



Glossaire de l'évaluation d'impact

Version n° 1

Date : décembre 2009

Pour toutes remarques ou suggestions concernant ce glossaire, écrire à :
3ie@3ieimpact.org.

Citation recommandée : *Glossaire de l'évaluation d'impact, 3ie (2009). Initiative internationale pour l'évaluation d'impact : New Delhi, Inde.*

Ampleur de l'effet

Mesure de la relation entre deux variables (particulièrement entre les variables et les résultats d'un programme). Voir aussi *Effet minimum détectable*.

Analyse coût-avantage

Comparaison de l'ensemble des coûts et avantages de l'*intervention*, auxquels une valeur monétaire est attribuée. L'intérêt de cette analyse par rapport à l'*analyse coût-efficacité* est qu'elle permet d'intégrer de multiples résultats et de réaliser une comparaison du retour sur investissement obtenu dans différents secteurs (et ainsi faciliter l'allocation efficiente des ressources de développement).

Analyse coût-efficacité

Analyse du coût permettant d'obtenir une variation d'une unité du *résultat*. L'intérêt de cette analyse par rapport à l'*analyse coût-avantage* est qu'elle évite la valorisation monétaire du résultat, souvent controversée. Elle peut être utilisée pour comparer l'efficacité relative des programmes pour obtenir le résultat visé.

Analyse de régression

Méthode statistique déterminant l'association entre la *variable dépendante* et une ou plusieurs *variables indépendantes*.

Appariement

Méthode de constitution de *groupes de contrôle* dans laquelle des groupes ou des individus sont appariés à ceux du *groupe de traitement* sur la base des caractéristiques jugées pertinentes pour le(s) *résultat(s)* de l'*intervention*.

Appariement sur les scores de propension (PSM)

Protocole quasi expérimental appliqué pour l'estimation de l'impact d'une *intervention*. Les résultats constatés pour le *groupe de traitement* sont comparés aux résultats pour un *groupe de contrôle*, ce dernier étant construit par *appariement* sur la base de scores de propension. Le score de propension est la probabilité de participation à l'*intervention*, donné par une régression probit sur les caractéristiques observées. Ces caractéristiques ne doivent pas être affectées par l'*intervention*. Le PSM permet donc d'apparier sur de multiples

caractéristiques en résumant celles-ci en un seul chiffre (le score de propension).

Assignation aléatoire

Protocole d'*intervention* dans lequel les membres de la *population éligible* à l'intervention sont assignés au hasard soit au groupe de traitement, soit au *groupe de contrôle (assignation aléatoire)*. C'est-à-dire que le fait pour une personne d'être placée dans le groupe de traitement ou dans le groupe de contrôle n'est qu'une question de hasard et ne dépend d'aucune autre de ses caractéristiques (observées ou non observées).

Attribution (imputation)

Mesure dans laquelle la variation du *résultat* observée est produite par l'*intervention*, compte tenu de tous les autres facteurs susceptibles de faire varier le résultat.

Attrition

Abandon des *participants* du *groupe de traitement* en cours d'*intervention* ou échec du recueil de données auprès d'une unité lors de cycles ultérieurs d'une *enquête par panel*. Ces deux formes d'attrition peuvent être à l'origine d'estimations d'impact *biaisées*.

Avant-après

Voir *Simple différence*.

Base de sondage

Liste complète de la *population* étudiée. Celle-ci n'est pas nécessairement la population entière du pays ou de la zone étudiée ; elle est limitée à la *population éligible*, par exemple familles avec des enfants de moins de cinq ans, ou ménages dont le chef de famille est une femme. Pour une *enquête sur les équipements*, la base de sondage serait constituée de tous les équipements dans la zone étudiée. En l'absence de base de sondage récente, il faut en construire une au moyen d'un inventaire sur le terrain.

Bénéficiaire

Individu, entreprise, équipement, village, etc. bénéficiant directement ou indirectement de l'*intervention*.

Biais

Écart entre la valeur estimée de l'*impact* et sa valeur réelle dû à des problèmes du protocole d'évaluation ou de la conception de l'*échantillon* (c'est-à-dire hors *erreur d'échantillonnage*).

Biais de sélection

Biais pouvant exister dans une étude par la sélection de personnes de différents types dans les groupes de traitement et de comparaison. De ce fait, les différences de résultats peuvent être expliquées par les différences préexistantes entre les groupes plutôt que par le traitement lui-même.

Calcul de puissance

Calcul de l'*échantillon* requis pour l'*évaluation d'impact*, qui dépend de l'*effet minimum détectable* et du *niveau de confiance* nécessaire.

Caractéristique non observable

Caractéristique qui ne peut être observée ni mesurée. La présence de caractéristiques non observables peut engendrer un *biais de sélection* dans les protocoles quasi expérimentaux.

Contamination

Fait, pour des membres du *groupe de contrôle*, d'être affectés par l'*intervention* (voir *Effet de diffusion*) ou par une autre *intervention* qui affecte également le *résultat* visé. La contamination est un problème courant car de multiples *interventions* de développement sont conduites dans la plupart des communautés.

Contrefactuel

Valeur du *résultat* pour le *groupe de traitement* en l'absence d'*intervention*.

Différence de différence Voir *Double différence*.

Données de panel

Données recueillies lors d'enquêtes successives dans le cadre desquelles les observations sont collectées sur l'*échantillon* de répondants à chaque cycle. Les données de panel peuvent subir un phénomène d'*attrition*, pouvant être à l'origine de *biais*.

Données de référence

Voir *Enquête de référence*.

Données primaires

Données recueillies expressément par le chercheur pour le projet de recherche.

Données secondaires

Données recueillies à d'autres fins, mais qui peuvent être ré-analysées dans une étude ultérieure.

Double différence

Différence entre la variation du *résultat* observée dans le *groupe de traitement* et la variation observée dans le *groupe de contrôle* ou, ce qui est équivalent, variation de la différence de résultat entre traitement et contrôle. La méthode de la *double différence* élimine le *biais de sélection* qui résulte de caractéristiques non observables invariantes dans le temps. Également appelée *Différence de différence*. Comparer avec *Simple différence* et *Triple différence*.

Échantillon

Sous-ensemble de la *population* étudiée. L'échantillon est constitué aléatoirement à partir de la *base de sondage*. Dans un échantillon aléatoire simple, tous les éléments de la base ont une probabilité égale d'être sélectionnés, mais on utilise généralement des plans d'*échantillonnage* plus complexes, qui requièrent l'utilisation de *poids d'échantillonnage*.

Échantillon par grappes

Échantillon constitué par une méthode d'*échantillonnage* en plusieurs étapes, dans laquelle on constitue d'abord un *échantillon* de zones géographiques (par exemple sous-districts ou villages), puis un *échantillon* de ménages, d'entreprises, d'équipements, etc. à partir des zones géographiques

sélectionnées. Cette méthode produit des erreurs types plus importantes que dans un *échantillon* aléatoire simple, mais elle est souvent utilisée pour des raisons de coûts.

Effet minimum détectable

Le plus petit effet que le chercheur estime nécessaire de détecter dans l'*évaluation d'impact*. Il intervient dans le *calcul de puissance* nécessaire pour déterminer la *taille de l'échantillon*.

Effet moyen du traitement

Valeur moyenne de l'*impact* sur le *groupe bénéficiaire* (ou *groupe de traitement*). Voir aussi *Effet de l'intention de traiter* et *Effet du traitement sur les traités*.

Effet de diffusion

Impact (positif ou négatif) de l'*intervention* sur des unités n'appartenant pas au *groupe de traitement*. La non-prise en compte des effets de diffusion biaise l'estimation d'*impact*. Si des effets de diffusion sont présents, le groupe de *bénéficiaires* est plus large que le groupe de *participants*. Lorsque les effets de diffusion affectent des membres du *groupe de contrôle*, on est en présence d'un cas particulier de *contamination*.

Enquête

Recueil d'informations utilisant (1) une stratégie d'*échantillonnage* préalablement définie et (2) un *instrument d'enquête*. Une enquête peut recueillir des données auprès d'individus, de ménages ou d'autres unités telles que des entreprises ou des établissements scolaires (voir *Enquête sur les équipements*).

Enquête sur les équipements

Enquête réalisée sur un *échantillon* d'équipements (habituellement de santé ou d'éducation, mais il peut aussi s'agir de commissariats, d'équipements de formation, etc.) qui vise à déterminer le niveau et la qualité de tous les éléments requis pour fournir des services. L'*unité d'observation* est l'équipement, bien que des données puissent être aussi recueillies sur le personnel des équipements dans une enquête distincte (par exemple une enquête sur les enseignants). Si une enquête sur les équipements est réalisée en parallèle à une enquête auprès des ménages, il est important, pour les besoins de l'analyse des données, que les *instruments d'enquête* comprennent des informations permettant de lier les ménages aux équipements qu'ils utilisent.

Erreur d'échantillonnage

Erreur se produisant lorsque les estimations résultent de données issues d'un *échantillon* et non de la *population* tout entière.

Essai contrôlé randomisé (ECR)

Protocole d'*évaluation d'impact* dans lequel les membres de la *population éligible* ont été affectés ou non à l'*intervention* par *assignation aléatoire*. Puisqu'il ne doit pas y avoir de corrélation entre les caractéristiques des *participants* et le *résultat*, toute différence de *résultat* entre le traitement et le contrôle peut être intégralement imputée à l'*intervention*, c'est-à-dire qu'il n'y a pas de *biais de sélection*. Cependant, les ECR peuvent être sujets à plusieurs types de *biais* et doivent donc suivre de stricts protocoles. Synonyme :

protocole expérimental, expérimentation contrôlée par assignation aléatoire.

Enquête de référence

Enquête visant à recueillir des données avant le début de l'*intervention*, dites données de référence. L'analyse en *double différence* requiert des données de référence, qui doivent être recueillies auprès du groupe de *traitement* et du groupe de *contrôle*.

Évaluation d'impact

Étude de l'*attribution* de variations du *résultat* à l'*intervention*. Les évaluations d'impact suivent soit un *protocole expérimental* soit un *protocole quasi expérimental*.

Évaluation prospective

Voir *Protocole d'évaluation ex ante*.

Facteurs confondants

Autres variables ou déterminants qui affectent le résultat.

Groupe de comparaison

Groupe d'individus présentant des caractéristiques similaires à celles du *groupe de traitement* (ou *participants*) mais qui ne reçoit pas l'*intervention*. Lorsque les conditions du protocole permettent à l'évaluateur de s'assurer qu'aucun *facteur confondant* n'affecte le groupe de comparaison, celui-ci est appelé *groupe de contrôle*.

Groupe de contrôle

Type particulier de *groupe de comparaison*, dont l'évaluateur peut contrôler l'environnement, ce qui lui permet de limiter les *facteurs confondants*.

Groupe de traitement

Groupe de personnes, d'entreprises, d'équipements, etc. qui reçoivent l'*intervention*. Synonyme : *participants*.

Hétérogénéité de l'impact

Variation de l'*impact* résultant de différences dans le contexte, les caractéristiques des *bénéficiaires* ou la mise en œuvre de l'*intervention*.

Hypothèse

Proposition relative aux relations entre deux variables. Dans une *évaluation d'impact*, l'hypothèse a généralement trait à l'*impact* attendu de l'*intervention* sur le *résultat*.

Impact

Effet de l'*intervention* sur le *résultat* pour la *population bénéficiaire*.

Instrument d'enquête

Formulaire (questionnaire) préétabli utilisé pour recueillir des données lors d'une enquête. Une enquête fait généralement appel à plusieurs *instruments d'enquête*, par exemple une enquête auprès des ménages et une *enquête sur les équipements*.

Test en simple/double ou triple aveugle

L'approche en simple aveugle consiste à dissimuler aux sujets leur appartenance au *groupe de traitement* ou au *groupe de comparaison*. Dans une approche en double aveugle, ni les sujets ni ceux qui conduisent l'étude ne savent qui sont dans les différents groupes, tandis que dans une étude en triple aveugle, on cache également à ceux qui analysent les données quel est le groupe de comparaison et quel est le groupe de traitement. En général, cette approche n'est pas praticable pour les *interventions* de développement socioéconomique, ce qui introduit des *biais*.

Intervention

Projet, programme ou politique qui fait l'objet de l'*évaluation d'impact*.

Méta-analyse

Analyse systématique d'un ensemble d'évaluations conduites sur des programmes similaires visant à tirer des conclusions générales, étayer les hypothèses ou produire une estimation des effets d'ensemble du programme.

Méthode de régression sur la discontinuité (RDD)

Protocole d'*évaluation d'impact* dans lequel les groupes de traitement et de contrôle sont proches de part et d'autre de la valeur seuil d'une variable. Cette variable peut être un score ou une caractéristique observée (par exemple, âge ou propriété foncière) utilisée par le personnel du programme pour déterminer la *population éligible*, ou ce peut être une variable trouvée pour distinguer les *participants* des *non-participants* lors de l'analyse des données. La RDD est un exemple de *protocole quasi expérimental*.

Méthode échelonnée (*pipeline approach*)

Protocole d'*évaluation d'impact* dans lequel le *groupe de contrôle* est composé de ceux qui ne participent pas encore à l'*intervention* mais vont en bénéficier. L'hypothèse est qu'il n'y aura pas de *biais de sélection*, car le *groupe de contrôle* et le *groupe de traitement* sont destinés à recevoir l'*intervention*. Il convient cependant de vérifier la qualité de l'*appariement*, car les *participants* à un stade ultérieur peuvent être différents des *participants* traités plus tôt.

Méthodes mixtes

Emploi de méthodes quantitatives et qualitatives dans un protocole d'*évaluation d'impact*. Parfois appelé Q au carré ou Q2.

Modèle logique

Modèle décrivant le principe de fonctionnement d'un programme en présentant la chaîne causale des moyens aux résultats, en passant par les activités et les produits. Si les modèles logiques présentent une théorie sur le résultat attendu du programme, ils ne démontrent pas si le programme est à l'origine du résultat observé. Une approche basée sur la théorie examine les hypothèses sous-jacentes aux liens dans le modèle logique.

N

Nombre de cas. La lettre « N » majuscule renvoie au nombre de cas dans la *population*, la lettre « n » minuscule au nombre de cas dans l'*échantillon*.

Niveau de confiance

Degré de certitude que la valeur réelle de *l'impact* (ou de toute autre estimation statistique) entrera dans une fourchette spécifiée.

Participant

Individu, équipement, entreprise, village, etc. recevant *l'intervention*.

Synonyme : *groupe de traitement*.

Poids d'échantillonnage

Technique employée pour garantir que les statistiques générées à partir de *l'échantillon* sont représentatives de la *population* sous-jacente à partir de laquelle *l'échantillon* est constitué. Il faut en principe utiliser des poids d'échantillonnage, même si ce qu'il convient de faire lorsqu'on utilise *l'appariement* des scores de propension, qui est un autre système de pondération, fait encore débat.

Population éligible

Population répondant aux critères pour être *bénéficiaire* de *l'intervention*. Elle peut être composée d'individus, d'équipements (écoles ou cliniques par exemple), d'entreprises, etc.

Protocole d'évaluation ex ante

Protocole d'évaluation d'impact préparé avant *l'intervention*. Les protocoles *ex ante* sont plus fiables que les protocoles *ex post* car ils permettent d'envisager une *assignation aléatoire* et de recueillir des *données de référence* auprès des *groupes de traitement* et de *contrôle*. Synonyme : *évaluation prospective*.

Protocole d'évaluation ex post

Protocole d'évaluation d'impact établi après le début de *l'intervention*, voire après sa conclusion. S'il n'y a pas eu *assignation aléatoire*, un *protocole quasi-expérimental* doit être employé.

Protocole expérimental

Voir *Essai contrôlé randomisé*.

Protocole factoriel

Essai contrôlé randomisé à groupes de traitement multiples, dans lequel un groupe reçoit le traitement A, un deuxième groupe le traitement B, et un troisième les deux traitements (A+B). Il peut également y avoir un *groupe de contrôle* ne recevant pas de traitement.

Protocole quasi expérimental

Protocole d'évaluation d'impact créant un *groupe de contrôle* par des procédures statistiques. L'intention est de garantir que les *groupes de traitement* et de *contrôle* présentent des caractéristiques identiques en tous points hormis *l'intervention*, comme ce serait le cas dans un *protocole expérimental*. Voir *Appariement des scores de propension* et *Méthode de discontinuité de la régression*.

Résultat(s)

Variable ou variables mesurant *l'impact* de *l'intervention*.

Effet de l'intention de traiter

Effet moyen du traitement calculé sur la totalité du *groupe de traitement*, indépendamment de la participation à l'*intervention*. Comparer avec *Effet du traitement sur les traités*.

Effet du traitement sur les traités

Impact (*effet moyen du traitement*) observé sur ceux qui ont effectivement reçu l'*intervention*. Comparer avec *Effet de l'intention de traiter*.

Revue systématique

Synthèse des résultats issus de la recherche sur un sujet particulier tel que l'efficacité de l'approvisionnement en eau et de l'assainissement, établie au moyen d'une recherche exhaustive de toutes les études pertinentes en utilisant des stratégies scientifiques pour réduire les erreurs associées à l'évaluation du protocole et des résultats des études. Une revue systématique est plus méticuleuse qu'une revue de littérature, mais elle n'emploie pas les techniques statistiques d'une *méta-analyse*.

Simple différence

Comparaison entre le *résultat* pour le *groupe de contrôle* après l'*intervention* et sa valeur de référence (également appelée *Avant-après*) ou comparaison ex post entre les *résultats* pour les groupes de traitement et de contrôle. Comparer avec *Double différence*.

Triple différence

Impact comparatif ou différentiel sur deux groupes, calculé comme la différence entre l'estimation de l'impact en *double différence* pour chaque groupe par rapport à un *groupe de comparaison* ne recevant pas de traitement. Une estimation en triple différence significative démontre la présence d'une hétérogénéité d'impact.

Unité d'analyse

Classe d'unités élémentaires constituant la *population* et les unités sélectionnées pour la mesure ; également, classe des unités élémentaires à laquelle les mesures sont généralisées.

Validité externe

Mesure dans laquelle les résultats de l'*évaluation d'impact* s'appliquent à un autre moment ou à un autre lieu.

Validité interne

Validité du protocole d'évaluation du point de vue de la gestion des questions telles que le *biais de sélection*, l'*échantillonnage*, les *effets de diffusion*, la *contagion* et l'*hétérogénéité de l'impact*.

Variable dépendante

Variable dont la valeur est prédite ou causée par une ou plusieurs autres variables (*variables indépendantes*). Ce terme est couramment employé en analyse de *régression*.

Variable dichotomique

Variable qui ne peut prendre que deux valeurs, par exemple le « sexe » (masculin = 0, féminin = 1). La *variable dépendante* dans l'équation *probit* de la

participation estimée pour l'*appariement des scores de propension* est une variable dichotomique pour laquelle « a participé »=1, « n'a pas participé » = 0.

Variable indépendante

Variable produisant des modifications de la *variable dépendante*, habituellement appliquée en *analyse de régression*.

Variable muette

Variable dichotomique couramment utilisée en *analyse de régression*. L'*évaluation d'impact* fait souvent appel à une variable muette de participation au programme (« a participé »=1, « n'a pas participé »=0) en tant que *variable indépendante* d'une régression dans laquelle la *variable dépendante* est le *résultat*.