccupations du commanditaire, les objectifs de l'évaluation ou des indications sur la méthodologie ou l'approche à adopter, des lignes directrices pour le rapport ou la communication des résultats
- Sur les ressources : le budget, le calendrier, la composition de l'équipe souhaitée
- Contractuels : différents engagements par exemple sur la confidentialité, l'éthique, etc.

Le document est approuvé par le(s) responsable(s) d'évaluations et les personnes chargées de conduire l'évaluation.

Les TDR peuvent être structurés de plusieurs façons. L'Encadré 12.1 donne à titre d'exemple les éléments constitutifs des TDR élaborés par le PNUD pour évaluation à mi-parcours du Programme Pays (CPD) PNUD Mali 2015-2019⁸.

Encadré 12.1 : Eléments constitutifs des TDR élaborés par le PNUD pour l'évaluation du CPD

Ci-dessous les sections des TDR élaborés par le PNUD pour évaluation à mi-parcours du Programme Pays (CPD) PNUD Mali 2015-2019 :

- | | | |
|---|---|---|
| 1. Contexte et Justification de la Mission | 6. Déroulement du processus | 11. Durée de la mission |
| 2. Justification de l'évaluation | 7. Résultats attendus | 12. Éthique d'évaluation |
| 3. Portée de l'évaluation et les principaux objectifs | 8. Livrables | 13. Planning indicatif du processus |
| 4. Questions d'évaluation | 9. Composition de la mission | 14. Modèle de rapport d'évaluation |
| 5. Méthodologie | 10. Critères de sélection des consultants | 15. Qualifications du consultant international et des consultants nationaux |

⁸ Le document complet est accessible à <https://erc.undp.org/evaluation/documents/download/11191> (consulté le 4 avril 2021).

1.2. Etape 2 : Contractualisation de l'évaluation

L'évaluation peut être conduite soit sous une forme dite interne ou une forme dite externe. En ce qui concerne l'évaluation interne, l'organisation la réalise elle-même à partir de ses référentiels et avec les ressources humaines disponibles. A l'opposé, l'évaluation externe est réalisée par un tiers ou par une équipe externe à l'organisation qui demande l'évaluation. Le choix de l'évaluateur externe se fait généralement avec beaucoup de précautions.

Chaque forme d'évaluation présente des avantages et des inconvénients dont quelques-uns sont résumés dans le Tableau 12.2. Plus généralement, l'évaluation interne a l'avantage d'être moins coûteuse et surtout d'aller généralement plus rapidement du fait de la connaissance préalable des évaluateurs sur le fonctionnement de l'organisation et sur le secteur d'intervention. Cette forme d'évaluation est aussi très utile pour assurer le suivi des recommandations formulées à la suite de l'évaluation, car les évaluateurs restent disponibles. Néanmoins, l'indépendance des évaluateurs est parfois plus difficile à garantir à l'interne, et l'évaluation doit se contenter des disponibilités des évaluateurs et des compétences qu'ils ont.

A l'opposé l'évaluation externe, le biais organisationnel est relativement faible, bien que la relation entre un client payeur et un évaluateur indépendant dont le salaire dépend de cette relation peut également amener des dérives. L'évaluation externe reste généralement beaucoup plus coûteuse et peut nécessiter des négociations relativement longues.

Le choix d'un consultant pour l'évaluation doit obéir à une démarche rigoureuse pour éviter les critiques. Une bonne pratique est de mettre en place un jury pour l'analyse des propositions soumises. Le jury peut être composé de personnes ayant (i) des connaissances et une expérience de l'évaluation, (ii) des connaissances du domaine concerné par l'évaluation, et (iii) des connaissances de la culture organisationnelle. Le jury devra analyser tous les dossiers de façon objective, équitable et transparente au regard des critères préalablement définis pour choisir la meilleure proposition, c'est-à-dire celle qui est la plus appropriée pour le travail demandé.

Lors du processus de soumission, il est conseillé de demander une offre technique et une offre financière dans deux documents séparés. L'équipe chargée d'évaluer les offres analysera d'abord les offres techniques et ensuite examinera les offres financières des institutions ayant les offres techniques les mieux classées. Ces évaluations doivent reposer sur des critères préalablement définis par le gestionnaire de l'évaluation pour classer les offres. L'examen des propositions financières ne doit se faire que la qualité de l'offre technique est acceptable ; c'est-à-dire si l'offre technique vérifie un niveau minimum d'éligibilité. Par exemple, on peut décider d'examiner l'offre financière que lorsque la note de l'offre technique est d'au moins 70/100.

Dans l'évaluation de l'offre technique, il faut prendre en compte (i) la compréhension de la mission ; (ii) la pertinence et la crédibilité de la méthodologie proposée ; (iii) la cohérence et la crédibilité des ressources humaines proposées ; (iv) l'expérience du consultant/cabinet ; (v) le respect des contraintes explicites fixées dans la mission. L'équipe responsable de l'évaluation des offres ne doit pas hésiter à rechercher des informations complémentaires sur le ou les cabinets postulants et traquer les incohérences internes et externes.

Pour définir la meilleure offre, il faut clairement définir la pondération entre l'offre technique et l'offre financière dans la note globale. Par exemple, une pondération 80% pour la qualité de l'offre technique et 20% pour l'offre financière peut être considérée. Cette pondération permet d'éviter le piège du moins-disant. Le moins-disant est la recherche d'une proposition qui offre les coûts les plus faibles. Il existe des situations dans lesquelles les soumissionnaires peuvent intentionnellement baisser les prix pour obtenir le marché sans disposer des moyens pour accomplir la mission.

Tableau 12.10: Avantages et inconvénients des évaluations interne et externe

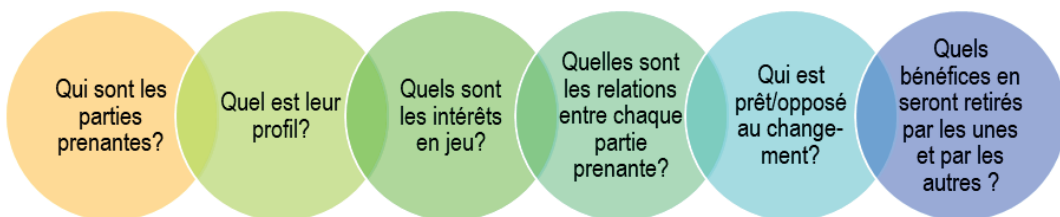
	Evaluation interne	Evaluation externe
Avantages	<ul style="list-style-type: none"> • Connait l'organisation, ses programmes et ses opérations • N'est pas vu comme un adversaire ou un outsider • A plus de chance de pouvoir suivre les recommandations • Connait bien les procédures internes et la façons d'obtenir les documents et les données • Est souvent moins cher • Ne nécessite pas de longues négociations de passation de marché • A plus d'occasions de développer les capacités d'évaluation du pays hôte 	<ul style="list-style-type: none"> • Peut être perçue comme plus indépendante • Peut amener de nouvelles perspectives, des idées, une expérience plus vaste, et des connaissances récentes et de pointe • Est plus facilement embauché pour un travail intensif et immédiat • Peut servir d'arbitre ou de facilitateur entre les parties prenantes dans des situations conflictuelles
Inconvénients	<ul style="list-style-type: none"> • A tendance à avoir moins de recul sur les interventions de l'organisation, à accepter les hypothèses de l'organisation • Est généralement plus occupé, et il faut planifier à l'avance son implication dans un dossier • Peut être gêné par un conflit de rôles à l'intérieur de l'organisation 	<ul style="list-style-type: none"> • Peut ne pas connaître l'organisation, ses politiques et ses procédures ou régulations • Peut ignorer ce qui pourrait empêcher la faisabilité des recommandations et formuler des recommandations maladroites • Peut être considéré comme un adversaire, ce qui cause une anxiété inutile • Peut-être coûteux (• Demande plus de temps pour les négociations de contrat, l'orientation et la supervision • Ne peut généralement pas faire le suivi des recommandations • Peut ne pas bien connaître l'environnement politique, culturel et économique local (consultant international)

1.3. Etape 3 : Identification des parties prenantes

Les parties prenantes sont les personnes, les groupes de personnes, les organismes et institutions concernés d'une manière ou d'une autre par l'intervention. Leur identification est utile pour déterminer ou comprendre la logique d'intervention. Il s'agit d'identifier l'ensemble des acteurs en lien avec l'action ; d'examiner leurs rôles respectifs, les intérêts en jeu, leur capacité à participer ou à freiner le projet, d'estimer le degré de collaboration ou les frictions potentielles entre les différentes parties.

Pour identifier les parties prenantes, l'évaluateur devra répondre aux questions résumées par la Figure 12.1. En général, l'on distingue trois catégories de parties prenantes. Il y a les parties prenantes primaires qui sont directement concernées par les interventions. Il s'agit des publics cibles ou bénéficiaires. Ensuite, il y a les parties prenantes secondaires indirectement touchées par les interventions. Cette catégorie bien que n'étant pas la cible principale des interventions peut affecter (ou être affectée par) les interventions. Enfin, la dernière catégorie est composée des parties prenantes clés, incontournables pour le processus de changement. Il peut s'agir, par exemple, des autorités publiques.

Figure 12.1: Questions clés pour l'identification des parties prenantes



Idéalement, l'évaluateur doit rencontrer toutes les parties prenantes afin de disposer d'une bonne cartographie de l'intervention et connaître la diversité des points de vue. Par ailleurs, il faut classer les parties prenantes en fonction de leurs responsabilités ou relation avec les interventions. Par exemple, les deux premières catégories (primaires et secondaires) ont une relation avec le projet ou les interventions tandis que les acteurs de la troisième catégorie ont une forte influence sur le projet. Dans les évaluations participatives, l'implication de l'ensemble des parties prenantes est

encore plus cruciale, car c'est en dégagant des consensus entre elles que l'on pourra élaborer le jugement évaluatif.

La rencontre avec les parties prenantes peut tourner autour de plusieurs points, dont (i) la nature et le contexte de la demande exprimée par le commanditaire de l'évaluation, et (ii) le besoin, le problème ou l'intérêt menant à la demande d'évaluation. En outre, les questions d'importance capitale que l'évaluation doit aborder ainsi que les délais cruciaux pour les conclusions de l'évaluation doivent, si possible être présentés et discutés avec les parties prenantes afin de solliciter leur appui. Ces rencontres peuvent permettre d'avoir accès à des informations importantes pour l'évaluation. On reste parfois limités dans le temps et par les façons de faire du commanditaire de l'évaluation, mais l'équipe d'évaluation doit le plus possible proposer d'inclure ces tâches dans sa méthodologie.

Comme illustration, le Tableau 12.3 présente la liste des parties prenantes d'un projet d'OXFAM. Pour chaque partie prenante, l'évaluateur spécifie les besoins d'informations, les questions spécifiques d'évaluation et la priorité qui y est associée. Cela permet de se faire une idée des groupes et organisations importants et de planifier au mieux leur rôle dans l'évaluation.

Tableau 11 : Parties prenantes d'un projet d'OXFAM

Parties prenantes de l'évaluation	Besoins d'informations	Questions spécifiques d'évaluation	Priorité : haute, moyenne, basse
Bénéficiaires du projet			
Gestionnaire de projet d'OXFAM			
Personnel sur le terrain d'OXFAM			
Gestionnaire de projet d'ODA (Bangkok)			
Bureau de pays d'ODA (Londres)			
Personnel sur le terrain du projet SAVE			
SAVE UK			
Conseillers pour le développement social d'ODA			

Comme pour tout le processus d'évaluation, les évaluateurs doivent toutefois maintenir un degré d'indépendance vis-à-vis des parties prenantes, quelle que soit leur position ou leur influence. L'équipe d'évaluation doit être en mesure de travailler librement avec chacune d'entre elles, et être perçue le plus possible comme un joueur neutre. Elle doit pouvoir compter sur la coopération des parties prenantes pour avoir accès à toute information pertinente.

1.4. Etape 4 : Élaboration d'un plan des ressources, des tâches et du budget

L'évaluation doit être un processus bien planifié. L'un des facteurs pouvant affecter sa mise en œuvre est une mauvaise programmation des activités ou une insuffisance de ressources. Il est donc important de dresser une liste exhaustive des tâches à exécuter assortie d'un diagramme de Gantt ou un calendrier donnant les échéances. En plus, l'équipe d'évaluation doit allouer les ressources appropriées aux activités identifiées.

1.4.1. Planification des activités

Le Tableau 12.4 donne une illustration de la répartition des tâches, des noms de personnes responsables de l'activité, et des dates de démarrage et de fin. L'évaluateur ou l'équipe d'évaluation doit veiller au caractère réaliste du calendrier proposé pour chacune des activités.

Tableau 12: Illustration d'une cartographique assortie de délais de réalisation

Tâches	Noms	Date de démarrage	Date échu
Mener la revue de littérature et identifier les questions internes et externes y afférentes	Anna, Miguel et Kabir	01/7	04/7
Passer en revue les documents du programme / projet	Kabir	05/7	23/7
Programmer et tenir une rencontre avec le client	Anna	15/7	31/7
Identifier les parties prenantes-clés et programmer des rencontres	Kabir et Miguel	15/7	17/7
Rédiger le résumé des réunions avec le client ainsi que des décisions prises	Anna	01/8	31/8
Mener les réunions avec les parties prenantes, faire le résumé	Anna et Miguel	05/8	15/8
Faire une proposition de conception initial d'évaluation	Anna	01/7	31/8

Outre le calendrier, l'équipe de projet peut réaliser un diagramme de Gantt qui est un graphique montrant les interrelations entre les étapes/tâches de l'évaluation en fonction du temps (cf. Tableau 12.5). La colonne de gauche du diagramme énumère toutes les tâches à effectuer, et la ligne d'en-tête représente les unités de temps les plus adaptées au projet (jours, semaines, mois etc.). Chaque tâche est matérialisée par une barre horizontale, dont la position et la longueur représentent la date de début, la durée et la date de fin. Ce diagramme est visuel et permet d'un seul coup d'œil de disposer de l'information sur (i) les différentes tâches à envisager, (ii) les dates de début et de fin de chaque tâche, (iii) la durée escomptée de chaque tâche, (iv) le chevauchement éventuel des tâches, et la durée de ce chevauchement, et (v) la date de début et la date de fin du projet dans son ensemble.

Tableau 13: Exemple de diagramme de Gantt

Tâches	semaine 1	semaine 2	semaine 3	Semaine 4	Semaine 5	Semaine 6	Semaine 7
Procéder à une revue de littérature	▲ [Barre bleue] ▲						
Tenir des réunions avec les parties prenantes		▲ [Barre bleue] ▲					
Proposer une théorie du changement Draft theory of change			▲ [Barre bleue] ▲				

1.4.2. Comment établir le budget d'une évaluation ?

Le budget d'une évaluation constitue l'une des dernières étapes, mais non moins importante, de la conception d'une évaluation. Le budget doit être exhaustif tout en restant conforme aux standards internationaux et/ou nationaux. De façon spécifique, le caractère détaillé du budget permet de comparer les coûts unitaires proposés aux coûts unitaires d'évaluation similaires.

En se référant au manuel de Gertler et al. (2011), un budget comprend globalement les frais de personnel pour l'équipe d'évaluation. Aux frais de personnel s'ajoutent les frais de déplacements et de mission (hôtels et indemnités quotidiennes) ainsi que les frais de diffusion, souvent sous forme d'ateliers, de rapports et de publications académiques.

Les coûts les plus importants d'une évaluation sont les coûts relatifs au temps passé pour la collecte des données (y compris la conception et la mise en œuvre d'une enquête pilote), au matériel et aux équipements nécessaires à cette collecte, à la formation et au salaire journalier des enquêteurs si on en embauche (ce qui n'est pas toujours nécessaire) et à leur déplacement. Pour calculer le coût de ces intrants, il est nécessaire de faire quelques hypothèses, par exemple, sur le temps nécessaire pour réaliser un questionnaire ou le temps de déplacement entre les sites. Une

collaboration avec l'Institut ou l'Office National de la Statistique du pays concerné par l'étude peut s'avérer utile, voire indispensable pour disposer de coûts réalistes.

Le Tableau 12.6 présente une feuille de calcul permettant d'estimer les coûts d'une évaluation (à titre illustratif). La colonne nombre spécifie le nombre d'unités nécessaire pour l'opération. Par exemple, on précisera le nombre de personne(s) responsable de l'évaluation. Si le projet nécessite une personne, on écrira « 1 » dans la colonne « nombre » en face de la ligne « Personnel chargé de l'évaluation (gestionnaire de l'évaluation, etc.) ». Le taux ou l'unité est le coût unitaire de chaque poste. Toujours pour la(les) personne(s) responsable de l'évaluation, on précisera le coût unitaire pour chaque personne. Le coût unitaire peut être journalier, mensuel, annuel ou forfaitaire (c'est l'unité à préciser). Si le chargé de l'évaluation coûte 100 000 FCFA par jour, on écrira ce montant dans la colonne « taux/unité » tout en précisant « j » pour dire jour. Une fois l'unité est retenue, il faut préciser le nombre d'unités pour chaque personne ou entité. Par exemple, le chargé d'évaluation sera mobilisé sur dix jours pour l'opération, on écrira « 10 » dans la colonne « Nombre d'unités » pour la ligne « Personnel chargé de l'évaluation (gestionnaire de l'évaluation, etc.) ». Enfin, on peut déduire le coût total de chaque poste (ligne) en multipliant le nombre, le taux unitaire et le nombre d'unités. Pour notre exemple, la personne chargée de l'évaluation coutera 1 000 000 FCFA ($100\,000 \times 10$).

La feuille de calcul (Tableau 12.6) doit être adoptée selon le projet ou le pays. Par exemple, certaines institutions prévoient un poste « frais de structure ». Les frais de structure sont généralement utilisés pour la gestion du projet et couvrent la gestion financière (comptable, frais de transactions financières, etc.) ou la mise à disposition d'équipements pour le projet si cela n'est pas explicité dans le budget. Par exemple, on peut utiliser les salles de réunions et les ordinateurs de l'institution tout au long du processus d'évaluation sans prévoir de lignes spécifiques pour les comptabiliser dans le budget. De même, les frais annexes comme la quantité d'électricité ou d'eau utilisée par les personnes engagées uniquement dans l'évaluation sont difficilement quantifiables et dissociables des coûts traditionnels de l'entreprise. Ces coûts peuvent être couverts par les frais de structure.

Tableau 14: Feuille de calcul pour l'estimation du coût d'une évaluation

Tâches et ressources	Nombre	Taux/unité	Nombre d'unités	Total
Personnel				
Personnel chargé de l'évaluation (gestionnaire de l'évaluation, etc.)				
Consultants internationaux et/ou nationaux (chercheurs/responsable d'enquêtes)				
Assistant de recherche				
Statisticien				
Coordonnateur de terrain				
Déplacements				
Billets d'avion/voyages nationaux et internationaux				
Transports routiers				
Frais de mission (hôtels et indemnités journalières)				
Collecte de données				
Conception de l'instrument				
Enquête(s) pilote(s)				
Formation				
<i>Déplacements et indemnités journalières</i>				
Matériel et équipement pour l'enquête				
<i>Impression des questionnaires</i>				
Personnel de terrain				
<i>Enquêteurs</i>				
<i>Superviseurs</i>				
Transport (véhicules et carburant)				
<i>Chauffeurs</i>				
Saisie et apurement des données				
Analyse et diffusion des données				
Ateliers				
Articles, rapports				
Autres				
Bureaux				
Communication				
Logiciels				

Source : Gertler et al. (2011).

1.5. Etape 5 : Note méthodologique

La note méthodologique peut constituer la dernière étape de la préparation de l'évaluation. En pratique, la note méthodologique est élaborée dans le but d'atteindre efficacement les résultats fixés avec les ressources disponibles. C'est donc un plan de travail qui détaille l'ensemble des activités à conduire, et ce document est à élaborer que l'on soit un évaluateur interne ou externe.

La note méthodologique débute le plus souvent par une section relative à la compréhension des TDR. L'évaluateur explique dans cette section sa compréhension des TDR et présente succinctement comment il répondra aux attentes du commanditaire. Elle doit être présentée de sorte à répondre aux TDR. Dans le cas où le consultant ou l'évaluateur estime que des ajustements sont nécessaires par rapport aux TDR initiaux pour la bonne conduite des travaux, il est libre d'ajouter une section relative à ses observations. Il motivera ses observations et tirera les conclusions par rapport au travail à conduire.

Ensuite, la note comporte généralement une section méthodologie qui détaille très clairement les étapes nécessaires ou la démarche pour atteindre les objectifs assignés dans les TDR. La méthodologie peut reposer essentiellement sur une matrice d'évaluation puisque celle-ci comporte les critères, questions, méthodes, techniques, sources de données, et instruments de collecte. Chacune des composantes de la matrice doit être expliquée de façon détaillée pour permettre au lecteur de disposer d'un ensemble d'information sur la capacité de l'équipe d'évaluation à faire correctement le travail. Ainsi, la note méthodologique est un document de stratégie et démarche de mise en œuvre des activités planifiées.

Le document doit comporter une section sur les livrables, assortie d'un calendrier précis (voir étape 4). Le calendrier doit être conforme à la démarche présentée dans la section méthodologie de la note.

La note méthodologique aide également à préparer le budget de l'évaluation, les étapes 4 et 5 peuvent donc être réalisées en parallèle. De même, des contraintes budgétaires conduiraient à un ajustement de la note méthodologique afin de garantir une cohérence entre la démarche et les ressources disponibles.

II. Coordination des activités de l'évaluation

La coordination est une partie intégrante de l'activité d'évaluation dans la mesure où elle contribue grandement à son succès. Cette section se concentre sur le rôle du gestionnaire d'évaluation dans les organisations et sur la logistique d'une l'évaluation.

2.1. Gestionnaire d'évaluation

Le gestionnaire d'évaluation est une instance (constituée parfois d'une seule personne consacrant au moins 50% de son activité à l'évaluation) dont le rôle peut être de deux ordres (SEVAL, 2014). Le premier rôle est de contribuer à la qualité d'évaluations spécifiques, de garantir leur qualité et par conséquent leur crédibilité et leur utilité. En second, il est chargé de promouvoir la culture d'évaluation au sein de son organisation.

De façon plus détaillée, le gestionnaire d'évaluation s'occupe, tout d'abord, de développer une stratégie d'évaluation pour son organisation, c'est-à-dire qu'il tente de répondre aux questions suivantes : que faut-il évaluer ? Pourquoi évaluer ? Quand évaluer ? Comment évaluer ? Ensuite, il est chargé de piloter le processus d'adjudication, d'assister à chaque évaluation pour assurer la qualité et l'utilisation de ses résultats. Il développe également les outils et les structures assurant un apprentissage d'une évaluation à l'autre.

On voit que le gestionnaire d'évaluation joue un rôle vital pour garantir la qualité des évaluations, leur utilisation et leur valorisation dans une organisation, il est donc un responsable central pour une organisation qui gère beaucoup d'interventions, comme un ministère, une ONG ou une organisation internationale.

Avant la mission d'évaluation, le gestionnaire d'évaluation contribue généralement à la rédaction des TDR et participe activement à la phase de contractualisation si l'évaluation est externe. Pour une évaluation interne, le gestionnaire d'évaluation joue le rôle de chef d'équipe en facilitant le travail en équipe et en assurant une liaison entre l'équipe responsable de l'évaluation et les parties prenantes, tout en garantissant que l'équipe d'évaluateurs puisse faire son travail de façon indépendante. Il suit également le travail de l'équipe d'évaluation de bout en bout (rencontre à fréquence régulière) pour s'assurer du bon déroulement de l'opération.

A la fin de l'évaluation, il intervient pour faciliter la discussion des résultats avec les parties prenantes. Le gestionnaire finalise généralement les activités, archive les informations et s'assure de la validation des rapports. La validation des rapports se fait en plusieurs étapes. Le plus souvent, le gestionnaire et l'équipe projet formulent des observations de forme et de fond sur les différents rapports qui sont ensuite pris en charge par l'évaluateur. Par la suite, les rapports sont transmis aux parties prenantes pour recueillir leurs observations. Des ateliers techniques de présentation des rapports peuvent également être organisés pour recueillir l'avis des parties prenantes ou pour favoriser l'apprentissage et les échanges autour des résultats de l'évaluation. Lorsqu'elles sont pertinentes, les observations formulées doivent également être prises en charge par l'évaluateur.

A la suite de la validation des rapports, le gestionnaire doit entreprendre les actions nécessaires pour la suite des activités et éventuellement la clôture du projet.

2.2. Logistique de l'évaluation

Comme nous l'avons vu dans les sections précédentes, l'évaluation comporte une phase intense de collecte des données. C'est la raison pour laquelle cette section est dédiée à la logistique nécessaire pour assurer le bon déroulement de l'opération de collecte.

Si on exclut l'identification des parties prenantes qui est généralement réalisée dès la conception de l'intervention, l'opération de collecte des données pour l'évaluation se déroule en trois étapes. Premièrement, il faut informer l'ensemble des parties prenantes surtout les autorités administratives au niveau central (ex. ministère responsable de la sécurité intérieure et/ou de la gestion administrative) et local (préfet, sous-préfet, maire, etc.) de même que les autorités coutumières (chefs de villages) sur les dates et les objectifs des visites sur les sites. Ces séances d'information peuvent s'étendre aux organisations locales (non gouvernementales) qui peuvent être d'une grande utilité lors de la collecte des données. Idéalement, ce travail doit être effectué bien à l'avance, pour préparer l'arrivée de l'équipe d'évaluation sur les sites d'interventions. Deuxièmement, ce travail d'information auprès des autorités devra bien souvent être complété par un courrier ou un communiqué d'une page signé par un représentant du ministère hôte du projet, informant les communautés sur la prochaine visite des agents de collecte (encore une fois, les pratiques peuvent varier

d'un pays à l'autre, mais ce type de courrier reste très utilisé). Ce courrier ou le communiqué devra comprendre une brève description du but de la visite, les dates anticipées des visites sur les sites, et les noms des membres de l'équipe d'évaluation. Il pourra, si nécessaire, être traduit en langue locale. Troisièmement, il faut prévoir les outils indispensables au bon fonctionnement de la mission, comme la nourriture, le transport, l'espace pour le bureau, les équipements comme l'ordinateur, les dictaphones, des cahiers et l'imprimante, le logement et une boîte à pharmacie pour l'équipe. L'ensemble de ces éléments doivent être inclus au budget de la collecte.

2.3. Evaluation la qualité d'une évaluation

La qualité d'une évaluation porte sur plusieurs facteurs, dont (i) le respect des exigences en termes de livrable et de délai ; (ii) la cohérence des choix effectués ; (iii) la gestion des imprévus ; et (iv) l'analyse des biais éventuels entre les actions planifiées et celles réalisées. Il s'agit d'une vérification du processus ayant conduit au résultat final de l'évaluation.

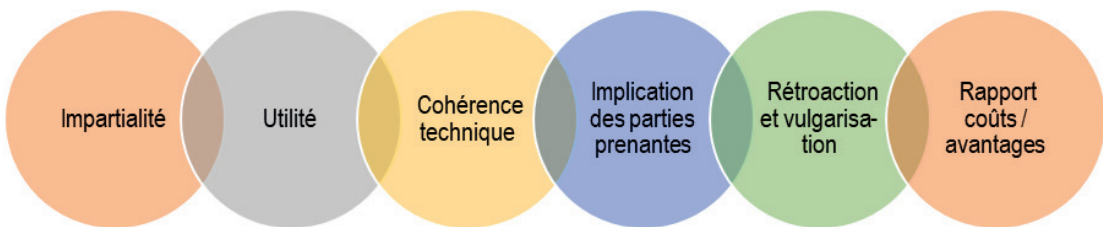
La qualité d'une évaluation prend en compte sa capacité à satisfaire les besoins et les exigences des parties prenantes ; c'est-à-dire à répondre aux questions soulevées au démarrage de la mission. L'évaluation doit être pertinente et réaliste, quant à son domaine d'intervention. Elle doit aussi utiliser des méthodes appropriées pour répondre aux questions d'intérêt. Elle doit se fonder sur des données fiables, exactes et valables. L'évaluation étant un exercice empirique, les données y jouent en effet un rôle central. C'est la raison pour laquelle, le rapport d'évaluation doit décrire avec une grande précision la stratégie de collecte des données et les traitements effectués.

L'évaluateur doit partager avec les lecteurs les normes de qualité qui régissent la collecte de données. Dans le cas d'une évaluation, le taux de non-réponses à un sondage et le taux d'attrition dans un échantillon peuvent être des indicateurs de mauvaise qualité de données, la formulation de question dans une entrevue, comme la présence de question orientée peut également révéler une méthodologie douteuse. Ces normes doivent constituer une partie importante de la formation des enquêteurs, s'ils sont utilisés et doivent être inscrites dans les manuels relatifs à l'opération de collecte. Il faut par exemple prévoir une stratégie très claire de remplacement (sous des conditions spécifiées) des individus qu'on ne retrouve pas afin de réduire les taux de non-réponse. En outre, il faut développer des procédures d'assurance-qualité à

toutes les étapes de la collecte des données ; c'est-à-dire applicables à la sélection des répondants, l'élaboration des outils de collecte, la préparation de la collecte, la collecte elle-même, (la saisie,) l'apurement des données et leur stockage. Des contrôles de qualité doivent être une priorité pendant les travaux de collecte des données sur le terrain et l'équipe de collecte doit être structurée de sorte à faciliter les différents contrôles de qualité de façon régulière.

A la fin de la collecte, les données doivent être analysées en utilisant des techniques appropriées en suivant une démarche rigoureuse et transparente. L'évaluateur doit faire preuve d'une impartialité dans l'analyse et la présentation des conclusions de l'étude. La Figure 12.2 présente sous une forme schématique différents éléments constitutifs de la qualité d'une bonne évaluation.

Figure 12-1: Qualités d'une bonne évaluation



Une autre approche permettant d'assurer une bonne qualité des données est d'utiliser une liste de contrôle ou de vérification. Les listes de contrôle sont de précieux outils d'évaluation lorsqu'elles sont soigneusement développées, validées et appliquées (Stufflebeam, 2000). Une bonne liste de contrôle pour l'évaluation (i) clarifie les critères à prendre en compte lors de l'évaluation ; (ii) aide l'évaluateur à ne pas oublier des critères importants ; et (iii) améliore l'objectivité, la crédibilité et la reproductibilité de l'évaluation.

Il existe plusieurs façons de développer des listes de contrôle. Néanmoins, il faut retenir que l'élaboration d'une liste de contrôle est un processus qui peut connaître des améliorations au fur et à mesure. Le lecteur est invité à consulter le manuel édité par Kellaghan et Stufflebeam (2003) ou les articles de Scriven (2007) et Stufflebeam (2000) pour l'élaboration et l'utilisation des listes de contrôle à des fins d'évaluation.

Messages clés :

- ✓ La planification d'une évaluation est un processus assez complexe, qui nécessite souvent l'élaboration d'un ensemble de documents contractuels entre un commanditaire (gestionnaire d'évaluation) et un évaluateur. Une évaluation bien planifiée a de meilleures chances d'apporter satisfaction aux deux parties.
- ✓ On distingue les évaluations internes, réalisées dans l'organisation responsable de la mise en œuvre de l'intervention, et les évaluations externes, prises en charge par un évaluateur indépendant. Chaque forme a ses avantages et inconvénients, et elles ont chacune leur place. Même si les évaluations internes et externes sont différentes à de nombreux points de vue, elles se gèrent de façon souvent similaire.
- ✓ Les responsabilités de gestion d'une mission d'évaluation se partagent entre le commanditaire de l'évaluation, souvent représenté par un gestionnaire d'évaluation pour l'organisation (il rassemble les préoccupations autour de l'intervention, rédige les documents de contractualisation, met en place une procédure de suivi et de contrôle qualité) et l'évaluateur, qui gère le projet en fonction des ressources et du temps qu'il a à sa disposition.

Bibliographie :

- Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. (2011). *L'évaluation d'impact en pratique*. The World Bank.
- Kellaghan, T., Stufflebeam, D. L. (2003). *International handbook of educational evaluation*. Springer.
- OCDE (2010). *Lignes directrices et ouvrages de référence du CAD : les normes de qualité pour l'évaluation du développement*. Organisation de Coopération et de Développement Économiques.
- Société Suisse d'Evaluation (SEVAL) (2014). Compétences requises pour les gestionnaires d'évaluation. Rapport du groupe de travail « compétence en évaluation ».
<https://www.bj.admin.ch/dam/bj/fr/data/staat/evaluation/netzwerk/2014-06-23-seval-f.pdf.download.pdf/2014-06-23-seval-f.pdf>
- Scriven, M. (2007). Key evaluation checklist. Consulté le 15 avril 2021 à partir de
<https://wmich.edu/sites/default/files/attachments/u350/2014/key%20evaluation%20checklist.pdf>
- Stufflebeam, D. L. (2000). Guidelines for developing evaluation checklists: the checklists development checklist (CDC). Consulté le 15 avril 2021 à partir de
https://wmich.edu/sites/default/files/attachments/u350/2014/guidelines_cdc.pdf

MODULE 13

EVALUATION DES INTERVENTIONS COMPLEXES

Objectifs d'apprentissage :

A la fin de la session, les participants seront capables de :

- Identifier les situations de complexité dans les interventions de développement
- Prendre en compte la complexité dans le dispositif d'évaluation
- Comprendre le déficit de capacités des acteurs impliqués dans l'évaluation des interventions complexes

Thèmes couverts

- V. Dimension des évaluations complexes : Caractéristiques, Principes et Stratégie d'intervention
- VI. Concept d'Evaluation du développement
- VII. Renforcement des Capacités d'évaluation en situation de complexité : Nécessité et Démarche

Pour ce dernier chapitre, il apparaît intéressant de se pencher sur l'évaluation d'un point de vue pratique, notamment dans le contexte des politiques publiques et des programmes qui font généralement l'objet d'évaluation en Afrique francophone. Au nombre de ceux-ci, on retrouve de nombreuses interventions visant le développement. Ce dont on se rend bien compte, c'est que pour l'utilisation de certains outils ou méthodes présentés dans ce manuel, la réalité est souvent plus difficile à appréhender et il est par exemple difficile de : démêler le rôle de nombreuses parties prenantes impliquées dans la mise en œuvre d'une intervention, de tracer une théorie de changement linéaire, d'identifier le réel problème à la base de l'intervention, de trouver des indicateurs d'effet pertinents, de planifier des collectes de données fonctionnelles ou encore de faire des recommandations qui peuvent avoir une réelle valeur ajoutée. Et ces difficultés s'ajoutent également à la nécessité de respecter des budgets limités, des échéanciers serrés, et de travailler dans des dynamiques politiques changeantes et incertaines.

Pour décrire ces situations, de nombreux auteurs ont réfléchi aux caractéristiques des interventions appelées *complexes*, et de leur influence sur la discipline de l'évaluation. Ce chapitre vise donc à définir cette notion de complexité, et ce qu'elle implique, pour pouvoir ainsi présenter des pistes de réflexion et de solution pour les évaluateurs confrontés à de telles interventions.

Mots-clés : complexité, développement, évaluation dans la vie réelle

I. Définition et caractéristiques d'une intervention complexe

Le sociologue Edgar Morin est souvent cité comme la référence de la pensée sur la complexité. Cité dans Pagani, et al. 2017, il explique que la réalité est « conçue comme essentiellement complexe, tout dont les éléments hétérogènes constitutifs s'enchevêtrent dans des maillages inextricables rendant son comportement, certes intelligible, mais non totalement algorithmique et prédictible ». Cette définition introduit parfaitement l'idée que les réalités complexes sont composées de multiples éléments différents, confondus et dont le fonctionnement d'ensemble est imprévisible. Dans cette compréhension, le tout est plus que la somme de ses éléments, c'est donc un objet d'étude en soi, ce qui n'exclut pas non plus d'analyser les éléments et leur comportement individuel. De nombreux auteurs de différentes disciplines ont donc étudié la complexité, et des liens intéressants peuvent être faits avec la théorie des systèmes voire le concept de chaos.

Plus concrètement, pour les interventions publiques qui nous intéressent, on peut citer la définition suivante : « la complexité d'une intervention réside dans le nombre de composantes qui agissent à la fois, de manière indépendante et interdépendante, le nombre et la difficulté des comportements requis par ceux qui fournissent et reçoivent l'intervention, le nombre et la variabilité des résultats, le nombre de groupes et de niveaux organisationnels ciblés par l'intervention, le degré de flexibilité ou d'adaptabilité de l'intervention » (Medical Research Council., 2000). Ce rapport du Medical Research Council est d'ailleurs une des sources les plus citées pour la réflexion sur la complexité des interventions en santé publique, et les implications de cette complexité dans leur gestion et leur évaluation.

On peut donc retenir que les interventions complexes se distinguent à la fois :

- des interventions simples qui s'attaquent à des problèmes simples et qui utilisent des théories de changement linéaires ;
- des interventions compliquées, qui regroupent généralement un ensemble d'interventions simples qui requièrent donc plus de coordination entre différentes parties prenantes et d'expertise de gestion sans pour autant être imprévisibles. (Marchal, Van Belle, De Brouwere, & Witter, 2013)

La complexité s'exprime ainsi à la fois comme une propriété de l'intervention et de ses composantes internes, mais également comme une propriété du système dans lequel l'intervention est mise en application, que l'on appelle généralement son contexte. Une intervention peut-être complexe ou peut être mise en œuvre dans un contexte complexe, et dans les cas des politiques de développement qui nous occupent souvent elle est les deux. On peut toutefois préciser les caractéristiques suivantes des interventions complexes :

- une diversité d'effets, prévus et non prévus, parfois complémentaires ou contradictoires qu'il est souvent difficile à circonscrire ou prédire ;
- une multiplicité d'acteurs, de parties prenantes et de bénéficiaires, qui rend le traçage des rôles et des responsabilités délicates, et la probabilité de consensus faible ;
- une théorie du changement diffuse, avec des flèches et des relations qui vont dans de nombreux sens (y compris des rétroactions et des cycles), et des liens de cause à effet flous ;
- des limites temporelles confuses, avec des continuums d'activités moins précis que le début et la fin d'un petit programme plus structuré
- un contexte d'implantation qui a une réelle influence sur l'intervention même, c'est-à-dire que l'intervention ne se comprend pas sans son contexte, et que ce contexte fait même partie de l'intervention.

Des exemples d'interventions complexes peuvent être des stratégies nationales comme un plan de lutte contre la pauvreté, ou sectorielles comme une politique de santé ou d'éducation. Il peut aussi s'agir d'interventions plus ciblées sur une thématique comme un plan d'action sur l'eau, une politique nationale sur le genre, ou une stratégie de prévention et d'adaptation aux changements climatiques. Parfois, c'est en voulant regrouper un ensemble d'intervention pour évaluer la performance globale que la complexité se manifeste. C'est par exemple le cas de l'évaluation d'un portefeuille d'interventions d'un bailleur de fonds dans un pays, tel que l'évaluation des plans-cadres des Nations Unies pour l'aide au développement (UNDAF).

II. Quelques pistes pour l'évaluation d'une intervention complexe

L'absence de prise en compte du caractère complexe d'une intervention dans son évaluation mène souvent à de nombreux échecs : difficultés à circonscrire la performance, conclusions erronées ou rejetées, et sous-utilisation des résultats par les décideurs. Plusieurs évaluateurs peuvent même, en appliquant des méthodes et des angles d'attaque différents arriver à des jugements en apparence contradictoires, même s'ils ne sont que des facettes de la même réalité. (Contandriopoulos, Rey, Brousselle, & Champagne, 2011). Il est donc nécessaire pour les évaluateurs de se familiariser avec ces concepts et d'adopter une pratique réflexive et évolutive de l'évaluation dans ces situations. On ne peut toutefois recommander une seule méthodologie d'évaluation qui fonctionnerait à coup sûr pour toutes les interventions complexes.

Certaines balises peuvent cependant être partagées pour guider les évaluateurs face à ces difficultés :

- Il est nécessaire de mieux réfléchir aux méthodes d'évaluation des effets dans le cadre de liens de causalité moins linéaires. Les procédures expérimentales sont bien souvent peu applicables en l'absence de non-bénéficiaires équivalents (Pagani, et al., 2017) et on préférera des méthodes plus qualitatives, visant l'analyse de la contribution d'une intervention à un changement, à la différence que cette intervention fait sur le terrain (Mayne , Making causal claims, 2012).
- L'usage de l'outil de la théorie de changement peut être maximisé : en ajoutant de nombreuses hypothèses et de facteurs externes, la dynamique complexe d'une intervention peut être mieux représentée. (Mayne , 2008) Il ne faut pas oublier qu'il s'agit d'un outil adaptable dans le temps et que des cases peuvent être ajoutées au fur et à mesure des investigations. Cet outil peut aussi être tracé avec différentes versions si différentes parties prenantes comprennent l'intervention de manière distincte. Une approche d'évaluation participative pourrait ensuite aider les acteurs à développer une vision plus complète et consensuelle de l'intervention.

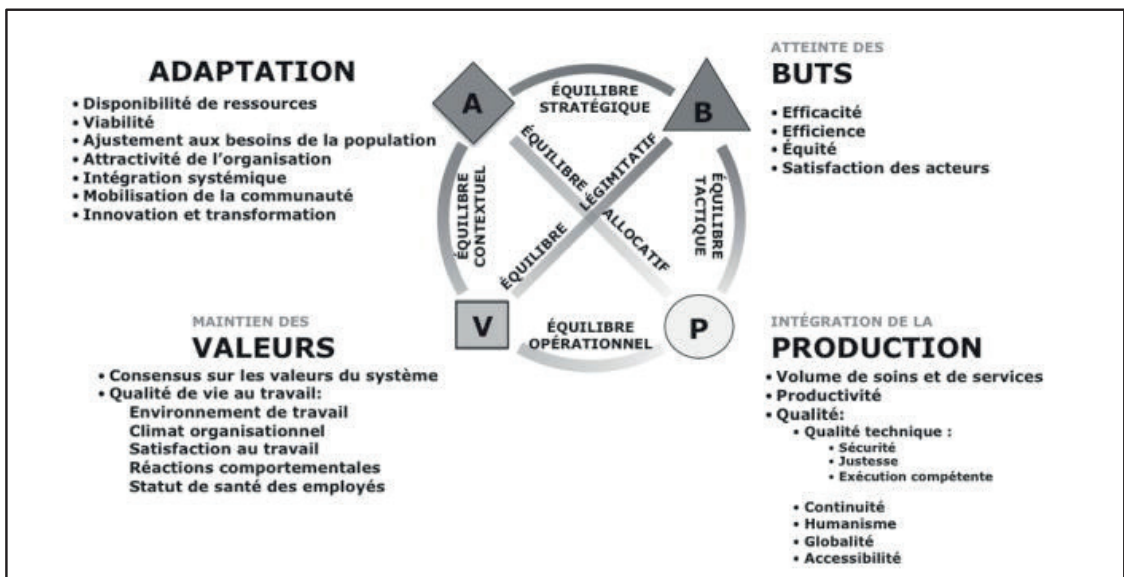
- Des théories de changement différentes peuvent aussi refléter des contextes d'implantation différents. Puisque le contexte influe fortement sur les réalisations, et les effets, l'analyse des succès et des échecs dans différents lieux de mise en œuvre peut permettre de retenir de bonnes pratiques de gestion à partager. Les approches réalistes d'évaluation et les études de cas peuvent ici s'avérer utiles.
- Les évaluateurs sont aussi invités à adopter une approche globale et holistique des interventions pour comprendre l'ensemble d'un système avec ses paradoxes et contradictions. Cela peut apporter des modifications dans le travail même d'évaluateur, qui peut être intégré à une équipe-projet, puisque les besoins d'information pour la gestion n'arrivent pas uniquement à mi-parcours ou à la fin de l'intervention. Les approches d'évaluation évolutive sont dans ce sens intéressantes, puisqu'elles permettent de sortir des résultats d'évaluation au besoin et en continu, ce qui correspond mieux à l'incertitude entourant les interventions complexes.
- Parfois, au-delà des approches utilisées par l'évaluateur, il s'agit aussi de développer les questions les plus pertinentes pour la gestion d'une intervention complexe, ce qui peut différer d'une intervention plus simple. On peut notamment s'intéresser à la collaboration et à l'approche partenariale qui a pu se mettre en place, à l'efficacité de la gouvernance d'ensemble, au degré de consensus entre les acteurs sur les visées de l'intervention.
- Une attitude générale plus souple est également nécessaire lorsqu'on se lance dans une évaluation dans un contexte complexe parce qu'elle fait souvent face à des contraintes d'ordre logistique (temps et budget pour les collectes de données par exemple) ou politique (comme le difficile consensus de parties prenantes) (Bamberger & Mabry, 2019).

On peut également pousser la réflexion un peu plus et s'inspirer des théories de la complexité pour développer des angles d'attaque d'évaluation intéressants et pertinents. Parmi cette littérature, on peut citer le modèle intégré d'évaluation des interventions complexes (Contandriopoulos, Rey, Brousselle, & Champagne, 2011), où les auteurs soutiennent qu'une intervention complexe performante doit remplir quatre fonctions :

- l'adaptation : c'est-à-dire qu'elle évolue en fonction de son contexte pour rester fonctionnelle à court terme et pertinente à long terme ;
- l'atteinte des buts : soit le fait que l'intervention puisse remplir ses objectifs ;
- l'intégration des processus de production : c'est la fonction plus habituelle de la performance, sa productivité ;
- le maintien et le développement des valeurs et de la culture : c'est-à-dire que l'intervention garde la cohésion de ses acteurs, qu'elle fait sens pour eux.

Ces quatre fonctions étant nécessaires à l'évaluation d'une intervention complexe, il y a lieu de maintenir un certain équilibre entre eux, comme le montre la figure ci-dessous :

Figure 13.1 : le modèle intégré d'évaluation des interventions complexes de Contandriopoulos, Rey, Brousselle, et Champagne (2011)



Ainsi, ce modèle peut nourrir la construction de la stratégie d'évaluation d'une intervention, en aidant l'évaluateur à se poser des questions particulières à la dynamique de la complexité.

Messages clés :

- ✓ La complexité est une composante inévitable des interventions publiques contemporaines, particulièrement dans un contexte humanitaire ou de développement. C'est pourquoi il est utile pour les évaluateurs de comprendre ses caractéristiques et ses conséquences pour le travail d'évaluation.
- ✓ Les outils et les méthodes d'évaluation peuvent parfois apparaître moins applicables face à des interventions complexes qui sont incertaines et multidimensionnelles. Toutefois, plusieurs pistes de solution existent autant via l'utilisation d'approches ou d'outils adaptés, ou via un usage différent des outils existants : de nouvelles questions d'évaluations, plusieurs théories de changement plutôt qu'une...
- ✓ De nombreux auteurs ont réfléchi à des approches renouvelées et adaptées de l'évaluation pour les interventions complexes. Les évaluateurs sont donc encouragés à être ouverts à ces innovations, et à toujours poursuivre leur apprentissage et leurs lectures sur le sujet.

Bibliographie :

- Bamberger, M., & Mabry, L. (2019). Real World Evaluation. SAGE Publications.
- Contandriopoulos, A.-P., Rey, L., Brousselle, A., & Champagne, F. (2011). Évaluer une intervention complexe : enjeux conceptuels, méthodologiques et opérationnels. *Revue canadienne d'évaluation de programme*, 26, pp. 1-16.
- Marchal, B., Van Belle, S., De Brouwere, V., & Witter, S. (2013). Studying complex interventions: reflections from the FEMHealth project on. *BMC Health Services Research*, 13, p. 469.
- Mayne, J. (2008). Contribution analysis: An approach to exploring cause and effect. *ILAC Briefs* (16).
- Mayne, J. (2012). Making causal claims. *ILAC Brief* (26).
- Medical Research Council. (2000). A framework for development and evaluation of RCTs for complex interventions to improve health.
- Pagani, V., Kivits, J., Minary, L., Cambon, L., Claudot, F., & Alla, F. (2017). La complexité : concept et enjeux pour les interventions de santé publique. *Santé Publique*, 29, pp. 31-39.



