

Qu'est-ce qu'une évaluation d'impact?

En collaboration avec le Centre pour le développement et la coopération de l'EPFZ (NADEL)



L'objectif des évaluations d'impact est d'**améliorer l'efficacité de l'aide**. Une évaluation d'impact analyse les effets **pouvant être attribués** à un projet ou programme de développement donné.

© ETH NADEL

En quoi consiste une évaluation d'impact?

L'évaluation d'impact (positif ou négatif, intentionnel ou non) a pour but d'analyser l'impact d'un projet, d'un programme ou d'une politique sur le groupe cible et de mesurer l'ampleur de cet impact. Une telle évaluation permet d'établir **un lien de cause à effet** entre un projet, un programme ou une politique et un ou plusieurs résultats au niveau des *outcomes*. Cela implique de reconstituer une situation contrefactuelle. A titre d'exemple, pour évaluer l'impact d'un projet ou programme de développement visant à améliorer la santé des enfants en construisant des pompes à eau communes, on analysera la réduction de l'incidence de la diarrhée par rapport à l'incidence attendue si aucune pompe n'avait été installée. Les évaluations d'impact indiquent dès lors si des changements mesurables dans la vie des personnes concernées **peuvent être attribués** à un projet ou programme de développement spécifique. Il ne s'agit pas de mesurer la contribution d'un projet ou programme aux grands objectifs de déve-

loppement avec lesquels le projet ou programme n'a qu'un lien indirect, comme l'atténuation du changement climatique ou la réduction de la pauvreté. Lorsqu'on parle d'évaluation d'impact, il est ici question spécifiquement d'un lien de cause à effet direct entre d'une part un projet, un programme ou une politique et d'autre part un *outcome* spécifique (comme la réduction de l'incidence de la diarrhée).

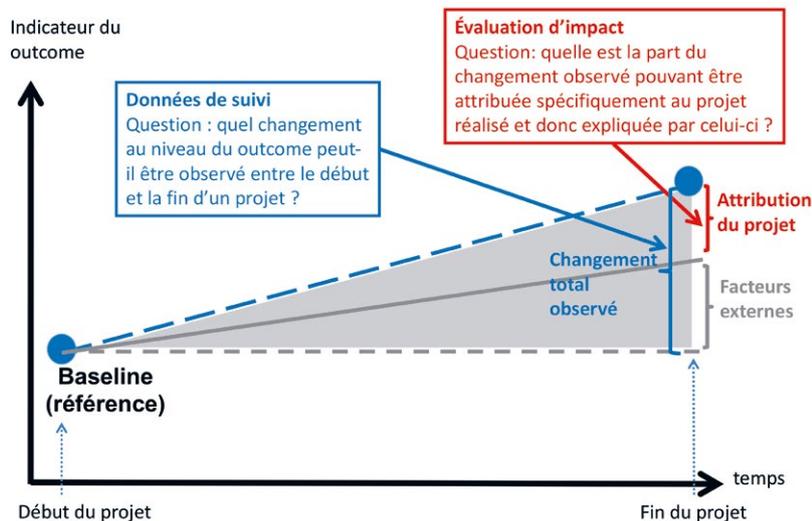
Quelle est la différence entre une évaluation d'impact et une évaluation de projet ?

Alors que la plupart des évaluations de projet ou de programme analysent en priorité le succès des opérations réalisées ou des services fournis, les évaluations d'impact tentent d'établir s'il existe un lien de cause à effet entre le projet ou programme et un changement dans la vie des personnes concernées, et d'expliquer la présence ou l'absence d'effet. Dans l'exemple présenté ci-dessus, une évaluation de projet ou programme chercherait à savoir si les pompes ont bien été construites et combien de per-

sonnes les utilisent. L'évaluation d'impact chercherait à constater si la construction de pompes a eu un effet bénéfique sur la santé. De nombreuses évaluations de projets et de programmes ont recours à des données de suivi pour mesurer, par exemple, le nombre de personnes utilisant une pompe. Certaines évaluations recueilleront même des données de suivi sur la prévalence de la diarrhée. Toutefois, les **données de suivi** et les comparaisons avant-après sont souvent insuffisantes pour établir un lien de cause à effet entre un projet ou programme de développement et le *outcome* observé, parce que les facteurs externes ne sont pas pris en compte.

Quelles raisons pour mener une évaluation d'impact ?

Les évaluations d'impact contribuent à **l'élaboration de politiques basées sur les faits**. Avec ces évaluations, l'attention autrefois focalisée sur les **activités** se porte davantage sur les **résultats**. A l'aide de cet outil, une organisation peut décider de réaliser des projets, dont l'impact positif a été



Les **données de suivi** servent à mesurer le changement au niveau des *outcomes* au fil du temps.

L'**évaluation d'impact** sert à quantifier la part du changement observé pouvant être attribuée au projet ou programme évalué.

démonstré, à une échelle plus importante (**scaling-up**) et d'**abandonner** les projets ou programmes n'ayant pas eu d'impact. Les évaluations d'impact peuvent contribuer à **améliorer** la conception des projets, programmes et politiques de développement et/ou à **comparer** leur efficacité. Elles sont ainsi avant tout un outil pour l'amélioration des projets et programmes de développement et pour l'**apprentissage**.

Quelles sont les défis liés à une évaluation d'impact ?

L'idéal serait de juger de l'efficacité d'un projet ou programme de développement en comparant les *outcomes* pour les individus ayant participé au projet ou programme à **ceux auxquels ces mêmes individus seraient parvenus s'ils n'avaient pas participé**, et ce au même instant. Cet *outcome* hypothétique est appelé le **contrefactuel**. Dans le monde réel, il est rarement observable. Dans l'exemple évoqué, où il s'agit d'établir la part de réduction de l'incidence de la diarrhée attribuable à la pompe (en comparaison avec une situation sans pompe), nous sommes confrontés au défi suivant : s'il est possible de mesurer l'incidence de la diarrhée dans le groupe de personnes ayant reçu une pompe (« groupe de traitement »), il n'est pas possible de mesurer simultanément l'incidence de la diarrhée dans un monde où le même groupe de personnes n'aurait pas bénéficié d'une pompe à ce moment-là (contrefactuel). L'un des principaux enjeux des évaluations d'impact est donc de se rapprocher du contrefactuel en trouvant un groupe d'individus, aussi proche que possible des individus qui y participent, n'ayant

pas participé au programme (« groupe de comparaison », ou « groupe témoin »). Les *outcomes* mesurés dans le groupe de comparaison servent d'indicateur sur ce qu'aurait été la situation des participants en l'absence du programme. **Le groupe de comparaison est déterminant pour la qualité d'une évaluation d'impact.**

Comment définir un groupe de comparaison ?

Différentes méthodes peuvent être utilisées pour définir un groupe de comparaison. Les plus fréquentes dans le contexte des politiques de développement sont la méthode des doubles différences, l'appariement statistique, la régression par discontinuité et les essais randomisés (voir les références pour davantage de détails), mais il y en a bien d'autres. Utilisées correctement, ces méthodes sont des outils puissants pour définir un groupe de comparaison pertinent et garantir ainsi que l'impact mesuré n'est pas dû à un biais – autrement dit, qu'il n'est pas induit par des **erreurs systématiques**.

Que qualifie-t-on de résultat biaisé ou d'erreur systématique ?

Un biais est possible lorsque les différences observées entre les individus ayant participé à un projet ou programme et le groupe de comparaison sont attribuées de **manière erronée à ce projet ou programme**, alors qu'elles peuvent être expliquées par d'autres facteurs. Par exemple, une simple comparaison de l'incidence de la diarrhée dans le même groupe **avant et après** l'installation de pompes ne per-

met pas toujours d'obtenir une estimation fiable de l'impact. De nombreux facteurs sans liens avec le programme en question (comme l'éducation à l'hygiène, une augmentation des revenus ou une amélioration générale de l'infrastructure sanitaire) pourraient expliquer le changement constaté en matière de santé au cours de la même période. De même, une simple comparaison de l'incidence de la diarrhée dans les villages **équipés ou non** d'une pompe peut s'avérer problématique. Un biais peut notamment s'expliquer par des différences préexistantes entre les villages, par exemple lorsque les villages qui bénéficient d'une pompe sont plus proches d'une route principale. Ce type d'erreurs systématiques se traduit par des estimations d'impact biaisées et donc incorrectes.

Qu'en est-il de la taille des échantillons ?

Pour qu'une évaluation d'impact soit possible, il est nécessaire de disposer d'un **échantillon suffisamment grand** pour garantir que **l'effet d'un projet, d'un programme ou d'une politique puisse être mesuré**. Le nombre de participants et le nombre d'individus dans le groupe de comparaison devraient également être suffisamment grands pour permettre de conclure que l'effet observé ne relève pas de la simple coïncidence, mais qu'il est bien dû au programme. Dans l'exemple de la pompe, l'on pourrait comparer cinq ménages qui ont utilisé une pompe à cinq ménages qui ont puisé leur eau dans un étang. Toutefois, il est concevable que ces dix ménages ne soient pas représentatifs de l'ensemble des utilisateurs de pompes et de ceux qui puisent leur eau dans

un étang. Par exemple, une différence constatée dans l'incidence des diarrhées dans les deux groupes pourrait ne pas être pertinente si, par hasard, les individus des ménages du groupe « pompe » ont une santé plus fragile que ceux des ménages appartenant au groupe « étang ». La **taille nécessaire de l'échantillon** dépend de l'**impact attendu** du projet ou programme en question. Plus cet impact est faible, plus la taille de l'échantillon devra être importante pour que l'impact puisse être détecté. La taille de l'échantillon dépendra également de la **variance générale du outcome**. Plus cette variance est grande, plus la taille de l'échantillon devra être importante. A titre d'exemple, le revenu varie davantage d'un ménage à l'autre que la quantité d'eau consommée. Par conséquent, l'échantillon nécessaire pour analyser l'impact d'une pompe à eau sur le revenu des ménages devrait être plus grand que pour examiner l'impact de la pompe sur la quantité d'eau consommée.

Quel est le moment opportun pour une évaluation d'impact ?

Pour constituer un groupe de comparaison approprié, une évaluation d'impact devrait être préparée dès la conception d'un projet, d'un programme ou d'une politique, et donc avant sa mise en œuvre. Autrement, il pourrait être trop tard pour obtenir des données de qualité et mettre en place un groupe de comparaison approprié. Toutefois, la qualité d'une évaluation d'impact passe également par une bonne compréhension du programme. En règle générale, il est dès lors conseillé de planifier l'évaluation d'impact une fois que les aspects opérationnels ont été réglés, mais avant le déploiement du programme.

Les évaluations d'impact demandent beaucoup de temps et de ressources. En conséquence, il n'est pas recommandé d'évaluer l'impact de chaque projet, programme ou politique. Cet outil devrait être réservé aux cas où le potentiel d'apprentissage est important. Une réponse affirmative à au moins l'une des questions ci-dessous laisse supposer qu'un projet ou programme gagnerait à faire l'objet d'une évaluation d'impact¹ :

- Est-il envisageable de réaliser le projet ou programme à une échelle plus importante (scaling-up) ?
- Est-ce que le projet ou programme est stratégiquement pertinent pour l'organisation par rapport à des objectifs de réplication ou d'apprentissage ?
- Est-ce qu'il s'agit d'un projet ou programme novateur qui nécessiterait des évidences sur son efficacité ?
- Les résultats peuvent-ils être utilisés pour appuyer l'élaboration de politiques au niveau national ou global ?

De même, il n'est pas possible d'analyser l'impact de tous les types de projets et programmes de développement. En d'autres termes, l'impact ne peut **pas** être évalué pour **tous les projets et programmes d'une organisation. Trois aspects sont ici primordiaux** :

- Le nombre de bénéficiaires du projet ou du programme permet-il de constituer un échantillon d'une taille suffisante pour l'évaluation d'impact ?
- Est-il possible de constituer un groupe de comparaison approprié ?

¹ Voir aussi le site de la Banque mondiale sur les évaluations d'impact : <http://www.worldbank.org/en/research/dime> (en anglais)

→ Les *outcomes* sélectionnés pour l'analyse peuvent-ils être mesurés dans un délai raisonnable ?

Lectures conseillées :
International Initiative for Impact Evaluation : Synthèse des évaluations d'impact existantes ; en anglais
www.3ieimpact.org/en

J-PA L : Introduction aux évaluations d'impact aléatoires
www.povertyactionlab.org/fr/research-resources/introduction

Banque mondiale : Impact Evaluation in Practice, Second Edition (Les évaluations d'impact dans la pratique, deuxième édition ; en anglais)
openknowledge.worldbank.org/handle/10986/25030

Contact:

Direction du développement et de la coopération DDC
 Evaluation et Contrôle de Gestion
 3003 Berne
sektion.evaluation-controlling@eda.admin.ch

Cette publication peut être téléchargée sous www.ddc.admin.ch

Avril 2017



© DDC