



**USAID**  
FROM THE AMERICAN PEOPLE

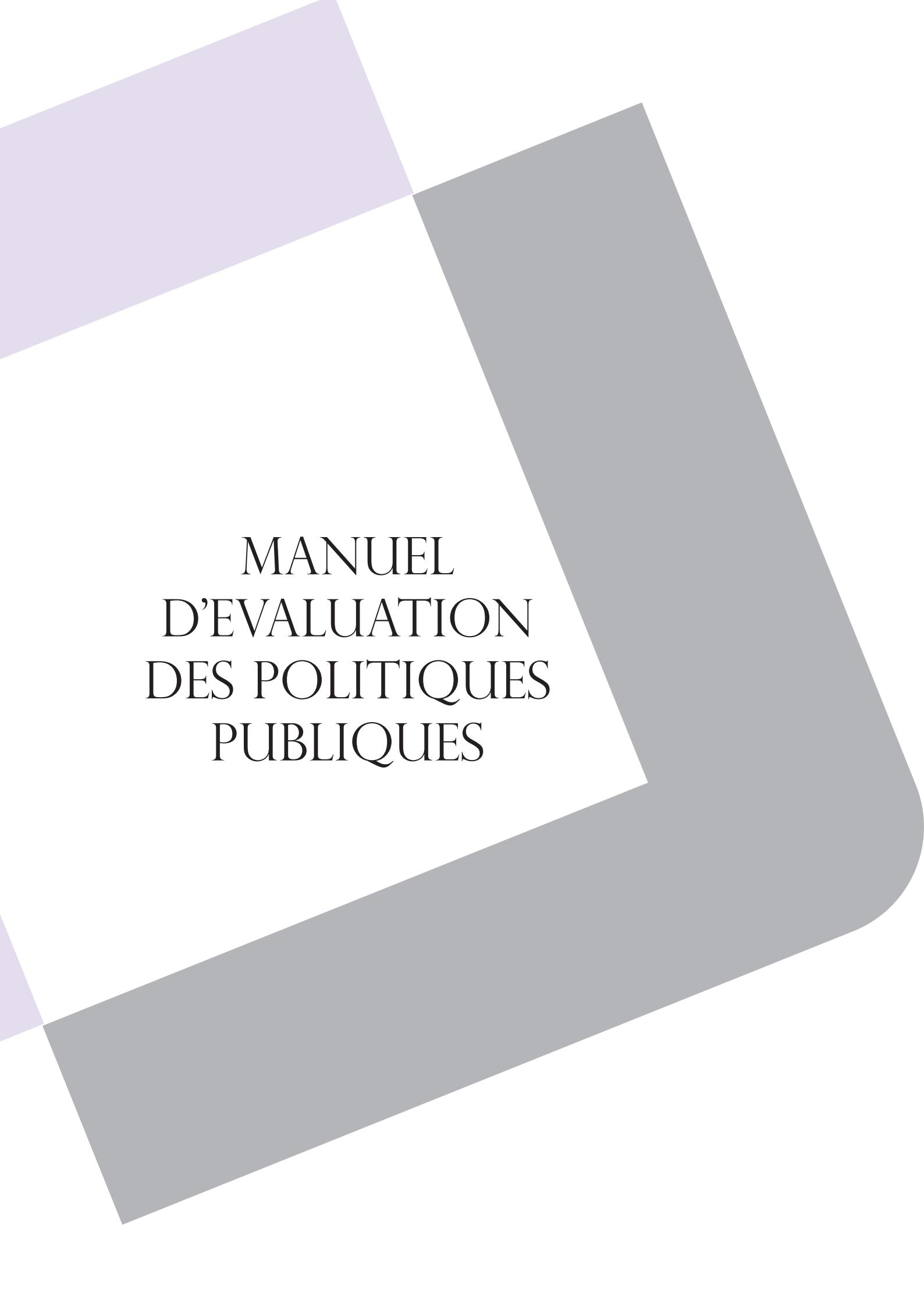


# MANUEL D'EVALUATION DES POLITIQUES PUBLIQUES

SECTION DES COMPTES DE LA COUR SUPREME DU MALI

CONTRÔLE GENERAL DES SERVICES PUBLICS DU MALI

OCTOBRE 2019



# MANUEL D'EVALUATION DES POLITIQUES PUBLIQUES

# TABLES DES MATIÈRES

Pages  
8-20

## PREMIÈRE PARTIE : CONNAITRE L'ÉVALUATION

### LES DÉFINITIONS

- La politique publique
- Le programme
- Le projet
- L'évaluation des politiques publiques
- Le socle de l'évaluation des politiques publiques : le programme
- Articulation entre les éléments constitutifs du programme et l'évaluation
- Le référentiel du programme et de l'évaluation
- La chaîne des résultats
- La typologie de l'évaluation
  - La typologie selon le moment de la réalisation
    - L'évaluation ex ante ou a priori
    - L'évaluation à mi-parcours
    - L'évaluation finale
    - L'évaluation ex post
  - La typologie selon les protagonistes
    - L'évaluation interne (ou auto-évaluation)
    - L'évaluation externe
    - L'évaluation mixte
    - L'évaluation participative
- L'évaluation des résultats/ L'évaluation d'impact : les enjeux
- Les critères d'évaluation
- L'évaluation versus l'audit, le contrôle de gestion, l'inspection
- Les finalités de l'évaluation
- L'évolution de l'évaluation
- La gestion axée sur les résultats et l'évaluation des politiques publiques
- Les enjeux éthiques
- Le processus évaluatif

3

Pages  
21-30

## DEUXIÈME PARTIE : PRÉPARER UNE ÉVALUATION

### LA COMMANDE (DEMANDE) DE L'ÉVALUATION

- Planifier l'évaluation
- Définir les rôles et responsabilités dans le processus évaluatif
- Explorer l'intervention
- Définir l'objectif de l'évaluation
  - Les questions d'évaluation :
    - *Les questions descriptives*
    - *Les questions normatives*
    - *Les questions de cause à effet*
- Définir les critères d'évaluation
- Choisir la méthode d'évaluation (évaluation des résultats)
  - Les méthodes quantitatives d'évaluation :
    - *La méthode expérimentale*
    - *La méthode quasi expérimentale*
    - *La méthode pré expérimentale*
    - *Les méthodes statistiques*
  - Les méthodes qualitatives d'évaluation :
    - *Le changement le plus significatif (MSC)*
    - *La cartographie des incidences*
  - Le mixte quantitatif/qualitatif
- Résumer la stratégie d'évaluation
- Rédiger ou amender les termes de référence
- Choisir l'équipe d'évaluation

# TABLES DES MATIÈRES

## TROISIÈME PARTIE : FAIRE UNE ÉVALUATION

### LA MISE EN ŒUVRE DE L'ÉVALUATION

Pages  
31-33

- Collecter les données
  - *les données secondaires*
  - *les données primaires*
- Collecter les données quantitatives :
  - *le questionnaire*
- Collecter les données qualitatives :
  - *l'entrevue (les entretiens)*
  - *l'entretien libre ou non-directif*
  - *l'entretien semi-directif*
  - *l'entretien directif*
  - *l'observation directe*
  - *le focus group*
- Traiter et analyser les données
- Dégager les conclusions

## QUATRIÈME PARTIE : PRÉSENTER ET COMMUNIQUER LES RÉSULTATS DE L'ÉVALUATION :

### LE RAPPORT

Pages  
35-36

- Rediger le rapport
  - *le résumé analytique ou résumé exécutif*
  - *l'objet et le but de l'évaluation*
  - *les questions et critères d'évaluation*
  - *la méthodologie*
  - *les résultats*
  - *les conclusions*
  - *les recommandations*
  - *les annexes*
- Diffuser les résultats
  - *identifier les difficultés de communication et de compte rendu*
  - *définir l'objectif de communication*
  - *choisir les meilleures méthodes de communication*
  - *élaborer une stratégie de communication*

### Figures

*Figure 1 : Schéma de l'articulation logique entre politique, programme et projet*

*Figure 2 : Schéma de la chaîne des résultats*

*Figure 3 : Schéma du cycle de l'évaluation*

*Figure 4 : Schéma de la théorie du changement*

### Tableaux

*Tableau 1 : Différences entre le contrôle, l'audit, le contrôle de gestion, l'inspection et l'évaluation*

*Tableau 2 : Récapitulatif des principales tâches à accomplir dans le cadre de la préparation d'une évaluation et des outils et/ou techniques adaptés à chacune des tâches identifiées.*

*Tableau 3 : Récapitulatif des principales tâches à accomplir dans le cadre de la mise en œuvre d'une évaluation et des outils et/ou techniques adaptés à chacune des tâches identifiées.*

### Annexe 1

Exemple de matrice d'évaluation portant sur une question évaluative

### Annexe 2 :

*Rôles clés et responsabilités dans le processus d'évaluation*

### Annexe 3 :

*Deux exemples de logiciel de traitement des données :*

### Annexe 4

*Les méthodes expérimentales et quasi expérimentales*

### Annexe 5

*Quelques outils de support à l'évaluation*

### Annexe 6

*Techniques d'échantillonnage*

### Annexe 7

*TDR*

# INTRODUCTION

La Section des Comptes de la Cour Suprême du Mali compte, depuis 2016, l'évaluation des politiques publiques parmi ses prérogatives de contrôle non juridictionnel.

En effet, la loi n°2016-046 du 23 septembre 2016, qui prend en compte la dynamique de réforme des finances publiques de l'UEMOA à travers ses directives adoptées en 2009, attribue à la Section des Comptes une nouvelle fonction qui consiste en l'évaluation des politiques publiques.

L'article 335 de cette loi en fixe les contours : « L'évaluation des politiques publiques a pour objet d'en apprécier les résultats à partir de ses divers impacts et des liens de causalité propres à les expliquer. Sa finalité est d'éclairer la prise de décision et d'informer les citoyens. »

L'objectif de ce manuel d'évaluation des politiques publiques pour le Contrôle Général des Services Publics est de permettre aux contrôleurs qui ne sont pas des spécialistes en évaluation, d'une part, une compréhension des concepts et principes généraux de l'évaluation, et, d'autre part, une appropriation de la démarche pratique de conduite d'une évaluation.

Les contours de l'évaluation des politiques publiques ainsi définis pour la Section des Comptes, l'approche d'évaluation des résultats ainsi que les méthodes, les outils et les techniques qui s'y rattachent sont mis en avant dans le présent manuel d'évaluation.

La volonté politique d'étoffer le dispositif national d'évaluation des politiques publiques du Mali est en cohérence avec le contexte de la gestion axée sur les résultats, d'adoption de l'approche « budget-programme » et de mise en œuvre de politiques publiques majeures en lien avec « La vision Mali 2025 » et le Cadre Stratégique de Lutte contre la Pauvreté (CSLP). Ces différents chantiers justifient le besoin de renforcer l'outillage en évaluation des politiques publiques des acteurs évaluatifs nationaux. C'est dans cet objectif que s'inscrit ce manuel qui a été conçu de manière participative.

## LISTE DES SIGLES ET ABREVIATIONS

<b>AFD</b>	Agence Française de Développement
<b>CSLP</b>	Cadre Stratégique de Lutte contre la Pauvreté
<b>GAR</b>	Gestion Axée sur les Résultats
<b>FIDA</b>	Fonds International pour le Développement Agricole
<b>LOLF</b>	Loi Organique Relative aux Lois des Finances
<b>MSC</b>	Most Significant Change
<b>OCDE</b>	Organisation de coopération et de développement économiques
<b>PNUD</b>	Programme des Nations Unies pour le Développement
<b>SCCSP</b>	Section des Comptes de la Cour Suprême
<b>SFE</b>	Société Française de l'Evaluation
<b>UEMOA</b>	Union Economique et Monétaire Ouest Africaine
<b>UNEG</b>	United Nations Evaluation Group
<b>USAID</b>	United States Agency for International Development

# 1<sup>ERE</sup> PARTIE : CONNAITRE L'ÉVALUATION

# LES DÉFINITIONS

## LA POLITIQUE PUBLIQUE

On entend par politique publique une décision prise par les institutions de l'État central et/ou les collectivités dans le souci de l'intérêt général et dotée de moyens institutionnels, légaux, juridiques et financiers, en vue d'intervenir de manière durable et structurante dans l'espace public.

La politique publique peut être perçue comme un ensemble de programmes complémentaires soutenus par une série de projets structurés autour de diverses actions.

8

## LE PROGRAMME

Un programme est un ensemble cohérent et structuré d'objectifs, de ressources (humaines, financières, matérielles et informationnelles) et d'activités permettant la production de biens et services spécifiques répondant à un ou plusieurs besoins d'une population ciblée. La signification du terme programme s'étend à tout projet, service, initiative ou plan d'actions gouvernemental visant à résoudre un problème ou à répondre à un besoin de l'ensemble ou d'une partie de la société.

## Message 1

Il n'existe pas de définition universellement admise des notions et concepts qui entourent l'évaluation des politiques publiques. La compréhension de la réalité qu'en-globe chaque notion et concept est le principal enjeu.

## LE PROJET

Le terme projet est un mot courant du vocabulaire auquel on associe une signification plus ou moins claire relative à un ensemble d'actions envisagées pour atteindre un but (projet d'études, de vacances, de carrière, de mariage, etc.). Dans le domaine des politiques et programmes de développement, il se définit comme une opération ponctuelle qui consiste à des actions coordonnées, ayant un début et une fin, mettant en œuvre des ressources humaines et matérielles, faisant l'objet d'une budgétisation et devant aboutir à un ensemble de livrables.

Un projet est processus qui consiste à un ensemble d'actions coordonnées et maîtrisées comportant des dates de début et de fin dans le but d'atteindre un objectif conforme à des exigences spécifiques telles que des contraintes de délais, de coûts et de ressources. (ISO 10006).

En conclusion, il est fréquent que la politique publique soit déclinée en un ou plusieurs programmes qui sont mis en œuvre à travers des projets spécifiques. Ainsi, le projet renvoie aux mêmes objectifs que le programme, à savoir répondre à un ou plusieurs besoins d'une population, de sorte qu'un programme peut être composé de plusieurs projets mis en œuvre dans des espaces, pour des durées et des cibles plus réduits.

# L'ÉVALUATION DES POLITIQUES PUBLIQUES

L'évaluation d'une politique publique consiste en une appréciation objective d'une politique publique avant sa mise en œuvre, au cours de sa réalisation ou après son achèvement.

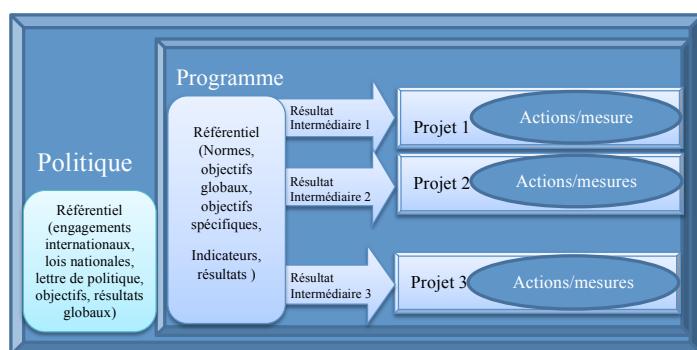
L'évaluation permet de répondre à des questions précises liées à la conception, à la mise en œuvre ou aux résultats des politiques publiques. Elle est donc un examen systématique de la valeur d'une politique publique et un moyen pour indiquer les succès et les échecs, ainsi que les leçons apprises et les façons d'améliorer une intervention .

Contrairement au suivi qui est continu et interne, les évaluations sont la plupart du temps périodiques et effectuées à un moment donné, généralement par des spécialistes extérieurs à la politique publique observée. Elles peuvent toutefois être effectuées de façon continue.

L'évaluation des politiques publiques revêt trois (3) aspects essentiels :

- L'évaluation est systématique, c'est-à-dire qu'elle est effectuée selon un plan, lequel est idéalement conçu en même temps que l'élaboration du programme public qui fait office de politique publique;
- L'évaluation se fonde sur des preuves, valides et fiables de la valeur et de la performance du programme public;
- L'évaluation produit des enseignements qui aident à prendre des décisions publiques par rapport au programme public évalué et des décisions opérationnelles (mesures de correction, etc.).

*Figure 1 : Schéma de l'articulation logique entre politique, programme et projet*



9

## Message 2

Le programme se présente comme nouveau socle de l'évaluation des politiques publiques dans le contexte de la gestion axée sur les résultats et de la réforme des finances publiques dans l'espace UEMOA inspirée par la LOLF adoptée en France en 2001 et marquée par l'approche dite « budget-programme ».

# LE SOCLE DE L'ÉVALUATION DES POLITIQUES PUBLIQUES: LE PROGRAMME

*Les contours du programme et l'évolution récente de la politique à programme*

Selon la LOLF de 2001, un programme regroupe les crédits destinés à mettre en œuvre une action ou un ensemble cohérent d'actions relevant d'un même ministère et auquel sont associés des objectifs précis, définis en fonction de finalités d'intérêt général, ainsi que des résultats attendus et faisant l'objet d'une évaluation. LOLF, art. 7).

Ainsi, chaque programme devra comporter, en annexe, la présentation des actions, des coûts associés, des objectifs poursuivis, des résultats obtenus et attendus pour les années à venir, mesurés au moyen d'indicateurs précis dont le choix est justifié (art.51).

La LOLF semble ainsi souligner « l'évaluabilité » du programme par opposition à la politique publique dans son sens large.

10

En France, les prémisses de l'évaluation remontent à la fin des années 1970 avec la mise en place de la rationalisation des choix budgétaires (RCB). Après l'échec de cette tentative, la fin des années 1980 a pu sembler décisive pour l'évaluation avec la politique de renouveau du service public décidée par le gouvernement de Michel Rocard, suite au rapport de Patrick Viveret intitulé L'évaluation des politiques et des actions publiques.

Michel Rocard lance un dispositif d'évaluation interministériel par le décret du 22 janvier 1990, en instituant un Conseil Scientifique de l'Évaluation (CSE). Ce dispositif n'a pas répondu aux attentes, d'où une relance de l'évaluation à la fin des années 90, avec la création du Conseil National de l'Évaluation (Décret de 1998) supprimant le Conseil Scientifique de l'Évaluation, et la création, en 1999, de la Mission d'Évaluation et de Contrôle instituée par la Commission des Finances ainsi que la création de la Société Française d'Évaluation (SFE) afin de promouvoir l'activité d'évaluation et la profession d'évaluateur.

Une distinction hiérarchique est établie entre la mission rattachée à une politique publique définie, le programme qui est l'unité intermédiaire que surplombe la mission et qui domine l'action, l'action étant l'unité concrète de la dépense publique.

L'ensemble des documents budgétaires sont dorénavant présentés par programmes, ce qui permet de mieux mesurer les résultats de l'action publique. Pour chaque programme sont identifiés : une finalité d'intérêt général, des acteurs, des objectifs et des indicateurs permettant de mesurer les résultats.

## Message 3

Avec la réforme des finances publiques dans l'espace UEMOA, fortement inspirées par la LOLF de 2001 en France, les documents budgétaires du Mali sont maintenant présentés dans une logique de programme (budget-programme).

## ARTICULATION ENTRE LES ÉLÉMENS CONSTITUTIFS DU PROGRAMME ET L'ÉVALUATION

Dans la LOLF de 2001 qui consacre la gestion axée sur les résultats (GAR), la notion de programme et ses composantes structurent l'évaluation. La GAR est une approche globale qui vise à définir des résultats mesurables et des méthodes et outils à utiliser pour obtenir et mesurer ces résultats. Il s'agit de planifier, gérer et apprécier une intervention en se concentrant sur les résultats escomptés.

Le terme programme couvre un ensemble de projets concourant à un même objectif, intégrés et organisés de manière transversale ou interactive. Ainsi, il existe plusieurs définitions du programme qui décrivent toutes les mêmes éléments constitutifs : des ressources organisées, des activités ou des services réalisés par des acteurs, destinés à des cibles ou bénéficiaires situés sur un territoire, des objectifs liés à un problème, un contexte donné et des résultats.

C'est dans la phase de conception que l'évaluation trouve sa légitimité ou son référentiel, les ressources et la fréquence de sa mise en œuvre, ainsi que ses finalités d'amélioration, de reddition des comptes et d'apprentissage. En effet, les différentes phases qui entrent en jeu dans la planification du programme font appel à l'évaluation, à savoir :

- L'appréciation initiale : elle permet de comprendre la situation en cours, la nature du problème ou du besoin qui nécessite l'intervention, la politique publique à travers un programme; elle peut être basée sur une étude évaluative dite évaluation ex ante.

• La planification : elle est la phase de définition des résultats attendus de l'intervention ou du programme destiné à prendre en charge le problème ou le besoin (objectifs), les apports et les activités nécessaires pour produire ces résultats, les indicateurs servant à mesurer ces résultats et les hypothèses clés qui peuvent influer sur l'obtention des résultats. La planification prend en compte les besoins, les résultats, les ressources, les mandats et les capacités de l'institution ou organisation responsable et des parties prenantes. Cette planification inclue la planification des évaluations du programme.

• La mise en œuvre : pendant la mise en œuvre, des moyens sont mobilisés et des activités déployées afin d'atteindre les résultats. Il est recommandé de mettre en place un ou des systèmes de suivi durant la phase de planification, de sorte à pouvoir recueillir des informations et des données sur les progrès accomplis dans le sens de l'atteinte des résultats. Dans ce cas, les rapports intermédiaires, l'évaluation à mi-parcours ou les évaluations périodiques (annuelles) peuvent guider la décision sur la nécessité ou non de modifier ou d'adapter l'intervention à mesure que la situation évolue.

• Les résultats : La phase d'évaluation doit être conçue dès la planification, dans le but de déterminer la pertinence et l'accomplissement des objectifs, la productivité en termes de développement, l'efficacité, l'impact et la viabilité, selon les critères de l'OCDE/CAD. Un programme peut faire l'objet de plusieurs évaluations, selon sa nature, son référentiel et la façon dont il a été conçu, programmé et mis en œuvre .

La structuration du programme selon la loi et les modalités de sa mise en œuvre définies depuis les phases de conception et de planification détermine ainsi les objectifs et l'utilité de son évaluation. Cette structuration du programme implique (au minimum) la définition des résultats attendus, des indicateurs de ces résultats selon les normes et critères de l'évaluation, des moyens choisis pour mesurer ces résultats et des périodes de collecte des données et informations devant servir de preuve aux progrès réalisés.

#### Message 4

L'évaluation doit être prise en compte dès la planification du programme afin de disposer de suffisamment de temps pour la préparation et la réalisation, et aussi pour en assurer le financement.

La planification de l'évaluation durant la phase d'élaboration du programme permet de renforcer la conception même du programme par une approche plus rigoureuse des résultats escomptés et des activités à mettre en œuvre dans cette optique. La planification précoce de l'évaluation renforce aussi le système de collecte et de stockage des données de suivi à mesure que se déroule le programme, ainsi que la communication avec les parties prenantes dont l'intérêt pourra être très tôt porté sur le plan d'évaluation.

## LE RÉFÉRENTIEL DU PROGRAMME ET DE L'ÉVALUATION

Le référentiel du programme à évaluer est le repère initial dans la conduite d'une évaluation. Ce référentiel définit les critères et les normes auxquels devraient être confrontés les résultats observés au cours de l'évaluation.

Le référentiel d'une évaluation est constitué d'un ensemble de références explicites et organisées, réunies ou non en un document, et devant servir d'instrument pour l'évaluation. Ce référentiel peut être constitué des missions confiées à l'institution en charge de la politique publique, des exigences que cette institution s'impose, des objectifs de la politique publique, des normes réglementaires des procédures internes et/ou valeurs qui servent de référence pour attester des conformités et performances dans le cadre de l'évaluation.

Élaborer le référentiel d'une évaluation implique une relecture des objectifs de la politique publique à évaluer, voire leur interprétation et leur hiérarchisation en fonction des priorités politiques de l'heure.

Ces objectifs doivent surtout être articulés pour faire apparaître les liens logiques entre les différents niveaux d'objectifs (généraux, spécifiques, opérationnels, intermédiaires) et permettre leur traduction en des critères d'évaluation observables ou mesurables au moyen d'indicateurs précis. Ainsi, la constitution du référentiel implique :

- a) La connaissance des missions de l'institution en charge de la politique ou du programme à évaluer ;
- b) La définition des contours précis de la politique à évaluer ;
- c) La reconstitution de la logique d'intervention du programme ou de l'ensemble des mesures à étudier ;
- d) L'identification de la finalité de l'évaluation et du type d'évaluation à mener ;
- e) La sélection des questions auxquelles l'évaluation devra répondre (les questions évaluatives) ;
- f) La détermination des hypothèses de départ de l'évaluation ;
- g) Le choix des normes et critères de référence de l'évaluation ;
- h) Le choix des modes de communication autour de l'évaluation.

12

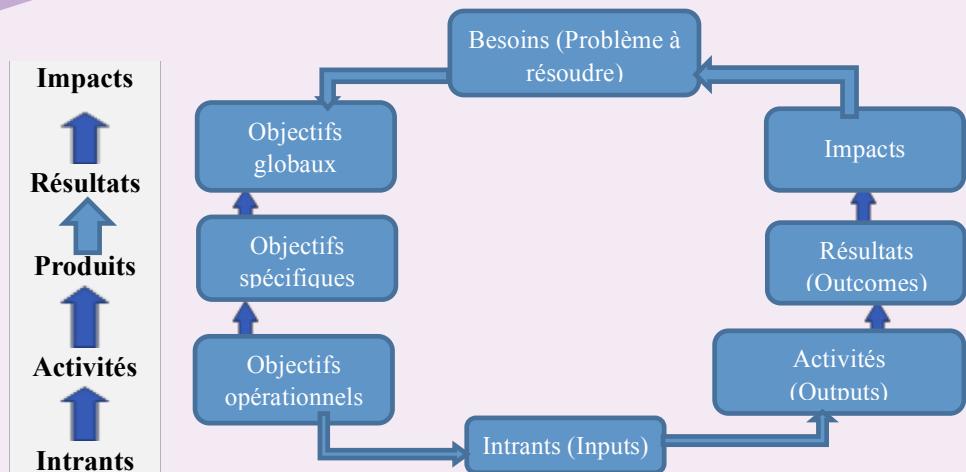
### Message 5

Le programme, comme unité de structuration des politiques publiques, offre à l'évaluation une base de travail tangible à travers les éléments de sa conception.

## LA CHAINE DES RÉSULTATS

La chaîne des résultats est le processus par lequel des intrants sont mobilisés pour développer des activités en vue de la réalisation de produits, lesquels doivent produire des résultats à court terme et des impacts à moyen et long terme pour répondre à des besoins initialement identifiés.

Figure 2 : Schéma de la chaîne des résultats



# LA TYPOLOGIE DE L'ÉVALUATION

Deux critères de classification peuvent être utilisés pour établir une typologie des évaluations : il s'agit, d'une part, du moment de réalisation et, d'autre part, des protagonistes.

## La typologie selon le moment de la réalisation

L'évaluation peut s'inscrire à différents moments de la mise en œuvre d'une politique publique. On distingue : l'évaluation ex ante, l'évaluation en cours d'exécution (souvent à mi-parcours et parfois annuelle), l'évaluation finale et l'évaluation ex post.

### *L'évaluation ex ante ou a priori*

L'évaluation ex ante intervient lors de la phase de conception, pour définir le cadre d'intervention pertinent au regard des besoins du territoire et de la population et identifier les effets et les impacts attendus d'une politique publique envisagée. Elle met en exergue les différents scénarios de mise en œuvre possible et d'optimiser les ressources à mobiliser.

L'évaluation ex ante peut notamment s'appuyer sur les expériences comparatives dans d'autres contextes (benchmark). L'OCDE/DAC définit l'évaluation ex ante comme une appréciation globale de la pertinence, de la faisabilité et de la durabilité probable d'une action de développement avant que la décision de financement ne soit prise.

### *L'évaluation à mi-parcours*

L'évaluation à mi-parcours (ou évaluation intermédiaire, évaluation concurrente, évaluation concomitante ou évaluation on-going) se situe pendant la mise en œuvre d'une politique publique. Elle est aussi dite évaluation formative ou évaluation in itinere dans la mesure où elle se déroule pendant que le programme est en cours.

L'évaluation à mi-parcours vise à connaître :

- si les objectifs seront atteints, et s'il existe ou non des effets collatéraux, indésirables ou non prévus;
- si les hypothèses utilisées lors de la formulation et la logique d'intervention sont toujours valides au fur et à mesure que l'intervention se déroule;
- quel résultat global peut-on attendre de l'exécution de la politique ?
- l'organisation et les moyens de gestion, les ressources humaines et matérielles ainsi que les relations interinstitutionnelles;
- l'obtention et le mode d'utilisation des ressources nécessaires pour l'intervention;
- les avancements réalisés en termes de services rendus aux bénéficiaires;
- les progrès au niveau des infrastructures et leur qualité;
- les prévisions existantes sur la future intervention de la politique.

13

## Message 6

Tout comme les notions et les concepts qui entourent l'évaluation des politiques publiques, il n'y a pas de typologie unanimement partagée en rapport avec l'évaluation.

L'évaluation à mi-parcours doit permettre de faciliter l'adoption à temps des réorientations et de faire le point entre les responsables de l'exécution et ceux qui financent l'intervention, et elle est souvent utilisée dans la décision de ces derniers pour procéder à des transferts de nouveaux fonds.

#### *L'évaluation finale*

L'évaluation finale se déroule à la fin de l'intervention. Elle examine les résultats et tire des observations sur les effets à court terme. Elle favorise l'examen des différents éléments de la politique publique, vérifie les effets et les impacts possibles en accordant une attention particulière à la durabilité des bénéfices produits et à la rentabilité de l'intervention en des termes aussi bien économiques que sociaux.

#### *L'évaluation ex post*

L'évaluation ex post d'une politique publique se déroule nettement après la fin de l'intervention et permet d'en apprécier les effets à moyen et long terme à travers le critère de l'utilité et s'intéresse aux facteurs de succès et d'échec ainsi qu'à la durabilité des effets et des impacts. Elle tire des conclusions qui pourraient être généralisées à d'autres politiques de développement. Elle est alors utilisée pour décider du renouvellement ou non d'une politique publique ou pour définir les améliorations à y apporter.

L'évaluation ex post doit inclure une analyse de la soutenabilité avérée de la politique publique et de sa capacité à consolider les transformations obtenues. Le moment précis et l'opportunité pour réaliser ce type d'évaluation dépendent essentiellement des caractéristiques de la politique publique à évaluer.

## La typologie selon les protagonistes

#### *L'évaluation interne (ou auto-évaluation)*

14

L'évaluation interne est conduite par un service et/ou des personnes qui dépendent de l'action de développement au sein des organisations chargées de la mise en œuvre ou chez le bailleur de fonds, chez ses partenaires (CAD/OCDE). Elle est réalisée par les propres responsables de la gestion de la politique publique. Les évaluations internes s'effectuent habituellement pendant la phase d'exécution de l'intervention.

Généralement moins coûteuses que les évaluations externes, elles présentent également l'avantage de favoriser plus aisément le processus d'apprentissage.

#### *L'évaluation externe*

L'évaluation externe d'une politique publique est conduite par les services et/ou des experts extérieurs à l'organisation responsable de la mise en œuvre directe de l'intervention.

Ce modèle est plus communément utilisé, du fait de sa présumable objectivité dérivée de son caractère indépendant et de la possibilité de disposer de compétences et d'expertise de haut niveau. Le coût élevé de l'évaluation externe est un facteur à prendre en compte.

#### *L'évaluation mixte*

L'évaluation mixte est menée conjointement par les responsables de l'intervention (équipe interne) et une équipe externe pour obtenir une évaluation globale qui intègre les deux perspectives.

L'idée qui est souvent à la base de l'évaluation mixte est de combiner les points de vue internes à l'intervention avec ceux d'évaluateurs externes. Cette approche contribue au renforcement des capacités institutionnelles.

#### *L'évaluation participative*

L'évaluation participative est une approche novatrice d'évaluation selon laquelle les représentants des agences d'aide et des autres parties (y compris les bénéficiaires) collaborent pour concevoir et conduire une évaluation et en tirer les conclusions.

## L'ÉVALUATION DES RÉSULTATS/ L'ÉVALUATION D'IMPACT : LES ENJEUX

Selon la Charte de la Société Française d'Évaluation (SFE), l'évaluation vise à produire des connaissances sur les actions publiques, notamment quant à leurs effets dans le double but de permettre aux citoyens d'en apprécier la valeur et d'aider les décideurs à en améliorer la pertinence, l'efficacité, l'efficience, la cohérence et les impacts.

Dans le contexte de la GAR, l'évaluation des résultats des politiques publiques revêt tout son sens. Elle crée les conditions de production de résultats à travers les politiques publiques et répond ainsi à une attente citoyenne légitime vis-à-vis de ces politiques.

L'évaluation d'impact s'inscrit dans la tendance internationale de l'élaboration de politiques fondées sur des preuves et de la gestion axée sur les résultats, une tendance internationale en pleine croissance. Cette tendance internationale accorde une attention particulière aux résultats, au-delà des intrants et des moyens mobilisés.

L'évaluation d'impact étudie les changements causés par une politique publique ou toute autre action de développement. Contrairement aux évaluations générales qui peuvent répondre à plusieurs types de questions, l'évaluation d'impact est structurée autour d'un type particulier de question : quel est l'impact (ou effet causal) d'une politique publique sur un résultat donné?

Ainsi, dans une évaluation d'impact, les résultats attendus de l'intervention publique ont, certes, une grande importance, mais les résultats imprévus sont aussi tenus en compte. Les impacts sont considérés généralement comme se produisant plus tard que les effets immédiats dont ils résultent.

La dimension causale est primordiale dans l'évaluation d'impact et porte sur les changements au niveau des résultats causés directement par la politique publique. En effet, les impacts sont définis comme les effets positifs ou négatifs, prévus ou imprévus, directs ou indirects, primaires ou secondaires produits par une intervention (CAD/OCDE).

Selon l'USAID, l'évaluation d'impact en tant que mesure du changement basé sur des modèles de cause à effet requiert un scénario contrefactuel crédible et rigoureusement défini pour contrôler les facteurs

### Message 7

Pour la Section des Comptes, « L'évaluation des politiques publiques a pour objet d'en apprécier les résultats à partir de ses divers impacts et des liens de causalité propres à les expliquer. Sa finalité est d'éclairer la prise de décision et d'informer les citoyens. » (article 335 de la loi n°2016-046 du 23 septembre 2016)

15

### Message 8

L'évaluation des résultats est l'approche d'évaluation qui fait le plus appel à des techniques, des méthodes et des outils spécifiques permettant de mettre en évidence les liens de causalité, la preuve, le résultat.

autres que l'intervention, qui pourraient expliquer le changement.

L'évaluation d'impact vise ainsi à déterminer quels changements peuvent être directement et exclusivement attribués à l'intervention.

L'évaluation d'impact est indispensable pour :

- Orienter les décisions politiques, notamment pour décider de la continuation ou de l'expansion d'une politique publique ayant fait les preuves de son utilité, de l'arrêt d'une intervention inefficace, de l'ajustement des bénéfices ou de la sélection entre différentes options programmatiques;
- Analyser l'efficacité d'une politique publique, notamment à travers la comparaison d'une catégorie ciblée par l'intervention (groupe de traitement) avec une autre catégorie qui n'y a pas participé (groupe témoin).

L'évaluation d'impact s'avère particulièrement utile dans le cadre d'interventions pilotes qui mettent en œuvre de nouvelles approches prometteuses.

L'évaluation d'impact se justifie surtout lorsque la politique considérée est :

- novatrice, c'est-à-dire qu'elle permet de tester une nouvelle approche,
- reproductible, c'est-à-dire lorsque l'intervention peut être reproduite dans un autre contexte,
- stratégiquement pertinente, notamment dans la mesure où la politique publique nécessite des ressources importantes, couvre un grand nombre de bénéficiaires, permet de faire des économies importantes et/ou peut produire des changements de comportement, d'attitudes ou de pratiques décisifs,
- influente, c'est-à-dire lorsque les résultats permettront d'orienter des décisions politiques majeures.

16

#### Message 9

L'évaluation des résultats (effets) fait appel aux critères en lien avec l'intervention (efficacité, efficience de l'intervention, etc.). Ces mêmes critères ont leurs pendants opérationnels qui servent à l'évaluation de la mise en œuvre d'une intervention (efficacité, efficience opérationnelle, etc.).

## LES CRITÈRES D'ÉVALUATION

Quelques critères d'évaluation :

### • La pertinence

Une politique publique se justifie par l'identification d'un besoin à satisfaire ou d'un problème à résoudre auquel les pouvoirs publics se sentent tenus de faire face. Une politique est considérée comme pertinente si ses objectifs ou sa théorie de changement sont adaptés à la nature du (des) besoins ou problème(s) qu'elle est censée prendre en charge.

Ainsi, le critère de pertinence permet de s'intéresser à la raison d'être de la politique publique, à sa conformité aux besoins exprimés, aux problèmes à l'origine, aux objectifs et au contexte.

Il vise à évaluer l'adéquation des objectifs de l'intervention aux besoins identifiés et s'adresse ainsi au fondement même de la politique publique évaluée. L'analyse de la pertinence d'une politique publique peut donc entraîner une remise en cause forte de l'intervention et des préconisations pouvant réorienter le dispositif ou en entraîner l'arrêt de la politique publique.

Le critère de la pertinence répond à une question telle que : dans quelle mesure les objectifs répondent-ils aux besoins du territoire et des populations?

### • La cohérence

Une politique publique est constituée d'un ensemble d'objectifs, de dispositifs réglementaires et organisationnels et de moyens humains et financiers. On peut distinguer la cohérence interne de la cohérence externe.

La cohérence interne permet d'évaluer si les actions et les moyens mis à disposition sont en mesure de permettre l'atteinte des objectifs. Ce critère permet d'étudier la complémentarité des actions évaluées avec les autres actions menées au sein du territoire ou ciblant les mêmes populations. La cohérence externe permet d'évaluer la complémentarité entre la politique publique évaluée et celle d'autres acteurs publics ou privés portant sur la même thématique.

Le critère de cohérence d'évaluation est particulièrement intéressant pour évaluer les interventions publiques transversales qui couvrent plusieurs directions ainsi que les politiques publiques impliquant de nombreux partenaires. Il répond à une question telle que : dans quelle mesure les actions et les moyens mis en œuvre permettent d'atteindre les objectifs?

### • L'efficacité

Une politique publique produit des effets. Le critère d'efficacité de l'intervention s'intéresse aux effets ou résultats produits par l'intervention sur les cibles en les comparant aux objectifs initiaux.

L'efficacité d'une politique publique s'analyse toujours en rapport aux objectifs fixés et résultats attendus. Lorsqu'une politique publique a des effets négatifs ou des effets positifs qui ne correspondent pas aux objectifs initiaux, on peut en déduire un défaut de conception de la politique publique. Ce critère pose une question telle que : dans quelle mesure les objectifs ont-ils été atteints au regard des produits réalisés?

- **L'efficience**

Toute politique publique a un coût, et on ne peut porter un jugement d'ensemble sur sa réussite et son bien-fondé qu'en comparant ses résultats aux ressources financières utilisées.

Ainsi, le critère d'efficience compare les résultats obtenus aux moyens humains, financiers et matériels utilisés pour la mise en œuvre des actions. Il permet d'évaluer si une utilisation différente, une augmentation ou une diminution des moyens permettraient d'atteindre les résultats visés ou des résultats équivalents, voire les améliorer.

Pour ce faire, l'évaluation de l'efficience est basée sur des analyses coûts-avantages et coûts-efficacité. Ce critère de l'efficience répond à une question telle que : dans quelle mesure les résultats produits ont-ils été atteints avec une utilisation optimale des moyens?

- **La performance**

Le critère de l'efficience est souvent associé au critère d'efficacité, ces deux critères permettant de qualifier la performance de la politique publique en établissant une relation entre les objectifs, les moyens et les résultats. En France, la loi organique relative aux lois de finances du 28 juin 2001 (LOLF) a fait émerger un nouveau discours dans lequel la performance constitue la pierre angulaire, entraînant une révolution dans les pratiques administratives.

Dans cette démarche de modernisation de l'administration, l'évaluation des politiques publiques s'appuie sur la notion de performance. En effet, la LOLF entend faire passer le budget de l'État d'une logique de moyens à une logique de résultats et instaurer, pour les finances publiques, un pilotage d'après les performances.

## L'ÉVALUATION VERSUS L'AUDIT, LE CONTRÔLE DE GESTION, L'INSPECTION

17

La différence entre l'évaluation, l'audit, le contrôle et l'inspection peut être établie à travers l'objet observé ou la finalité de l'observation, les références auxquelles sont comparées ce que l'on observe et les conséquences qui en découlent.

- L'audit a pour objet de réduire les risques d'une institution ou d'une organisation en référence aux pratiques professionnelles. Les conséquences d'un audit sont des recommandations ou des rappels à l'ordre que les personnes responsables de l'institution ou de l'organisation peuvent accepter et suivre, ou rejeter ou ignorer partiellement ou totalement, notamment lorsqu'elles considèrent que les avantages recherchés valent les risques encourus et qu'elles sont prêtes à les assumer.
- L'objet du contrôle est de suivre la mise en œuvre et de vérifier la réalisation des actions qui avaient été programmées lors de leur conception, la conséquence pouvant être une rectification de la trajectoire.
- L'objet de l'inspection est de vérifier la conformité à la réglementation. La norme de référence est donc constituée des textes règlementaires, la conséquence étant, en cas d'entorse à ces règles, le refus d'un quitus, une amende ou une poursuite judiciaire.

Tableau 1 : Différences entre le contrôle, l'audit, le contrôle de gestion, l'inspection et l'évaluation

	<b>L'audit</b>	<b>Le contrôle de gestion</b>	<b>Le contrôle (ou inspection)</b>	<b>L'évaluation</b>
<b>Finalité</b>	Réduire les risques	Suivre l'exécution des actions	Vérifie la conformité à la réglementation	Observer et expliquer les écarts entre les effets attendus, réalisés et les
<b>Référentiel</b>	Normes, recommandations	Programme fixé ex ante	Lois, règlements (cadre comptable), conventions, obligations	Références à construire selon les finalités de l'évaluation et les objectifs
<b>Conséquences</b>	Alerter sur les risques en cours	Rectifier la trajectoire	Sanctionner un écart, amendes, poursuites judiciaires	Aider à la décision stratégique et/ou opérationnelle.

## LES FINALITÉS DE L'ÉVALUATION

L'évaluation poursuit quatre finalités non exclusives:

- ***Une finalité cognitive***

L'évaluation permet la production de connaissances sur une intervention publique pour mieux comprendre le sens et les effets. Elle permet d'avoir une lisibilité sur les publics concernés, les réalisations et les mécanismes de mise en œuvre d'une politique publique.

L'évaluateur se fonde ainsi ses travaux sur des informations fiables qu'il récolte au moyen de différentes méthodes des sciences sociales (analyse documentaire, analyse des données d'activité, enquête quantitative, enquête qualitative, observation évaluative...)

- ***Une finalité normative***

L'évaluation d'une politique publique permet d'apprécier la valeur de l'intervention à travers ses effets à partir de critères normatifs. Les données collectées sont analysées afin de produire une appréciation sur la conduite et les impacts de l'intervention publique, d'où le croisement des données récoltées auprès de différents acteurs aux intérêts parfois divergents et l'autonomie de l'évaluateur par rapport aux commanditaires et aux autres parties prenantes de l'intervention.

- ***Une finalité décisionnelle***

L'évaluation d'une politique publique est destinée à éclairer la décision politique ou administrative. Les constats posés et les préconisations alimentent ainsi la décision publique et peuvent déboucher sur des décisions d'amélioration, de prolongation, de réorientation ou d'arrêt de l'intervention publique.

- ***Une finalité démocratique :***

18

L'évaluation permet une optimalisation de l'utilisation des ressources publiques en améliorant la pertinence et l'efficience des politiques publiques menées. Elle implique ainsi la consultation, voire l'association des bénéficiaires, usagers ou citoyens.

## L'ÉVOLUTION DE L'ÉVALUATION

Jusqu'à la fin des années 1980, les évaluations dites managériales étaient réalisées au sein des organisations, sans l'implication d'acteurs extérieurs. Elles s'interdisaient, pour garantir l'indépendance de l'évaluateur et la qualité de son jugement, d'associer d'autres acteurs que l'administration responsable du programme évalué.

Le modèle de l'évaluation participative est devenu dominant en France au cours des 1990, est appliqué à tous les types de politiques. La dernière génération d'évaluations tend à rendre les citoyens acteurs de la conduite de l'action publique en positionnant la société civile au cœur de la démarche. Ainsi, se développe la Nouvelle Gestion Publique fondée sur l'exigence de reddition des comptes qui était en vigueur dans les pays d'Amérique du Nord depuis plusieurs décennies et qui a progressivement gagné les pays européens dont la France.

Selon BARON et MONNIER, il existe deux générations d'évaluations participatives : la première consiste simplement à impliquer des parties prenantes à la démarche.

La seconde qui s'est largement développée aux États-Unis, et au Canada notamment, va beaucoup plus loin puisqu'elle tend à impliquer totalement la société civile dans la démarche depuis la formulation des questions évaluatives jusqu'à la rédaction des préconisations. Ce sont les évaluations dites émancipatrices (empowerment evaluations). En France, une génération plus récente d'évaluations a cherché à intégrer au processus les parties prenantes à la politique évaluée : les associations, les opérateurs, les représentants de la société civile mais également les citoyens ou usagers. Le but est d'étendre l'analyse aux points de vue des bénéficiaires des politiques évaluées afin d'augmenter la pertinence de la démarche.

Les politiques d'aide au développement sont l'un des domaines où l'évaluation a cherché à s'imposer depuis longtemps comme outil de pilotage et de compte rendu. Au plan international, la Banque Mondiale a joué un rôle décisif à partir des années 1970 avec la mise en place d'une fonction d'évaluation indépendante de la direction de la Banque dès 1975. L'évaluation y a pris une ampleur considérable au fil des années et a concerné des centaines de projets chaque année.

<sup>8</sup> BARON, Gaëlle et MONNIER, Eric : «Une approche pluraliste et participative : coproduire l'évaluation avec la société civile», in Revue Informations sociales, Septembre 2003.

<sup>9</sup> Paris Declaration on aid effectiveness. Ownership, harmonization, alignment, results and mutual accountability. DAC, Editor, 2005,DAC/OECD, Paris.

Aujourd’hui, avec des milliers de programmes concernés chaque année, et plusieurs milliers de dollars d’engagement, la Banque Mondiale joue un rôle moteur pour l’ensemble des activités d’évaluation de l’aide au développement. En France, des missions d’évaluation de l’aide au développement ont été créées dès les années 1980 au sein du ministère de la Coopération, et particulièrement avec la mise en place de l’Agence Française de Développement (AFD) en 1998.

## LA GESTION AXÉE SUR LES RÉSULTATS ET L’ÉVALUATION DES POLITIQUES PUBLIQUES

Une démarche pertinente de gestion axée sur les résultats est un processus continu dans le cycle de vie des politiques publiques, à travers le retour d’information, l’apprentissage et l’amélioration. Les politiques publiques en cours sont régulièrement modifiées grâce aux enseignements tirés de l’évaluation et les politiques à venir sont élaborées à partir de ces enseignements.

Ainsi, les évaluations devraient être réalisées non seulement pour améliorer les politiques alors qu’elles sont en cours d’exécution, mais également pour apporter les informations nécessaires à la planification de nouvelles politiques. Le processus itératif visant à faire apprendre et améliorer est au cœur de la gestion axée sur les résultats (GAR).

En mars 2005, les pays membres de l’OCDE se sont engagés à améliorer l’efficacité de l’aide au développement à travers la Déclaration de Paris. L’aide au développement devra désormais prendre en compte son appropriation par les pays bénéficiaires, son alignement et son harmonisation avec les politiques nationales et les engagements internationaux, la gestion axée sur les résultats et la responsabilité mutuelle des pays donateurs et pays bénéficiaires. Pour atteindre les objectifs de la Déclaration de Paris de 2005, les pays de l’OCDE ont décidé de soutenir la mise en œuvre de la gestion axée sur les résultats (GAR) et d’avoir recours systématiquement à l’évaluation des politiques et programmes de développement.

En Afrique, la GAR s’est progressivement installée à travers des réformes visant la modernisation de l’action publique et en conformité avec la Déclaration de Paris, notamment par l’usage de cadres logiques d’intervention dans la planification des politiques publiques.

Cependant, des efforts importants restent à faire en ce qui concerne l’évaluation, car si les modes de planification se sont transformés et la GAR bien installée, les systèmes de reddition des comptes et les pratiques d’évaluation n’ont pas beaucoup évolué.

Or, de plus en plus, les décideurs, les bailleurs et la société civile demandent et exigent que les responsables des programmes de développement rendent compte de la performance de leurs interventions.

## LES ENJEUX ÉTHIQUES

«*Nous avons besoin d’une culture de l’évaluation, une évaluation réalisée de manière indépendante, à temps réel et en toute transparence*» Antonio Guterres, Secrétaire Général des Nations Unies.

Les enjeux éthiques et d’ordre déontologique sont à la base de normes et règles d’évaluation promues par les institutions des Nations Unies, l’OCDE, l’Union Européenne, l’Association Africaine d’Évaluation et toutes les associations régionales et nationales concernées par l’évaluation.

Parmi les normes générales promues par l’UNEG, figurent :

- Utilité

La finalité de toute évaluation commandée et réalisée doit être claire : les constats et l’analyse, les conclusions et les recommandations doivent servir à orienter les décisions et les actions à entreprendre.

On entend par l’utilité d’une évaluation, sa capacité à contribuer de manière pertinente et opportune à l’apprentissage organisationnel, à l’élaboration de processus décisionnels éclairés et à la responsabilité de rendre compte des résultats.

#### • Crédibilité

Les évaluations doivent être crédibles. La crédibilité repose sur l'indépendance, l'impartialité et des méthodes rigoureuses.

Les principaux critères de crédibilité incluent des processus d'évaluation transparents, des stratégies profitant à tous et impliquant les parties prenantes concernées, ainsi que des systèmes d'assurance qualité fiables. Pour garantir leur crédibilité, il est nécessaire que les évaluations soient réalisées de manière éthique par des évaluateurs possédant les compétences culturelles et professionnelles requises.

#### • Indépendance

L'indépendance de l'évaluation est essentielle à sa crédibilité; elle influence la manière dont l'évaluation est utilisée et garantit aux évaluateurs impartialité et liberté face à toute pression éventuelle, tout le long du processus d'évaluation. L'indépendance de la fonction d'évaluation comporte deux aspects clés : l'indépendance comportementale et l'indépendance organisationnelle.

L'indépendance comportementale concerne la capacité à procéder à l'évaluation sans subir l'influence indue d'une quelconque partie.

Les évaluateurs doivent disposer de l'entièvre liberté de mener leurs travaux en toute impartialité, le déroulement de leur carrière ne devant avoir en à subir aucun contrecoup; ils doivent pouvoir être en mesure d'exprimer librement leur opinion. L'indépendance de la fonction d'évaluation sous-tend le libre accès des évaluateurs aux renseignements sur l'objet de l'évaluation.

L'indépendance organisationnelle exige que la fonction d'évaluation centrale soit indépendante des autres fonctions de gestion, qu'elle se charge de définir le programme de l'évaluation et qu'elle dispose de ressources appropriées pour mener à bien ses travaux.

20

#### • Impartialité

L'impartialité est caractérisée par l'objectivité et l'intégrité professionnelle ainsi que l'absence de parti pris. L'impartialité doit se retrouver à tous stades du processus d'évaluation, notamment la planification de l'évaluation, la communication avec les parties prenantes, la réalisation de l'évaluation et la formulation des constats et recommandations.

Les évaluateurs doivent se montrer impartiaux, ce qui signifie que les membres de l'équipe d'évaluation ne doivent pas être responsables de la mise en place de la politique, de la conception ou de la gestion dont fait l'objet l'évaluation, ni être pressentis de l'être prochainement.

#### • Éthique

L'évaluation doit être réalisée selon les exigences les plus strictes en matière d'intégrité et de respect des croyances, des us et coutumes des environnements sociaux et culturels dans lesquels elle a lieu, dans le respect des droits de l'homme et de l'égalité des sexes, et conformément au principe de ne pas nuire, dans le cadre de l'aide humanitaire.

Les évaluateurs doivent respecter le droit des particuliers et des institutions de fournir des renseignements à titre confidentiel, et faire en sorte que des données qualifiées de sensibles soient protégées et ne permettent pas de remonter à leur source. Ils doivent, en outre, valider les déclarations présentes dans les rapports auprès de leurs auteurs. Lorsqu'ils souhaitent utiliser des informations personnelles, les évaluateurs doivent obtenir l'accord éclairé des personnes concernées.

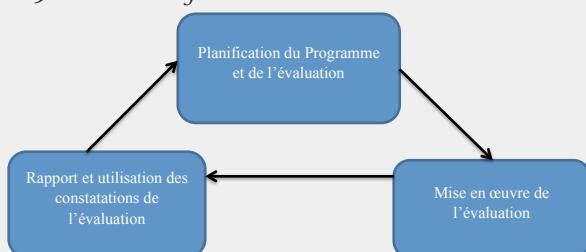
#### • Transparence

La transparence est un élément majeur d'une évaluation qui favorise la confiance, renforce la participation des parties prenantes et accroît la reddition de compte. Les produits de l'évaluation doivent être accessibles au public.

## LE PROCESSUS ÉVALUATIF

Le processus évaluatif est généralement structuré autour des trois (3) grandes étapes que sont la préparation de l'évaluation, la mise en œuvre de l'évaluation et la présentation des résultats de l'évaluation. Chacune de ces trois étapes comporte des tâches essentielles dont les principales sont représentées sous forme de cycle.

Figure 3 : Schéma du cycle de l'évaluation



Source : Réseau international de l'information financière : Guide détaillé pour l'évaluation des programmes en matière d'éducation financière. Gouvernement du Canada/OCDE, p.5, non daté, 17 pages.

#### Message 10

Un processus évaluatif doit suivre la même rigueur dans sa gestion qu'un projet quelconque. Il est fortement recommandé d'utiliser des outils de gestion de projet quand on gère une évaluation.

## 2<sup>ÈME</sup> PARTIE : PRÉPARER UNE ÉVALUATION

## LA COMMANDE (DEMANDE) DE L'ÉVALUATION

L'élément déclencheur d'une évaluation de politiques publiques est la commande de celle-ci. Cette commande peut provenir de différents acteurs. Les articles 335 et 336 de la loi n°2016-046 du 23 septembre 2016 fixent les différentes sources d'où peut provenir la commande d'évaluation de politiques publiques pour la SCCS. Il s'agit de la Section des Comptes « de sa propre initiative », du Président de la République, du Premier ministre, du Président de l'Assemblée nationale et du Président du Haut Conseil des Collectivités.

### PLANIFIER L'ÉVALUATION

La planification de l'évaluation inclut la réflexion sur les différentes tâches à accomplir en vue de l'évaluation (la détermination de son objectif et sa portée au moyen de consultations avec les parties prenantes et les personnes-ressources, la définition des questions importantes auxquelles l'évaluation devra répondre, le choix du type d'évaluation qui convient le mieux, la détermination des méthodes qui seront utilisées pour recueillir les données, la démarche appropriée pour rédiger le rapport de l'évaluation, la confection du budget de l'évaluation, etc.).

22

### DÉFINIR LES RÔLES ET RESPONSABILITÉS DANS LE PROCESSUS ÉVALUATIF

La gouvernance de l'évaluation est matérialisée par la mise en place de comités tels que le comité de pilotage et le comité de gestion de l'évaluation.

Le Commissaire de l'évaluation assure la coordination du Comité de pilotage de l'évaluation, lequel est responsable de la formulation de la demande d'évaluation. Ce Comité est composé des plus hautes autorités concernées par le programme, des partenaires clés et éventuellement des représentants des collectivités territoriales et de la société civile selon la nature du programme à évaluer.

L'unité de gestion du programme et les parties prenantes clés commanditaires de l'évaluation, doivent nommer un responsable de l'évaluation qui assumera les responsabilités quotidiennes ayant trait à la gestion de l'évaluation, tout en jouant le rôle de liaison avec les différents acteurs clés. Dans la mesure du possible, ce rôle est assumé par un spécialiste du suivi et évaluation au sein de l'unité de gestion du programme.

Chaque évaluation doit disposer d'un groupe de référence composé des parties prenantes clés qui collaborent étroitement avec le spécialiste responsable de l'évaluation.

### Message 11

Il n'y a pas un ordre établi entre les différentes tâches des différentes phases du processus évaluatif comme cela peut être le cas pour certains domaines voisins à l'évaluation.

En résumé, le processus de l'évaluation implique directement différents groupes d'acteurs dont le comité de pilotage de l'évaluation qui est responsable de la formulation de la demande d'évaluation, l'unité de gestion du programme dont le responsable du suivi et évaluation est souvent le gestionnaire de l'évaluation, le groupe de référence et le groupe des évaluateurs.

Le Commissaire de l'évaluation (au sein du Comité de l'évaluation) a pour rôle, entre autres, la détermination des lignes directrices de l'évaluation, l'allocation des fonds nécessaires et la surveillance du respect des principes de transparence du processus de l'évaluation.

Le responsable ou Gestionnaire de l'évaluation assure la gestion quotidienne du processus de l'évaluation, depuis l'élaboration des termes de référence jusqu'à la diffusion du rapport final en passant par la sélection du groupe des évaluateurs, la contractualisation avec l'équipe d'évaluation, la fourniture de l'appui technique et des informations disponibles aux évaluateurs sélectionnés, ainsi que la liaison avec le Comité de l'évaluation et le groupe de référence.

Le Groupe de référence définit et confirme le profil, les compétences, les rôles et responsabilités du responsable de l'évaluation et propose des candidats à cette fonction, participe à la rédaction et à la révision de l'ébauche des TDR.

De même, le groupe de référence prête assistance à la collecte des données requises, supervise les progrès et la conduite de l'évaluation, révise le rapport du projet d'évaluation et assure que le rapport final répond aux normes de qualité.

Le Groupe ou Équipe des évaluateurs a pour responsabilités clés essentielles de réaliser les dispositions contractuelles en accord avec les normes éthiques et déontologiques, de développer la matrice de l'évaluation comprenant les critères et les questions de l'éva-

luation, les indicateurs, les données à collecter, les sources de données, les méthodes d'analyse, de rédiger et soumettre les rapports requis, jusqu'au rapport final de l'évaluation.

Durant ce processus, l'équipe des évaluateurs se doit de communiquer avec le commissaire, le responsable et le groupe de référence de l'évaluation sur les progrès de la mission d'évaluation, les contraintes et toutes autres questions liées à la bonne mise en œuvre de l'évaluation et de prendre en considération les remarques et suggestions du responsable et des parties prenantes de l'évaluation.

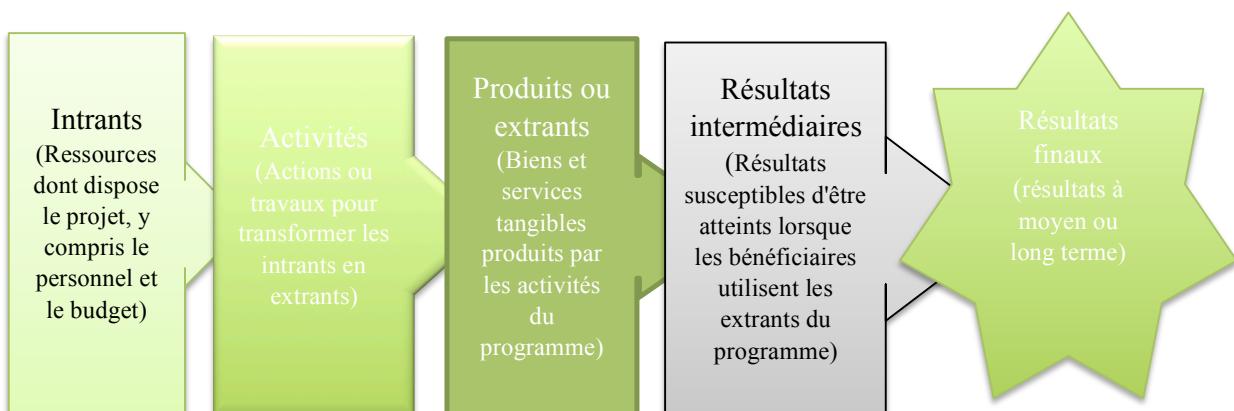
Finalement, le groupe des évaluateurs doit s'assurer que tous les livrables, y compris les bases de données, le rapport final, les listes de personnes interrogées et autres documents sont soumis au responsable de l'évaluation.

## EXPLORER L'INTERVENTION

Explorer l'intervention (le programme) consiste à approfondir sa connaissance à travers ses intentions (les éléments de son référentiel) :

- **La raison d'être** : la nature du problème ou du besoin qui nécessite l'intervention, la politique publique à travers un programme ;
- **La théorie du changement** : une théorie du changement est ce qui explique comment une intervention (ou un ensemble d'interventions) est censée conduire à un changement précis sur le plan du développement grâce à des liens de cause à effet. Le terme théorie du changement est parfois utilisé pour désigner la chaîne de résultats qui se présente sous la forme d'une série de cases allant des intrants vers les produits, les réalisations et les impacts, ou d'un cadre logique qui présente les informations sous forme de matrice. La théorie du changement peut aussi comporter les facteurs contextuels qui favorisent ou contrecarrent ce changement, ainsi que les hypothèses sur lesquels il repose. Il existe donc différentes manières d'élaborer et de représenter la théorie du changement.

Figure 4 : Schéma de la théorie du changement



Source : Réseau international de l'information financière : Guide détaillé pour l'évaluation des programmes en matière d'éducation financière. Gouvernement du Canada/OCDE, p.5, non daté, 17 pages.

Une théorie du changement explique comment un programme est censé fonctionner. Elle peut appuyer une évaluation d'impact en permettant d'identifier :

- *Les questions d'évaluation spécifiques, en particulier en ce qui concerne les éléments de la théorie du changement pour lesquels il n'existe pas encore de données probantes,*
- *Les variables pertinentes devant être utilisées dans la collecte de données,*
- *Les indicateurs devant opérationnaliser les variables pertinentes,*
- *Les effets intermédiaires pouvant servir être incluses dans la collecte des données,*
- *Les aspects de la mise en œuvre devant être examinés,*
- *Les facteurs contextuels pertinents devant être pris en considération dans la collecte et l'analyse des données, afin de rechercher des constantes.*

- **les parties prenantes** : l'analyse des parties prenantes est une démarche minutieuse et structurée qui permet de synthétiser les informations essentielles sur l'ensemble des personnes, des groupes, des organismes ou autres qui entrent en jeu dans le programme en tenant compte de leurs intérêts, de leur influence et de leurs positions de pouvoir. L'analyse des parties prenantes permet de prendre les mesures idoines pour gérer leurs intérêts et leurs attentes afin que le programme fonctionne aussi bien que possible et atteigne les résultats visés. Les parties prenantes sont des personnes ou des groupes qui sont directement ou indirectement affectés par un projet ainsi que ceux ayant des intérêts dans un projet et/ou la capacité d'influencer sur ses résultats, que ce soit positivement ou négativement. Les parties prenantes peuvent comprendre les communautés ou les individus localement affectés ainsi que leurs représentants officiels et non officiels, les autorités gouvernementales locales ou nationales, les politiciens, les responsables religieux, des groupes et organisations de la société civile avec leurs intérêts spéciaux, le monde de l'enseignement ou d'autres entreprises.

24

#### Message 12

Faire une évaluation,  
c'est répondre à des  
questions d'évaluation.

## DÉFINIR L'OBJECTIF D'ÉVALUATION

La définition de l'objectif de l'évaluation est un préalable en réponse à l'opportunité de l'évaluation en tant que mesure du changement à réaliser, du changement en cours de réalisation ou du changement qui a été réalisé par un programme.

La définition de l'objectif de l'évaluation découle ainsi du questionnement autour du programme concerné en vue de formuler un jugement sur le changement attendu ou réalisé.

A quels types de questions répond l'évaluation?

D'une manière générale, l'évaluation s'attèle à répondre à trois types de questions :

### • Les questions descriptives

Les questions descriptives visent à décrire les processus, les conditions qui prévalent, les relations organisationnelles et les points de vue des parties prenantes d'une intervention.

A ce niveau, l'évaluation cherche à montrer ce qui se passe ainsi que les attitudes en rapport avec la politique publique. Ces questions permettent de décrire les apports, les activités et les résultats, et peuvent être utilisées pour recueillir les opinions des bénéficiaires.

Exemples de questions descriptives :

- *quel est le niveau des différences systématiques entre les sites dans la manière dont s'exécute l'intervention?*
- *quels sont les qualifications (ou les apports) des fournisseurs (ou prestataires de service) ?*
- *quelle est la proportion de femmes ayant participé au programme?*
- *à quel point les participants (et les participantes) ont-ils trouvé utile le programme?*
- *quels sont les objectifs (ou les résultats) du programme du point de vue des différents acteurs (ou bénéficiaires)?*
- *quelles sont les premières réalisations du programme?*
- *qui a bénéficié de quels services?*

### • Les questions normatives

Les questions normatives comparent ce qui se passe à ce qui devrait se passer en comparant la situation actuelle au niveau d'une cible, en rapport avec l'objectif ou un indice de référence spécifique.

Elles consistent à étudier les activités et à estimer si les objectifs ont été atteints ou non et elles peuvent concerner les intrants, les activités et les extrants (produits, résultats intermédiaires).

Exemples de questions normatives :

- *les dépenses ont-elles été effectuées conformément au budget prévu?*
- *le budget a-t-il été dépensé de manière efficace?*
- *80% des enfants ont-ils été scolarisés conformément à l'objectif fixé?*
- *le processus de sélection des participants (et des participantes) a-t-il été juste et équitable?*
- *à quel point le programme reflète-t-il le principe de parité homme/femme?*

### Message 13

L'évaluation des résultats des politiques publiques, qui est attendu de la Section des Comptes de la Cour Suprême du Mali, de par ses textes, doit se fonder sur des questions de type « cause à effet ».

- **Les questions de cause à effet**

Les questions de cause à effet se concentrent sur les résultats et cherchent à déterminer dans quelle mesure l'intervention a produit des changements. Elles visent donc à déterminer la différence faite par l'intervention et à savoir si les résultats ou les changements ont été le fruit de l'intervention.

#### Exemples de questions de cause à effet :

- *la stratégie de partenariat entre les acteurs institutionnels et les acteurs privés a-t-elle préservé la biodiversité de la zone affectée tout en assurant les moyens de substances durables aux populations locales?*
- *les participants à la formation pour l'emploi ont-ils obtenu des emplois décents et plus rémunérateurs que ceux auxquels ils pouvaient prétendre avant la formation?*
- *le programme d'appui aux micro-entreprises a-t-il contribué à réduire la pauvreté dans les régions cibles?*

Les questions évaluatives initiales ou préalables donnent des indications sur les types ou méthodes d'évaluation (voir 4), sur le moment et le but de l'évaluation. Elles permettent de déterminer s'il s'agit d'une évaluation ex ante (liées à la conception du programme), d'une évaluation « in itinere » destinée ainsi à fournir des connaissances et des repères pour un réajustement des efforts déployés (évaluation à but formatif) ou d'une évaluation ex post qui s'oriente vers les effets (résultats et impacts) obtenus par le programme (évaluation à but sommatif).

## DÉFINIR LES CRITÈRES D'ÉVALUATION

Pour arriver à un jugement sur le changement à réaliser ou réellement effectué par un programme dans un contexte donné, une palette de critères sont définis en vue de fonder ce jugement sur des références ou des mesures fiables. Les critères d'évaluation sont généralement autour de cinq (5), notamment la pertinence, la cohérence, l'efficacité, l'efficience et la performance.

Le Comité d'Aide au Développement de l'OCDE (CAD/OCDE) recommande aussi cinq (5) critères d'évaluation que sont : la pertinence, l'efficacité, l'efficience, l'impact et la viabilité. L'impact est défini comme étant l'ensemble des effets positifs ou négatifs, directs ou indirects, intentionnels ou non induits par une intervention. La viabilité « permet de mesurer si les bienfaits d'une intervention de développement ont des chances de perdurer une fois que le donneur aura achevé de la financer ».

Les critères d'évaluation permettent de formuler les questions évaluatives qui peuvent être des questions descriptives, normatives ou de cause à effet.

Lors de la planification du programme, les aspects suivants concernant l'évaluation doivent être pris en compte :

- Les exigences et les attentes du commanditaire de l'évaluation ;
- Les grandes lignes des termes de référence de l'évaluation ;
- Une budgétisation provisoire de l'évaluation.
- Un premier calendrier approximatif ;
- Une description de la qualification des experts.

Le plan de l'évaluation doit comprendre tous les aspects liés à la préparation et la conduite de l'évaluation, ainsi que les modalités de gouvernance de l'évaluation, d'élaboration et de diffusion des termes de référence, de contractualisation éventuelle avec des experts, de mise en relations des experts avec les parties prenantes, de rédaction du rapport provisoire et du rapport final, de l'utilisation des résultats et de partage des conclusions.

### Message 14

Le choix d'une méthode d'évaluation doit être dicté par ce que l'évaluation projette de mettre en évidence mais aussi par les ressources et les expertises disponibles.

## CHOISIR LA MÉTHODE D'ÉVALUATION (ÉVALUATION DES RÉSULTATS)

### Les méthodes quantitatives d'évaluation

#### • La méthode expérimentale

La méthode expérimentale consiste à tester la validité d'une hypothèse causale en reproduisant un phénomène et en faisant varier un paramètre. Le paramètre que l'on fait varier est impliqué dans l'hypothèse. Le résultat de l'expérience valide ou non l'hypothèse.

Le protocole d'expérimentation regroupe la description des conditions et du déroulement d'une expérience ou d'un test. La description doit être suffisamment claire afin que l'expérience puisse être reproduite à l'identique et il doit faire l'objet d'une analyse critique pour notamment détecter d'éventuels biais.

La méthode expérimentale représente l'approche la plus rigoureuse pour établir des relations causales entre un programme et ses résultats. Elle est caractérisée par la répartition aléatoire des participants éventuels entre le groupe expérimental et le groupe témoin afin d'assurer l'équivalence entre les deux.

La répartition aléatoire (ou randomisation) signifie qu'il y a pour chaque membre de la population cible une probabilité connue qu'il soit choisi pour faire partie du groupe expérimental ou du groupe témoin. Ces probabilités sont en principe égales, auquel cas chaque membre a des chances égales d'être choisi pour faire partie d'un groupe ou de l'autre.

La méthode expérimentale permet d'éliminer la plupart des obstacles à la validité interne dans l'évaluation des résultats et des impacts des programmes. Son principal inconvénient est qu'elle est difficile à appliquer.

#### • La méthode quasi expérimentale

Tout comme la méthode expérimentale basée sur la randomisation, la méthode quasi expérimentale (qui ne permet pas une randomisation) consiste à tester les hypothèses causales.

Que la méthode soit expérimentale ou quasi expérimentale, la politique publique à évaluer est considérée comme intervention qui consiste à évaluer, au moyen d'un ensemble d'indicateurs prédéfinis, la capacité d'un traitement (à savoir les éléments de la politique à l'étude) à remplir ses objectifs.

Toutefois par définition, la méthode quasi expérimentale n'a pas recours à la randomisation. La répartition des participants (entre le groupe expérimental et le groupe de contrôle ou groupe témoin) repose sur le principe de l'auto-sélection (les participants choisissent eux-mêmes l'intervention) et/ou sur le choix des administrateurs (décideurs politiques, fonctionnaires, etc.).

La méthode quasi expérimentale identifie un groupe témoin qui doit être aussi proche que possible du groupe expérimental au niveau des caractéristiques initiales (caractéristiques préalables à l'intervention).

Le groupe témoin qui doit ressembler autant que possible au groupe expérimental doit permettre de faire des inférences causales, c'est-à-dire d'identifier les résultats qui seraient survenus si la politique publique n'avait pas été mise en œuvre (situation contrefactuelle) et d'attribuer les résultats observables chez le groupe expérimental comme attribuables au programme.

Cette méthode permet de prouver que la politique publique est bien à l'origine des différences de résultats entre le groupe expérimental et le groupe témoin.

Le groupe témoin peut être soit un groupe témoin créé de toutes pièces qui n'a pas été exposé au programme, soit un groupe réflexif c'est-à-dire le groupe expérimental lui-même avant son exposition au programme.

Les modèles quasi expérimentaux utilisés dans l'évaluation d'impact sont les suivants :

- Les modèles à mesure avant et après programme.
- Les modèles à mesure à séries temporelles ou modèle chronologique.
- Les modèles à mesure pris après le programme seulement.

#### • La méthode pré expérimentale

La méthode pré expérimentale ne présente ni situation de contrôle, ni groupe témoin. L'effet de l'intervention publique (variable indépendante) est mesuré par rapport à une dimension donnée (variable dépendante).

La méthode pré expérimentale peut être appliquée avant et après l'intervention : dans le cas d'une application avant et après, il peut concerner une population cible qui subit ainsi un test après l'intervention; dans le cas de l'application unique à la fin de l'intervention, la méthode est dite one shot.

Par exemple, la méthode pré expérimentale peut s'appliquer chez un groupe de personnes qui reçoit un enseignement avec une évaluation des connaissances avant et après l'enseignement ou uniquement à la fin de cet enseignement.

### **Message 15**

Le recours aux méthodes expérimentale et quasi expérimentale, en particulier, nécessite la maîtrise des techniques d'échantillonnage (voir Annexe 4 et Annexe 6)

#### • Les méthodes statistiques

Les méthodes statistiques sont basées sur l'utilisation de données (catégories) qualitatives ou quantitatives en vue de décrire des phénomènes et de procéder à des inférences quant aux relations entre variables.

Les données en question peuvent être soit objectives et concrètes soit subjectives et abstraites, mais les unes et les autres doivent être décrites et organisées de façon systématique.

La plupart des études évaluatives ou analytiques font appel à la méthode statistique. La première raison est que l'utilisation de la méthode statistique permet de faire une description, notamment à travers des tableaux statistiques.

La seconde raison est qu'elle permet de faire des inférences pour vérifier les rapports entre les variables étudiées ou pour généraliser des constatations en les appliquant à une population plus étendue ayant fait l'objet d'un échantillonnage.

Dans le cadre d'un rapport d'évaluation, les constatations sont souvent présentées succinctement sous forme de tableaux faisant ressortir des moyennes et des variances. Une statistique telle que l'écart type peut servir à résumer l'étendue de la distribution. Il correspond à la mesure dans laquelle les résultats individuels se rapprochent de la moyenne arithmétique (la normale).

Plus l'écart type est petit par rapport à la normale, moins la distribution est étendue.

Les graphiques peuvent aussi aisément présenter des données et des statistiques. Les graphiques à barre sont usuellement construits pour les distributions et les graphiques circulaires et les boîtes illustrent des proportions relatives.

Ces présentations visuelles faciles à produire avec des logiciels statistiques peuvent être très utiles pour résumer des données statistiques, puisqu'elles sont souvent plus faciles à lire qu'un tableau et n'exigent pas nécessairement une compréhension de tous les aspects statistiques.

La méthode statistique de l'analyse de régression peut servir à établir l'importance de la corrélation entre variables à l'étude. Dans ce cas, le rapport entre les résultats, la participation et la non participation au programme et le sexe ou l'âge du (de la) participant (e) pourrait être déterminé. Il pourrait ainsi donner lieu à des conclusions permettant de cerner l'importance des rapports entre les résultats d'un programme et les caractéristiques de ses participants.

D'une manière générale, les méthodes statistiques peuvent résumer les constatations d'une évaluation de façon claire, précise et fiable.

Cependant, les résultats d'un programme ne peuvent pas tous être analysés par des méthodes statistiques, et la façon de classer les données peut bien fausser les résultats que révèlent les différences constatées.

### **Message 16**

Le recours aux méthodes statistiques nécessite la maîtrise des techniques d'analyses statistiques et de logiciels tels que SPSS (voir Annexe 3). Il est fortement recommandé d'avoir un ou plusieurs spécialistes en statistiques dans les équipes d'évaluation ou, à défaut, faire appel à des spécialistes en statistiques externes à l'équipe d'évaluation.

## Les méthodes qualitatives d'évaluation

Quelques méthodes qualitatives d'évaluation :

### • Le changement le plus significatif (MSC)

Le changement le plus significatif (Most Significant Change) est une méthode qui propose une évaluation de type participatif sur la base de témoignages recueillis auprès des différents acteurs. Ces données de type qualitatif permettent une représentation plus subtile et concrète de la réalité qu'une série d'indicateurs prédéfinis.

Cette approche est née au cours des années 1990 en réaction aux méthodes d'évaluation normées et quantitatives, et elle est aujourd'hui reconnue comme une méthode alternative d'évaluation qui cherche moins à synthétiser des informations qu'à mettre en évidence le particulier et les divergences de points de vue des acteurs.

Le changement le plus significatif est particulièrement utilisé dans les évaluations de programmes à base de données qualitatives, dans la récolte d'informations sur les activités et dynamiques en cours et pour une meilleure compréhension du programme par les gestionnaires et les décideurs.

La détermination du changement le plus significatif procède par plusieurs étapes, à savoir :

- La détermination des aires ou domaines de changement par l'ensemble des acteurs,
- La détermination de la période de recueil des témoignages et du nombre des participants,
- Le choix de la question type pour l'identification du changement le plus significatif,
- La définition du processus de la collecte et du mode d'élection des témoignages les plus significatifs,
- La confrontation des points de vue des différents acteurs,
- La vérification de la véracité des témoignages.

Le changement le plus significatif offre une meilleure compréhension du vécu des participants au programme, permet le partage des expériences, la mise en évidence des phénomènes marginaux et exceptionnels et la participation d'un grand nombre d'acteurs à l'évaluation.

### • La cartographie des incidences

La cartographie des incidences concerne essentiellement l'analyse des effets d'une intervention sur son environnement.

Elle se concentre sur les changements qui s'inscrivent clairement dans la sphère d'influence d'un programme et donc sur les résultats obtenus en amont de l'impact d'une intervention, qu'elle évalue non pas en se basant sur la pertinence de la politique, mais sur les modifications de comportement, les relations, les agissements ou les activités des personnes, groupes et organisations avec lesquels le programme travaille directement.

La cartographie des incidences est un système permettant de mener une réflexion holistique et stratégique sur la manière dont le programme à l'intention d'obtenir des résultats et qui promeut l'apprentissage et le perfectionnement.

La cartographie des incidences aide les utilisateurs à comprendre les résultats obtenus par une organisation tout en tenant compte du rôle essentiel des contributions des autres acteurs dans l'obtention du type d'amélioration durable à grande échelle du bien-être humain et de l'état de l'environnement.

La cartographie des incidences tient compte du fait que les divers partenaires limitrophes d'un programme agissent selon des systèmes logiques et de responsabilisation différents. Ainsi, elle n'est pas fondée sur un système de relations de cause à effet, mais reconnaît plutôt que les événements multiples, non linéaires mènent au changement. Elle prend pour acquis les contributions au changement et ne vise pas d'en attribuer la responsabilité.

La cartographie des incidences est applicable au stade de la conception, à l'évaluation à mi-parcours et à l'évaluation ex post.

## Le mixte quantitatif/qualitatif

Les méthodes mixtes consistent à intégrer les méthodes qualitatives aux méthodes quantitatives.

Le mixte quantitatif/qualitatif est généralement utilisé dans les évaluations du fait des possibilités qu'elles offrent dans la mesure, l'analyse et la présentation de données quantitatives et qualitatives contribuant ensemble à la matérialisation des changements réalisés.

### Message 17

Le recours aux méthodes qualitatives nécessite la maîtrise de logiciels tels que NVIVO (voir Annexe 3). Il est fortement recommandé d'avoir un ou plusieurs spécialistes en analyse qualitative dans les équipes d'évaluation ou, à défaut, faire appel à des spécialistes en analyse qualitative externes à l'équipe d'évaluation.

## RÉSUMER LA STRATÉGIE D'ÉVALUATION

La conception méthodologique de l'évaluation nécessite souvent la construction d'une matrice comprenant les questions d'évaluation alignées selon les critères d'évaluation et décomposables en variables, les indicateurs de résultat pour chaque variable et les sources de données et méthodes de collecte de données. En effet, la matrice d'évaluation facilite l'élaboration du plan de collecte des données pour répondre aux questions et permet d'établir un lien entre les questions, les approches et les outils de l'évaluation.

La matrice d'évaluation est en définitive l'outil qui permet de résumer la stratégie d'évaluation. La matrice d'évaluation est un tableau bidimensionnel qui identifie les interactions entre les activités de l'évaluation figurant sur un axe (indicateur, méthode de collecte des données, sources de données, devis de mesure) et la ou les variable(s) induite(s) par les questions de l'évaluation sur l'autre axe.

Dans ce tableau, on peut inscrire les interactions entre les variables et les activités à mener dans les cases ou intersections pour donner une vue synoptique de l'ensemble des opérations de collecte et d'analyse des données.

La conception méthodologique de l'évaluation nécessite souvent la construction d'une matrice comprenant les questions de l'évaluation alignées selon les critères d'évaluation et décomposables en sous-questions, les indicateurs de résultat pour chaque sous-question et les sources de données et méthodes de collecte de données.

## RÉDIGER OU AMENDER LES TERMES DE RÉFÉRENCE

Les termes de référence de l'évaluation définissent la portée, les exigences et les attentes de l'évaluation et servent d'orientation et de repères tout le long du processus de l'évaluation.

Si le premier projet de termes de référence émane de l'unité de gestion du programme ou du commanditaire de l'évaluation, les parties prenantes de l'évalu-

ation devraient se consulter pour élaborer, à leur tour, des termes de référence qui assurent que leurs préoccupations sont abordées, ou commenter et amender les termes de référence de l'unité de gestion dans le sens de leurs intérêts. Les termes de référence de l'évaluation devraient être explicites et ciblés pour donner une mission claire à l'équipe d'évaluation.

## CHOISIR L'ÉQUIPE D'ÉVALUATION

Lorsque l'option de mener une évaluation par une équipe interne est prise, celle-ci est choisie sur la base de la connaissance et/ou de l'expérience en évaluation. Lorsque l'option pour une évaluation externe est prise, se pose alors la question du choix de l'évaluateur ou de l'équipe d'évaluation. Généralement, un appel d'offres basé sur les termes de référence de l'évaluation est publié en direction des experts susceptibles d'avoir les compétences pour la réalisation de l'évaluation. L'appel d'offres indique la forme que doit prendre le dossier de candidature, notamment la proposition technique qui doit renseigner sur la compréhension par le soumissionnaire des enjeux de la demande d'évaluation (Note conceptuelle de compréhension des termes de référence) ainsi que la méthodologie générale qu'il propose pour la conduite de l'évaluation et les compétences et aptitudes dont il dispose et références qui permettent de s'assurer de la véracité de ces compétences et aptitudes. Le dossier de candidature doit aussi comprendre en pli séparé, une offre financière du soumissionnaire indiquant le détail des honoraires, débours, compensations et autres frais entrant dans le cadre de la prise en charge financière de l'évaluation. Durant la phase de la conception de la proposition de réponse à l'appel d'offres, tout soumissionnaire potentiel pourra adresser des questions d'éclaircissement au Comité de Gestion de l'évaluation en toute transparence et obtenir des réponses en toute transparence.

La soumission de plusieurs offres entraîne une étude des dossiers de candidature par un Comité de Sélection préalablement constitué à cet effet, et composé des personnes compétentes et ayant un intérêt pour la bonne exécution de l'évaluation, et disposant des mêmes critères et barèmes pour leur appréciation des propositions reçues en bonne et due forme. Ces propositions techniques et offres financières sont jugées sur la base de ces critères et barèmes en vue d'identifier celle qui répond le plus à la demande d'évaluation du point de vue des qualités techniques et l'offre financière (rapport qualité/prix).

Le Comité de sélection pourra, durant ses délibérations, estimer poser des questions de clarification aux soumissionnaires en toute transparence.

En général, entre trois (3) et cinq (5) des meilleures propositions permettront de mener des négociations en commençant par la meilleure offre qui, lorsqu'elles n'aboutissent pas, permettront d'envisager la contractualisation avec les offres suivantes, une à une selon le classement, jusqu'à l'aboutissement du processus de sélection qui permettra de sélectionner l'évaluateur ou l'équipe d'évaluation avec lequel (laquelle) le contrat d'évaluation sera signé.

Tableau 2 : Récapitulatif des principales tâches à accomplir dans le cadre de la préparation d'une évaluation et des outils et/ou techniques adaptés à chacune des tâches identifiées.

Tache	Outil/Technique
Définir les rôles et responsabilités dans le processus évaluatif	Voir Annexe 2: Tableau des rôles et responsabilités dans le processus de l'évaluation
Explorer l'intervention	Référentiel de l'intervention Analyse des parties prenantes
Définir l'objectif de l'évaluation	Technique d'élaboration des questions d'évaluation
Définir les critères d'évaluation	Supports du jugement attendu de l'évaluateur
Choisir une méthode quantitative d'évaluation des résultats	Méthode expérimentale Méthode quasi expérimentale Méthode pré expérimentale Méthodes statistiques Techniques d'échantillonnage et de collecte de données quantitatives Techniques et outils de traitement statistique
Choisir une méthode qualitative d'évaluation des résultats	Le changement le plus significatif (MSC) La cartographie des incidences (entre autres méthodes) Techniques de collecte et traitement de données qualitatives Approches participatives
Résumer la stratégie d'évaluation	La matrice d'évaluation Voir Annexe 1 : Exemple de matrice d'évaluation portant sur une question évaluative
Rédiger ou amender les termes de référence	Technique de rédaction de termes de référence (Voir Annexe 7)

## 3ÈME PARTIE : FAIRE UNE ÉVALUATION

# LA MISE EN ŒUVRE DE L’ÉVALUATION

Lors de la phase de mise en œuvre de l’évaluation, sont mises en exergue les principales actions parfois simultanées qui impliquent le groupe commanditaire composé des Comités de pilotage et de gestion de l’évaluation et du groupe de référence d’un côté et le groupe prestataire composé de l’équipe d’évaluateurs de l’autre.

La mise en œuvre de l’évaluation comprend essentiellement la collecte des informations quantitatives et qualitatives, secondaires et primaires susceptibles de répondre aux questions de l’évaluation, l’analyse de ces informations en vue de tirer des conclusions et la production de conclusions. Au cours de la mise en œuvre de l’évaluation, le groupe commanditaire participe dans la mise à disposition des informations disponibles et reçoit, commente et approuve les rapports d’étapes et le rapport préliminaire, le rapport provisoire et final de l’évaluation, entre autres actions d’accompagnement de l’équipe d’évaluation.

La mise en œuvre de l’évaluation nécessite l’élaboration d’une méthodologie qui indique le type d’évaluation, le contexte, les objectifs, les critères, les questions de l’évaluation ainsi que les données à collecter, les indicateurs qui permettent de mesurer les écarts entre les résultats projetés et les effets (résultats et impacts) ainsi que les sources de collecte des données et les méthodes d’analyse.

32

## Collecter les données

Les données sont quantitatives ou qualitatives, primaires ou secondaires. Les données secondaires sont collectées à partir des documents disponibles, dont ceux relatifs à la conception et la mise en œuvre du programme évalué, et les données primaires collectées par l’équipe d’évaluation sur le terrain dans le cadre d’enquêtes et d’observations. La plupart des évaluations sont conçues de manière à utiliser les méthodes quantitatives et les méthodes qualitatives (méthode mixte : quantitative/qualitative). Les principales tâches liées à la mise en œuvre d’une évaluation consistent à :

### LES DONNÉES SECONDAIRES

Les données secondaires sont constituées des données quantitatives et qualitatives tirées de sources documentaires. Le référentiel de l’évaluation est un document ou un ensemble de documents à prendre en compte parmi les données secondaires, ainsi que les documents de rapports de suivi.

En dehors de ceux mis à la disposition de l’équipe des évaluateurs par le Gestionnaire de l’évaluation, d’autres sources de données secondaires comme les documents de politiques connexes et les statistiques officielles sanitaires, démographiques, sociales, peuvent être identifiées, consultées et fournir des données utiles et fiables, pour renseigner sur le contexte et son évolution.

### LES DONNÉES PRIMAIRES

Les données primaires sont constituées de données quantitatives et qualitatives tirées de l’observation des sites du programme et des entretiens avec les groupes de cibles.

Cependant, l’équipe d’évaluateurs peut collecter des données auprès de groupes qui n’ont pas été ciblés par l’intervention. Par exemple, dans une évaluation expérimentale, comme dans une évaluation quasi expérimentale, il sera nécessaire de déterminer le groupe de traitement et le groupe de contrôle ou de comparaison.

### COLLECTER LES DONNÉES QUANTITATIVES

Les données quantitatives sont collectées par le moyen du questionnaire.

La conception du questionnaire comprend l’élaboration des questions, les choix de réponses ainsi que le pré-test. Cette conception est précédée d’une définition des objectifs et des hypothèses générales, et qui, en définitive, doivent aider à répondre aux questions et sous-questions en lien avec celles qui sont posées dans la matrice de l’évaluation. Ces objectifs et hypothèses contribuent à définir les aspects du problème qui seront recueillis par les différentes questions .

Le questionnaire précise l’information recherchée par items. Un item est la plus petite unité d’information pouvant être mesurée ou observée. Les items peuvent être regroupés en domaines quand il s’agit de problèmes à plusieurs composantes.

Dans la rédaction du questionnaire, le concepteur doit utiliser un vocabulaire compréhensible pour toutes les personnes de la population ciblée, et les questions doivent être standardisées et avoir le même sens pour toutes ces personnes.

Le questionnaire ne doit pas être une juxtaposition de questions indépendantes, mais plutôt être planifié comme un entretien dans lequel les questions se suivent dans un ordre logique.

La présentation du questionnaire aux répondants est une étape décisive pour l’obtention de réponses sincères et réfléchies. La longueur du questionnaire peut provoquer une lassitude chez les répondants et entraîner des entorses dans les réponses ou une interruption de l’entretien.

Pour toutes ces raisons, le questionnaire doit être testé sur un nombre restreint de personnes issues de la population ciblée avant d’être administré à l’échelle de l’échantillon.

## COLLECTER LES DONNÉES QUALITATIVES

Les données qualitatives sont collectées par les moyens que sont principalement l'entrevue, l'observation et le focus group.

- **L'entrevue (les entretiens)**

- *l'entretien libre ou non-directif*

L'entretien libre est principalement utilisé pour les récits de vie. Il procède par une question initiale ou peu de questions pour laisser le sujet s'exprimer sans être arrêté ou orienté. Mais l'enquêteur peut être amené à relancer le sujet s'il ne parvient pas à continuer son récit.

- *l'entretien semi-directif*

Les questions sont ouvertes, mais en nombre restreint suivant un plan précis.

- *l'entretien directif*

L'entretien directif consiste en de nombreuses questions courtes, précises et enchaînées qui se répondent souvent par oui ou par non.

Ces entretiens peuvent être réalisés en face à face, par téléphone ou internet, et les supports peuvent être oraux ou textuels, et se dérouler dans le contexte résidentiel, sur le lieu de travail, de loisirs, etc.

- **L'observation directe**

L'observation directe consiste à collecter des données de terrain en utilisant divers supports, sans intervenir de manière significative sur le terrain. Elle consiste à suivre et à étudier méthodiquement le comportement dans une situation naturelle et permettent d'éviter le biais que pourrait induire l'interviewé lorsqu'il décrit son comportement.

L'observation directe est particulièrement adaptée si l'on s'intéresse aux actions proprement dites et aux dynamiques ainsi qu'aux interactions entre les individus dans une situation donnée.

L'observation directe peut être une alternative ou utilisée conjointement ou en complément à d'autres

modes de recueils tels que le questionnaire, l'expérimentation, les entretiens individuels, les entretiens de groupe, les études de cas.

Elle peut varier selon le contexte et être renforcée par l'observation numérique (forums participatifs en ligne, journaux en ligne, etc.).

- **Le focus group**

Le focus group ou entretien de groupe réunit idéalement entre 6 et 12 individus engagés à discuter extensivement autour d'une à 3 questions progressivement introduites par l'enquêteur qui n'intervient pas pour orienter les réponses.

Les données qualitatives ainsi recueillies sont retranscrites en vue de leur analyse de contenu. Le codage est un processus qui a pour but d'explorer ligne par ligne, étape par étape, les textes d'interviews.

Il consiste à décrire, classer et transformer les données qualitatives textuelles issues des échanges en fonction d'une grille d'analyse. La grille d'analyse de contenu est construite à partir d'une catégorisation faite sur la base de critères et d'indicateurs appelés catégories d'analyse.

## TRAITER ET ANALYSER LES DONNÉES

Le traitement des données qualitatives et quantitatives est une étape importante de la mise en œuvre de l'évaluation.

Ce traitement des données qualitatives et quantitatives vise à donner du sens aux données en les transformant en informations utiles à la compréhension des phénomènes observés dans le cadre de l'évaluation. Il peut se faire par le moyen de logiciels spécialisés dans les deux types de données (qualitatives et quantitatives).

## DÉGAGER LES CONCLUSIONS

Les conclusions doivent être déduites de l'analyse et de la discussion sur les résultats de l'analyse et doivent constituer des réponses aux questions de l'évaluation.

Tableau 3 : Récapitulatif des principales tâches à accomplir dans le cadre de la mise en œuvre d'une évaluation et des outils et/ou techniques adaptés à chacune des tâches identifiées.

Tâche	Outil/Technique
Collecter les données secondaires	Exploiter les données existantes
Collecter les données primaires	Collecter ses propres données (l'évaluateur)
Collecter les données quantitatives	Questionnaire
Collecter les données qualitatives	L'entrevue (les entretiens) L'observation directe Le focus group
Traiter et analyser les données	Les logiciels (Spss pour les données quantitatives, entre autres logiciels, Nvivo pour les données qualitatives, entre autres logiciels) Voir Annexe 3 : Exemples de logiciel de traitement de données



## 4ÈME PARTIE : PRÉSENTER ET COMMUNIQUER LES RÉSULTATS DE L'ÉVALUATION

# LE RAPPORT

## REDIGER LE RAPPORT

La structure du rapport d'évaluation est basée sur des normes et standards qui comprennent les éléments suivants :

### Le résumé analytique ou résumé exécutif

Le résumé analytique ou résumé exécutif tient environ sur trois (3) pages et donne un aperçu des éléments essentiels du rapport.

Le résumé analytique ou résumé exécutif doit être précis et complet pour pouvoir être lu de manière autonome. Il peut comprendre :

- une vue d'ensemble de l'objet de l'évaluation ;
- les objectifs de l'évaluation ;
- la méthodologie de l'évaluation ;
- les résultats et les conclusions ;
- les principales recommandations.

### L'objet et le but de l'évaluation

L'évaluation devrait décrire clairement l'objet qui est en train d'être évalué, notamment le modèle logique et/ou la chaîne des résultats attendus, les stratégies d'exécution et l'intervention, la couverture des interventions et les principales hypothèses.

Le but de l'évaluation doit être clairement défini, notamment la raison pour laquelle cette évaluation s'imposait à ce moment précis, ainsi que les informations recherchées, qui en a besoin et comment seront-elles utilisées.

### Les questions et critères d'évaluation

Le rapport d'évaluation doit préciser les questions d'évaluation ainsi que les critères d'évaluation utilisés.

### La méthodologie

Le rapport doit contenir une description claire de la méthodologie et une description complète de la participation des acteurs dans l'évaluation.

Elle détaillera quelles données ont été collectées, comment elles ont été collectées et par qui, ainsi que les limites possibles de l'évaluation.

### Les résultats

Les résultats de l'évaluation devraient porter sur les critères de l'évaluation ainsi que les critères transversaux (genre, droits humains, gestion axée sur les résultats, lutte contre la pauvreté, protection de l'environnement, etc.).

Les résultats seront objectivement présentés avec évidence. Les limites ou gaps dans les données ainsi que les résultats non anticipés seront également présentés et discutés.

Les raisons de la non réalisation des droits, des échecs et particulièrement des contraintes récurrentes doivent être identifiées autant que possible.

Enfin, les résultats doivent être présentés avec clarté, logique et cohérence.

### Les conclusions

Les conclusions devraient porter sur les objectifs d'évaluation et les questions clés, s'appuyer sur des faits et être logiquement liées aux résultats de l'évaluation et donner un aperçu des problèmes importants identifiés et/ou de leurs solutions.

Les enseignements tirés sont des contributions à la connaissance générale. Ils doivent être bien soutenus par les résultats et les conclusions présentés.

### Les recommandations

Le rapport doit décrire le processus de l'élaboration des recommandations, notamment la consultation avec les parties prenantes et doit identifier clairement le groupe cible pour chaque recommandation formulée.

### Les annexes

Les annexes devraient contenir les termes de références, la bibliographie, la liste des personnes interviewées et des sites visités, des informations supplémentaires sur la méthodologie notamment sur ses limites, les outils de collecte de données, les biographies des évaluateurs, la justification de la composition de l'équipe et le code de conduite éthique signé par les évaluateurs.

En dehors de ces annexes requises, l'évaluateur et le commanditaire de l'évaluation peuvent y ajouter tout autre item utile pour une meilleure compréhension du rapport de l'évaluation.

35

### Message 18

Il est nécessaire de s'inscrire dans une perspective de mété évaluation (l'évaluation de l'évaluation) afin de garantir la qualité des évaluations et des rapports d'évaluation sur lesquels se base le travail de mété évaluation.

## DIFFUSER LES RESULTATS

La diffusion du rapport et des résultats de l'évaluation peut se faire par plusieurs canaux, selon les objectifs de l'évaluation et l'intérêt des parties prenantes et du public.

Un atelier de partage des résultats de l'évaluation est souvent organisé par les commanditaires de l'évaluation, avec la participation des parties prenantes ainsi qu'une couverture médiatique de l'événement, animé par l'équipe d'évaluation.

Il est recommandé de préparer en amont, un plan de diffusion des résultats de l'évaluation en quatre (4) étapes :

1. Identifier les difficultés de communication et de compte rendu.
2. Définir l'objectif de communication.
3. Choisir les meilleures méthodes de communication.
4. Élaborer une stratégie de communication.

Le détail de ce plan doit prendre en compte les points suivants :

### Identifier les difficultés de communication et de compte rendu

- anxiété générale par rapport à l'évaluation ;
- manque de planification de la communication en amont ;
- faiblesse de la culture organisationnelle des parties prenantes.

### Définir l'objectif de la communication

- déterminer le besoin d'être informé pour chaque individu, groupe ou organisation ;
- sélectionner les résultats les plus intéressants pour chaque individu, groupe ou organisation ;
- Impliquer les individus, les groupes et les organisations les plus concernés ou intéressés.

36

### Choisir les meilleures méthodes de communication

- évaluer le niveau de connaissance des parties prenantes pour définir la meilleure méthode de communiquer avec elles sur les résultats ;
- renforcer leur expérience et leur intérêt pour l'utilisation des résultats des évaluations ;
- utiliser les langues qui leur sont accessibles.

### Élaborer une stratégie de communication

#### Message 19

Il est important de s'entendre avec le commanditaire de l'évaluation sur les modalités de diffusion et de communication des résultats de l'évaluation.

#### Message 20

La communication est un enjeu majeur pour l'évaluation. L'utilisation des résultats de l'évaluation, qui est l'objectif ultime visé, est tributaire de la communication efficace de ces résultats.

Technique	Objectifs	Moyens
<i>Avant l'évaluation</i>		
Convaincre	Convaincre de la nécessité de l'évaluation	Arguments persuasifs sur les avantages de l'évaluation
<i>Pendant l'évaluation</i>		
Communiquer	Améliorer l'engagement pour l'évaluation	Utiliser les données disponibles et extrapolier sur les avantages de l'évaluation
<i>Après l'évaluation</i>		
Persuasion	Convaincre pour le changement	Dialogue, preuve de l'utilité de l'évaluation
Information	Aider à prendre la bonne décision sur la diffusion des résultats de l'évaluation	Documentation sur l'évaluation, visite des sites du programme, couverture médiatique.
Consultation	Croiser les points de vue avec les individus, les groupes et les organisations qui manifestent une indifférence ou des réticences	Visites, réunions.
Codécision	Anticiper les résistances	Réunions, groupes de discussion
Délégation	Partage des responsabilités	Partenariat ou ententes de diffusions

## GLOSSAIRE

**Analyse coûts-avantages :** Analyse comparant les avantages que procure un programme aux coûts associés à son exécution. Une valeur monétaire est attribuée aux avantages et aux coûts.

**Analyse coûts-efficacité :** Analyse comparant les coûts d'un programme à ses retombées. Dans cette analyse, les retombées ne sont pas transformées en valeur monétaire.

**Analyse statistique :** Manipulation des données numériques ou catégoriques afin de prévoir des phénomènes, de tirer des conclusions sur des rapports entre variables ou de généraliser des résultats.

**Attribution :** Estimation de la mesure dans laquelle les résultats sont attribuables à un programme, ce qui signifie que le programme a eu des effets incrémentiels.

**Biais :** Facteur de distorsion, d'influence indue ou d'erreur affectant la validité interne de la mesure.

**Données qualitatives :** Observations catégoriques plutôt que numériques portant souvent sur les attitudes, les perceptions et les intentions.

**Données quantitatives :** Données numériques.

**Ecart type :** Le degré de regroupement des mesures individuelles autour de la moyenne, pour un ensemble de mesures observables sur une échelle d'intervalles.

**Echantillonnage aléatoire :** Sélection des unités d'une population fondée sur le principe de la répartition au hasard. Il existe pour chaque unité de la population une probabilité calculable (différente de 0) d'être choisie.

**Echantillonnage non aléatoire :** Choix des unités d'un échantillon effectué de façon à ce que chaque unité de la population n'ait pas une probabilité calculable différente de zéro d'être choisie pour faire partie de l'échantillon.

**Echantillonnage stratifié :** Technique d'échantillonnage aléatoire suivant laquelle une population est divisée en couches relativement homogène appelées strates. Des échantillons appropriés sont choisis dans chaque strate.

**Enquête :** Méthode de collecte des données qui suppose une démarche planifiée en vue de recueillir les données requises d'un échantillon (ou au moyen d'un recensement complet). La population visée est composée des personnes ou des entités touchées par le programme (ou de personnes ou entités semblables).

**Entretien non structuré :** Technique d'entrevue utilisant une conversation normale menant à des questions spontanées souvent dans le cadre de l'observation régulière des activités d'un programme.

**Groupe expérimental :** Groupe de sujets qui bénéficie d'un programme. Le groupe expérimental est aussi appelé groupe de traitement ou groupe exposé au programme.

**Groupe témoin :** Dans les modèles expérimentaux, groupes de sujets qui éprouve toutes les influences, sauf celles du programme, exactement de la même façon que le groupe exposé à un traitement (qu'on appelle aussi groupe expérimental).

**Modèle quasi expérimental :** Structure d'étude utilisant des groupes de comparaison pour faire des inférences causales, mais sans recours à la randomisation pour constituer un groupe expérimental et un groupe témoin. Le premier groupe est ordinairement acquis ; le groupe expérimental est choisi pour correspondre le plus possible avec lui, de façon à permettre des inférences sur les effets incrémentiels du programme.

**Randomisation :** Utilisation d'un plan de probabilité pour établir un échantillon.

**Taille de l'échantillon :** Nombre d'unités à enquêter.

## Annexe 1

Exemple de matrice d'évaluation portant sur une question évaluative

<b>Question d'évaluation :</b> <b>Le programme d'appui aux micro-entreprises a-t-il contribué à réduire la pauvreté dans les régions ciblées ?</b>				
Variables à l'étude	Indicateur	Méthode de collecte des données	Sources de données	Devis de mesure
<i>Variable 1 : Taux de pauvreté dans les régions ciblées</i>	Indicateur 1.1. : Perception du niveau de pauvreté par les populations des régions ciblées entre l'avant et l'après intervention du programme.	Entrevues	Groupe de bénéficiaires A	Grille d'analyse des données qualitatives Logiciel Nvivo
		Questionnaire	Groupe de Bénéficiaires A et Groupe de contrôle.	Devis quasi expérimental
<i>Variable 2 : Couverture géographique du programme d'appui aux micro-entreprises</i>	Indicateur 2.1 : Nombre de régions touchées par le programme	Entrevues	Groupe de bénéficiaires A	Grille d'analyse des données qualitatives
		Étude documentaire	Documents de conception du programme Documents de suivi du programme	Analyse de contenu
	Indicateur 2.2 : Taux de pauvreté dans les régions ciblées avant et après l'intervention.	Étude documentaire	Statistiques officielles.	Analyse statistique

## Annexe 2

### Rôles clés et responsabilités dans le processus d'évaluation

Personne/organisation	Rôles et responsabilités
<i>Le Commissaire de l'évaluation</i>	<ul style="list-style-type: none"> <li>➤ Détermine les résultats du programme et grandes lignes de l'évaluation</li> <li>➤ Fournit des conseils au responsable de l'évaluation</li> <li>➤ Veille à la transparence et à l'indépendance de l'évaluation tout le long du processus</li> <li>➤ Alloue les fonds et les ressources humaines nécessaires.</li> </ul>
<i>Le Responsable de l'évaluation</i>	<ul style="list-style-type: none"> <li>➤ Dirige l'élaboration des termes de références de l'évaluation</li> <li>➤ Assure la sélection et le recrutement des experts évaluateurs</li> <li>➤ Gère les dispositions contractuelles, le budget et le personnel impliqué dans l'évaluation</li> <li>➤ Fournit aux évaluateurs un soutien administratif</li> <li>➤ Met à la disposition des évaluateurs les données requises</li> <li>➤ Assure la liaison avec l'équipe d'évaluation, l'unité de gestion du programme et le groupe de référence et/ou les partenaires clés</li> <li>➤ Garantit l'approche globale et la transparence</li> <li>➤ Révise le rapport initial, le rapport du projet d'évaluation, les rapports de la mission d'évaluation et le rapport final</li> <li>➤ Assure que les normes de qualité ont été prises en considération dans le rapport final.</li> </ul>
<i>Le groupe de référence (représentants des parties prenantes et des bénéficiaires)</i>	<ul style="list-style-type: none"> <li>➤ Définit et confirme le profil, les compétences, les rôles et responsabilités du responsable de l'évaluation et proposent des candidats à cette fonction</li> <li>➤ Participe à la rédaction et à la révision de l'ébauche des TDR</li> <li>➤ Prête assistance à la collecte des données requises</li> <li>➤ Supervise les progrès et la conduite de l'évaluation</li> <li>➤ Révise le rapport du projet d'évaluation</li> <li>➤ Assure que le rapport final répond aux normes de qualité.</li> </ul>
<i>Le Groupe des évaluateurs</i>	<ul style="list-style-type: none"> <li>➤ Réalise les dispositions contractuelles en accord avec les normes éthiques et déontologiques</li> <li>➤ Développe la matrice de l'évaluation comprenant les critères et les questions de l'évaluation, les indicateurs, les données à collecter, les sources de données, les méthodes d'analyse</li> <li>➤ Rédige et soumet les rapports requis, jusqu'au rapport final de l'évaluation</li> <li>➤ Communique avec le commissaire, le responsable et le groupe de référence de l'évaluation sur les progrès de la mission d'évaluation, les contraintes et toutes autres questions liées à la bonne mise en œuvre de l'évaluation</li> <li>➤ Prend en considération les remarques et suggestions du responsable et des parties prenantes de l'évaluation.</li> <li>➤ S'assure que tous les livrables, y compris les bases de données, le rapport final, les listes de personnes interrogées et autres documents sont soumis au responsable de l'évaluation.</li> </ul>

Source : PNUD : Guide de la planification, du suivi et de l'évaluation axés sur les résultats de développement, PNUD, New York, 2009, p. 150; Site web du guide : [www.undp.org/oe/handbook](http://www.undp.org/oe/handbook)

## Annexe 3

### Deux exemples de logiciel de traitement des données :

#### - Un logiciel de traitement des données quantitatives : SPSS

La première version du logiciel Statistical Package for Social Sciences (SPSS) a été mise en vente en 1968 et fait partie des programmes utilisés pour l'analyse statistique en sciences sociales. Il est utilisé par des chercheurs en économie, en science de la santé, par des compagnies d'études, par le gouvernement, des chercheurs de l'éducation nationale, etc.

En plus de l'analyse statistique, la gestion des données (sélection de cas, reformatage de fichier, création de données dérivées) et la documentation des données (un dictionnaire de métadonnées est sauvegardé avec les données) sont deux autres caractéristiques du logiciel.

SPSS impose des contraintes concernant la structure interne des fichiers, les types des données, le traitement des données et les fichiers de correspondance. Les séries de données de SPSS ont une structure en tableau à deux dimensions où les rangées représentent typiquement les cas (tels des individus ou des foyers) et où les colonnes représentent les mesures (telles que âge, sexe ou revenu du foyer).

Il n'y a que deux types définis de données : nombres ou chaînes de caractères. Tous les traitements de données se déroulent de façon séquentielle les cas à la suite les uns des autres dans l'ordre du fichier. Les fichiers peuvent être comparés un à un, un envers les autres, mais pas plusieurs d'un coup envers les autres.

Différentes versions de SPSS existent pour Windows, Mac OS X et Unix. La version pour Windows est mise à jour plus fréquemment et possède plus de caractéristiques que les versions pour les autres systèmes d'exploitation.

SPSS peut lire et écrire des données à partir et dans des fichiers texte en caractères ASCII (incluant les fichiers hiérarchiques), d'autres paquetages statistiques, des feuilles de calcul et des bases de données (des tables de base de données relationnelles externes via ODBC et SQL).

#### - Un logiciel de traitement des données qualitatives : NVivo

Les données qualitatives proviennent de l'observation et des entretiens. Elles peuvent être codées et soumises à un traitement manuel ou statistique à l'aide de logiciels spécifiques dont le NVivo7. Le NVivo7 est le mariage entre le N6 et NVivo2. Il permet de combiner les actions de codage, la construction quantitative de liaisons, le façonnage et la modélisation en vue de questionner et d'interpréter en profondeur des données en faveur de la compréhension des phénomènes.

## Annexe 4

### Les méthodes expérimentales et quasi expérimentales

Méthodologie	Description	Groupe de comparaison	Hypothèses requises	Données requises
<i>Avant et après</i>	Mesure l'amélioration ou le changement connu par les participants au fil du temps.	Les participants eux-mêmes avant de participer au programme.	Le programme a été le seul facteur ayant influé sur les changements dans le résultat mesuré au fil du temps.	Données avant et après pour les participants au programme
<i>Différence simple</i>	Mesure la différence entre la participation au programme et les non participants après l'achèvement du programme.	Personnes n'ayant pas participé au programme (quelle que soit la raison) mais pour qui des données ont été recueillies après le programme.	Les non participants identiques aux participants à l'exception de la participation au programme et ils avaient les mêmes chances d'entrer dans le programme avant son démarrage.	Données après pour les participants au programme et les non participants.
<i>Divergence dans la différence</i>	Mesure l'amélioration ou le changement au fil du temps des participants au programme par rapport à l'amélioration ou changement des non participants	Personnes n'ayant pas participé au programme (quelle que soit la raison) mais pour qui des données ont été recueillies après le programme.	Si le programme n'existe pas, les deux groupes auraient eu des trajectoires identiques au cours de cette période.	Données avant et après pour les participants au programme et les non participants
<i>Régression à variables multiples</i>	Les personnes ayant reçu le traitement sont comparées à celles qui ne l'ont pas reçu et d'autres facteurs qui pourraient expliquer les différences sont neutralisés.	Personnes n'ayant pas participé au programme (quelle que soit la raison) mais pour qui des données ont été recueillies après le programme. Dans ce cas, les données n'incluent pas seulement des indicateurs de résultat mais aussi d'autres variables	Les facteurs qui ont été exclus (parce qu'ils ne sont pas observables ou n'ont pas été mesurés) ne faussent pas le résultat parce que soit ils ne sont pas corrélés avec le résultat, soit parce qu'ils ne sont pas différents entre les participants et les non participants.	Résultats et variables de contrôle pour les participants et les non participants au programme.
<i>Appariement statistique</i>	Les personnes du groupe de contrôle sont comparées à des personnes similaires dans le groupe expérimental.	<i>Correspondance exacte</i> : pour chaque participant, au moins un non participant qui est identique pour certaines caractéristiques. <i>Appariement sur les scores de propension</i> : les non participants qui ont un mélange de caractéristiques prédisant qu'ils seraient susceptibles de participer en tant que participants.	Les facteurs qui ont été exclus (parce qu'ils ne sont pas observables ou n'ont pas été mesurés) ne faussent pas le résultat parce que soit ils ne sont pas corrélés avec le résultat, soit parce qu'ils ne sont pas différents entre les participants et les non participants.	Résultats et variables d'appariement pour participants et non participants.
<i>Méthode de la régression sur la discontinuité</i>	Les personnes sont classées en fonction de critères spécifiques mesurables. Une limite d'inclusion détermine si une personne est admissible à participer. Les participants sont ensuite comparés aux non-participants et le critère d'admissibilité est neutralisé.	Personnes qui ne sont pas proches de la limite d'inclusion mais qui tombent du « mauvais » côté et ne reçoivent donc pas le programme.	Après avoir neutralisé les critères (et d'autres mesures retenues), les différences restantes entre les personnes directement en dessous et au-dessus de la valeur limite ne sont pas statistiquement significatives et ne faussent pas le résultat. Pour ce faire, la condition nécessaire mais suffisante et que les critères d'inclusion soient scrupuleusement respectés.	Résultats mesurés sur les critères (et tous les autres contrôles).

## Annexe 5

### Quelques outils de support à l'évaluation

Description	Usage	
<b>OUTILS D'EXPLORATION ET D'INVESTIGATION</b>		
<i>L'enquête structurée</i>	Les questions sont fixées avant l'entretien sur un questionnaire standardisé et destiné à un échantillon représentatif de la population étudiée. Les questions sont formulées en Quoi? Combien?	Analyse statistique des données.
<i>L'enquête semi structurée</i>	Les questions sont modifiées et les réponses appellent d'autres questions au cours de l'entretien, et adressées à un échantillon qui évolue au fur et à mesure des entretiens, en vue d'explorer un maximum de situations. Les questions sont formulées en Pourquoi? Comment?	Analyse qualitative des données.
<i>L'observation évaluative (observation directe, observation participante)</i>	Sont observés les objets, les évènements, les processus et les relations, ainsi que les comportements, les pratiques et les attitudes en partant de questions simples et informelles posées aux personnes rencontrées.	Analyse qualitative des données sous forme de diagramme.
<i>Le focus group</i>	Sur la base d'un guide de discussion, une discussion ouverte est organisée entre 6 à 12 porteurs d'information ayant des positions sociales différentes et des intérêts dissemblables, pouvant déboucher à de synthèses.	Analyse qualitative des perceptions, attitudes et des croyances sur une base égalitaire de communication.
<b>OUTILS DE CLASSIFICATION</b>		
<i>L'échelle de notation</i>	Des notes inscrites sur une échelle de 0 à 5, 0 à 7 ou 0 à 10 sont attribuées par les acteurs des variables (exemple de l'échelle de Likert). L'échelle de notation peut aussi être figurée en des images exprimant les opinions des acteurs (exemple de la carte des scores).	Analyse statistique des attitudes des niveaux d'accord et de désaccord, de satisfaction ou d'insatisfaction, etc.
<i>Le calendrier</i>	Les différentes variables qui affectent les acteurs dans le temps sont représentées par semaine, mois, semestre, année...	Analyse qualitative de variables socio-temporelles.
<i>L'arbre à problèmes</i>	A partir d'un problème principal défini de manière consensuelle, clairement formulé et figuré sous la forme d'un tronc d'arbre, chacune des causes principales est indiquée au niveau racinaire et les conséquences correspondant à chacune des causes inscrites au niveau du feuillage. Les solutions peuvent être figurées au niveau du feuillage sous la forme de fruits de l'arbre.	
<b>OUTILS D'ANALYSE ORGANISATIONNELLE</b>		
<i>Le diagramme de VENN</i>	Le diagramme de VENN reflète la manière de faire et de vivre, les relations, les réseaux, les conflits et les systèmes de solidarité au sein de l'organisation ou de l'institution.	Analyse institutionnelle.
<i>Le diagramme de polarisation</i>	Plusieurs variables en termes d'accès (à la santé, à l'éducation, à l'eau et autres services de base), à l'administration, ainsi que les relations avec l'extérieur sont représentées selon leur position géographique sur un site.	Analyse de l'accessibilité et de la couverture en service.

## Annexe 6

### Techniques d'échantillonnage

Lorsqu'il s'avère irréalisable ou inefficient de faire étudier toute la population comme dans le cadre d'un recensement, il est alors possible de procéder à un échantillonnage.

L'échantillonnage dépend de la population étudiée et de la méthode utilisée dans l'évaluation d'impact. Par exemple, l'approche quasi expérimentale par comparaison nécessite une représentativité de chaque groupe pris individuellement.

En principe, un échantillon est représentatif si les unités qui le constituent ont été choisies par un procédé tel que tous les membres de la population ont la même probabilité de faire partie de l'échantillon. Si ce n'est pas le cas, on dira que l'échantillon est biaisé, du fait que certains individus avaient plus de chances que d'autres d'être choisis, et que par conséquent, les catégories auxquelles ils appartiennent occuperont plus de place qu'elles ne devraient dans l'échantillon. Cette situation entraîne que les caractéristiques de l'échantillon sont systématiquement différentes de celle de la population.

Il existe différentes manières d'échantillonner, dont le sondage simple, le sondage stratifié, le sondage en grappes, le sondage par quotas et le sondage par effet boule de neige.

Ainsi, la nature et la portée de l'échantillonnage doivent satisfaire aux trois exigences suivantes :

- Les constatations doivent être généralisables à l'ensemble de la population visée ;
- La méthode choisie doit satisfaire aux exigences minimales de précision et le degré de précision ainsi que le niveau de confiance attendus de l'enquête indiqués ;
- Les coûts d'échantillonnage doivent respecter les limites du budget.

**Le processus d'échantillonnage s'effectue selon les six étapes suivantes :**

- Définir la population de manière détaillée en indiquant la date, le lieu et les caractéristiques socioéconomiques pertinents ;
- Préciser la base de sondage, soit une liste des éléments de la population ;
- Préciser l'unité d'échantillonnage, c'est-à-dire l'unité employée dans l'enquête (lieu, ménage, entreprise...) ;
- Préciser la méthode d'échantillonnage, c'est-à-dire la méthode utilisée pour choisir les unités d'échantillonnage (échantillonnage systématique, stratifiée) ;
- Déterminer la taille de l'échantillon, notamment le nombre d'unités d'échantillonnage, ainsi que le pourcentage inclue dans l'échantillon ;
- Choisir l'échantillon.

## Annexe 7

### Termes de référence

#### Exemple de plan de termes de référence :

- L'introduction des termes de référence doit porter sur la structure commanditaire, ses missions, le programme à évaluer et la provenance des fonds mobilisés pour sa mise en œuvre ;
- L'objet ou le but de l'évaluation doit permettre aux experts de connaître les objectifs de l'évaluation (ex ante, in itinere, ex post, etc.) ;
- La description du programme à évaluer donne un aperçu du contexte et des composantes du programme, de l'origine des fonds, des limites géographiques, des parties prenantes clés, des cibles, de la stratégie, des activités et des résultats escomptés ;
- Le contexte et la justification de l'évaluation indiquent les principales raisons à la base de la décision de l'évaluation, l'origine de la commande et le type d'évaluation ;
- La portée de l'évaluation indique la couverture géographique et les principales thématiques en jeu dans l'évaluation ;
- Les parties prenantes que sont les groupes, institutions, structures ou personnes impliqués sont indiquées selon qu'elles sont rangées parmi les partenaires clés, les cibles/bénéficiaires ou prestataires de service ;
- Les critères et les questions de l'évaluation sont indiqués, par exemple les questions relatives à la cohérence, à la pertinence, à l'efficacité, à l'efficience et à la durabilité,
- Les produits escomptés, dans le cadre de l'évaluation (ou livrables), sont cités avec des détails sur leur date de livraison et modalités d'approbation ;
- Le calendrier et la durée sont indiqués éventuellement avec des détails sur les réunions de cadrage, de débriefing et de présentation du rapport provisoire et du rapport final ;
- L'expertise et les compétences requises sont mentionnées.

**La conception de l'évaluation a des implications majeures dans l'implication des décideurs, des parties prenantes clés et des participants au programme, ainsi que dans la mobilisation des ressources en temps et en argent, et dans l'utilisation des résultats de l'évaluation.**

## REFERENCES

- Loi n° 2016-046/ du 23 septembre 2016 portant loi organique fixant l'organisation, les règles de fonctionnement de la Cour Suprême et la procédure suivie devant elle
- Ordinance Numéro 00-051/P-RM Du 27 SEP. 2000 Portant création du Contrôle Général des Services Publics
- Loi Numéro 00-067/Du 30 NOV. 2000 Portant ratification de l'Ordinance Numéro 00-051/P-RM Du 27 SEP. 2000 Portant création du Contrôle Général des Services Publics
- Décret Numéro 01-067/P-RM Du 12 FEV. 2001 Fixant l'organisation et les modalités de fonctionnement du Contrôle Général des Services Publics
- RÉFÉRENTIEL D'AUDIT COMPTABLE ET FINANCIER : Code d'éthique et de déontologie Secteur public Rapport Définitif – 4 Juillet 2009
- PRÉSENTATION DE LA SECTION DES COMPTES DE LA COUR SUPREME DU MALI- juillet 2018- (Par M. Abdoulaye SOW, Président de la Chambre de contrôle et de jugement des opérations financières des Institutions de la République et des administrations de l'Etat)
- Directive n° 5-97/CM/Uemoa du 16 décembre 1997 portant lois de finances et ses modificatifs ; VU 1
- Directive n° 01/2009/CM/Uemoa du 27 mars 2009 portant Code de transparence dans la gestion des finances publiques au sein de l'UEMOA
- DIRECTIVE N°06/2009/CM/Uemoa PORTANT LOIS DE FINANCES AU SEIN DE L'UEMOA
- Loi de Finances 2019
- « La vision Mali 2025 »
- Cadre Stratégique de Lutte contre la Pauvreté (CSLP).

Banque Mondiale : L'évaluation d'impact en pratique, 2011

Banque Mondiale, Suivi et évaluation : Quelques outils, méthodes et approches, 2004

BARON, Gaëlle et MONNIER, Eric : «Une approche pluraliste et participative : coproduire l'évaluation avec la société civile», in Revue Informations sociales, Septembre 2003.

Fondation Européenne pour la Formation : Guide d'analyse des politiques publiques, 2018

CHARBE-FERRET, Sylvain, DUPONT-COURTADE, Laura, TREISCH, Nicolas : Évaluation des politiques publiques : Expérimentations randomisées et méthodes quasi expérimentales, INRA Paris, 2017

CICR, Planification de projet/programme : Manuel d'orientation, Genève 2010.

FIDA, Manuel d'évaluation, 2ème Edition, 2015

Guide détaillée pour l'évaluation des programmes en matière d'éducation financière. Gouvernement du Canada/ OCDE.

Guide de l'évaluation, Ministère des Affaires Étrangères, Direction générale de la coopération internationale et du développement, Service de la stratégie, des moyens et de l'évaluation, Bureau de l'évaluation : Paris, 2005.

Le POLE Stratégies de développement et finances, Initiative conjointe France-PNUD : Guide méthodologique d'élaboration et de suivi-évaluation des politiques publiques, Tome 2, janvier 2016

Mesure et attribution des résultats des programmes, 3ème édition, Direction des affaires publiques, Secrétariat du Conseil du Trésor, Canada : 2010.

OCDE/CAD: Normes de qualité pour l'évaluation du développement, 2012

OCDE, Glossaire des principaux termes relatifs à l'évaluation et à la gestion axée sur les résultats, 2002

PNUD : Guide de la planification, du suivi et de l'évaluation axée sur les résultats du développement, New York, 2009

UNICEF, Méthodes et modèles quasi expérimentaux, 2014

ROGERS, Patricia : Présentation de l'évaluation d'impact; UNICEF, Note méthodologique n°1, septembre 2014

UNEG: Normes et règles d'évaluation, New York, UNEG New York 2016

WANLIN, Phillip : L'analyse de contenu comme méthode d'analyse qualitative d'entretiens : Une comparaison entre les traitements manuels et l'utilisation de logiciels. Association pour la Recherche qualitative, Université Luxembourg 2007

WHITE, Howard et SHAGUN, Sabarwal, Evaluation d'impact, Note méthodologique n°8, UNICEF, Septembre 2014.

**Par Abdoul Aziz Sall**  
*Spécialiste en évaluation des politiques publiques*