GARANTIR LA QUALITÉ DE L'ÉLABORATION ET DE LA MÉTHODOLOGIE DE L'ÉVALUATION

CHAPITRE 7

Ce chapitre décrit les principales composantes du modèle d'évaluation de la qualité, ainsi que certains éléments des rapports d'évaluation de la qualité, afin d'aider les responsables du PNUD, les responsables de l'évaluation et les partenaires à mettre en œuvre un **contrôle de qualité efficace des processus et des produits de l'évaluation.** Le présent chapitre est destiné à améliorer la connaissance des méthodes et outils disponibles, en vue d'assurer la conformité des principaux produits d'évaluation — notamment les termes de référence, le modèle d'évaluation et les rapports d'évaluation — avec les critères de qualité tels que définis par les normes, dispositions réglementaires et politiques applicables. Ce chapitre vise également à permettre aux évaluateurs externes d'appréhender les normes de qualité devant être appliquées dans le cadre des évaluations réalisées au sein du PNUD.

Si les évaluateurs externes sont chargés d'affiner la méthodologie et de mettre en œuvre l'évaluation, le modèle général et le cadre méthodologique sont largement déterminés par les dispositions établies dans les termes de référence dédiés à l'évaluation. Par conséquent, les personnes ayant la responsabilité d'élaborer ces termes de référence pourront se reporter au présent chapitre pour obtenir des informations concernant les principales composantes du modèle, le rôle des parties prenantes, la définition du contexte et la finalité de l'évaluation, ainsi que le champ de l'évaluation, avant de rédiger et de finaliser ces mêmes termes de référence. Certains aspects relatifs à l'assurance qualité de la méthodologie d'évaluation (section 7.5) sont également traités dans ce chapitre.

7.1 VUE D'ENSEMBLE

Concevoir un modèle d'évaluation de la qualité exige une compréhension approfondie de l'objet soumis à l'évaluation (**un programme donné et son contexte**), nécessite de procéder à un certain nombre d'arbitrages portant sur les principaux éléments suivants, et implique de déterminer dans quelle mesure chacun de ces éléments contribuera à valider les résultats utiles de l'évaluation :

- La **finalité** de l'évaluation ;
- Le **champ** de l'évaluation, c'est-à-dire les **questions essentielles** auxquelles l'évaluation tente de répondre ;

- Les **sources et méthodes** utilisées afin d'obtenir une information crédible et défendable ;
- Les procédures qui seront mises en œuvre afin d'**analyser et d'interpréter les données et les résultats des rapports** ;
- Les **normes** devant impérativement être respectées afin de considérer un programme comme réussi ;
- Les **indicateurs** qui seront utilisés afin de mesurer les performances du programme et d'établir ses résultats (produits et effets) ;

RÔLE DES PARTIES PRENANTES

Les parties prenantes jouent un rôle important dans l'élaboration et la mise en œuvre de l'évaluation de la qualité. Elles incluent les individus et les groupes ayant un intérêt direct dans le programme ou dans les résultats de l'évaluation. Leur implication à tous les stades de l'évaluation — y compris dans la définition du champ de l'évaluation, dans l'élaboration des questions à traiter, dans l'identification des sources d'indicateurs crédibles, dans l'analyse des conclusions et dans la contribution à leur interprétation — renforce la crédibilité, ainsi que l'utilité potentielle et la durabilité des résultats de l'évaluation. En règle générale, les parties prenantes se subdivisent en trois principales catégories, pouvant le cas échant se recouper :

- **Les parties prenantes impliquées dans l'exécution du programme :** par exemple, les donateurs, les collaborateurs, les partenaires stratégiques, les administrateurs, les responsables et les membres du personnel ;
- **Les parties prenantes destinataires du programme ou affectées par celui-ci :** par exemple, les bénéficiaires prévus, les organisations et agences concernées, les responsables des pouvoirs publics, les groupes de pression, les sceptiques, les opposants et le personnel des organisations exécutantes ou concurrentes ;
- **Les premiers utilisateurs de l'évaluation :** par exemple, les personnes spécifiquement en position d'effectuer ou de décider une action relative au programme, notamment les donateurs, les homologues des programmes du PNUD (Comité de programme ou de résultats) et les partenaires associés à l'évaluation conjointe.

Le niveau d'implication des parties prenantes variera selon les évaluations. Au moment d'élaborer une évaluation, il est important que l'unité de programme mandatée puisse déterminer les parties prenantes en amont et qu'elle fasse appel à leur savoir pendant l'étape d'élaboration de l'évaluation, en commençant par leur implication significative dans le préparation des termes de référence. Ce point est particulièrement crucial pour les évaluations conjointes pour lesquelles les partenaires impliqués dans l'évaluation doivent participer à toutes les étapes de l'élaboration du modèle de l'évaluation.

7.2 DÉTERMINER LE CONTEXTE

Les évaluations du PNUD soutiennent la stratégie de développement humain de celui-ci pour « aider les populations à construire une vie meilleure » en créant des connaissances sur ce qui fonctionne, les raisons de ce fonctionnement et ses circonstances. Par conséquent, les évaluations de qualité ne se bornent pas simplement à la réalisation des produits et des effets, elles évaluent également la manière avec laquelle les initiatives s'adaptent aux contextes dans lesquels elles opèrent et expliquent comment et pourquoi elles contribuent à la réalisation des produits et des effets.

Les évaluations doivent être conçues et élaborées avec une connaissance approfondie de l'initiative à mettre en place et du contexte dans lequel celle-ci va opérer. L'unité mandatée du PNUD et les principales parties prenantes engagées dans l'élaboration du projet de mandat (voir annexe 3) sont tenues d'énoncer clairement les informations nécessaires aux évaluateurs pour qu'ils aient une bonne compréhension de l'initiative, du contexte de l'évaluation, des points d'intérêt et du but de l'évaluation et, les questions-clés qui seront traitées dans celle-ci.

Tableau 27. Aspects-clés de l'initiative

Aspect-clé	Questions à se poser
Demande	Quels sont les besoins ou les demandes de l'initiative? Quel problème ou possibilité de développement l'initiative compte-t-elle traiter ?
Bénéficiaires	Qui sont les bénéficiaires ou cibles de l'initiative? Qui sont les individus, groupes ou organisations, ciblés ou non, qui bénéficient directement ou indirectement de l'initiative de développement ?
Portée	Quelle est la portée de l'initiative en termes de limites géographiques et nombre de bénéficiaires espérés ?
Produits et effets	Quels changements (effets) ou produits et services tangibles (produits) sont anticipés en tant que résultat de l'initiative ? Que doit accomplir le projet, programme ou stratégie pour être considéré une réussite ? Comment les effets espérés sont-ils liés aux priorités nationales, aux priorités du PCNUAD et aux objectifs du plan stratégique institutionnel ?
Activités	Quelles activités, stratégies, ou actions, planifiées ou pas, le programme exécute-t-il en vue d'un changement ?
Théorie du changement ou modèle de résultats	Quelles sont les principes et hypothèses sous-jacents ou quelle est la théorie définissant la relation ou chaîne de résultats qui mènent les stratégies de l'initiative aux effets souhaités ? Quelles sont les hypothèses, les facteurs ou les risques inhérents à l'élaboration pouvant influencer la réussite ou l'échec de l'initiative ?
Ressources	Quelles ressources en matière de temps, d'expérience, de technologie, d'informations et de finances sont allouées à cette tâche ?
Stratégie de partenariat et des parties prenantes	Quels sont les principaux acteurs et partenaires impliqués dans le programme ou projet et ayant un intérêt direct ? Quels sont leurs rôles, participation et contributions –en ressources financières, contributions en nature, leadership et plaidoyer - y compris pour les organisations des nations Unies et autres ? Comment la stratégie de partenariat a-t-elle été divisée ? Comment opère-t-elle ?
Etape de mise en œuvre	Quelle est le degré de maturité du programme ou du projet, c'est-à-dire à quel étape ou année en est la mise en œuvre ? La mise en œuvre suit-elle le cours planifié de l'initiative ? Le programme est-il engagé de manière générale dans les activités de planification ou de mise en œuvre ?
Modifications du modèle original	Quels changements dans les plans et les stratégies de l'initiative, le cas échéant, ont-ils eu lieu ? Quelles sont les implications potentielles pour la réalisation des résultats escomptés ?
Evaluabilité	Le projet ou programme, tel que définit, peut-il être réellement évalué ? Les résultats escomptés (produits et effets) sont-ils adéquatement déterminés, appropriés et cités en termes mesurables et, les résultats sont-ils vérifiables ? Les systèmes de suivi et d'évaluation qui fourniront les données exploitables et fiables sont-ils mis en place ?
Questions transversales	A quel point les questions transversales et les valeurs prônées par les Nations Unies ont-elles été intégrées et traitées dans l'élaboration, la mise en œuvre et les résultats ?

COMPRENDRE L'INITIATIVE

Afin de pouvoir fournir des informations crédibles et utiles aux responsables de la prise de décision, les évaluations doivent être élaborée avec une claire compréhension de l'initiative, sa manière d'opérer, comment était-elle censée opérer, pourquoi opère-t-elle de cette manière et quels sont les résultats qu'elle produit. Il ne suffit pas de savoir ce qui a fonctionné et ce qui ne l'a pas (c'est-à-dire si les résultats ont été atteints ou pas). Pour une action en bonne connaissance de cause, les évaluations doivent fournir des informations crédibles sur la raison pour laquelle une initiative a produit de tels résultats et déterminer également les facteurs qui ont contribué à ces résultats (aussi bien les résultats positifs que négatifs). Comprendre parfaitement ce qui a été mis en œuvre, et pourquoi, établit les bases pour comprendre l'importance ou la signification des résultats d'un programme ou d'un projet.

Par conséquent, les évaluations doivent être élaborées sur la base d'une compréhension approfondie de l'initiative évaluée, y compris la chaîne de résultats escomptés (apports, produits et effets espérés), sa stratégie de mise en œuvre, son étendue et, les hypothèses-clés et risques sous-jacents au modèle de résultats ou à la théorie du changement. Les questions posées dans le tableau 27 devraient être comprises par les évaluateurs lorsqu'ils mènent une évaluation.

LE CONTEXTE DE L'ÉVALUATION

Le contexte de l'évaluation concerne deux groupes de facteurs corrélés qui influent sur la précision, la crédibilité et l'utilité des résultats de l'évaluation[48] :

- **Facteurs sociaux, politiques, démographiques et institutionnels, aussi bien internes qu'externes,** qui influencent le comment et le pourquoi de la réalisation des résultats (positifs ou négatifs) et leur durabilité.
- **Facteurs sociaux, politiques, démographiques et institutionnels contenus au sein de l'environnement et du calendrier fixé pour l'évaluation** et qui influencent la précision, l'impartialité et la crédibilité des résultats de celle-ci.

Encadré 38. Questions-indications pour définir le contexte[49]

- Quel est l'environnement opérationnel du projet ou programme ?
- De quelle manière des facteurs tels que l'histoire, la géographie, les politiques, les conditions sociales et économiques, les tendances séculaires et les efforts fournis par les organisations concernées ou en compétition, influencent-ils la mise en œuvre de la stratégie de l'initiative, de ses produits, de ses effets ?
- Comment le contexte au sein duquel l'évaluation est menée (comme par exemple les mœurs culturelles, la langue, le cadre institutionnel, les perceptions des populations, etc.) influence-t-il l'évaluation ?
- Comment le projet ou le programme collabore-t-il et se coordonne-t-il avec d'autres initiatives et celles d'autres organisations ?
- Comment le programme est-il financé ? Les fonds sont-ils suffisants ? Le projet ou le programme possède-t-il un financement assuré pour le futur ?
- Dans quel environnement normatif et politique le projet ou le programme opère-t-il ? Comment les alternatives politiques actuelles et émergeantes peuvent-elles influencer les produits et les effets de l'initiative ?

48 Les résultats de l'évaluation sont le produit final de l'évaluation – la somme des informations que l'évaluation génère pour les utilisateurs, y compris les résultats, les conclusions, les recommandations et les enseignements tirés.

49 Les questions posées dans ce chapitre ont un simple but d'illustration et ne sont pas exhaustives.

Etudier les facteurs internes et externes au sein desquels une initiative pour le développement opère, aide à expliquer la raison pour laquelle celle-ci a été mise en œuvre d'une certaine manière et pourquoi certains produits ou effets ont été réalisés et d'autre ne l'ont pas été. L'examen du contexte de l'initiative peut également révéler des facteurs pouvant entraver la réalisation de produits et d'effets anticipés ou compliquer l'estimation de leur réalisation, ou de la contribution que les produits apportent aux effets. De plus, comprendre les cadres politique, culturel et institutionnel de l'évaluation peut fournir d'importantes indications sur la meilleure manière d'élaborer et de mener l'évaluation afin de garantir l'impartialité, la crédibilité et l'utilité de ses résultats.

7.3 L'OBJET DE L'ÉVALUATION

Toutes les évaluations commencent par un objet qui met en place les directives à suivre. En l'absence d'un énoncé clair et complet des objets, une évaluation court le risque d'être inutile et de manquer de crédibilité. Les évaluations peuvent répondre à nombre de différents besoins. L'énonciation des objets doivent répondre clairement aux points suivants :

- Pourquoi l'évaluation est-elle menée et pourquoi à ce moment bien précis?
- Qui exploitera les informations?
- De quelles informations a-t-on besoin?
- Comment cette information sera-t-elle utilisée?

L'objet et le calendrier d'une évaluation devrait être déterminé au moment de son élaboration (voir chapitre 3 pour de plus amples informations). L'énonciation de l'objet de l'évaluation peut par la suite être élaborée en même temps que le projet des termes de référence de l'évaluation afin de documenter l'élaboration de l'évaluation.

CONSEIL **Modèle d'énonciation de l'objet –** « Ce projet d'évaluation est mené à la demande du gouvernement national et du PNUD afin de fournir des informations sur le statut de la mise en œuvre du projet et permettre ainsi de garantir la redevabilité vis-à-vis des dépenses engagées à ce jour et la réalisation des produits, pour que les responsables puissent prendre les mesures correctives à mi-parcours nécessaires. »

7.4 AJUSTER L'ÉVALUATION

PORTÉE DE L'ÉVALUATION

La portée de l'évaluation réduit le centre d'intérêt de l'évaluation en établissant les limites de ce que l'évaluation couvrira et ce qu'elle ne couvrira pas dans le cadre de l'exécution de son objet. La portée détermine les aspects de l'initiative et leur contexte situés dans les limites de l'évaluation. Ainsi, la portée détermine par exemple :

- Les unités d'analyse à couvrir par l'évaluation tels que, un système de programmes associés, les politiques ou stratégies, un seul programme impliquant un ensemble de projets, un seul projet, ou une sous-composante ou processus au sein d'un projet
- Le temps ou étape(s) de la mise en œuvre
- Les fonds dépensés au moment de l'évaluation par rapport au montant total alloué

- La couverture géographique
- Les groupes ciblés ou bénéficiaires à inclure

La portée aide donc à bien déterminer la sélection de questions liées à l'évaluation dans le cadre des limites prédéfinies.

OBJECTIFS DE L'ÉVALUATION ET CRITÈRES

Les objectifs de l'évaluation sont des formulations sur ce que l'évaluation fera pour remplir ses objets. Ils sont basés sur de minutieuses considérations sur : le type de décisions que les utilisateurs de l'évaluation prendront ; les questions auxquelles ils auront à trouver une réponse pour prendre ces décisions et ; ce que l'évaluation aura à réaliser pour contribuer à ces décisions. Une évaluation donnée peu poursuivre un ou plusieurs objectifs. Le point le plus important reste néanmoins que les objectifs sont directement issus de l'objet de l'évaluation et servent à la centrer sur les décisions à prendre.

CONSEIL **Proposition de projet d'objectifs d'une évaluation –** « évaluer le statut des produits; évaluer la manière avec laquelle ces produits sont réalisés; évaluer l'efficacité avec laquelle les produits sont réalisés. »

Les critères de l'évaluation aident à mettre au point les objectifs en déterminant les normes d'évaluation de l'initiative. Les évaluations du PNUD appliquent généralement les critères suivants : la pertinence, l'efficacité, l'efficience, la durabilité et les impacts des efforts du développement.[50]

La pertinence fait référence au degré de compatibilité qu'une initiative de développement et ses produits et effets escomptés entretien avec les politiques nationales et locales et avec les priorités et les besoins des bénéficiaires visés. Elle tient compte par ailleurs du degré de réactivité de l'initiative face aux priorités de développement humain et de plan institutionnel du PNUD, en matière de redevabilité des populations et des questions d'égalité des sexes. Elle vérifie également le niveau de compatibilité entre la perception des besoins tels que planifiés par les chargés de la planification de l'initiative et la réalité des besoins du point de vue des bénéficiaires visés. Enfin, elle englobe le concept de capacité de réaction, c'est-à-dire le degré de réactivité du PNUD face aux changements et émergence des priorités et besoins de développement.

L'une des importantes sous-catégories de la pertinence est le critère de **justesse**, qui traite de l'acceptation culturelle ainsi que de la faisabilité des activités ou méthode d'une initiative de développement.

Alors que la pertinence étudie l'importance de l'initiative par rapport aux besoins et aux priorités des bénéficiaires visés, la justesse étudie quant à elle l'acceptation et la faisabilité de l'initiative telle qu'elle opère au sein d'un contexte local. Ainsi par exemple, une initiative peut être pertinente parce qu'elle traite un besoin que les bénéficiaires ciblés estiment important mais

50 '*Critères du CAD d'évaluation de l'aide au développement*', Comité d'Aide au Développement (CAD), OCDE / 'DAC Criteria for Evaluating Development Assistance', Development Assistance Committee. Disponible sur le site : http://www.oecd.org/document/22/0,2340,en_2649_34435_2086550_1_1_1_1,00.html.

elle peut être également inadéquate parce que la méthode d'exécution utilisée est incompatible avec la culture locale ou infaisable au vu des réalités géographiques ou contextuelles. Grâce à l'application du critère de pertinence, les évaluations devraient être à même de déterminer jusqu'à quel point la planification, l'élaboration et la mise en œuvre des initiatives tiennent compte du contexte local.

L'efficacité est la mesure du niveau de réalisation des résultats (produits ou effets) escomptés d'une initiative donnée ou la mesure des progrès réalisés pour l'atteinte des produits ou effets.

L'évaluation de l'efficacité dans les évaluations d'un projet implique une appréciation des causes et des effets, c'est-à-dire l'attribution des changements observés aux activités du projet et à ses produits. Par exemple, le niveau de changement du nombre de votants pouvant être **attribué** à un projet d'éducation électorale. L'appréciation de l'efficacité des effets des évaluations se penchera plus sur l'examen des contributions apportées par le PNUD et par d'autres partenaires à la réalisation des effets escomptés. L'évaluation d'un effet pourra par exemple étudier jusqu'à quel point les produits observés d'un projet d'éducation électorale, ainsi que d'autres produits bénéficiant du soutien du PNUD tels que la professionnalisation de l'administration électorale, ont contribué à la réalisation des effets formulés en rapport avec une participation inclusive et mesurés par des observateurs internationaux et des experts réputés.

L'évaluation de l'efficacité implique trois étapes fondamentales :

1. Mesurer le changement dans les produits et les effets observés
2. Attribuer ces changements ou les progrès réalisés à l'initiative (évaluation du projet) ou déterminer les contributions du PNUD à ces changements observés
3. Juger la valeur du changement (positive ou négative)

L'efficience mesure la manière avec laquelle les ressources ou les apports (tels que les fonds, la compétence et le temps) sont transformés de façon économe en résultats. Une initiative est considérée efficiente lorsqu'elle utilise les ressources correctement et de façon économe pour atteindre les produits souhaités. L'efficience est importante pour garantir le bon usage des ressources et mettre en relief d'autres usages efficaces de ces ressources.

Etant donné que la nature et les objets premiers des évaluations des projets et des effets diffèrent, l'application du critère à utiliser diffère également. Ainsi, lors de l'évaluation de l'efficience par exemple, une évaluation de projet peut être amenée à étudier le degré d'utilisation des ressources pour réaliser les produits escomptés et comment ces ressources peuvent-elles être utilisées plus efficacement pour réaliser les résultats prévus. D'autre part, une évaluation des effets, peut impliquer des estimations de l'investissement global du PNUD (tous les projets et l'appui-conseil) destiné à un effet de développement donné. L'application de ce critère, particulièrement dans les évaluations du PNUD, constitue un certain défi vu que la nature même des initiatives du PNUD (par exemple l'appui-conseil), ne se prêtent pas toujours aux indicateurs conventionnels d'efficience. Dans ces cas, il faut donc considérer de leur appliquer certaines analyses de taux d'exécution, de raisons pour lesquelles certaines initiatives sont mises en œuvre plus rapidement que d'autres et, de ratios de gestion au niveau du programme. Il est également important d'évaluer la manière avec laquelle la stratégie de partenariat a influencé l'efficience des initiatives du PNUD à travers des mesures de partage de frais et d'activités complémentaires.

La durabilité mesure le degré auquel les bénéfices des initiatives perdurent après que l'aide au développement en provenance de l'extérieur ait touché à sa fin. Evaluer la durabilité implique une appréciation du niveau auquel les pertinentes conditions sociales, économiques, politiques et

autres sont réunies et, toujours sur la base de cette appréciation, d'effectuer des projections sur la capacité nationale à maintenir, diriger et garantir les résultats du développement dans le futur.

Une évaluation de la durabilité peut par exemple étudier le degré auquel :

- Une stratégie de durabilité, y compris le renforcement de la capacité des parties prenantes nationales, a été développée ou mise en œuvre.
- Il existe des mécanismes financiers et économiques mis en place pour assurer la continuité du flux de bénéfices en cours une fois que l'aide touchera à sa fin.
- Les modalités organisationnelles appropriées (dans le secteur public ou privé) ont été appliquées.
- Les cadres politiques et de régulation ont été mis en place et soutiendront la continuité des bénéfices.
- La capacité institutionnelle requise (systèmes, structures, personnel, compétence, etc.) existe.

L'impact mesure les changements dans le développement humain et dans le bien-être des populations apportés, directement ou indirectement, prévus ou non, par les initiatives de développement. Plusieurs organisations de développement évaluent l'impact parce qu'il génère des informations utiles à la prise de décision et soutien la redevabilité dans la réalisation des résultats. Par moments, l'évaluation d'impact fait face à certains défis: ainsi, confirmer si les bénéfices apportés aux bénéficiaires peuvent être directement attribués au soutien du PNUD peut s'avérer difficile, et particulièrement lorsque le PNUD est l'un des nombreux donateurs. Evaluer l'impact des initiatives du PNUD doit néanmoins avoir lieu à chaque fois que des bienfaits directs sur les populations sont perceptibles.

De manière générale, l'application des critères les plus utilisés - pertinence, efficacité, efficience, durabilité et impact – en les associant, peut aider à s'assurer que l'évaluation couvrira les domaines les plus importants de l'initiative. Tous les critères ne sont cependant pas applicables, ou applicables de manière identique, à chaque évaluation. Différents critères peuvent être appliqués dans des cas uniques. A l'heure de choisir le critère à appliquer, il faut tenir compte du type d'évaluation et les contributions des informations relatives à l'objet du coût (utilisation de l'évaluation des ressources). Les évaluations des programmes humanitaires ou en cas de conflit peuvent faire appel à des critères de connexion, de cohérence, de couverture et de coordination.[51] L'encadré 39 passe brièvement en revue les questions pouvant aider à déterminer le critère d'évaluation et les questions sur l'évaluation qui y sont associées.

Encadré 39. Questions-indications pour la détermination des critères de l'évaluation

- Jusqu'à quel point le critère renseigne-t-il sur l'objet de l'évaluation ?
- De quelle quantité et de quel type d'informations les utilisateurs potentiels ont-ils besoin ?
- L'accent devrait-il être mis de la même manière sur chacun des critères ou est-ce que certaines informations seront plus utiles ?
- Ce critère représente-t-il une mesure utile ou adéquate à l'évaluation en particulier ?
- Quel critère génèrera le plus d'informations utiles au vu des ressources disponibles ?

51 Pour plus de détails, consulter: Beck T, 'Evaluating Humanitarian Action Using OECD/DAC Criteria', 2006.

QUESTIONS SUR L'ÉVALUATION

Apporter des réponses aux questions sur l'évaluation permet d'offrir aux utilisateurs de celle-ci les informations qu'ils recherchent pour leurs prises de décision, prises de mesures ou l'enrichissement de la base de connaissances. Les questions sur l'évaluation permettent également d'affiner la mise au point de l'évaluation en rendant plus explicites les aspects de l'initiative pris en compte au moment d'apprécier sa performance.

Les questions sur l'évaluation reflètent par ailleurs la chaîne d'hypothèses sous-jacente sur la manière avec laquelle l'initiative est supposée opérer au sein de son contexte conformément aux produits et effets escomptés. Le choix de ces questions doit découler d'une compréhension approfondie des opérations, des intentions et du contexte de l'initiative et selon le rôle que ces questions vont jouer dans l'exécution de l'objet, des objectifs et des critères pertinents de l'évaluation.

Théoriquement, un nombre indéfini de questions peut être posé pour chaque critère d'évaluation, en pratique cependant, les évaluations ont des limites en termes de temps, de budget et de ressources. Il est par conséquent important de déterminer de manière stratégique les informations les plus indispensables et d'établir des priorités pour les questions sur l'évaluation. Il est plus judicieux de répondre plus solidement à un petit nombre de questions plutôt que de répondre de manière superficielle à un grand nombre d'entre elles. Etablir un ensemble clair et concis des questions les plus pertinentes, garantit des évaluations bien ciblées, plus simples à mener, rentables et utiles.

Afin de s'assurer que les questions-clés choisies pour l'évaluation sont les plus pertinentes et les plus à même de générer les informations les plus significatives aux utilisateurs, les unités de programme du PNUD doivent solliciter des apports et négocier des accords parmi les partenaires et autres parties prenantes, y compris l'équipe d'évaluation. Les bureaux mandatés doivent garantir que la matrice d'évaluation contenue dans le rapport préliminaire de l'évaluation, établit des liens bien déterminés avec les critères d'évaluation, les questions sur l'évaluation et les besoins en informations des utilisateurs ciblés (voir annexe 3 pour plus de détails).

APPROCHE AXÉE SUR L'ÉGALITÉ DES SEXES, LA SENSIBILITÉ À L'EXCLUSION ET LES DROITS

En accord avec les efforts de développement du PNUD, les évaluations de ce dernier sont guidées par les principes d'égalité des sexes, d'approche axée sur les droits et de développement humain.[52] Par conséquent, les évaluations du PNUD mesurent le degré auquel les initiatives : ont abordé les questions relatives à l'inclusion selon les sexes et le niveau social, l'égalité entre les hommes et les femmes et la redevabilité ; ont contribué à renforcer l'application de ces principes aux différents efforts de développement dans un pays donné ; et ont assimilé l'engagement du PNUD aux approches axées sur les droits et à l'intégration des deux sexes dans l'élaboration de l'initiative.

L'intégration des deux sexes est le processus par lequel sont évaluées les implications des femmes et des hommes dans toute action planifiée, y compris législative, politique ou de programmes, dans tous les domaines et à tous les niveaux. Il s'agit d'une stratégie destinée à faire des questions et expériences liées à l'égalité des sexes une partie intégrante de l'élaboration,

52 PNUD, 'la politique d'évaluation du PNUD', Document du Conseil d'administration DP/2005/28, mai 2006. Disponible sur le site: http://www.undp.org/eo/documents/Evaluation-Policy.pdf.

de la mise en œuvre, du suivi et de l'évaluation des politiques et des programmes, et ce dans toutes les sphères économiques, politiques et sociales, et apporter ainsi les mêmes bénéfices aussi bien aux hommes qu'aux femmes et évitant les inégalités entre les deux sexes. Les évaluations du PNUD doivent déterminer le degré auquel les initiatives de ce dernier ont tenu compte d'intégrer la question de l'égalité des sexes dans l'élaboration, la mise en œuvre et résultat de l'initiative et si aussi bien les hommes que les femmes peuvent prétendre de manière égale, et au même niveau présumé, aux bénéfices apportés par l'initiative. De la même manière, les évaluations devraient également déterminer jusqu'à quel point le PNUD a défendu les principes d'égalité et de développement et a contribué à renforcer et répondre aux besoins des populations les plus désavantagées et vulnérables dans une société donnée.

L'approche axée sur les droits dans les efforts de développement, suppose le besoin de garantir que les stratégies de développement répondent aux revendications des ayants droit et que les garants de ces droits respectent leurs obligations. Cette approche met également en valeur besoin crucial d'examiner les causes sous-jacentes et structurelles immédiates pour lesquelles ces droits n'ont pas été respectés. Le concept d'engagement civique, en tant que mécanisme de revendication des droits, est un aspect important du cadre général de travail. Les évaluations doivent, lorsque cela s'avère nécessaire, d'estimer le point auquel l'initiative a permis aux ayants droits de revendiquer leurs droits et aux garants de ces droits de respecter leurs obligations.

Les évaluations doivent enfin aborder également d'autres questions transversales, telles que le degré auquel le PNUD a intégré et encouragé dans son initiative la coopération sud-sud, la gestion du savoir, le volontarisme et la réforme des Nations Unies et ce, selon l'orientation de l'évaluation.

7.5 MÉTHODOLOGIE DE L'ÉVALUATION

L'élaboration de l'évaluation doit contenir un plan de travail détaillé étape par étape, précisant les méthodes que l'évaluation va adopter pour la collecte de l'information nécessaire au traitement des critères d'évaluation et pour répondre aux questions sur l'évaluation, ainsi que les méthodes à suivre pour l'analyse des données, l'interprétation des conclusions et le rapport des résultats.

Les méthodes d'évaluation doivent donc être choisies selon leur rigueur à produire des données empiriques destinées à aborder les critères d'évaluation et répondre aux questions d'évaluation. Le rapport initial d'évaluation doit contenir une matrice d'évaluation qui expose, pour chaque critère d'évaluation, les questions et sous-questions auxquelles l'évaluation apportera une réponse et, pour chaque question, les données à collecter qui renseigneront celle-ci ainsi que les méthodes à suivre pour la collecte de ces informations (voir encadré 40). Le rapport initial d'évaluation doit en outre mettre explicitement en relief, la théorie ou hypothèses sous-jacentes sur la manière avec laquelle chaque donnée contribuera à comprendre les résultats de développement – imputabilité, contribution, processus, mise en œuvre et ainsi de suite – et la raison de la collecte d'information et des méthodologies choisies pour l'analyse et la rédaction de rapports.

MÉTHODES DE COLLECTE DE DONNÉES

Les données à collecter et la méthode à suivre à cet effet seront déterminées par : les éléments nécessaires pour répondre aux questions de l'évaluation ; l'analyse utilisée pour traduire ces données en conclusions significatives pour répondre aux questions de l'évaluation et ; les jugements sur quelles données il faudra collecter au vu des contraintes de temps et de ressources.

Encadré 40. Questions pour les évaluateurs

Le bureau commissaire doit garantir au minimum que les méthodes exposées dans le rapport initial des évaluateurs répondent à chacune des questions suivantes :

- Quel élément est nécessaire pour répondre aux questions de l'évaluation ?
- Quelle(s) méthode(s) de collecte de données seront utilisées pour répondre aux critères et aux questions de l'évaluation ? Pourquoi ces méthodes ont-elles été choisies ? Les ressources allouées sont-elles suffisantes ?
- Qui sera chargé de la collecte des données?
- Quel est le cadre choisi pour la modélisation ? Quel principe suit-il ?
- Qui sera chargé d'impliquer les participants et autres parties prenantes?
- Quels systèmes de gestion de données seront utilisés ? C'est-à-dire, quelle est la logistique, y compris les procédures, le calendrier et l'infrastructure physique, mises en place pour réunir et traiter l'information?
- Comment les informations collectées seront-elles analysées et les conclusions interprétées et rapportées?
- Quelles sont les questions de méthodologie à revoir pour garantir la qualité?

Les évaluations du PNUD puisent énormément dans les données (indicateurs de performance) générées à travers le suivi pendant le cycle de mise en œuvre du programme ou projet. Les indicateurs de performance sont des moyens simples et fiables de documenter les changements des conditions de développement (effets), de la production, ou de la réalisation de produits et services (résultats), liés à une initiative de développement (voir chapitre 2).

Les indicateurs de performance sont utiles mais ont des limites : ils indiquent, ils n'expliquent pas. Les indicateurs ne répondront pas à l'ensemble des questions auxquelles l'évaluation cherche à répondre. Ainsi, ils pourront mesurer le progrès, par exemple, mais ne pourront donner une explication sur la raison de ce progrès ou déterminer les facteurs ayant contribué à ce progrès. Les évaluations du PNUD font en général appel à un mélange d'autres sources de données, collectées selon différentes méthodes, pour donner un sens aux informations fournies par les indicateurs de performance à propos d'une initiative.

Les données de base consistent en un ensemble d'informations observées par les évaluateurs ou collectées directement auprès des parties prenantes, et concernent leur premier contact avec l'initiative. Ces données constituent en général un ensemble d'observations sur les valeurs, les croyances, les attitudes, les opinions, les comportements, les motivations et les connaissances des parties prenantes, réunies à travers des questionnaires, des enquêtes, des entretiens, des groupes d'experts, d'informateurs-clés, de panels d'experts, d'observations directes et d'études de cas. Ces méthodes permettent d'effectuer une étude plus approfondie et fournissent plus d'information pouvant apporter une meilleure compréhension des changements observés dans les effets et les produits (aussi bien voulus que non), et des facteurs ayant contribué à un renforcement du contexte opérationnel des résultats.

Les données secondaires sont les données de base qui ont été collectées puis qui ont été compilées et publiées par quelqu'un d'autre. Ces données peuvent adopter plusieurs formes mais consistent en général en une série d'éléments de documentation ayant un intérêt direct pour les objets de l'évaluation. Les sources de ces éléments comprennent : des données démographiques locales, régionales ou nationales ; des rapports publiés à niveau national ou international ; des indicateurs économiques, sociaux ou sanitaires ; des plans de programme ou projet ; de rapports de suivi ; d'examen, d'évaluations et autres rapports antérieurs, de plans

Tableau 28. Récapitulatif des méthodes de collecte de données communément appliquées pour les évaluations du PNUD[53]

Méthode	Description	Avantages	Défis
Systèmes de suivi et d'évaluation	Utilisent les indicateurs de performance pour mesurer le progrès, particulièrement les résultats actuels face aux résultats escomptés.	Peuvent être une méthode fiable, efficiente et objective pour l'évaluation du progrès des produits et des effets.	Dépend de systèmes de suivi et d'évaluation viables ayant mis en place des indicateurs et des cibles de référence et ont collecté des données fiables liées aux cibles dans le temps et aux indicateurs d'effets.
Rapports et documents existants	Documentation existante, y compris des informations quantitatives et descriptives sur l'initiative, ses produits et ses effets, telles que la documentation issue des activités de renforcement des capacités, les rapports des donateurs, et autres éléments.	Rentables	Les éléments de documentation peuvent être difficiles à coder et à analyser en tant que réponse aux questions. Difficulté de vérifier la fiabilité et la validité des données.
Questionnaires	Fournissent une approche standardisée pour l'obtention des informations sur un large éventail de sujets de la part d'un grand nombre de diverses parties prenantes (en employant généralement des techniques d'échantillonnage) pour l'obtention d'informations sur leurs attitudes, croyances, opinions, perceptions, niveau de satisfaction, etc. concernant les opérations, les apports, les produits, et les facteurs contextuels de l'initiative du PNUD.	Efficaces pour réunir rapidement et relativement à moindre coût, des données descriptives sur un large éventail de sujets Faciles à analyser Respecte l'anonymat des participants.	Les auto-rapports peuvent affecter leur objectivité. Les données peuvent fournir une vue générale mais manquer d'approfondissement. Susceptibles de ne pas fournir des informations adéquates sur le contexte. Sujets à des erreurs d'échantillonnage.
Entretiens	Requièrent des réponses de personne à personne à des questions prédéterminées pour l'obtention d'informations approfondies sur les impressions ou expériences d'une personne, ou d'apprendre plus sur leurs réponses à des questionnaires ou enquêtes menés.	Une couverture, un éventail et approfondissement plus importants des informations sur un sujet.	Peuvent nécessiter beaucoup de temps. Peuvent être difficiles à analyser. Peuvent être coûteux. Peuvent donner lieu à un parti pris dans les réponses.
Observations sur le terrain	Impliquent l'utilisation d'un formulaire d'observation détaillé afin d'enregistrer les plus pertinentes informations sur le terrain sur le fonctionnement d'un programme (activités en cours, processus, discussions, interactions sociales et résultats tels qu'observés au cours d'une initiative).	Peuvent suivre les opérations d'un programme en temps réel. Peuvent s'adapter aux évènements en temps réel.	Peuvent être difficiles à catégoriser ou compliquer l'interprétation des comportements observés. Peuvent être coûteux. Peuvent donner lieu à un parti pris dans le choix du site.
Entretiens de groupes	Entretiens avec de petits groupes (6 à 8 personnes) pour étudier en profondeur les opinions des parties prenantes, les points de vue semblables ou divergents, ou les jugements portés sur une initiative de développement ou politique ainsi que les informations sur leurs comportements, compréhension et perception face à une initiative ou la collecte d'informations sur les changements tangibles et intangibles résultants d'une initiative.	Moyen rapide et fiable d'obtenir des impressions communes sur différentes parties prenantes. Moyen efficace d'obtenir un ensemble important d'informations approfondies dans une courte durée.	L'analyse des réponses peut être difficile. Requièrent un animateur expérimenté. Peuvent être difficiles à mettre en place.

53 Les méthodes décrites ne sont pas exhaustives et illustrent les types de méthodes applicables dans le contexte d'évaluation du PNUD.

Tableau 28. Récapitulatif des méthodes de collecte de données communément appliquées pour les évaluations du PNUD

Méthode	Description	Avantages	Défis
Informateurs-clés	Des entretiens qualitatifs approfondis, souvent en tête à tête, avec un large éventail de parties prenantes qui ont une connaissance de première main des opérations de l'initiative et du contexte. Ces communautés d'experts peuvent fournir une connaissance et une compréhension particulières des problèmes et faire des recommandations pour des solutions.	Peuvent fournir un aperçu de la nature des problèmes et offrir des recommandations pour des solutions. Peuvent fournir différentes perspectives sur une même question ou sur plusieurs.	Sujets à tendre vers un échantillonnage. Doivent avoir les moyens de vérifier ou corroborer les informations.
Panels d'experts	Un examen par les pairs ou par un groupe de référence, composé d'experts externes permettant de fournir un apport en termes techniques ou autres sujets couverts par l'initiative.	Augmentent la crédibilité. Peuvent servir en tant que source (experts) supplémentaire d'informations approfondies. Peuvent vérifier ou justifier les informations et les résultats dans le domaine du sujet.	Coût de la consultation et frais inhérents, le cas échéant. Doivent garantir l'impartialité et l'absence de conflits d'intérêts.
Etudes de cas	Impliquent une étude détaillée, grâce au croisement de comparaisons de cas, pour l'obtention d'informations approfondies avec, pour objectif, de comprendre pleinement les dynamiques opérationnelles, les activités, les produits, les effets et les interactions d'un projet ou programme de développement.	Utiles pour une étude complète des facteurs contribuant aux produits et aux effets.	Requièrent un temps considérable et des ressources souvent indisponibles pour les évaluations mandatées. Peuvent être difficiles à analyser.

stratégiques de pays et ; de rapports de recherches pouvant avoir un intérêt pour l'évaluation. Les éléments de documentation sont particulièrement utiles lorsque le projet ou programme manque d'indicateurs et de cibles de référence pour mesurer le progrès des produits et des effets. Malgré que ce ne soit pas la méthode la plus conseillée, les données secondaires peuvent être utilisées pour recréer des données et des cibles de références. Elles viennent enfin en complément et en supplément aux données de base mais ne remplacent pas la collecte de données de sources primaires.

Au vu de la nature et du contexte des évaluations du PNUD au niveau décentralisé, et en tenant compte des limitations de temps et de ressources, les évaluateurs sont souvent amenés à utiliser un mélange de méthodes, y compris des indicateurs de performance, des éléments de documentation pertinents supplémentaires issus de sources secondaires et, de données de qualité collectées par différents moyens.

Le tableau 28 présente de brèves descriptions des méthodes de collecte de données qui sont le plus communément appliquées pour les évaluations du PNUD aussi bien pour les évaluations de projet que pour les évaluations des effets.

Les bureaux mandatés doivent garantir que les méthodes et les instruments (questions, enquêtes, protocoles, listes de vérifications) utilisées pour la collecte ou l'enregistrement des

données sont : cohérentes avec les normes de qualité en termes de validité et de fiabilité,[54] sensibles à la culture et en adéquation avec les populations concernées et, pertinentes et appropriées aux différents types d'informations recherchées et aux questions sur l'évaluation auxquelles elles doivent apporter une réponse. Dans le cas de situations en milieux de conflits, les facteurs tels que les questions de sécurité, le manque d'infrastructures, l'accès limité aux populations détenant les informations ainsi que les sensibilités et considérations éthiques en termes de travail avec des populations vulnérables doivent alors être prise en ligne de compte pour déterminer les méthodes de collectes de données les plus adéquates.

QUESTIONS SUR LA QUALITÉ DES DONNÉES

Les bureaux mandatés du PNUD doivent s'assurer que l'évaluation collecte des données liées aux objets fixés et qu'elle emploie les méthodologies et procédures de collecte d'information rigoureuses et défendables et qu'elles donnent lieu à des éléments empiriquement vérifiés, qui soient valides, fiables et crédibles.

La fiabilité et la validité sont également des aspects importants de la qualité d'une évaluation. **La fiabilité** fait référence à la cohérence des mesures effectuées – s'assurer par exemple qu'un instrument de collecte de données particulier, tel qu'un questionnaire, donnera lieu aux mêmes réponses s'il est mené dans des conditions similaires. **La validité fait référence à la précision des mesures effectuées – s'assurer par exemple** qu'un instrument de collecte de données particulier mesure bien ce qu'il était censé mesurer. C'est également le degré auquel les déductions et conclusions tirées des données sont fondées et justifiables. **La crédibilité** concerne le degré perception des parties prenantes, et particulièrement les utilisateurs des résultats de l'évaluation, de la validité, fiabilité et impartialité des éléments d'évaluation et des résultats. Il existe trois grandes stratégies pour améliorer la fiabilité et la validité qu'une bonne évaluation devrait aborder :

- Améliorer la qualité de l'échantillonnage
- Améliorer la qualité du rassemblement des données
- Utiliser un mélange de collecte de données et de stratégies constructives (comme par exemple la 'triangulation' ou croisement des données ou l'utilisation de sources multiples de données) afin de vérifier ou recouper les données en utilisant plusieurs éléments plutôt qu'un seul.

Améliorer la qualité de l'échantillonnage

Les évaluations du PNUD rassemblent souvent les éléments d'informations d'un échantillon de populations ou de lieux. Si cet échantillon n'est pas représentatif de la portion d'une population, il y a un risque que les conclusions tirées sur la population en question soient fausses. Ainsi, si un entretien de groupe par exemple, ne réunit que les personnes vivant en ville et pouvant donc facilement accéder au lieu ou se tient l'entretien, les intérêts et expériences de ceux vivant en dehors de la ville pourraient ne pas être tenu en compte. L'échantillon doit donc être sélectionné sur la base d'un raisonnement ou d'un but directement lié aux objets de l'évaluation et est censé garantir une précision dans l'interprétation des conclusions et de l'utilité des résultats de l'évaluation. Les bureaux mandatés doivent donc s'assurer que l'élaboration de l'évaluation retrace clairement les caractéristiques de l'échantillon, la manière avec laquelle il sera

54 Voir la discussion sur la validité et la fiabilité dans la section 'Questions sur la qualité des données' du présent chapitre.

sélectionné et les raisons de cette sélection et enfin les limites de l'échantillon pour l'interprétation des résultats de l'évaluation. Si un échantillon n'est pas utilisé, les raisons et les répercussions sur l'évaluation doivent être discutées.

Garantir la consistance du rassemblement des données

Que ce soit à l'heure d'utiliser des questionnaires, de fixer des entretiens, de suivre des protocoles d'observation ou autres outils de rassemblement de données, les équipes d'évaluation doivent tester tout d'abord les outils de collecte de données et s'assurer qu'ils rassemblent des éléments à la fois précis et cohérents. Quelques démarches à suivre seraient :

- La formation de collecteurs de données grâce aux protocoles d'observation, afin de s'assurer qu'ils enregistrent tous leurs observations de la même manière
- Vérifier la signification de mots-clés utilisés dans les questionnaires et les calendriers des entretiens, particulièrement s'ils ont été traduits, afin de s'assurer que les personnes sondées comprennent bien ce qui leur est demandé.
- Prendre en ligne de compte comment les caractéristiques des personnes chargées de mener un entretien (particulièrement l'âge, le sexe et s'ils sont connu des informateurs) peuvent améliorer ou réduire la précision des informations fournies.

'Trianguler' les données pour vérifier leur précision: utiliser plusieurs sources de données

Un bon élément d'évaluation doit être à la fois cohérent et précis. Concevoir des stratégies destinées à vérifier les données améliorera la fiabilité et garantira des résultats valides.

- Utiliser un mélange de méthodes pour collecter l'information plutôt que de s'en tenir à une seule source ou à un seul élément. Recouper par exemple l'élément d'une source (telle que l'entretien de groupe) avec un autre élément sur les expériences des personnes vivant en milieu rural (pouvant être un élément documentaire en provenance de rapports ou d'entretiens avec des informateurs-clés crédibles et bien informés sur la situation).
- Faire appel à des experts pour examiner et valider les faits.

Le défi pour les évaluations du PNUD est d'utiliser des méthodes d'élaboration d'évaluations rigoureuses qui donneront lieu à des informations utiles et basées sur des faits crédibles et défendables face au défi que représente la précision et la validité des déductions tirées de ces faits.

CONSIDÉRATIONS ÉTHIQUES

Les évaluations doivent être élaborées et menées dans le respect et la protection des droits et du bien-être des populations et des communautés auxquelles elles appartiennent, conformément à la Déclaration Universelle des Droits de l'Homme des Nations Unies[55] et autres conventions sur les droits de l'Homme. Les évaluateurs doivent donc respecter la dignité et la diversité des participants aux évaluations au moment de sa planification, de son déroulement et de la rédaction des rapports, en partie grâce à l'utilisation des instruments d'évaluation appropriés au milieu culturel. Plus encore, les éventuels participants à l'évaluation doivent être considérés comme autonomes, que le temps et les informations nécessaires leur soient donnés pour décider de leur participation ou pas, et qu'ils puissent prendre une décision libre et sans aucune pression.

55 Nations Unies, 'Déclaration Universelle des Droits de l'Homme'. Disponible sur le site: http://www.un.org/en/documents/udhr/.

Les responsables de l'évaluation et les évaluateurs doivent être conscients des implications liées à la conduite d'une évaluation en zones de conflit. Les évaluateurs doivent particulièrement savoir que leur manière d'agir, y compris dans les messages explicites et implicites qu'ils transmettent, peut avoir des répercussions sur la situation et faire courir de plus grands risques aux personnes avec lesquelles ils interagissent.[56] Lorsque les évaluateurs doivent s'entretenir avec des groupes vulnérables, ils doivent s'assurer que les personnes interrogées sont conscientes des implications potentielles de leur participation à l'évaluation et qu'ils ont reçu suffisamment d'informations leur permettant de décider de leur participation à cet exercice. Tous les évaluateurs mandatés par les unités de programme du PNUD doivent s'accorder à signer le Code de conduite des évaluateurs du système des Nations Unies.[57] Pour de plus amples informations sur les questions déontologiques dans le cadre des évaluations, veuillez consulter: *'UNEG Ethical Guidelines for Evaluation'*.[58]

Encadré 41. Droits de l'Homme et égalité des sexes dans l'élaboration de l'évaluation

Les évaluations au sein du PNUD sont guidées par les principes du respect des Droits de l'Homme et de l'égalité des sexes. Ceci a des répercussions sur l'élaboration et la conduite des évaluations et requiert une compréhension mutuelle de ces principes et une attention particulière de part des évaluateurs, des responsables des évaluations et des parties prenantes. Ainsi, lors de la collecte de données par exemple, les évaluateurs doivent s'assurer que les femmes et les groupes les moins avantagés sont représentés de manière adéquate. Afin de faire apparaître les groupes exclus ou les moins avantagés, les données doivent être ventilées par sexe, âge, handicap, ethnicité, caste, richesse et toute autre différence pertinente.

De plus, les données doivent être analysées, à chaque fois que cela est possible, sous différents points de vue, y compris le sexe, le groupe socio-économique, l'ethnicité et le handicap. Les groupes marginalisés sont souvent sujets à différentes formes de discriminations, et il est donc important de comprendre comment ces discriminations se croisent et privent ces groupes de leurs droits.

ANALYSE ET SYNTHÈSE DES DONNÉES

La collecte de données implique la gestion de questionnaires, la conduite d'entretiens, l'observation d'opérations de programmes et l'examen ou l'intégration de données de sources déjà existantes. L'analyse des données est un processus systématique pour lequel il s'agit d'organiser et de classer les informations collectées sous forme de tableaux, de les résumer et de comparer les résultats avec ceux d'autres informations adéquates afin d'en extraire des déductions pouvant répondre aux questions de l'évaluation et à ses objets. C'est également un processus qui permet de déchiffrer les faits issus d'un ensemble de témoignages grâce à une codification et un rassemblement systématique des données collectées, en veillant à s'assurer de leur précision, en effectuant toute analyse statistique nécessaire et en transposant ces données sur des formats ou unités d'analyse liés à chacune des questions sur l'évaluation.

56 Guide de l'OCDE pour l'évaluation des activités de prévention des conflits et de construction de la paix – Document de travail pour la période d'application, 2008.

57 UNEG, 'Code of Conduct', Juin 2008. Disponible en anglais sur le site: http://www.uneval.org/papersandpubs/documentdetail.jsp?doc_id=100.

58 UNEG, 'Ethical Guidelines for Evaluation', Juin 2008. Disponible en anglais sur le site: http://www.uneval.org/search/index.jsp?q=ethical+Guidelines.

L'analyse des données cherche à identifier des modèles dans les faits, soit en isolant d'importantes déductions (analyse), soit en combinant des sources d'informations pour atteindre une plus grande compréhension (synthèse). Le mélange des méthodes d'évaluation requiert donc une analyse séparée de chaque élément et une synthèse de toutes les sources afin d'examiner les schémas, la convergence ou la complexité.

Plan d'analyse

L'analyse et la synthèse des données doivent provenir d'un plan d'analyse devant être établi dans l'élaboration de l'évaluation et son plan de travail, détaillé dans le rapport initial de l'évaluation. Le plan d'analyse est un outil d'évaluation fondamental permettant de déterminer la manière avec laquelle l'information collectée sera

Organisée, classée, mise en corrélation, comparée et présentée selon les questions de l'évaluation, y compris ce qui sera fait pour intégrer les multiples sources, et particulièrement celles qui fournissent des informations sous forme narrative, et toute méthode statistique à laquelle on fera appel pour l'intégration ou la présentation des données (calculs, sommes, pourcentages, analyses des coûts, et ainsi de suite). Les éventuelles limites de l'analyse des données et les défis auxquels elle peut faire face doivent être décrits dans le plan d'analyse. Ce dernier doit être rédigé en même temps que les méthodes et instruments de collecte de données plutôt qu'après.

Interprétation des résultats de l'analyse

L'interprétation des déductions est le processus par le biais duquel un sens est donné aux conclusions de l'évaluation tirées de l'analyse. Il puise dans l'ensemble des synthèses d'informations provenant de faits, de formulations, d'opinions et de documents pour transformer les déductions faites des données en jugements sur les résultats de développement (conclusions). Des recommandations pour les actions à venir seront alors faites sur la base de ces conclusions. L'interprétation est l'opération qui permet de comprendre la les déductions et de saisir la signification des éléments rassemblés dans une évaluation et de ses applications pour l'efficacité du développement.

Tirer des conclusions

Une conclusion est un jugement raisonné se basant sur une synthèse de déductions empiriques ou formulations factuelles selon des circonstances bien déterminées. Les conclusions ne sont pas des déductions ; ce sont des interprétations donnant un sens aux déductions. Les conclusions sont considérées valables et crédibles lorsqu'elles sont directement liées au fait et peuvent être justifiées selon des méthodes d'analyse et de synthèse résumant les déductions. Les conclusions doivent :

- Réfléchir à d'autres moyens de comparer les résultats (comme par exemple la comparaison avec les objectifs du programme, un groupe de comparaison, les normes nationales, les performances passées ou les besoins)
- Générer d'autres explications pour les déductions et indiquer la raison pour laquelle il ne faut plus tenir compte de ces explications.
- Créer la base pour la recommandation d'actions ou de décisions en cohérence avec les conclusions
- Se limiter aux situations, périodes de temps, personnes, contextes et objets pour lesquels les déductions sont applicables.[59]

59 Basé sur Joint Committee on Standards for Educational Evaluation, 'Programme Evaluation Standards: How to Assess Evaluations of Educational Programmes', 2nd ed, Sage Publications, Thousand Oaks, CA.

Faire des recommandations

Les recommandations sont des propositions basées sur les faits en vu de prise de mesures de la part des utilisateurs de l'évaluation. Elles doivent être fondées sur les conclusions. Faire des recommandations est néanmoins un élément différent de l'évaluation qui requiert plus d'informations que celles qui sont nécessaires aux conclusions. Elaborer des recommandations implique la prise en compte d'alternatives efficaces, de politiques, de priorités de financement entre autres, dans un contexte plus large. Cet exercice requiert également une connaissance approfondie de contexte, particulièrement le contexte organisationnel au sein duquel les décisions sur les politiques et les programmes seront prises ainsi que le contexte politique, social et économique dans lequel l'initiative va opérer.

Les recommandations devraient être formulées de manière à faciliter le développement d'une réponse de gestion (voir chapitre 6 et l'annexe 6 sur le système de réponse de gestion). Elles doivent également être réalistes et refléter une bonne compréhension de l'organisation mandatée et les principales contraintes à suivre. Chaque recommandation devrait clairement identifier ses groupes ciblés et stipuler l'action recommandée et sa raison.

Leçons tirées

Les leçons tirées d'une évaluation comprennent les nouvelles connaissances acquises de la circonstance particulière (initiative, effets du contexte et même des méthodes d'évaluation) qui sont applicables et utiles à d'autres contextes similaires. Les leçons tirées mettent souvent en valeur les forces et les faiblesses de la préparation, de l'élaboration et de la mise en œuvre qui ont une influence sur la performance, l'effet et l'impact. Quelles leçons ont été tirées et pourquoi ces leçons sont significatives et peuvent être généralisées à des programmes ou des projets similaires ou à venir, sont des questions pouvant être soulevées par l'évaluation.