

FAIRE LE LIEN ENTRE LE SUIVI-ÉVALUATION ET L'ÉVALUATION D'IMPACT

de Burt Perrin



Ceci est la seconde d'une série de quatre notes sur l'évaluation d'impact, élaborées par InterAction avec l'appui financier de la Fondation Rockefeller. Les autres notes de cette série sont : *Introduction à l'évaluation d'impact* ; *Introduction aux méthodes mixtes dans l'évaluation d'impact* et *Utilisation des résultats d'évaluations d'impact*. On peut trouver la série complète sur le site Internet d'InterAction : www.interaction.org/impact-evaluation-notes.

Photo: Manuel Meszarovits

Introduction	1
1. Comment le suivi et d'autres formes d'évaluation peuvent-ils appuyer l'évaluation d'impact ?	2
1.1. Principales caractéristiques du suivi, de l'évaluation et de l'évaluation d'impact	2
1.2. Comment le S&E peut contribuer à l'évaluation d'impact	7
2. Comment intégrer l'évaluation dans la réflexion et les pratiques de S&E	8
2.1. Formuler la théorie du changement	8
2.2. Identifier les priorités pour entreprendre une évaluation d'impact	11
2.3. Identifier les besoins d'informations/ de données	11
2.4. Commencer avec ce que l'on a	14
2.5. Concevoir et mettre en œuvre l'évaluation d'impact, analyser et interpréter les résultats	15
2.6. Utiliser les résultats	17
2.7. Revoir, réfléchir et mettre à jour	17
3. Impliquer toutes les parties de l'organisation	17
3.1. S&E : Une fonction centrale de gestion qui a besoin du leadership et de l'appui de la direction	18
3.2. Le personnel des programmes doit avoir un rôle actif	18
Résumé	19
Références et autres ressources utiles	20
Annexe 1 – Analyse de la contribution	21

Introduction

L'objectif des organisations non gouvernementales (ONG) et autres acteurs du développement est d'aider les populations à améliorer leur vie, leur communauté, leur société et leur environnement. Les ONG veulent généralement contribuer à un changement social, par exemple, l'amélioration de la justice ou de la sécurité économique et sociale, tout en reconnaissant que ce sont des effets à long terme qui dépassent la réalisation de résultats à court terme.

Comme l'indique la première Note d'orientation, c'est de cela qu'il s'agit dans l'impact et dans l'évaluation d'impact : identifier de manière systématique et empirique les effets *résultant* d'une intervention, qu'ils soient prévus ou imprévus, directs ou indirects. Il est généralement admis que les impacts se manifestent plus tard que les effets directs, et sont des résultats de ceux-ci. De plus en plus, on se rend compte que les bonnes intentions ne suffisent pas. Dans l'évaluation d'impact, on ne se contente pas de voir ce que les agences font, mais aussi *ce qui se passe du fait de ces activités* et dans quelle mesure ces interventions font réellement une différence dans la vie des personnes et dans leurs conditions de vie.

La plupart des ONG font diverses activités de suivi-évaluation (S&E). L'importance de ces activités peut être très variable. Très souvent, ce que l'on appelle S&E est composé principalement de suivi et des comptes rendus, même si certaines organisations s'impliquent de plus en plus dans toute une gamme d'activités d'évaluation, examinant divers aspects de leur performance. Par contre, beaucoup moins s'intéressent à l'évaluation d'impact.

Cette note d'orientation illustrera la relation entre le S&E de routine et l'évaluation d'impact. Elle va en particulier indiquer comment les activités de suivi et les activités d'évaluation peuvent participer à une évaluation d'impact valide et significative et même la rendre possible. Les évaluations d'impact sont généralement externes et faites, entièrement ou en partie, par un expert indépendant extérieur à l'agence.¹ Cependant, le S&E a un rôle crucial à jouer, comme :

- Identifier quand, et dans quelles circonstances il serait possible et approprié d'entreprendre une évaluation d'impact.
- Fournir des données essentielles pour faire une évaluation d'impact, comme différents types de données de base et des informations sur la nature de l'intervention.
- Fournir les informations nécessaires pour interpréter et appliquer les résultats de l'évaluation d'impact. Il s'agit d'informations sur

¹ Par exemple, cf. *Évaluation Policy* de l'USAID (janvier 2011). Les évaluations d'impact peuvent cependant être faites à l'interne ou avec une combinaison des deux, et cela se fait parfois. La Note d'orientation 1 en dit plus sur qui devrait faire une évaluation d'impact.

le contexte et de données telles que la qualité de la mise en œuvre, qui sont nécessaires pour comprendre *pourquoi* des changements particuliers se sont passés ou pas et ce que nous pouvons faire pour rendre nos efforts plus efficaces à l'avenir.

La Section 1 de cette note d'orientation présente quelques-unes des caractéristiques de l'évaluation d'impact et montre en quoi elle est différente du suivi et d'autres formes d'évaluation et plus précisément, comment ces approches différentes mais complémentaires peuvent être mises en lien pour aider à faire une évaluation d'impact significative.

La Section 2 donne des directives et des idées sur les diverses étapes et les approches qui peuvent être utilisées pour maximiser la contribution d'un S&E de routine à l'évaluation d'impact.

L'un des principaux publics de cette note d'orientation est le *personnel de S&E* des ONG, pour l'aider à identifier les étapes qu'il peut suivre pour mieux lier les activités de S&E à l'évaluation d'impact. Cette note peut aussi être intéressante pour le personnel d'autres types d'organisations, comme des fondations.

De plus, cette note sera aussi intéressante pour la *direction des ONG* ainsi que pour le *personnel des programmes*. Pourquoi ? D'abord, comme cela est noté dans la Section 3, faire une évaluation d'impact ne sert pas à grand-chose si toute l'organisation ne pense pas en termes d'impact et d'amélioration continue, en se demandant comment améliorer ce que l'on fait. L'organisation doit être prête à agir sur les implications de l'évaluation d'impact et d'autres formes d'évaluation.

Ensuite, le S&E n'est pas quelque chose qui doit être laissé aux seuls spécialistes de l'évaluation.

Le personnel des programmes peut et doit jouer un rôle très actif dans la collecte des données et l'enregistrement de ses observations et de son expérience, ce qui peut participer à garantir que les évaluations d'impact s'intéresseront aux questions qu'il faut et reflèteront correctement ce que font les programmes.

1. Comment le suivi et d'autres formes d'évaluation peuvent-ils appuyer l'évaluation d'impact ?

Le suivi et l'évaluation peuvent apporter des contributions essentielles à l'évaluation d'impact. En effet, une évaluation d'impact *significative*² n'est pas possible si les activités habituelles de S&E de l'organisation ne lui donnent pas un bon appui. Cette note a un contenu trop pointu pour pouvoir parler de la nature et de la diversité du suivi-évaluation, mais il est important de reconnaître certaines des différences significatives entre le « suivi » et l'« évaluation » qui contribuent de manière différente à l'évaluation d'impact. Il est donc utile d'étudier quelques caractéristiques de base, comme les différences et les possibilités de complémentarité avant de déterminer plus précisément comme le « S&E » peut contribuer à l'évaluation d'impact.

1.1. Principales caractéristiques du suivi, de l'évaluation et de l'évaluation d'impact

Le tableau qui suit résume certaines des caractéristiques les plus courantes et généralement acceptées du suivi, de l'évaluation et de l'évaluation d'impact. Ce tableau devrait être regardé avec prudence. Comme le suggèrent certains des mots (par ex. « habituel », « peut », « généralement)

² Une évaluation d'impact qui est « significative » représenterait de façon appropriée l'intervention en question et son contexte, elle est valide dans tous les domaines (par ex. elle permet à la fois une validité interne et une validité externe) et, par dessus tout, elle fournit des informations utiles qui pourront servir de bases à des directions futures.

Suivi, évaluation et évaluation d'impact : quelques caractéristiques de base

Suivi	Évaluation	Évaluation d'impact
<ul style="list-style-type: none"> • Périodique, utilise des données collectées systématiquement ou faciles à trouver, généralement à l'interne, s'intéresse habituellement aux activités et aux produits, même si des indicateurs d'effet/impact sont aussi parfois utilisés. • Part du principe que le programme, les activités, les objectifs et les indicateurs sont appropriés. • Suit généralement les progrès par rapport à un petit nombre de cibles ou d'indicateurs préétablis. • Généralement quantitatif. • Ne peut pas indiquer de causalité. • Peut difficilement permettre à lui seul d'évaluer l'impact. 	<ul style="list-style-type: none"> • Généralement épisodique, souvent externe. • Va plus loin que les produits pour évaluer les effets. • Peut mettre en doute la justification et la pertinence du programme, des objectifs et des activités. • Peut identifier à la fois les effets imprévus et les effets planifiés. • Peut répondre aux questions « comment » et « pourquoi ». • Peut conseiller des directions pour l'avenir. • Peut utiliser des données venant de sources différentes et obtenues avec des méthodes très diverses. 	<ul style="list-style-type: none"> • Une forme d'évaluation particulière. • Sporadique, peu fréquente. • Le plus souvent externe. • C'est généralement une étude séparée. • S'intéresse spécifiquement à l'attribution (causalité) d'une manière ou d'une autre, le plus souvent avec un scénario contrefactuel. • S'intéresse généralement aux changements à long terme, comme la qualité de vie des bénéficiaires prévus. • Doit prendre en compte ce qui a vraiment été fait (par ex. grâce au S&E) et aussi identifier les impacts.

ainsi que le texte en dessous, il y a souvent des variations dans la manière dont ces concepts sont vus et appliqués.

Le suivi implique généralement de surveiller les progrès par rapport à des plans ou objectifs identifiés auparavant, en utilisant de manière continue des données facilement saisies et mesurées. Le suivi utilise le plus souvent des données quantitatives, mais le suivi de données qualitatives est aussi possible et certaines agences le font régulièrement.

Le suivi est fait pour des raisons diverses et variées, qui n'ont généralement pas grand-chose à voir avec l'évaluation d'impact. Certaines des raisons les plus fréquentes du suivi sont :

- Une utilisation interne par les gestionnaires et le personnel des projets pour mieux comprendre et surveiller la manière dont les choses se passent, en particulier pour déterminer si un projet est sur la bonne voie ou non. Il s'agit de suivre les données sur les services qui sont réellement fournis, la qualité des services fournis et des questions de ce genre. Le suivi des données peut aussi souvent servir de système d'alerte précoce et, dans le cas de constatations négatives ou inattendues, peut suggérer qu'il est nécessaire de changer d'approche alors que le projet ou le programme est toujours en cours.
- Utilisation interne à l'organisation au niveau régional, national et/ou international, au siège, pour que l'agence puisse suivre l'état d'un projet ou d'une activité par rapport aux

plans et aux attentes ; pour la planification et la gestion ; pour répondre aux besoins de comptes-rendus au conseil d'administration de l'agence, aux financeurs et au public.

- Répondre à des exigences externes de respect des règles et de contrôle, par exemple les rapports et comptes-rendus demandés par les bailleurs de fonds.

Le suivi peut aussi prendre d'autres formes. Plutôt que de se concentrer particulièrement sur ce que fait l'organisation, par exemple, le suivi pourrait comprendre des notes données par les citoyens aux services publics,³ dans de nombreuses organisations, des services de plaidoyer faisant un suivi des pratiques corrompues, des approches environnementales, du travail des enfants dans les entreprises locales, etc.

Quelques définitions du suivi :

Processus continu de collecte systématique d'informations, selon des indicateurs choisis, pour fournir aux gestionnaires et aux parties prenantes d'une action de développement en cours, des éléments sur les progrès réalisés, les objectifs atteints et l'utilisation des fonds alloués. (OCDE/Comité d'assistance au développement [CAD] Glossaire.

La surveillance des produits et des effets du projet en tant qu'indicateurs de l'efficacité du projet ou la mesure dans laquelle le projet réalise ses objectifs déclarés. (USAID).

Il s'agit là de raisons légitimes et importantes pour faire un suivi, mais aucune d'entre elles ne vise à contribuer à l'évaluation d'impact. Par conséquent, les types de données collectées et la manière dont on les garde et dont on en rend compte les

³ Lecture supplémentaire : Samuel Paul. *Holding the State to Account : Citizen Monitoring in Action*. Bangalore : Public Affairs Centre, 2002.

rendent difficiles à utiliser pour une évaluation. Comme le suggère l'encadré à droite, on doit faire des plans à l'avance pour que les données du suivi soient utiles pour l'évaluation d'impact. Dans des circonstances idéales, les personnes faisant une évaluation d'impact peuvent aider à concevoir et à structurer le système de suivi d'une intervention.

En même temps, il est important de reconnaître les limites du suivi. Le suivi s'intéresse généralement au progrès par rapport à des objectifs et des indicateurs prédéfinis et suppose que ceux-ci sont appropriés. Mais le suivi ne suffit pas à lui seul pour tirer des conclusions d'attribution ni pour identifier les raisons pour lesquelles des changements se sont produits ou pas (par exemple, dans quelle mesure ces changements sont un résultat de l'intervention ou dus à d'autres causes). Il ne peut généralement pas identifier les effets imprévus, les besoins non satisfaits dans les services, etc. Pour cela, on a généralement besoin de l'évaluation.

Un scénario bien trop courant : Une évaluation d'impact prévue pour la fin d'un projet découvre que les informations de référence et autres données intéressantes n'ont pas été collectées ou n'ont pas été désagrégées de manière à pouvoir montrer les changements qui se sont produits ou même fournir des informations nécessaires sur les activités du programme, qui a été servi et d'autres informations essentielles. Il est donc impossible de déterminer quel a été l'impact ou de l'attribuer en aucune manière au programme.

L'évaluation suppose une investigation systématique, basée sur des indices, qui peut décrire et estimer *n'importe quel* aspect d'une politique, d'un programme ou d'un projet. L'évaluation utilise une grande variété de méthodes, quantitatives aussi bien que qualitatives, fournissant des informations

plus complètes sur ce qui se passe, pourquoi, et si c'est approprié ou non et aussi pour informer les futures directions.

L'évaluation peut être menée dans différents objectifs et prendre des formes variées. Certains des types d'utilisation suivants (la liste n'est pas exhaustive, il y a aussi de nombreuses autres approches d'évaluation) peuvent contribuer à l'évaluation d'impact dans certaines circonstances.⁴

- **L'évaluation des besoins** consiste en une estimation ou une évaluation des besoins ou de la situation problématique, souvent avant l'élaboration initiale du projet. Ce type d'évaluation identifie fréquemment des manières de répondre à ces besoins.
- **Les évaluations du processus (ou de la mise en œuvre)** décrivent la nature de l'intervention telle qu'elle est réellement mise en œuvre en pratique. Dans une certaine mesure, le suivi pourra peut-être fournir des données sur les activités du projet qui seront utiles pour l'évaluation d'impact. Cependant, les interventions sont rarement mises en œuvre exactement comme il était prévu à l'origine et changent fréquemment avec le temps, souvent pour des raisons valables. Il peut être étonnamment difficile de déterminer ce qui se passe vraiment, qui est servi, de quelles manières ou dans quelle mesure, et ce qu'il se passe d'autre. L'évaluation de processus peut rentrer plus dans les détails que ce n'est possible avec le suivi, en utilisant souvent explicitement

⁴ Le *Glossaire des principaux termes relatifs à l'évaluation et la gestion axée sur les résultats* <http://www.oecd.org/dataoecd/29/21/2754804.pdf> du Comité d'assistance au développement de l'OCDE est une mine de renseignements pour la définition des termes clés utilisés dans le développement pour l'évaluation, comme ceux qui sont utilisés dans la présente note. Cf. aussi : Michael Scriven. *Evaluation Thesaurus*. 4^{ème} édition. Sage Publications, 1991. La politique d'évaluation de l'USAID groupe toutes les formes d'évaluation autres que l'évaluation d'impact dans la catégorie « évaluation de la performance. »

comme point de départ des questions venant du suivi.

Les programmes doivent fréquemment changer, parfois de façon importante, en réponse aux informations venant des évaluations des besoins et d'autres formes de feedback du suivi-évaluation. Par exemple, un programme ciblant des orphelins et enfants vulnérables a commencé avec l'hypothèse que les enfants vivant seuls seraient les plus vulnérables mais s'est rapidement rendu compte qu'il y avait des ménages dont le chef était un adulte et où les enfants étaient encore plus vulnérables.

Si on ne comprend pas ce qu'est *réellement* le « programme », même l'évaluation d'impact la plus sophistiquée et la plus rigoureuse dans ses statistiques n'aura que peu de signification. Les évaluations qui représentent clairement le programme, avec les raisons qui expliquent les divergences par rapport à ce qui était prévu à l'origine, peuvent donner des informations inestimables pour comprendre comment les interventions du programme pourraient avoir eu un impact (ou peut-être pour indiquer des difficultés dans la mise en œuvre d'un programme ou certaines de ses hypothèses de base qui ont pu l'empêcher d'avoir un impact). Par exemple, si l'impact d'un programme a été limité, les données de l'évaluation de processus peuvent aider à établir si cela était dû à un problème de *théorie* de la manière dont le programme était censé fonctionner ou à des limites dans la manière dont il a été mis en œuvre.

Quelques définitions de l'évaluation :

Appréciation systématique et objective d'un projet, d'un programme ou d'une politique, en cours ou terminé, de sa conception, de sa mise en œuvre et de ses résultats. (Glossaire OCDE/CAD)

La collecte et l'analyse systématiques d'informations sur les caractéristiques et les effets de programmes et de projets pour servir de base pour des jugements, pour améliorer l'efficacité et/ou éclairer des décisions sur des programmes actuels et futurs. (USAID)

- **L'évaluation formative**, menée à mi-chemin de la mise en œuvre, doit améliorer la performance durant les étapes suivantes du programme ou du projet. Elle peut participer à identifier les effets immédiats, à quel moment il semble probable (ou non) que l'intervention aura un impact et ce qu'il faudra peut-être faire d'autre pour en améliorer l'efficacité. Cela peut donc rendre plus significative l'évaluation d'impact ou l'évaluation récapitulative⁵ faite plus tard.
- **L'évaluation organisationnelle** s'intéresse à l'efficacité de l'ensemble d'une organisation, ou peut-être d'une unité de l'organisation. Des facteurs organisationnels (par ex. la gouvernance, la gestion, les ressources humaines, les finances, les relations intra et inter organisationnelles, etc.) sont souvent plus importantes pour la réussite d'une intervention que sa conception – ces facteurs sont des informations essentielles qui doivent être prises en compte durant la conception et la mise en œuvre de l'évaluation d'impact.

⁵ L'évaluation récapitulative, souvent assimilée à l'évaluation d'impact, consiste à évaluer dans quelle mesure les effets anticipés ont été produits à la fin de l'intervention prévue et de porter un jugement sur la valeur de celle-ci.

De plus, le Comité d'assistance au développement (CAD) de l'OCDE, l'un des organismes les plus influents pour l'évaluation du développement, a identifié cinq critères (ou questions) de base pour l'évaluation : pertinence, efficacité, efficience, impact et durabilité.⁶ Notez que l'évaluation de *l'impact* ne représente que l'un de ces critères. Il est rarement possible de faire une évaluation d'impact se concentrant sur l'attribution sans avoir fait aussi d'autres formes d'évaluation pour mieux comprendre ce qui s'est passé. Il ne sert pas à grand chose de documenter qu'une intervention a un impact s'il n'est plus pertinent ou s'il peut y avoir des manières plus efficaces ou moins onéreuses de répondre au besoin de base ou si l'intervention et/ou ses résultats ont peu de chances de continuer ou d'être durables.⁷ La Note d'orientation 1 présente les circonstances dans lesquelles l'évaluation d'impact pourrait être utile.

L'évaluation d'impact a généralement les mêmes caractéristiques de base que d'autres formes d'évaluation et devrait être l'une des approches d'évaluation d'une ONG. Cependant, comme le suggère le tableau de la page 3, il y a habituellement des différences importantes, et c'est pourquoi il est nécessaire d'utiliser d'autres formes de S&E pour qu'une évaluation d'impact soit significative.

La Note d'orientation 1 observe que l'impact et l'évaluation d'impact ont parfois des définitions différentes. Cependant, un aspect important

⁶ Les *Critères du CAD pour l'évaluation de l'aide au développement* sont l'une des nombreuses sources contenant plus d'informations sur ces critères. On peut les trouver, avec de nombreuses autres ressources du CAD pour l'évaluation sur http://www.oecd.org/document/35/0,3746,en_21571361_34047972_31779555_1_1_1_1,00.html.

⁷ Souvent, les interventions soumises à des évaluations d'impact ou à d'autres recherches intensives, ont des ressources tellement plus élevées qu'un travail « normal » que ce n'est pas durable ou qu'elles ne peuvent pas facilement être portées à une plus grande échelle ou appliquées dans d'autres contextes.

de l'évaluation d'impact est l'*attribution*, le fait de monter un lien de cause à effet entre une intervention et des impacts qui peuvent être documentés. Il ne suffit pas cependant de savoir que des impacts se sont produits à cause d'une intervention. Pour pouvoir appliquer les résultats d'une évaluation d'impact à d'autres situations et/ou à d'autres groupes de bénéficiaires, on doit savoir *pourquoi* et *comment* les résultats donnés se sont produits et les caractéristiques de ceux qui en ont bénéficié (ou pas).

C'est là qu'un S&E sérieux peut être utile, à condition que l'on détermine à l'avance, aussi précisément que possible, les types d'informations dont on aura besoin et comment les diverses formes de S&E peuvent les procurer.

1.2. Comment le S&E peut contribuer à l'évaluation d'impact

Comme l'indique la présentation ci-dessus, il y a de nombreux types de données que le S&E peut, et même doit apporter pour permettre une évaluation d'impact. Le S&E peut apporter quatre catégories d'informations à l'évaluation d'impact.

1. Des informations sur la nature de l'intervention qui sont essentielles pour que l'évaluation d'impact établisse un lien causal entre ce qui a été fait et d'éventuels effets documentés. On a besoin d'informations de bases venant du S&E systématique sur : la nature réelle des services fournis ; les données de départ ainsi que les changements dans le temps, par rapport à qui est servi (ou à qui on s'adresse, comme dans les initiatives de plaidoyer) ; le nombre et les caractéristiques des bénéficiaires servis, avec les sous-groupes concernés ; les produits du programme et comment l'intervention a progressé. Ce sera peut-être différent de ce qui était prévu à l'origine. Comme on l'a déjà noté, l'intervention prévue

peut être très différente de ce qui a réellement été mis en œuvre, soit en réaction au feedback du S&E soit à cause de changements de situation. La population ciblée peut aussi changer sur la durée du programme, au fur et à mesure que les organisations savent mieux qui devrait être éligible pour un programme ou réussissent mieux à atteindre ces populations. Si l'on ne prend pas cela en compte, les résultats de l'évaluation d'impact n'auront aucun sens.

Une grande partie de ces données est de nature quantitative et peut venir du suivi, avec des descriptions qualitatives de la nature du programme ou du projet. L'évaluation de processus, en particulier, peut souvent fournir des informations plus approfondies montrant dans quelle mesure les services (ou mes messages, dans le cas d'un programme de plaidoyer) atteignent réellement les publics prévus, comment ils sont vus et comment ils varient selon les différents groupes de personnes.

2. Informations sur le contexte de l'intervention.

Quelles autres interventions et quels autres facteurs peuvent avoir une influence sur le fait qu'une intervention « fonctionne » ou pas ? Est-ce que différents groupes de personnes peuvent être affectés de façon différente ? (par ex. l'aide fournie après une catastrophe qui est distribuée de façon disproportionnée à ceux qui peuvent plus facilement y avoir accès ; les services et autres formes d'assistance pour lesquels les gens doivent aller sur un site, ce qui fait que ceux qui ont des problèmes de mobilité ne sont pas servis.) Ce genre de choses doit être pris en compte dans la conception des évaluations d'impact.

De la même façon, quels facteurs peuvent être spéciaux dans le contexte de cette intervention ? Par exemple, la bonne distribution de l'aide

alimentaire et les vies qui sont sauvées grâce à elle dépendent souvent de choses qui sont tout à fait en dehors du contrôle du programme, comme la logistique, le temps qu'il fait, l'appui (ou l'interférence) du gouvernement ou d'organisation terroristes, etc. Ce type d'informations est essentiel pour bien interpréter les données de l'évaluation d'impact et pour identifier les implications des résultats pour d'autres contextes ou situations. Les données de ce type sont généralement (mais pas forcément) qualitatives. Elles viennent parfois du suivi, mais souvent il faut que l'évaluation permette une exploration plus approfondie. La Note d'orientation 3 donne plus d'informations sur les méthodes qui peuvent être utilisées pour collecter ces données et d'autres formes de données.

3. Ce que nous savons déjà (ou au moins ce que nous suspectons) sur l'impact.

Quelquefois, le suivi et/ou l'évaluation donnent déjà des indices, ou au moins de forts soupçons, que des changements se produisent. Dans ces cas, une évaluation d'impact pourrait documenter et confirmer ce qui se passe réellement et dans quelle mesure ces changements sont le résultat des activités du programme. Par contre, le S&E pourrait soulever des questions sur la viabilité d'une évaluation d'impact, au moins actuellement, en indiquant qu'il est peut-être prématuré de chercher un impact. Une fois qu'une évaluation d'impact est en route, les données du suivi et autres formes d'évaluation devront bien sûr être prises en compte avec toutes les autres données collectées.

4. Quels autres indices pourraient influencer sur la décision de faire une évaluation d'impact.

Comme cela a été montré ci-dessus, les informations que donne l'évaluation sur la pertinence, l'efficacité et la durabilité d'une intervention devraient être prises en compte avant de fixer

des priorités pour les évaluations d'impact.

2. Comment intégrer l'évaluation dans la réflexion et les pratiques de S&E

Comme cela a été montré plus haut, le S&E est vital pour l'évaluation d'impact, mais ses contributions ne sont pas automatiques. Cette section donne des directives et des idées sur la manière de planifier et de mener le S&E pour qu'il *puisse* appuyer une évaluation d'impact significative. Elles sont organisées autour des étapes présentées dans le diagramme ci-dessous et discutées dans le texte qui le suit.

2.1. Formuler la théorie du changement

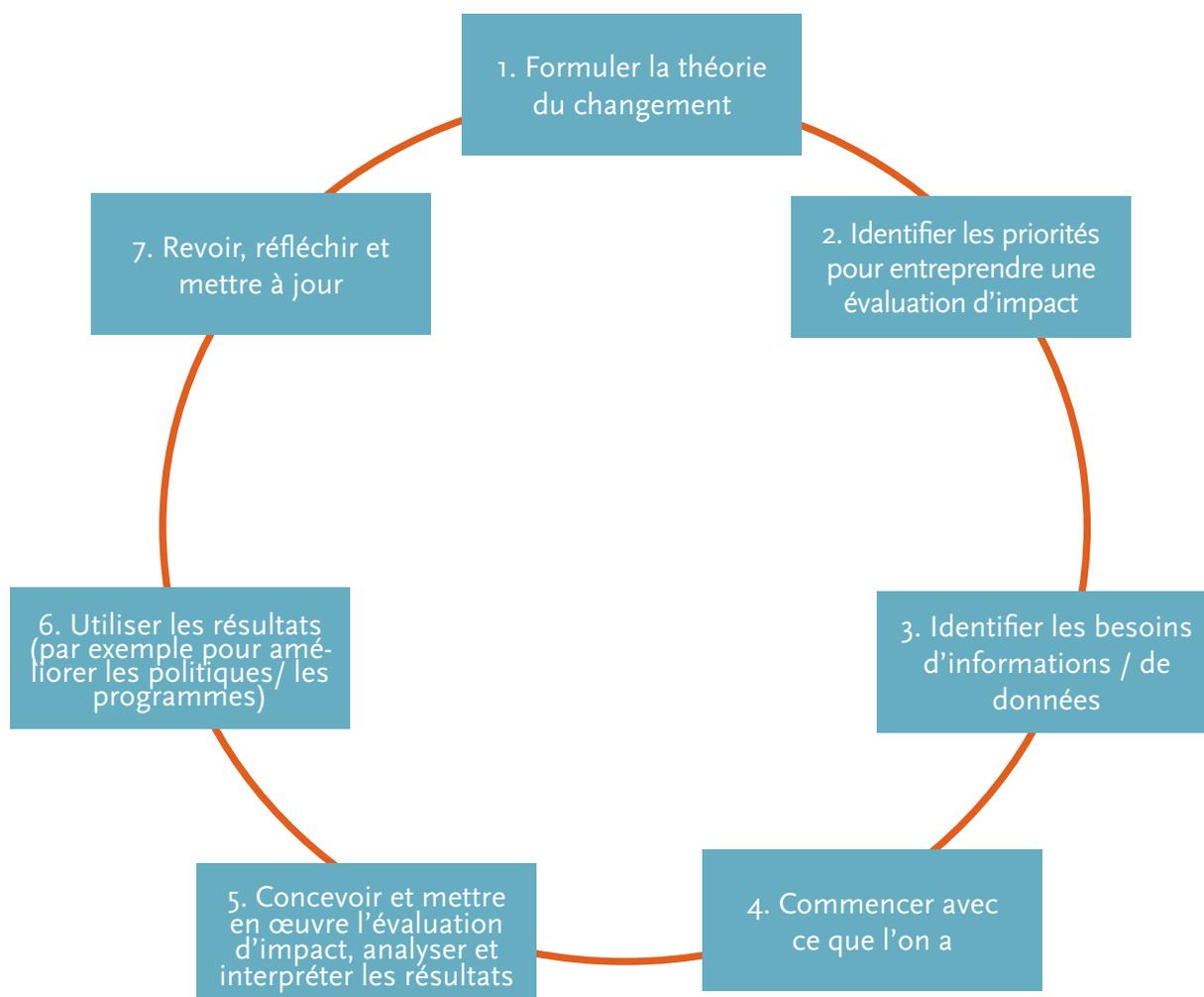
Comme l'indique la Note d'orientation 1, une *théorie du changement* est un modèle qui explique comment une intervention est censée conduire aux impacts prévus ou observés. Souvent appelée *théorie du programme*, *chaîne des résultats*, *modèle logique du programme* ou *logique d'attribution*, la théorie du changement illustre la série d'hypothèses et de liens qui montrent les relations présumées entre :

- Les intrants (par ex. le financement, le personnel, les volontaires, un appui tangible ou en nature venant d'autres personnes, etc.) ;
- Les activités et leurs produits immédiats (ce qui est fait, comme les sessions de formations qui sont menées) ;
- Les effets intermédiaires à différents niveaux (par ex. les enseignants formés, les enseignants appliquant dans leur classe ce qu'ils ont appris) ; et
- L'impact prévu (comme la réduction du travail des enfants, des familles et des communautés devenues auto-suffisantes financièrement, l'amélioration de l'état de santé, moins de morts dans les situations d'urgence, etc.).

La Note d'orientation 1 a identifié les éléments clés d'une théorie du changement et les différentes manières dont elle peut améliorer l'évaluation d'impact. Celui-ci a, lui, pour le S&E des implications très directes qui peuvent être résumées par les principales catégories ci-dessous.

1. Identifier la phase dans la trajectoire d'une intervention où il serait approprié d'entreprendre diverses formes d'évaluation – y compris l'évaluation d'impact – en identifiant les diverses étapes

intermédiaires et autres facteurs qui peuvent aussi influencer sur l'impact (par exemple l'intervention d'autres alliés, des variables contextuelles, etc.). Le S&E pourra peut-être rendre compte des progrès pour ce qui est des activités, produits et effets initiaux, ce qui indiquera quand il serait possible d'envisager une évaluation d'impact. Ceci pourrait éviter la situation bien trop courante où l'on essaie prématurément de mesurer l'impact, parfois même avant que le programme n'ait eu la possibilité de bien s'établir.



Une théorie du changement pour le travail de plaidoyer, similaire à celles qui sont élaborées pour les projets de services, devrait déterminer la succession des étapes intermédiaires (par ex. une prise de conscience, d'autres partenaires acceptent de mener des actions complémentaires, la communauté est mobilisée, un dialogue avec les fonctionnaires gouvernementaux est mis en route) et d'autres facteurs qui pourraient aussi intervenir (par ex. les activités d'autres personnes) grâce auxquels on devrait arriver à des changements dans les politiques et pratiques du gouvernement ou peut-être dans les pratiques communautaires.

Michael Woodcock de la Banque Mondiale a insisté sur le fait qu'il est important de déterminer le calendrier et les trajectoires d'impact auxquels on peut raisonnablement s'attendre. Il est crucial de reconnaître que des projets différents peuvent suivre des trajectoires très différentes. Comme il le fait remarquer, « Certains projets peuvent, de par leur nature, produire des impacts initiaux importants... tandis que d'autres peuvent prendre naturellement beaucoup plus longtemps, parfois même des décennies, pour montrer des résultats, non pas parce qu'ils « ne fonctionnent pas » après trois ans, mais simplement parce que c'est le temps que ça prend » (Woodcock, 2011). Il est aussi important de ne pas oublier que le personnel d'une organisation peut être trop optimiste par rapport au temps qu'il faut pour mettre en place un nouveau programme ou pour qu'il obtienne des résultats. Essayer d'évaluer les résultats prématurément, c'est programmer l'échec.

2. Décrire la manière dont l'intervention est censée fonctionner en utilisant une théorie du changement. Souvent appelée le *chemin vers les résultats* (ou *l'impact*), elle peut aider à suggérer les types de données de S&E nécessaires pour documenter la nature de l'intervention et comment

elle a évolué dans la pratique. La théorie du changement peut indiquer quels sont les aspects de la mise en œuvre dont on doit vérifier la qualité pour permettre de distinguer un échec de *mise en œuvre* d'un échec de *théorie*. Elle donne aussi une base pour voir à quel endroit sur le chemin vers l'impact (ou la chaîne causale) une intervention peut cesser de fonctionner. Ce type d'information est essentiel pour déterminer un lien de causalité entre l'intervention et d'éventuels effets ou impacts documentés. Il est aussi essentiel pour expliquer et interpréter la signification et les implications des résultats de l'évaluation d'impact.

3. Parvenir à une compréhension commune de l'intervention en impliquant le personnel et d'autres parties prenantes clés dans l'élaboration de la théorie du changement. Si une théorie du changement n'est élaborée que par un spécialiste du S&E en se basant surtout sur des documents, il risque de ne pas y avoir d'appropriation du modèle. Celui-ci peut aussi ne pas refléter ce qui se passe réellement. Il est conseillé, partout où c'est possible, d'impliquer le personnel des programmes, les employés sur le terrain et d'autres parties prenantes clés dans l'élaboration de la théorie du changement. Le spécialiste de S&E a un rôle crucial dans ce processus, en tant que facilitateur actif et pour établir le modèle proprement dit. Parfois, il peut y avoir des points de vue différents sur la manière dont on peut prévoir que le programme « fonctionnera », ce qui résultera en des théories du changement alternatives. Cela peut suggérer des questions qui devront être explorées par le S&E, peut-être avant de mettre en route l'évaluation d'impact elle-même.

De plus, si l'on adopte une approche participative, l'élaboration de la théorie du changement peut aider tous les participants à penser en termes d'effets. Cela peut permettre une appropriation et une

compréhension commune du but du programme et de ce qui est nécessaire pour qu'il soit efficace.

2.2. Identifier les priorités pour entreprendre une évaluation d'impact

De nombreux facteurs aident à établir des priorités pour entreprendre des évaluations d'impact, comme les intérêts des parties prenantes clés ou des bailleurs de fonds, le financement disponible ou les possibilités de collaborer avec d'autres. Mais une considération clé est de savoir quand l'évaluation d'impact aurait plus de chance d'apporter des informations utiles qui ne peuvent pas être obtenues par des moyens plus simples et moins coûteux et qui peuvent servir à guider les futures stratégies ou les futures approches de programmes.

Le S&E joue un rôle important dans ce processus. La théorie du changement peut indiquer quelles questions il est bon d'aborder à quel moment dans un programme. Les données de suivi-évaluation peuvent permettre d'indiquer jusqu'où les interventions ont progressé, à quelle phase elles sont « prêtes » pour une évaluation d'impact et quand il y a suffisamment de documentation de S&E pour que ce soit possible. Commencez par identifier les besoins d'information, puis voyez quelle quantité d'informations est déjà disponible grâce au S&E régulier avant de décider d'entreprendre une évaluation d'impact. À cette étape, on peut préparer une histoire de la performance pour voir ce qu'on connaît déjà, déterminer quels sont les manques d'informations et les questions auxquelles on ne peut pas répondre et déterminer à quel moment une évaluation d'impact serait plus valable. Cela peut être pour fournir des indices « solides » pour confirmer ce qui est suspecté ou, dans des situations où il y a de réelles questions

sur la valeur de l'intervention auxquelles on ne peut pas répondre par d'autres moyens.

2.3. Identifier les besoins d'informations/ de données

Quels types de données de S&E seront nécessaires pour aider à l'évaluation d'impact ? La réponse simple est : « ça dépend. » Il n'y a pas de prescription universelle. Ce serait aller droit à des données manquant de pertinence. C'est pourquoi cette note d'orientation a souligné l'importance de formuler la théorie du changement, de préférence en impliquant le personnel des programmes et d'autres parties prenantes dans le processus et en utilisant des techniques telles qu'une attitude réaliste par rapport à l'évaluation et une approche d'analyse de la contribution telles qu'elle est décrite ci-dessous. Ces techniques permettront d'identifier les questions pour l'évaluation qui pourront aider à déterminer quel sont les types de données nécessaires et les techniques et méthodes de S&E pour obtenir ces données.

La Note d'orientation 1 donne une liste d'exemples de questions clés pour l'évaluation d'impact. La Section 1 de cette note identifiait diverses manières dont les données de S&E peuvent contribuer à l'évaluation d'impact. Mais il n'y a pas de formule magique. Il faut identifier, très spécifiquement, quels types de données seront réellement nécessaires pour l'évaluation d'impact et dans quelle mesure les procédures de suivi et les pratiques d'évaluation qui existent devront peut-être être modifiées pour fournir ces informations (ou pour indiquer que ce n'est pas possible). Le tableau ci-dessous peut servir de liste de vérification récapitulant les types d'informations nécessaires que les approches de S&E devraient pouvoir fournir dans la plupart des situations.

Données de référence
<ul style="list-style-type: none"> • Identifiez l'état initial dans les dimensions intéressantes pour l'impact prévu du programme (la théorie du changement devrait pouvoir aider à identifier ces dimensions) • Dans les situations où l'on n'a pas établi de données de référence, y a-t-il des moyens de les recréer rétrospectivement ? *
Nature du programme/ de l'intervention telle qu'ils sont réellement mis en œuvre
<ul style="list-style-type: none"> • Identifiez les changements par rapport à la manière dont l'intervention devait à l'origine être faite et, autant que possible, les raisons de ces modifications. • Identifiez les diverses modifications et les changements en cours de route (et s'il n'y en a pas eu, le concept initial était-il réellement parfait à tous points de vue ou le programme n'a-t-il pas réussi à réagir aux changements de situations ?). • Identifiez d'autres informations descriptives sur le processus de mise en œuvre, comme l'échelonnement des diverses activités ou des variations selon les sites ou les types de bénéficiaires.
Qui le programme a-t-il servi ?
<ul style="list-style-type: none"> • Désagréger autant que possible les caractéristiques des bénéficiaires, et surtout selon les caractéristiques identifiées pendant l'élaboration de la théorie du changement et de l'histoire de la performance. • Il y a souvent des différences à suivre (par ex. le genre, l'âge, les différents groupes tribaux, culturels ou socioéconomiques) qui peuvent être évidents mais il y en a souvent d'autres qui peuvent être très pertinents pour une situation donnée (par ex. les enfants ayant des parents en vie ou non, vivant avec de la famille ou seuls).
Que s'est-il produit d'autre qui pourrait affecter l'impact du projet, dans un sens positif ou non, par ex. :
<ul style="list-style-type: none"> • Comment le programme a-t-il travaillé conjointement avec d'autres agences ou d'autres programmes, y compris les partenaires identifiés ? • Y a-t-il eu d'autres facteurs, comme l'intervention d'autres acteurs, des politiques gouvernementales, des initiatives du secteur privé (comme la création de possibilités d'emploi ou peut-être la fermeture d'un important employeur de la zone), des catastrophes naturelles ou causées par l'homme, etc., qui ont affecté positivement ou négativement la capacité du programme à progresser ou qui peuvent avoir demandé une approche un peu différente de ce qui était prévu à l'origine ? • Dans quelle mesure les facteurs ci-dessus ont-ils aidé ou perturbé le programme ?
Quels effets ou impacts peuvent être documentés, à court, moyen et long terme ?
<ul style="list-style-type: none"> • Ils devraient être en lien avec la théorie du changement, en tenant compte des types d'effets qui étaient prévus à des moments donnés. • Ceux-ci sont-ils prévus ou non ? Y a-t-il des données, ou même des évaluations subjectives, qui pourraient les expliquer ? • Les impacts ont-ils des chances d'être durables ? Qu'est-ce qui le montre ? • Quelle est la solidité des données ? Que faut-il d'autre pour monter de manière plus convaincante que l'impact écoule d'une manière ou d'une autre de l'intervention du programme.
Que s'est-il produit d'autre (effets imprévus/inattendus) ?
<ul style="list-style-type: none"> • Dans quelle mesure ceux-ci étaient-ils positifs ou négatifs ? • Dans quelle mesure ceux-ci étaient-ils potentiellement sous contrôle du programme ? • La conception du programme et la théorie du changement devaient-elles être révisées pour les prendre en compte pour l'avenir ?
Dans quelle mesure le programme tel qu'il était conçu et tel qu'il a été mis en œuvre était-il approprié et pertinent ?
<ul style="list-style-type: none"> • Par exemple, dans quelle mesure le programme a-t-il répondu aux demandes des bénéficiaires, aux besoins de la communauté ou du pays ou à d'autres priorités locales ou mondiales identifiées ? • Y a-t-il eu des changements dans les besoins et les considérations de contexte qui suggéreraient que le programme devrait être modifié d'une manière ou d'une autre ?
De quoi peut-on encore avoir besoin ?
<ul style="list-style-type: none"> • Y a-t-il d'autres besoins auxquels le programme lui-même devrait répondre ? • Y a-t-il d'autres besoins auxquels il vaudrait mieux que d'autres répondent (par ex. implications pour le plaidoyer) ?
Autres considérations ...
<ul style="list-style-type: none"> • ... telles qu'elles ont été identifiées initialement dans la théorie du changement • ...qui pourraient apparaître plus tard, par exemple en réponse à des constatations préliminaires, par des réunions de réflexion, etc. • ...qui répondent aux préoccupations ou aux perspectives d'autres parties prenantes, y compris des critiques du programme.
* Lectures supplémentaires : RealWorld Evaluation techniques, décrites dans Bamberger et al. 2012..

Les suggestions dans le tableau ci-dessus apparaissent peut-être comme une liste impressionnante de choses à suivre. Mais tous les programmes ne doivent pas collecter toutes les formes de données. Les questions ne doivent pas toutes être posées en même temps ou elles peuvent être échelonnées tout au long du cycle de vie du programme.

Il est aussi important de donner des priorités aux besoins d'informations et de se concentrer sur les informations vraiment nécessaires. Quand on élabore un plan de suivi, il faut savoir dès le début comment chaque information peut être utilisée. Si l'on essaie d'obtenir trop d'informations, cela peut avoir des conséquences imprévues. Par exemple, le personnel qui doit collecter les données peut être débordé et se trouver forcé à faire des raccourcis qui peuvent affecter la qualité des informations fournies et les rendre moins facile à utiliser.

Évaluation réaliste. Quand on détermine quelles sont les données nécessaires pour permettre une évaluation d'impact significative, il est utile de penser au principe d'*évaluation réaliste*,⁸ qui part du principe que dire qu'un programme « fonctionne » ou non est trop simplificateur et qu'une différence statistiquement significative entre deux groupes⁹ peut masquer des variations considérables. Il y a toujours des personnes qui y gagnent et d'autres non, et presque tous les programmes peuvent « fonctionner » pour certaines personnes dans certaines situations, mais pas pour d'autres. L'évaluation réaliste déclare que les effets surviennent par l'intermédiaire de combinaisons variées de contexte et de mécanismes. Par exemple,

⁸ Lecture complémentaire : Pawson et Tilley, *Realistic Evaluation* (1997).

⁹ Qui, statistiquement est basé sur la comparaison des résultats moyens de chaque groupe.

la poudre à canon ne s'enflammera que si elle est allumée d'une certaine manière. Une implication majeure est donc que l'évaluation d'impact doit déterminer ce qui fonctionne pour qui, dans quelles circonstances et par quels mécanismes.

Il y a de nombreux exemples d'approches (par ex. La mobilisation communautaire) qui doivent être adaptées selon les moyens d'existence du groupe cible, par exemple éleveurs ou populations sédentaires (comme les paysans). De la même façon, une approche pour améliorer la santé reproductive des jeunes peut fonctionner avec certains groupes culturels mais pas avec d'autres, ou dépendre d'autres facteurs.

Les principes d'évaluation réaliste ont des implications importantes pour la conception de systèmes de S&E, en particulier pour les types de données qui doivent être collectées et maintenues désagrégées. Par exemple, certaines approches de programme qui sont efficace pour les garçons peuvent, dans certaines situations, être nuisibles pour les filles. Certains programmes qui fonctionnent dans les zones urbaines peuvent ne pas fonctionner dans les zones rurales. La pertinence de certaines dimensions (par exemple le genre, l'âge ou le niveau de pauvreté) peut être évidente (mais pas forcément toujours suivie) tandis que d'autres seront peut-être moins évidentes. Ainsi, une approche éducative qui peut aider certains enfants peut ne pas le faire pour d'autres, par exemple des enfants qui ont subi un traumatisme, quel qu'il soit. En plus des caractéristiques des communautés et des bénéficiaires, il est aussi important de tenir compte du contexte ou, dans ce cas, quels services ont été fournis et dans quelles circonstances.

Les employés des ONG, étant donné leur

proche connexion avec le terrain et les bénéficiaires, sont souvent bien placés pour pouvoir observer, ou même suspecter, les facteurs qui peuvent faire que des approches fonctionnent pour certains mais pas pour d'autres. Documenter ce type d'informations grâce au S&E continu peut être une contribution importante à l'évaluation d'impact. Il est essentiel, non seulement de comprendre ce qui s'est réellement produit dans un programme particulier, mais aussi de tirer des leçons de cette expérience et d'être capable d'appliquer ces informations dans des situations différentes.

2.4. Commencer avec ce que l'on a

Avant de mettre en route une évaluation d'impact coûteuse et complexe, commencez avec les informations déjà disponibles, venant du S&E et d'autres sources. Peut-être qu'il existe déjà suffisamment d'indices et qu'il ne vaut pas la peine de faire une évaluation d'impact. Au minimum, cela peut aider à concentrer l'évaluation d'impact sur les questions et les problèmes auxquels on ne peut pas trouver de réponse autrement. Cela peut suggérer d'autres formes d'évaluation qu'il vaudrait peut-être mieux essayer d'abord. Voici quelques idées d'approches systématiques.

L'analyse de la contribution comporte les six étapes ci-dessous, qui seront expliquées plus en détail dans l'Annexe 1 :

1. Élaborer une théorie du changement.
2. Évaluer les indices qui existent sur les résultats.
3. Évaluer des explications alternatives.
4. Préparer l'histoire de la performance.
5. Chercher des indices supplémentaires.
6. Réviser et renforcer l'histoire de la performance.

En substance, il s'agit d'utiliser les indices venant du S&E réalisé pour voir ce qu'ils peuvent révéler sur les effets ou même sur l'impact d'une intervention, tout en voyant aussi ce qui, en dehors de l'intervention, pourrait également avoir causé les effets trouvés. À partir de là, on peut préparer une histoire provisoire de la performance pour voir dans quelle mesure il est raisonnable de faire l'hypothèse que les actions du programme pourraient avoir contribué aux effets observés et pour identifier les faiblesses et là où on pourrait avoir besoin de données additionnelles. C'est une manière efficace d'utiliser les données existantes pour déterminer ce que l'on sait et où on a besoin de données venant d'autres formes de S&E ou, si nécessaire, d'une évaluation d'impact pour obtenir une image plus convaincante.

Discussions de réflexion. Une approche d'analyse de la contribution n'est pas forcément un gros travail, surtout pour de petits projets ou des projets contrôlés. Des discussions de réflexion systématiques (ou des réunions de « revue » ou de « suivi ») sont un bon exemple de cette approche. Elles pourraient faire partie des réunions régulières du personnel ou pourraient être le point principal de réunions spéciales organisées périodiquement. Elles peuvent comporter une discussion informelle de ce qui semble fonctionner ou pas, de la base sur laquelle on base ces conclusions, en tenant compte d'explications concurrentes des effets observés ou prouvés (autrement dit, une version simplifiée de l'histoire de performance), puis identifier quelles autres formes d'indices seront nécessaires à l'avenir pour prouver de manière plus convaincante l'impact du programme. Dans certains cas, par exemple quand des employés et des parties prenantes venant de différentes zones sont réunis, il peut être bon d'utiliser une approche plus formelle et un facilitateur externe. Ce type de suivi quantitatif peut être inestimable pour comprendre

ce qui peut avoir mené aux effets ou aux impacts (et si ces mêmes résultats se seraient produits quand même en l'absence du projet), ce qui est essentiel comme informations pour appliquer les résultats des évaluations d'impact.

AWARD, un programme de bourses de recherche pour les femmes scientifiques dans la recherche et le développement agricoles en Afrique, utilise des sessions périodiques de réflexion pour donner aux boursières et à leurs mentors dans différents pays une possibilité de réfléchir aux progrès, aux réalisations jusqu'à présent et à ce qui a, ou non, bien marché. Cela s'ajoute aux données de suivi de l'équipe de gestion ainsi qu'à leur propre réflexion continue sur la performance. Cela donne des occasions de tester la théorie du changement du programme et aide aussi l'équipe de gestion à travailler en mode « d'amélioration continue » en déterminant quelles sont les mesures correctives ou les changements de direction nécessaires.

Éliminer les explications plausibles rivales. La 3^{ème} étape de l'approche d'analyse de la contribution consiste à apprécier des explications alternatives possibles. Il s'agit là d'un concept simple mais puissant, bien ancré dans la littérature de l'évaluation. Il a été exprimé en premier par Donald T. Campbell,¹⁰ qui a aussi dit qu'on devrait utiliser les approches les plus simples et les moins coûteuses qui peuvent permettre d'avoir la confiance nécessaire dans les constatations.

Commencez par voir ce qu'il y a comme autres explications possibles qui pourraient expliquer

¹⁰ Par exemple : Campbell, D. T. et J. C. Stanley. 1963. *Experimental and Quasi-Experimental Designs for Research*. Boston: Houghton-Mifflin. Campbell est peut-être mieux connu comme le père de l'évaluation moderne des programmes et comme un partisan des méthodes expérimentales et quasi expérimentales. Il a été le premier à identifier ce dernier concept, ainsi que les concepts de validité interne et externe.

comment l'impact a été produit, en dehors de l'intervention du programme ou du projet, puis fournissez les indices nécessaires pour montrer que ces explications alternatives ne sont pas probables. Il suffit de s'intéresser à d'autres explications *plausibles*. On peut supposer sans trop de risques que les changements dans les chiffres d'emploi des jeunes dans un district n'ont pas été causés par le fantôme d'Elvis ! Mais l'ouverture ou la fermeture d'une nouvelle entreprise, ou des changements dans les politiques d'une entreprise existante, peuvent être pertinents dans certaines situations. Il est important cependant de se rendre compte que les parties prenantes peuvent avoir une vision différente de ce qui est plausible ou non et qu'il faut en tenir compte quand on conçoit une approche de S&E.

Quelques approches pour exclure des explications rivales plausibles

- Anticiper des questions potentielles du/des public(s) de l'évaluation.
- Abordez les menaces à la validité interne et externe.
- Utilisez une approche basée sur la théorie.
- Déterminez quels types d'indices seraient convaincants.
- Fournissez plusieurs lignes d'indices.
- Assurez-vous de la validité d'aspect.*

* Il s'agit de présenter les résultats de manière à ce que les personnes qui ne sont pas des chercheurs aient l'impression de prime abord qu'ils sont logiques.

2.5. Concevoir et mettre en œuvre l'évaluation d'impact, analyser et interpréter les résultats

Comme cela a été indiqué auparavant, le S&E a un rôle important à jouer pour qu'une évaluation d'impact pose les bonnes questions et que sa conception tienne compte de manière réaliste des données intéressantes de S&E qui peuvent déjà

exister ou pourraient être fournies pratiquement. Il faudra peut-être des activités complémentaires d'évaluation, comprenant des approches qualitatives ainsi que quantitatives, ce qui contribuera à une approche de méthodes mixtes (présentée dans la Note d'orientation 3) qui pourrait permettre une triangulation des résultats de l'évaluation d'impact et plus tard aider à l'interprétation. Des spécialistes de S&E devront sans aucun doute être impliqués, au moins à certains moments du processus d'évaluation d'impact, pour faciliter la collecte de données additionnelles. Les spécialistes de S&E doivent aussi être parfaitement au courant des problèmes et des compromis qu'il faut faire pour pouvoir établir une validité quelconque.¹¹

Une hypothèse courante et qu'une fois que la conception de l'évaluation d'impact est faite, tout doit rester statique, souvent pendant des années. Mais en réalité, les programmes évoluent et changent avec le temps, et ils doivent le faire pour répondre aux changements de situation, au feedback et aux nouvelles informations. Les conceptions d'évaluations d'impact doivent pouvoir refléter les changements dans la nature de l'intervention et s'y adapter, faute de quoi les constatations risquent de n'avoir qu'une signification limitée. Comme l'indique la Note d'orientation 3, certaines approches de l'évaluation d'impact peuvent être plus souples et adaptables que d'autres. Néanmoins, il est important de garder à l'esprit que l'évaluation d'impact a besoin d'une coopération du programme pour réussir et qu'autant que possible, les changements importants devraient être discutés avec l'équipe d'évaluation.

¹¹ Une ressource très utile : Huey T. Chen, Stewart I. Donaldson, et Melvin M. Mark (éds.). *Advancing Validity in Outcome Evaluation : Theory and Practice. New Directions for Evaluation*. Jossey-Bass et American Evaluation Association. 2011.

Les spécialistes du S&E (et par leur intermédiaire, d'autres personnels de l'organisation et d'autres parties prenantes) devraient jouer un rôle important dans l'interprétation des résultats de l'évaluation d'impact. Il est utile de demander si les résultats ont un sens, et s'ils sont cohérents avec d'autres informations. Si ce n'est pas le cas, il serait bon de demander ce que cela peut signifier et ce qui pourrait expliquer les contradictions apparentes.

Comme le soulignait la Section 1, on a besoin du S&E pour fournir des informations sur les facteurs contextuels pertinents et sur la nature des interventions. Ces interventions seront essentielles pour montrer des relations de cause à effet. En particulier, on a besoin du M&E pour fournir des *explications* des constatations, pour indiquer les « pourquoi » et les « comment, » des informations qui sont essentielles pour agir sur la base des résultats de l'évaluation d'impact. Les hypothèses faites sur « l'intervention » sont-elles encore valides ? Les conceptions avec des scénarios contre-factuels ont souvent beaucoup de mal à expliquer les différences entre le groupe expérimental et le groupe de contrôle. Même s'il n'y a pas de différence entre les scores moyens du groupe expérimental et du groupe de contrôle, il est probable que certaines personnes ont quand même progressé et il est toujours possible qu'une majorité des gens pourraient faire mieux. Des informations de S&E complémentaires peuvent aider à entrer dans le « boîte noire » pour pouvoir comprendre et interpréter les différences, et ce qu'elles signifient pour les directions à l'avenir.¹² Ces réflexions

¹² Par exemple, les évaluations d'impact sont souvent entreprises dans l'objectif d'indiquer si une approche donnée devrait être continuée et/ou répliquée dans une situation différente (possibilité de généraliser, ou validité externe). Mais sans savoir comment ni pourquoi l'impact s'est produit (ou non), il est impossible de prendre des décisions en connaissance de cause pour ce type de choix.

pourraient suggérer d'autres analyses de sous-groupes qui seraient encore possibles ou au moins une aide pour bien interpréter constatations.

2.6. Utiliser les résultats

L'utilisation de l'évaluation d'impact fait l'objet de la Note d'orientation 4 et déborde largement du sujet de la présente note. Il est néanmoins intéressant de noter que le S&E peut avoir un rôle important à jouer dans ce processus. D'autres informations de S&E peuvent aider à identifier dans quelle mesure les résultats d'une évaluation d'impact peuvent être généralisés ou peuvent être applicables à d'autres situations ou à d'autres populations ou même au même endroit, était donné les changements inévitables dans le contexte, dans la nature de l'intervention et dans le groupe cible, au cours d'une évaluation d'impact pluriannuelle. Les spécialistes de S&E peuvent aider à animer des sessions impliquant le personnel et la direction pour voir quelles sont les implications des résultats de l'évaluation d'impact à la lumière d'autres informations.

2.7. Revoir, réfléchir et mettre à jour

Quels que soient les efforts fait pour la formulation initiale de la théorie du changement, pour l'élaboration du plan et la conception de l'évaluation et pour la mise en place des mécanismes de collecte des données, cela ne signifie pas qu'on puisse, ni qu'on doive ensuite ne plus y penser. Il est utile de voir régulièrement si les hypothèses qui sous-tendent l'approche de l'évaluation d'impact sont toujours valables ou s'il faut y faire des changements. Les programmes réactifs sont rarement statiques – ils adaptent leurs approches aux changements de contexte, pour répondre à de nouveaux besoins ou de nouvelles opportunités et en réponse au feedback de leurs bénéficiaires, d'autres parties prenantes et du S&E. Si ceux-ci sont identifiés suffisamment tôt, l'évaluation

d'impact pourra peut-être encore les prendre en compte. Il faudrait choisir de bonnes méthodologies d'évaluation d'impact qui soient suffisamment souples pour s'adapter aux changements de direction du programme qui sont, dans la plupart des cas, presque inévitables.

Il est intéressant de faire des revues intermédiaires (qui peuvent être aussi simple qu'une réunion ciblée du personnel) pour étudier tous les changements importants dans le programme et réexaminer la théorie du changement pour voir si les hypothèses d'origine sont toujours valables ou s'il faut les mettre à jour. Comme cela a été suggéré plus haut, cette revue peut être une occasion de préparer ou de mettre à jour une histoire de la performance et d'identifier les informations manquantes qu'il faudra se procurer d'une manière ou d'une autre. Les revues intermédiaires peuvent être une occasion de réfléchir à la manière dont le programme semble progresser mais aussi servir à revoir, et si nécessaire à mettre à jour, le plan de S&E lui-même.

Cette note insiste sur le rôle que peut jouer le S&E pour aider à faire une évaluation d'impact significative. Mais l'inverse est vrai aussi. L'évaluation d'impact peut toujours suggérer les types de données qui devront être suivies à l'avenir et poser des questions qui pourront être étudiées par d'autres formes d'évaluation. Il est intéressant, à la conclusion d'une évaluation d'impact, de considérer les implications pour les priorités futures du S&E.

3. Impliquer toutes les parties de l'organisation

Dans le S&E et l'évaluation d'impact, les spécialistes du S&E doivent évidemment jouer un rôle actif pour mettre en place des conceptions d'évaluation appropriées et pour garantir que tous

les aspects du processus sont faits de manière à maximiser la validité et l'utilité des données. Mais le personnel de gestion et des programmes peut aussi jouer des rôles importants qui ne sont pas toujours bien appréciés.

3.1. S&E : Une fonction centrale de gestion qui a besoin du leadership et de l'appui de la direction

Faire le S&E et réfléchir et agir en termes d'impact, sont des choses bien trop importantes pour être laissées aux seuls spécialistes du S&E. Tout d'abord, le S&E représente un outil de gestion qui peut aider à fixer les directions, à évaluer les progrès, à en apprendre plus sur les types d'approches qui semblent fonctionner ou non dans des circonstances changeantes, à prendre des décisions, et à beaucoup d'autres choses. Cela signifie que la direction des ONG doit se l'approprier, fournir un leadership et encourager tous les membres de l'organisation à réfléchir en termes d'effets ou de résultats en se demandant comment ils peuvent savoir que leur travail fait une différence et ce qui peut être fait pour l'améliorer. L'apprentissage et le changement supposent d'être prêts à reconnaître et à accepter le fait que, quelquefois, certaines approches n'ont pas bien fonctionné et doivent être modifiées. Mais pour cela, il faut que la direction encourage et aide les employés à innover et ne les punisse pas si tout ne se passe pas comme prévu – à condition qu'ils soient prêts à tirer des leçons de l'expérience et de l'évaluation. Il est essentiel d'avoir cet appui pour que l'évaluation d'impact, avec le S&E, ait un impact réel sur la manière dont les programmes sont mis en œuvre.

Les gestionnaires à tous les niveaux d'une agence devront fournir l'appui, les ressources et le leadership nécessaires pour permettre un S&E, et une évaluation d'impact, intéressants. La direction,

en particulier aura un rôle clé à jouer pour que ce ne soit pas seulement un travail sur le papier qui risquerait d'entraîner un cynisme et une obéissance passive, mais pour montrer au contraire leur engagement à agir et à utiliser toutes formes d'informations tirées des résultats. Faire une évaluation d'impact, même si un budget y est consacré, aura aussi des implications pour la charge de travail et peut-être aussi des implications de ressources pour le personnel des programmes et les spécialistes du S&E. La direction devrait être au courant de ces implications lorsqu'elle décide d'entreprendre une évaluation d'impact. La pertinence et la réussite de l'évaluation d'impact et la capacité du S&E habituel à contribuer à cette réussite dépendent en grande partie de l'attitude et de l'appui de la direction.

3.2. Le personnel des programmes doit avoir un rôle actif

Pour une évaluation d'impact significative, ainsi que pour le S&E habituel, il faut que le personnel des programmes soit activement impliqué. Dans la plupart des cas, c'est le personnel des programmes qui aura la responsabilité de collecter la plupart des données, de les enregistrer et d'en rendre compte. Il est important de reconnaître ce rôle clé et ce qu'il implique. Si l'on impose des exigences à des gestionnaires et à des employés de programme déjà très occupés, cela sera dans la pratique souvent considéré comme une tâche administrative supplémentaire, d'une valeur incertaine et il y aura une résistance. Par conséquent il risquera d'y avoir des problèmes, avec des données de moindre qualité ou incomplètes.

Faire que le S&E et l'évaluation d'impact soient utiles pour les employés des programmes et ne soient pas considérés comme une corvée

- Les impliquer dans le processus.
- S'intéresser aux questions qu'ils considèrent comme importantes et qui peuvent les aider.
- Fournir des occasions de participer à l'interprétation des données du S&E / de l'évaluation d'impact.
- Donner un feedback quand les données sont soumises (même un simple accusé de réception ou un « merci »).
- Féliciter les employés qui aident à l'évaluation d'impact.
- Faire savoir comment les données ont été utilisées.

Comme cela a été suggéré plus haut, la direction d'une organisation doit apporter un leadership convenable afin de préparer le terrain pour une bonne collecte des données. Pour que les employés des programmes coopèrent, comme cela est nécessaire, à la collecte des données, il faudra leur donner des incitations et les ressources nécessaires (temps, personnel et financement) pour permettre la collecte et l'enregistrement de données de qualité ainsi que des mécanismes appropriés pour le contrôle de la qualité. Pour que la collecte des données soit pertinente et significative, il faudra que le personnel des programmes soit impliqué dans l'identification des questions et des problèmes dont il pense qu'ils sont importants et pertinents pour lui et pour les communautés où il travaille. L'encadré qui suit donne quelques idées pour rendre le S&E et l'évaluation d'impact significatifs. Par exemple, les employés d'un projet d'éducation pour tous en Ouganda ont commencé à apprécier l'importance des données de suivi qu'on leur demandait de fournir et à améliorer leur qualité une fois que

certaines des rapports initiaux d'évaluation leur ont été communiqués.

Autant que possible, l'approche de S&E ou d'évaluation d'impact devrait inclure des réflexions sur le rôle et l'appui du personnel des programmes. Les employés des programmes qui devront collecter les données sont souvent les mieux placés pour identifier les informations qui seront faciles à obtenir et celles qui peuvent être plus problématiques ainsi que les demandes de données qui sont ambiguës et peuvent être interprétées de manières différentes ou risquent même d'être tellement peu fiables qu'elles seront inutilisables. Grâce à leur implication dans le processus, ils ont plus de chances de considérer le S&E comme un appui plutôt que comme une corvée inutile ou quelque chose « qu'on leur fait » et qui n'a pas de valeur visible.

Résumé

L'évaluation d'impact est parfois vue, par des chercheurs comme par de nombreux gestionnaires et employés des programmes, comme une activité très technique qui ne peut être réalisée que par des experts externes travaillant à une certaine distance du programme, de son personnel et de ses parties prenantes. Et dans certaines situations, c'est comme ça qu'elle est abordée.

Mais comme l'indique cette note, l'évaluation d'impact ne représente qu'une forme d'évaluation. Il n'arrive que rarement qu'une évaluation d'impact significative puisse être menée sans se baser sur des données venant d'autres activités continues de S&E. Mais la valeur et l'utilisation du S&E pour l'évaluation d'impact ne sont pas automatiques. Il faut planifier à l'avance pour identifier les types de données dont on aura besoin pour pouvoir déterminer l'impact de politiques,

de programmes et de projets et prévoir la manière dont on peut se procurer ces données. Et, comme le souligne cette note, quand tout le monde dans une agence pense en termes d'impact ou de résultats, cela peut suffire à améliorer l'efficacité des opérations de l'agence et aider à améliorer la vie de populations partout dans le monde.

Références et autres ressources utiles

- Bamberger, M., Rugh, J. et, L. (2012). *RealWorld Evaluation : Working Under Budget, Time, Data, and Political Constraints* (2ème éd.). Sage Publications.
- Chen, H., Donaldson, S., Mark, M. (éds.) (2011). *Advancing Validity in Outcome Evaluation : Theory and Practice. New Directions for Evaluation*. No130. Jossey-Bass et American Evaluation Association.
- Funnell, S. et Rogers, P (2011). *Purposeful Program Theory : Effective Use of Theories of Change and Logic Models*. Jossey-Bass Publications.
- Jones, H. (2011). « A guide to monitoring and evaluating policy influence. » Londres : Overseas Development Institute. <http://www.odi.org.uk/resources/docs/6453.pdf>
- Mayne, J. (2001). « Addressing attribution through contribution analysis : Using performance measures sensibly », *Canadian Journal of Program Evaluation*, 16(1), 1-24.
- Mayne, J. (2011). « Contribution Analysis : Advancing Cause and Effect. » In *Evaluating the Complex : Attribution, Contribution, and Beyond*, édité par Robert Schwartz, Mita Mara, et Kim Forss. Londres : Transaction Publishing.
- OCDE/CAD. (2011). *Evaluating Impacts : An Overview*. http://www.oecd.org/document/17/0,3746,en_21571361_34047972_46179985_1_1_1_1,00.html
- OCDE/CAD. (2002). *Glossaire des principaux termes relatifs à l'évaluation et la gestion axée sur les résultats*. <http://www.oecd.org/dataoecd/29/21/2754804.pdf>
- Malhotra, A. (2010). « Connecting Routine M&E with Impact Evaluation, » International Center for Research on Women. Présentation à InterAction/3ie Impact Evaluation Design Clinic, Washington, D.C., avril 2010. http://www.interaction.org/sites/default/files/18/Anju_Malhotra_-_Connecting_Routine_M&E_with_Impact_Evaluation.pdf
- Paul, S. (2002). *Holding the State to Account : Citizen Monitoring in Action*. Bangalore : Public Affairs Centre.
- Pawson, R. et Tilley, N. (1997). *Realistic Evaluation*. Sage Publications.
- Scriven, M. (1991). *Evaluation Thesaurus* (4ème éd). Sage Publications.
- Woolcock, M. (2011). « Guest Post on The Importance of Time and Trajectories in Understanding Project Effectiveness. » Development Impact. Blogs.worldbank.org. <http://blogs.worldbank.org/impacetevaluations/guest-post-michael-woolcock-on-the-importance-of-time-and-trajectories-in-understanding-project-effe>.

Annexe 1 – Analyse de la contribution

John Mayne, qui travaillait à l'époque pour le Bureau du Vérificateur général du Canada, a suggéré d'utiliser l'analyse de la contribution comme une approche grâce à laquelle on peut faire le meilleur usage des données existantes sur les effets et

l'impact. Il a résumé cette approche comme suit. On trouvera des informations plus détaillées sur une approche de l'analyse de la contribution dans Mayne (2001), où ce tableau a été copié mot pour mot, ou dans Mayne (2011).

<p>Étape 1 : Élaborer la chaîne des résultats</p>	<p>Élaborer le modèle de théorie du programme/ la logique du programme/ la chaîne des résultats décrivant comment le programme est censé fonctionner. Identifier aussi les principaux facteurs extérieurs en jeu qui peuvent expliquer les effets observés. Cette théorie de programme devrait mener à une association plausible entre les activités du programme et les effets recherchés. Certains maillons de la chaîne des résultats seront assez bien compris ou acceptés. D'autres seront moins bien compris ou il faudra expliquer les autres que le fait que le programme était la « cause. » De cette manière, vous reconnaîtrez que l'attribution est bien un problème.</p>
<p>Étape 2 : Évaluer les indices qui existent sur les résultats</p>	<p>La chaîne des résultats devrait donner une bonne idée des résultats attendus (produits, effets immédiats et finaux) qui ont pu être mesurés. Quels indices (informations tirées des mesures de performance et des évaluations) sont actuellement disponibles pour montrer que ces divers résultats se sont produits ? Les maillons de la chaîne des résultats doivent aussi être estimés. Lesquels sont solides (existence d'indices clairs, logique solide ou largement acceptés) et lesquels sont faibles (peu d'indices, logique faible ou peu d'accord parmi les parties prenantes).</p>
<p>Étape 3 : Évaluer des explications alternatives</p>	<p>Les effets sont, par définition, influencés non seulement par l'action du programme mais aussi par des facteurs externes – d'autres programmes ou des facteurs sociaux et économiques. En plus d'une estimation des indices qui existent sur les résultats, il est nécessaire de prendre explicitement en compte l'influence que peuvent avoir ces facteurs externes. Des indices ou une argumentation logique peuvent suggérer que certains n'ont qu'une petite influence et que d'autres peuvent avoir une influence plus importante sur les résultats prévus.</p>
<p>Étape 4 : Préparer l'histoire de la performance</p>	<p>Avec ces informations, vous pourrez démarrer votre histoire de la performance, expliquant pourquoi il est raisonnable de penser que les actions du programme ont contribué (d'une manière que vous pourrez essayer de caractériser) aux effets observés. À quel point cette histoire est-elle crédible ? Des personnes raisonnables sont-elles d'accord avec elle ? Le schéma des résultats observés valide-t-il la chaîne des résultats ? Quelles sont les principales faiblesses de l'histoire ? Il y aura toujours des faiblesses. Celles-ci montrent où il serait utile d'avoir des données ou des informations supplémentaires.</p> <p>S'il n'est pas possible (du moins pour l'instant) de se procurer des indices supplémentaires, c'est tout ce que vous pouvez dire sur le degré où le programme a fait une différence.</p>
<p>Étape 5 : Chercher des indices supplémentaires</p>	<p>Pour améliorer votre histoire de la performance, vous aurez besoin d'indices supplémentaires. Il pourra s'agir d'informations à la fois sur la mesure dans laquelle des résultats spécifiques apparaissent dans les chaînes de résultats et sur la force de certains maillons de la chaîne. Ce document présente un certain nombre de techniques de renforcement que vous pourriez adopter.</p>
<p>Étape 6 : Réviser et renforcer l'histoire de la performance</p>	<p>Avec les nouveaux indices, vous devriez pouvoir construire une histoire plus crédible, une histoire qu'une personne raisonnable aura plus de chance d'accepter. Elle ne sera sans doute pas infaillible, mais elle sera plus solide et plus crédible.</p>