



HAL
open science

L'évaluation basée sur la théorie

Agathe Devaux-Spatarakis

► **To cite this version:**

Agathe Devaux-Spatarakis. L'évaluation basée sur la théorie. LIEPP Fiche méthodologique n°27, 2023, 5 p. hal-04159070

HAL Id: hal-04159070

<https://sciencespo.hal.science/hal-04159070>

Submitted on 11 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

LIEPP FICHE MÉTHODOLOGIQUE n°27

2023

Cette fiche fait partie d'un ensemble de fiches méthodologiques publiées par le LIEPP. A ce titre, elle bénéficie du soutien apporté par l'ANR et l'État au titre du programme d'Investissements d'avenir dans le cadre de l'IdEx Université Paris Cité (ANR-18-IDEX-0001).

L'évaluation basée sur la théorie

Agathe DEVAUX-SPATARAKIS (Quadrant-Conseil)

adevaux@quadrant-conseil.fr



Partage selon les Conditions Initiales 4.0 International License

www.sciencespo.fr/liepp

Comment citer cette publication :

DEVAUX-SPATARAKIS, Agathe, **L'évaluation basée sur la théorie**,
LIEPP Fiche méthodologique n°27, 2023-07

This publication is also available in English:

DEVAUX-SPATARAKIS, Agathe, **Theory-based Evaluation**,
LIEPP Methods Brief n°28, 2023-07

EN BREF :

L'évaluation basée sur la théorie s'est développée en réponse aux limites des approches expérimentales et quasi-expérimentales, qui ne permettent pas de saisir les mécanismes par lesquels une intervention produit ses impacts. Cette approche consiste à ouvrir la « boîte noire » de l'action publique en décomposant les différentes étapes de la chaîne causale liant l'intervention à ses résultats finaux. Les hypothèses ainsi formulées sur les mécanismes en jeu peuvent ensuite être testées empiriquement.

Mots-clés : Méthodes qualitatives, théorie, chaîne causale, logique d'intervention, diagramme logique d'impact, logigramme, théorie du changement, chemin d'impact

I. En quoi consiste cette approche ?

Formalisée dans les années 1970 et 1980 aux États-Unis, l'évaluation basée sur la théorie (EBT) n'est pas une méthode, mais plutôt une logique de recherche évaluative, une démarche analytique susceptible de mobiliser différentes méthodes ou outils d'évaluation. Le terme « théorie » est ici à comprendre comme la décomposition et l'explicitation d'une chaîne causale liant l'intervention publique et les résultats et impacts escomptés sur les publics ciblés et bénéficiaires. Celle-ci prend alors la forme d'une représentation graphique qui devient alors « l'échafaudage » sous-tendant les investigations de l'évaluation. Selon les usages on peut nommer cet outil « logique d'intervention », « diagramme logique d'impact (DLI) », « logigramme », « théorie du changement » ou « chemin d'impact ». Cette démarche est largement répandue dans la pratique évaluative, cependant elle recouvre une diversité de pratiques de qualité parfois variable et mobilisant des approches plus ou moins techniques.

Une approche basée sur la théorie, contrairement à ce que pourrait évoquer cette appellation, va chercher à être au plus près de la réalité du terrain et du contexte de l'intervention. Le terme théorie prend ici un sens plus ou moins scientifique selon les usages. A minima, une approche basée sur la théorie s'intéresse à la théorie de mise en œuvre d'une intervention, c'est-à-dire aux actions que l'on met en place pour la bonne réalisation de l'intervention permettant d'atteindre les premiers résultats (logique d'intervention). Dans la plupart des cas, elle tend à faire un pas de plus et à expliciter les liens causaux hypothétiques entre des mécanismes délivrés par une intervention et leurs résultats escomptés allant jusqu'aux impacts finaux (théorie du changement). Cette analyse est généralement structurée en plusieurs catégories :

- Les réalisations, qui décrivent quelles actions sont mises en place par les autorités publiques,
- Les résultats, qui décrivent les premiers effets immédiats, directement liés aux réalisations et pouvant être observés sur les publics participants aux actions,
- Les impacts intermédiaires et finaux qui décrivent les impacts que devraient avoir les réalisations sur les bénéficiaires finaux dont on cherche à améliorer la situation.

La conduite de cette démarche débute systématiquement par un travail d'élaboration de la « théorie ». Celle-ci s'appuie généralement sur une analyse documentaire ainsi que des discussions avec les parties prenantes afin d'identifier les composantes clés de l'intervention, les publics ciblés, les principaux résultats escomptés ainsi que les impacts finaux ciblés par l'intervention. La théorie est ainsi souvent co-construite entre les équipes d'évaluation et les parties prenantes et présente graphiquement les principales étapes allant des réalisations aux résultats et impacts finaux en les connectant entre eux et en explicitant les hypothèses pour passer d'une étape à une autre. Ces hypothèses peuvent couvrir les conditions de succès, les risques liés au contexte, mais aussi dans le cadre d'une évaluation réaliste (cf

fiche séparée) les mécanismes psycho-sociaux à l'œuvre ou dans le cadre d'une analyse de contribution, préciser les hypothèses de contribution de l'intervention. À ce stade, le travail peut être consolidé par une analyse de la littérature et donc une mobilisation de théories cette fois académiques pour étayer certains liens logiques (par exemple dans une intervention ciblant le retour à l'emploi, informer une relation causale entre un niveau de formation et emploi).

Ensuite, cette théorie devient l'échafaudage de l'évaluation. Deux types d'investigations sont alors conduites :

- D'abord une analyse du changement : est-ce que les étapes identifiées ou souhaitées lors de la conception de la théorie se vérifient sur le terrain ? Pour qui et dans quels contextes ?
- Puis une analyse causale : Comment peut-on expliquer le passage ou non d'une étape à une autre ? Dans quelle mesure l'intervention contribue-t-elle à générer ces résultats ? Dans quelles conditions ? Quels mécanismes psychosociologiques peuvent l'expliquer ?

Les questions d'évaluation sont donc structurées autour de vérification de la théorie et de son explication. Dans cette tâche, l'équipe d'évaluation peut mobiliser une diversité de méthodes ou d'outils d'évaluation. Par exemple, on peut avoir recours à des méthodes d'estimation quantitatives des résultats ou des impacts et les articuler avec des méthodes qualitatives pour approfondir l'analyse causale.

In fine, les résultats de cette démarche prennent la forme d'une nouvelle théorie « vérifiée » de l'intervention montrant les processus causaux effectivement à l'œuvre, ceux ne fonctionnant pas comme prévu, et apportant des explications sur ces phénomènes. Ce type de résultat permet alors d'identifier à quelle étape, dans quel contexte et auprès de quels publics l'intervention a rencontré des difficultés et de proposer des pistes d'améliorations opérationnelles et stratégiques pour les prochaines programmations.

II. En quoi cette approche est-elle utile pour l'évaluation des politiques publiques ?

Mobiliser une théorie du changement dans une évaluation peut servir différents usages. Décomposer les étapes et les liens de cause à effet permet lors des investigations d'identifier quelles sont les étapes fonctionnelles de la théorie et celles plus problématiques. Une intervention publique n'est jamais complètement un échec ou un succès. Par exemple, si l'on constate que l'intervention n'a pas atteint les résultats souhaités, si l'on renseigne les différentes étapes de la théorie du changement lors de l'évaluation, on peut identifier si les difficultés sont liées à des problèmes liés à la mise en œuvre (des actions qui n'ont pas été correctement organisées ou une difficulté à mobiliser les publics ciblés) ou la capacité de l'intervention à avoir les effets escomptés (les actions ont bien été organisées et ont atteint les publics ciblés, mais ne parviennent pas à générer les effets escomptés sur tous les publics ou certaines catégories). Il est alors possible d'approfondir l'analyse de la causalité pour tenter de comprendre si ces difficultés sont liées au programme lui-même ou à des éléments de contextes extérieurs qui n'ont pas suffisamment été pris en compte lors de la conception ou la mise en œuvre du programme (concurrence avec d'autres dispositifs, contraintes freinant la participation des publics, mauvaise analyse des besoins, ou changement d'une conjoncture économique par exemple).

Cette démarche permet de répondre aux questions de type : « Dans quels cas l'intervention X a-t-elle contribué à générer les résultats Y et pourquoi ? ». Plus globalement elle peut être utilisée pour répondre à l'ensemble des registres de questionnements évaluatifs sur l'efficacité, l'impact, la pertinence, la cohérence interne ou externe ainsi que l'efficience. En effet, chacun de ces registres peut être investigué pour expliquer le succès ou l'échec dans la progression d'une étape à une autre de la théorie.

Elle peut être utilisée lors de la conception d'une intervention pour notamment réunir les parties prenantes, coconstruire cette théorie et en anticiper les risques en amont de la mise en œuvre et modifier la programmation pour renforcer les chances de succès de l'intervention. Dans ce cas, elle est d'ailleurs utile à l'appropriation d'une vision commune et partagée de l'intervention par les parties prenantes.

Elle est aussi particulièrement pertinente dans les évaluations chemin faisant ou *in itinere* afin de vérifier au cours de la mise en œuvre l'atteinte des différentes étapes et d'ajuster la mise en œuvre du projet pour favoriser l'atteinte des résultats. Enfin, dans le cas d'une évaluation *ex post*, en fin de projet elle permet de comparer la théorie initiale planifiée avec la théorie vérifiée dans le cadre de l'évaluation et de comprendre les raisons de ces écarts.

III. Un exemple de l'utilisation de cette approche : évaluation d'un programme de nutrition

Les approches d'évaluation basées sur la théorie sont mobilisées dans une diversité de contextes et sont autant adaptées aux projets simples qu'aux politiques publiques complexes. Un exemple de son utilisation dans le cadre d'un projet simple peut illustrer brièvement la valeur ajoutée de cette approche.

Dans un document pour l'International Initiative for Impact Evaluation, Howard White (2009) présente le cas de l'évaluation du projet de nutrition intégrée au Bangladesh (BINP). Ce projet repère les enfants en retard de croissance et propose une distribution de suppléments alimentaires ainsi que des conseils de nutrition aux mères de ces enfants. L'objectif final est une augmentation de la croissance des enfants.

Une première évaluation par groupe de comparaison par appariement de scores de propension n'a pas constaté d'impact de ce projet sur l'état nutritionnel des enfants, mais un impact positif uniquement sur les enfants les plus malnutris. Néanmoins, ce résultat en tant que tel ne permet pas d'expliquer ce résultat, de disposer d'enseignements sur les dysfonctionnements du projet ni sur les mesures pouvant être prises pour l'améliorer.

Une évaluation complémentaire basée sur la théorie a permis d'enrichir ces résultats et de proposer des orientations pour l'action. Cette démarche s'est d'abord intéressée à reconstituer la théorie du projet puis à expliciter et investiguer les hypothèses causales sous-tendant cette action.

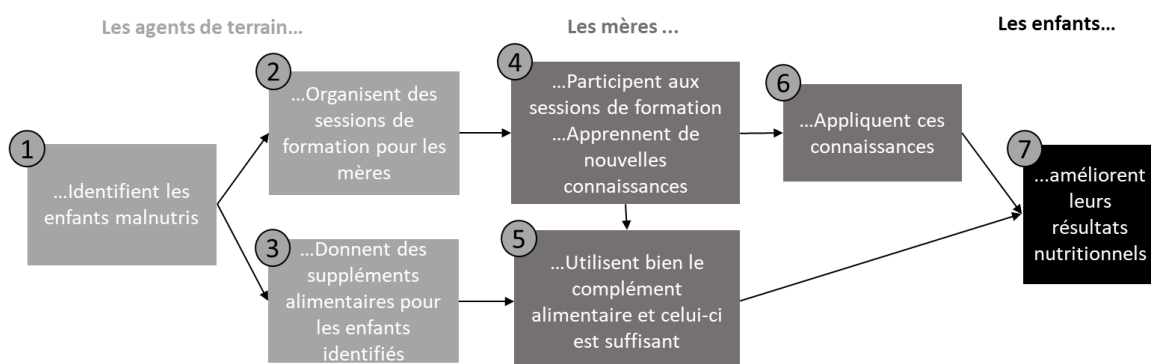


Figure 1 : Théorie du projet de nutrition intégrée (adaptation par l'auteur de cette fiche)

Une première étape est d'abord la bonne identification des enfants malnutris (1). Cette étape s'appuie sur l'hypothèse que les parents amènent bien leurs enfants aux centres de prévention et ensuite que les enfants malnutris soient bien repérés. Ici il semble que le programme parvenait à atteindre le public ciblé

(90% des femmes admissibles ont amené leurs enfants), néanmoins la sélection des enfants par les agents communautaires a montré plusieurs erreurs de type 1 (des enfants malnutris non sélectionnés) et de type 2 (des enfants sélectionnés alors qu'ils ne sont pas malnutris).

La chaîne causale 2-4-6-7 s'appuie sur l'hypothèse que les bons publics cibles ont été identifiés et que les modes d'action sont pertinents. Or, une étude anthropologique et des focus groupes ont ici révélé que les mères ciblées par ce dispositif n'avaient finalement que peu d'influence sur la nutrition de leur enfant, car les hommes avaient la responsabilité des courses alimentaires et pour la plupart les belles-mères étaient chargées de la préparation des repas. Ainsi, bien qu'elles participent aux formations, et apprennent de nouvelles connaissances les mères n'étaient en position de les appliquer que lorsqu'elles seraient belles-mères à leur tour.

La chaîne causale 3-5-7 s'appuie sur l'hypothèse que les enfants malnutris ont effectivement été sélectionnés, et que le complément alimentaire est bien utilisé en complément et non en substitution ou partagé entre les enfants et que celui-ci est bien suffisant pour améliorer les résultats nutritionnels des effets. Or il s'avère que les enquêtes menées ont montré que ces deux hypothèses n'étaient pas vérifiées dans de nombreux cas.

Cette évaluation a permis d'identifier des recommandations très opérationnelles pour améliorer les effets du projet notamment : La participation des belles-mères et des maris aux séances de conseils en nutrition ainsi qu'une meilleure formation des agents communautaires en charge de la sélection des enfants et un ciblage sur les enfants les plus malnutris.

IV. Quels sont les critères permettant de juger de la qualité de la mobilisation de cette approche ?

L'analyse de la qualité de la mobilisation de cette approche peut s'effectuer à deux moments de l'évaluation.

D'abord, la phase d'élaboration de la théorie doit aboutir à la théorie la plus plausible avant l'investigation de terrain. Dans l'idéal une bonne théorie est co-construite entre l'équipe d'évaluation et les différentes parties prenantes du dispositif et mobilise aussi des connaissances extérieures pouvant éclairer la plausibilité des hypothèses de liens de causalité entre les différentes étapes de cette théorie. Être plausible signifie qu'elle rend compte d'une ambition raisonnable de l'intervention dans la mesure des changements qu'elle est susceptible de générer. Enfin, une bonne théorie reste synthétique, simple à comprendre et indique clairement par une représentation graphique qui agit et qui est susceptible de changer de comportement ou de voir sa situation s'améliorer suite à la contribution de l'intervention étudiée.

Ensuite, les critères de qualité de la phase de test de la théorie s'inscrivent dans les standards généraux de la collecte et de l'analyse de données. Les outils de collecte doivent être diversifiés pour rendre compte d'une pluralité de sources de données (quantitatives et qualitatives) et de points de vue afin de consolider au maximum les résultats. La collecte et l'analyse sont idéalement conduites de manière itérative (par phases successives) afin d'affiner la théorie du changement et ses hypothèses tout au long de l'évaluation.

Les résultats de l'évaluation doivent ensuite être présentés autour de la vérification de la théorie et présenter ce qui a marché ou non pour quels types de publics pourquoi et dans quelles conditions.

V. Quels sont les atouts et les limites de cette méthode par rapport à d'autres ?

Le principal atout de cette approche est « d'ouvrir la boîte noire » de l'action publique. À la différence des approches expérimentales ou quasi-expérimentales qui permettent d'estimer des impacts sans analyser leur processus de production, l'évaluation basée sur la théorie permet d'identifier et de décomposer les mécanismes conduisant à la production de l'impact. Elle applique une démarche analytique qui décompose l'intervention en plusieurs étapes et en liens de causalité et qui permet de distinguer ce qui relève du contexte ou de l'intervention. Expliciter les hypothèses sous-jacentes à l'action oblige l'équipe d'évaluation à s'intéresser aux contextes de mise en œuvre, aux caractéristiques des publics ainsi qu'aux « paris » de l'intervention c'est-à-dire la manière dont elle est susceptible d'être un levier de changement. Les résultats d'évaluation sont donc particulièrement contextualisés et génèrent des enseignements même lorsque l'intervention n'est pas finalisée puisque chaque étape de la mise en œuvre peut être investiguée.

Cependant, cette approche peut être déstabilisante, car contrairement à d'autres méthodes, il n'y a pas une manière et un protocole distinct à suivre pour la mettre en œuvre. Ainsi, plusieurs praticiens ou praticiennes de l'évaluation peuvent revendiquer s'inscrire dans cette démarche sans pour autant correspondre aux standards de qualité de cet exercice.

Par ailleurs, on a pu reprocher à cette méthode d'accorder trop d'attention à l'analyse de la théorie escomptée et d'occulter l'identification d'explications alternatives des processus observés ou de résultats non anticipés. C'est en effet un écueil auquel l'équipe d'évaluation doit être attentive en tentant à la fois lors de la conception de la théorie et de sa vérification d'être aussi ouvert à la formulation et l'identification de ces explications alternatives. La mobilisation de la méthode d'analyse de contribution (voir la fiche séparée) permet notamment de concentrer l'analyse sur la place de l'intervention dans un paquet causal plus complexe et d'investiguer les explications alternatives.

Quelques références bibliographiques pour aller plus loin

Birckmayer, Johanna D., et Weiss, Carol H. 2000. « Theory-Based Evaluation in Practice, What Do We Learn? » *Evaluation Review* 24 (4) : 407-431.

Funnell, Sue., et Rogers, Patricia. 2011. *Purposeful Program Theory, Effective use of Theories of Change and Logic Models*. Jossey-Bass, Chichester

Rogers, Patricia., et Petrosino, Anthony. et al. 2000. « Program Theory Evaluation: Practice, Promise, and Problems ». *New directions for evaluation* 87, 5-13.

Van Es, Marjan, et Guijt, Irene, et Vogel, Isabel. 2015. *Theory of change thinking in practice : A stepwise approach, Hivos ToC Guidelines*, Hivos people unlimited.

Weiss, Carol. 1997. « Theory-based Evaluation: Past, Present, and Future ». *New directions for evaluation* 76, 41-55.

White, Howard. 2009. « L'évaluation d'impact basée sur la théorie : principes et pratiques », *Working paper* 3, International initiative for Impact Evaluation.