

Atelier de formation virtuel sur les Évaluations d'Impact Contrefactuelles (EIC)

C4ED – EUTF
Septembre 2021

Bienvenue à l'atelier de formation sur les Évaluations d'Impact Contrefactuelles (EIC)

Le contenu de cet atelier a été développé avec le soutien financier de l'Union Européenne. Ce contenu est la seule responsabilité de C4ED et ne reflète pas nécessairement les opinions de l'Union Européenne.

Bienvenue UE

Nathalie Vernhes

Bienvenue C4ED

Dr. Markus Olapade

- Nous sommes un centre de recherche indépendant à but non lucratif qui a l'objectif combler le fossé entre recherche et pratique dans le domaine du développement à l'échelle mondiale
- Nous conduisons des évaluations d'impact et nous testons des solutions innovantes.
- De plus, nous développons les compétences, conseillons et soutenons nos partenaires, et disséminons les preuves scientifiques de ce qui marche, pour qui, quand, pourquoi et dans quelles circonstances.

Protection des données

Inès Balança



Communication durant l'atelier

- Posez vos questions dans la zone de discussion
- Réagissez (« like ») aux questions posées par d'autres, pour nous indiquer qu'elles sont particulièrement pertinentes ou urgentes
- Nous répondrons aux questions au fur et à mesure
- Profitez des longues pauses pour poser plus de questions
- Suggérez des améliorations s'il vous est difficile de suivre ou si vous n'êtes pas d'accord (nous sommes ouverts aux critiques constructives et aux suggestions pour nous améliorer)
- Pour plus de questions et de feedback (en particulier pour la session Q&R):

Envoyez un email à Marc Gillaizeau
(m.gillaizeau@c4ed.org)

Jour 1 - Agenda

09:00 – 09:30	Bienvenue
9:30 – 10:40	Session 1: Introduction aux Evaluations d'Impact Contrefactuelles (EIC)
10:40 – 11:00	Pause
11:00 – 12:00	Session 2: Démêler les besoins du programme, Théorie du Changement, questions d'évaluation, et indicateurs
12:00 – 12:15	Session 3: Introduction aux Méthodes Expérimentales (Introduction)
12:15 – 13:00	Pause déjeuner
13:00 – 13:45	Session 3: Introduction aux Méthodes Expérimentales (suite) Conception de l'évaluation, Méthodes d'EIC
13:45 – 13:55	Pause
13:55 – 14:30	Session 3: Conception de l'évaluation, Méthodes d'EIC (suite et fin)
14:30 – 15:00	Q&R

Sommaire

- Qu'est-ce qu'une Evaluation d'Impact **rigoureuse** et quels en sont les bénéfices?
- Démêler les besoins du programme, Théorie du Changement, questions d'évaluation, et indicateurs
- Conception de l'évaluation, Méthodes d'EI(C)
 - Introduction aux Méthodes Expérimentales
 - Introduction aux Méthodes Quasi-Expérimentales
- Gérer les attentes – données (sources, indicateurs, et taille de l'échantillon) et échéances

Session 1: Introduction aux Evaluations d'Impact Contrefactuelle (EIC)

C4ED – EUTF
Septembre 2021

- Identification des objectifs qu'une EI peut atteindre
- Identification des questions d'évaluation appropriée pour les évaluations d'impact

Evaluation d'Impact (EI)

Qu'est-ce que l'évaluation d'impact?

- L'évaluation d'une intervention, d'une politique et/ou d'un projet
- La différence entre les résultats avec et sans le projet

Pourquoi évaluer l'impact?

- Pour déterminer si une intervention génère un changement dans les résultats qui est causal, attribuable, de quelle magnitude et comment (mécanisme causal)
- Pour comprendre quelle stratégie d'intervention marche le mieux
- Pour soutenir les décideurs politiques et informer leurs décisions sur la base de preuves rigoureuses

Comment mesurer l'impact

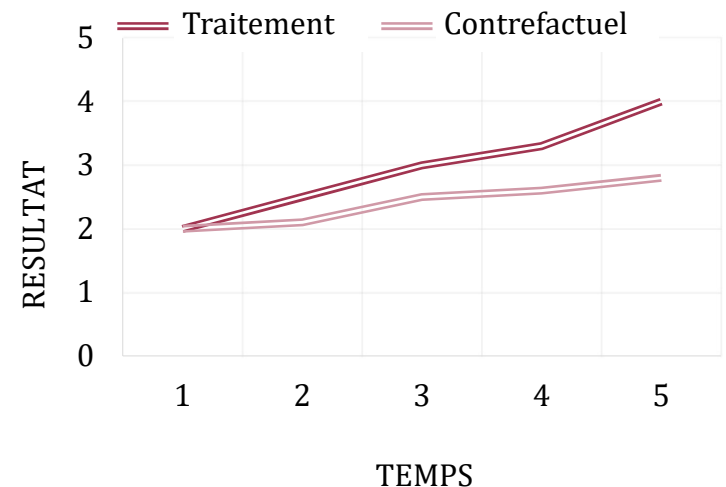
La différence entre les résultats avec et sans le projet

Contrefactuel: pour les bénéficiaires de l'intervention, c'est le résultat qui aurait été observé si ces mêmes bénéficiaires n'avaient pas bénéficié du programme

Problème fondamental: il est impossible de mesurer ou d'observer le contrefactuel

Solution: "imiter" le contrefactuel
→ créer un groupe de contrôle/comparaison

L'IMPACT D'UN PROGRAMME



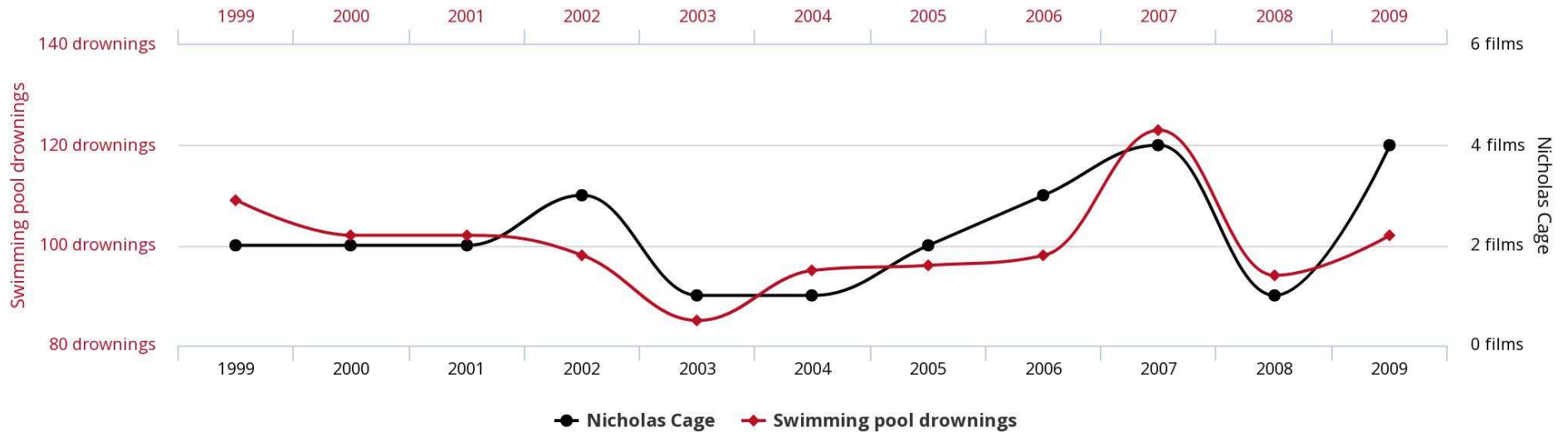
Pourquoi faire des Evaluations d'Impact Contrefactuelles?

Corrélations (parfois fallacieuses)

Le nombre de personnes qui se sont noyées en tombant dans une piscine

est corrélé avec

Le nombre de films dans lesquels Nicolas Cage apparaît



tylervigen.com

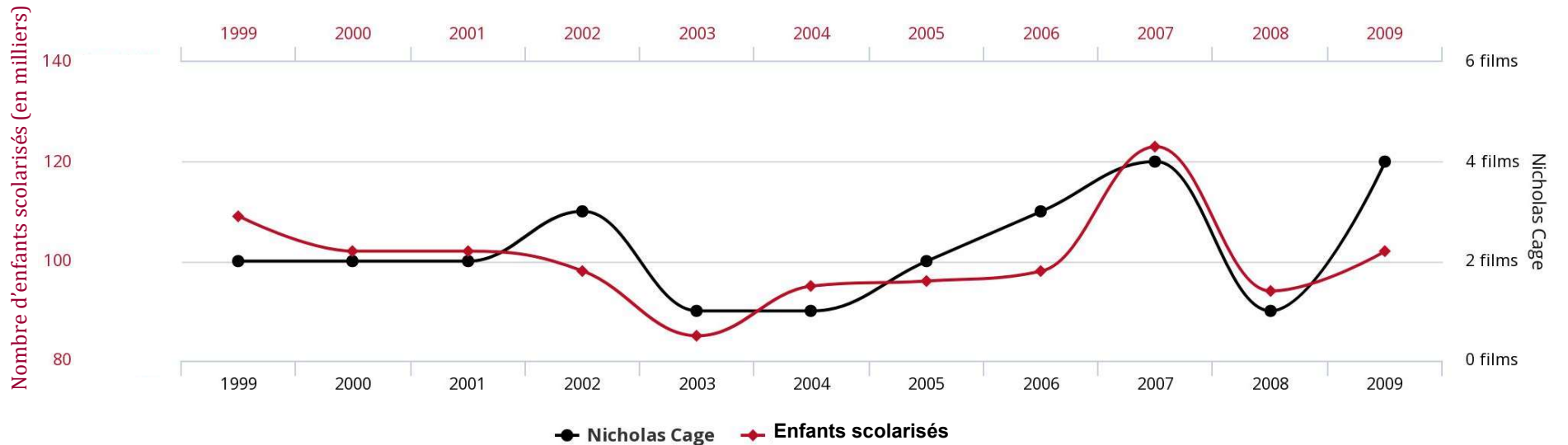
Source: Spurious Correlations at <http://tylervigen.com/spurious-correlations>

Corrélations (parfois fallacieuses)

Le nombre d'enfants allant à l'école dans le pays X

est corrélé avec

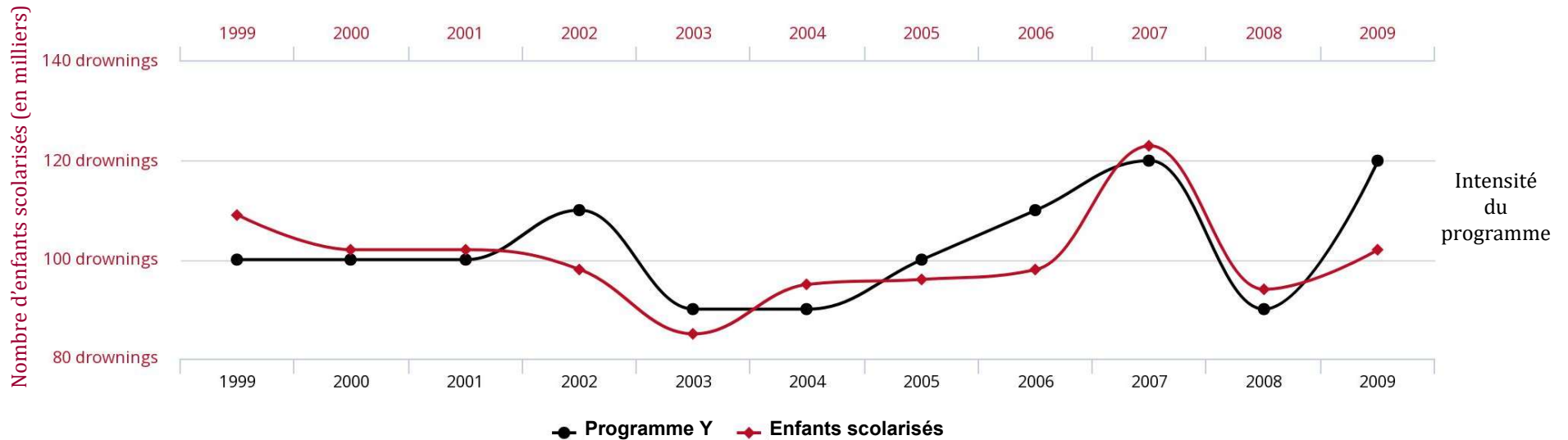
Le nombre de films dans lesquels Nicolas Cage apparaît



tylervigen.com

Idéalement un lien causal

Le nombre d'enfants allant à l'école dans le pays X est corrélé avec Le programme Y

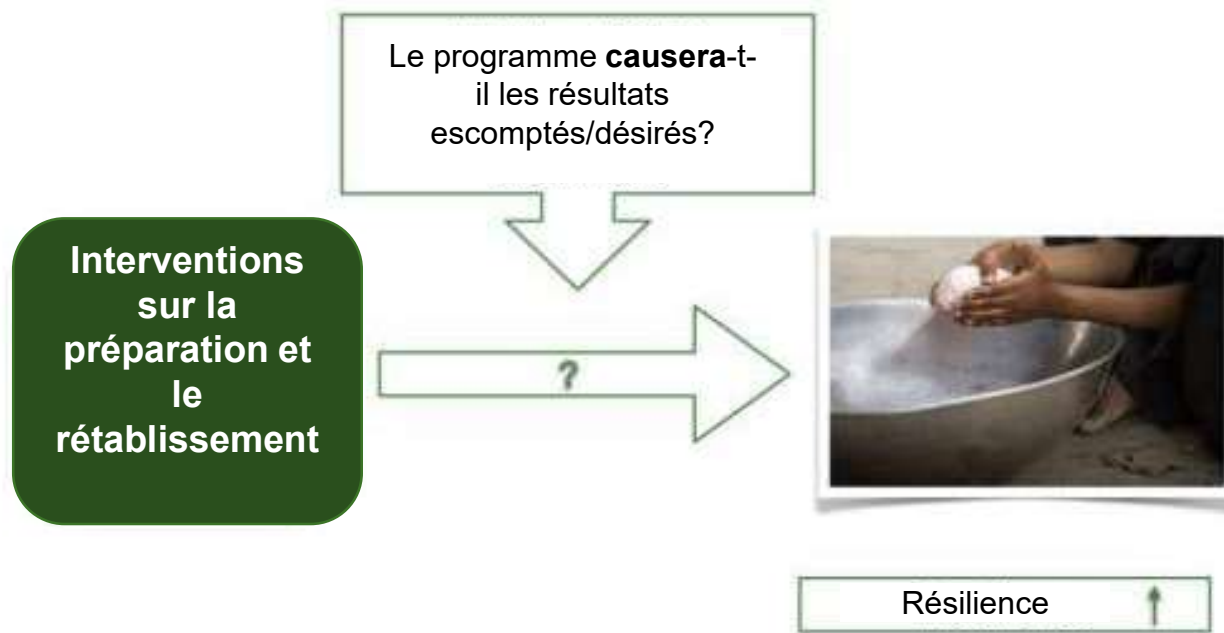


Question centrale

Quel est l'**impact ou l'effect causal** d'une politique/d'un programme/d'une certaine approche de mise en place sur la variable de résultat d'intérêt?

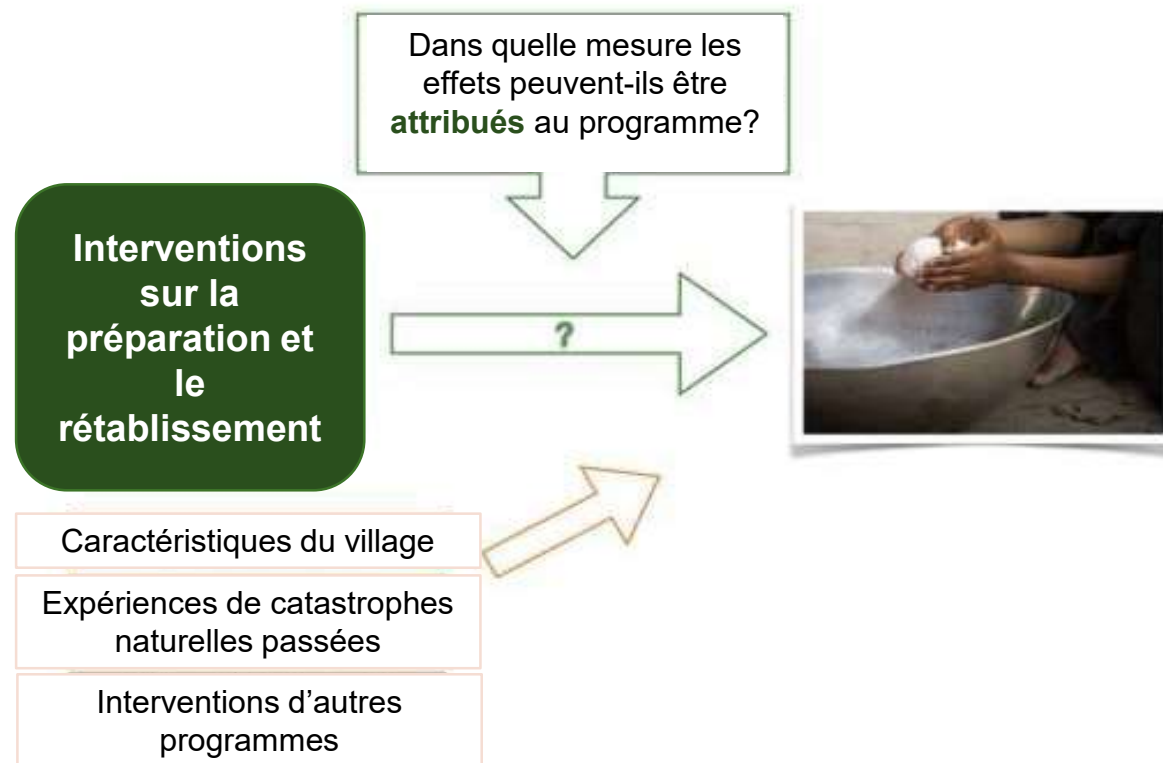
POURQUOI?

Mesurer les effets désirés



POURQUOI?

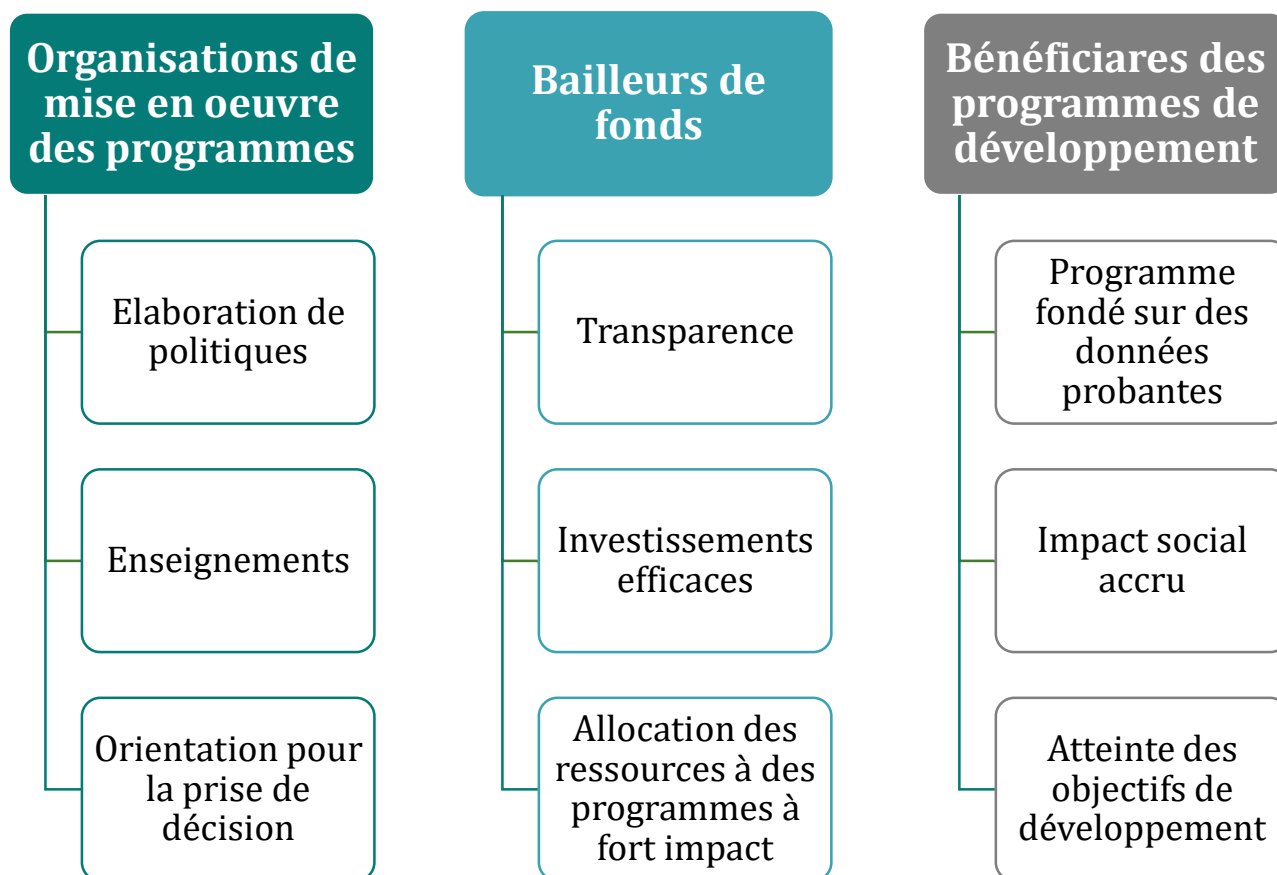
Mesurer les effets désirés



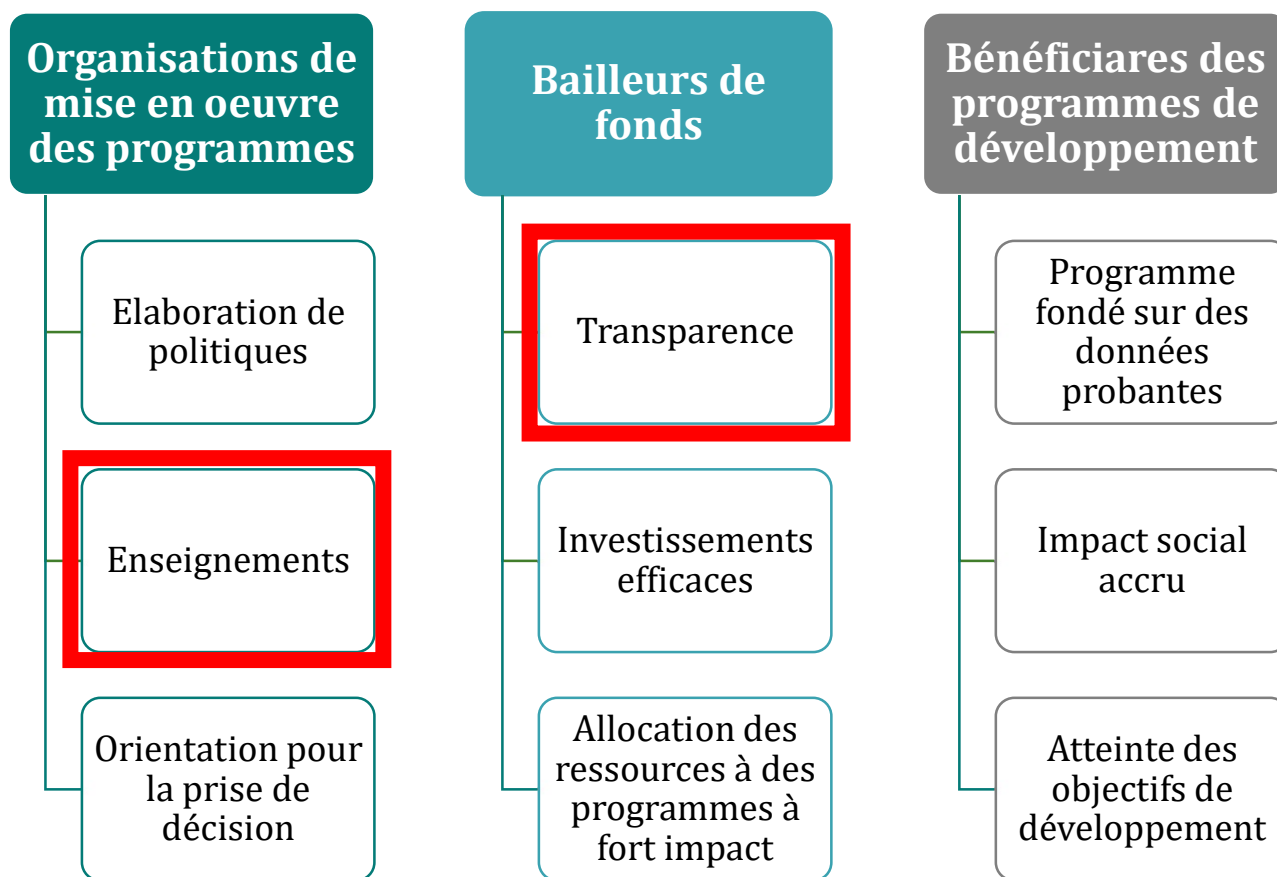
Les objectifs d'une évaluation d'impact contrefactuelle

- Une plus grande **transparence**
- **Apprentissage**

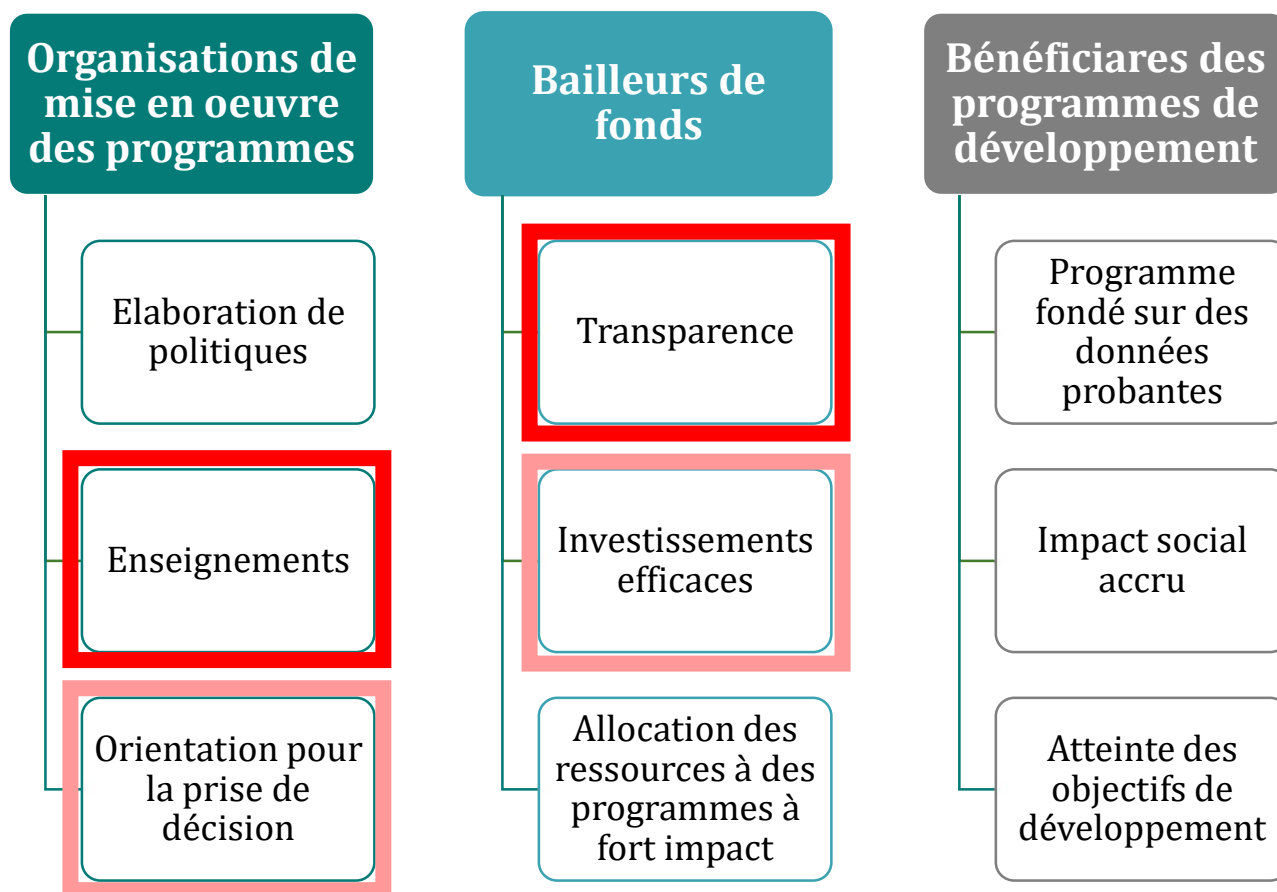
Les objectifs d'une évaluation d'impact contrefactuelle



Les objectifs d'une évaluation d'impact contrefactuelle

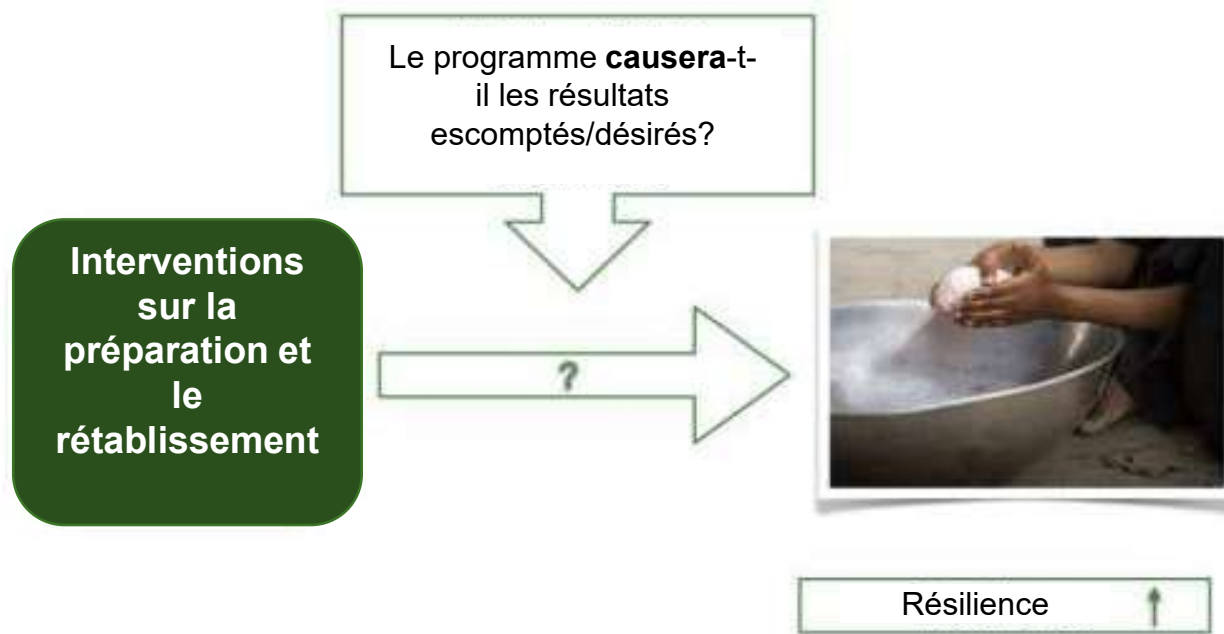


Les objectifs d'une évaluation d'impact contrefactuelle

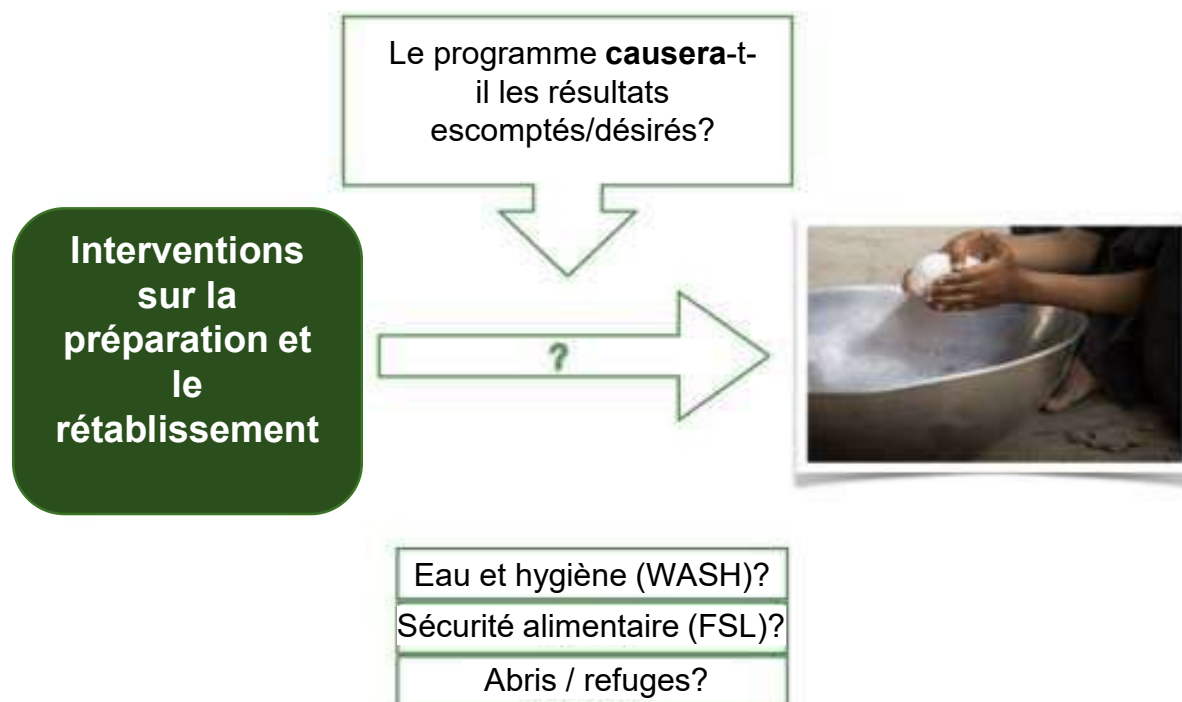


En pratique, quelle est la différence entre les questions d'évaluation portant sur la transparence et celles portant sur les enseignements dans la phase de mise en œuvre d'une EIC?

Transparence : est-ce que le programme marche ?



Enseignements: comment le programme marche-t-il le mieux?



Pour un meilleur apprentissage : avant de vous lancer dans une nouvelle EIC... vérifiez les preuves existantes

Les avantages de la synthèse des preuves existantes

Lorsque **suffisamment** de preuves scientifiques existent:

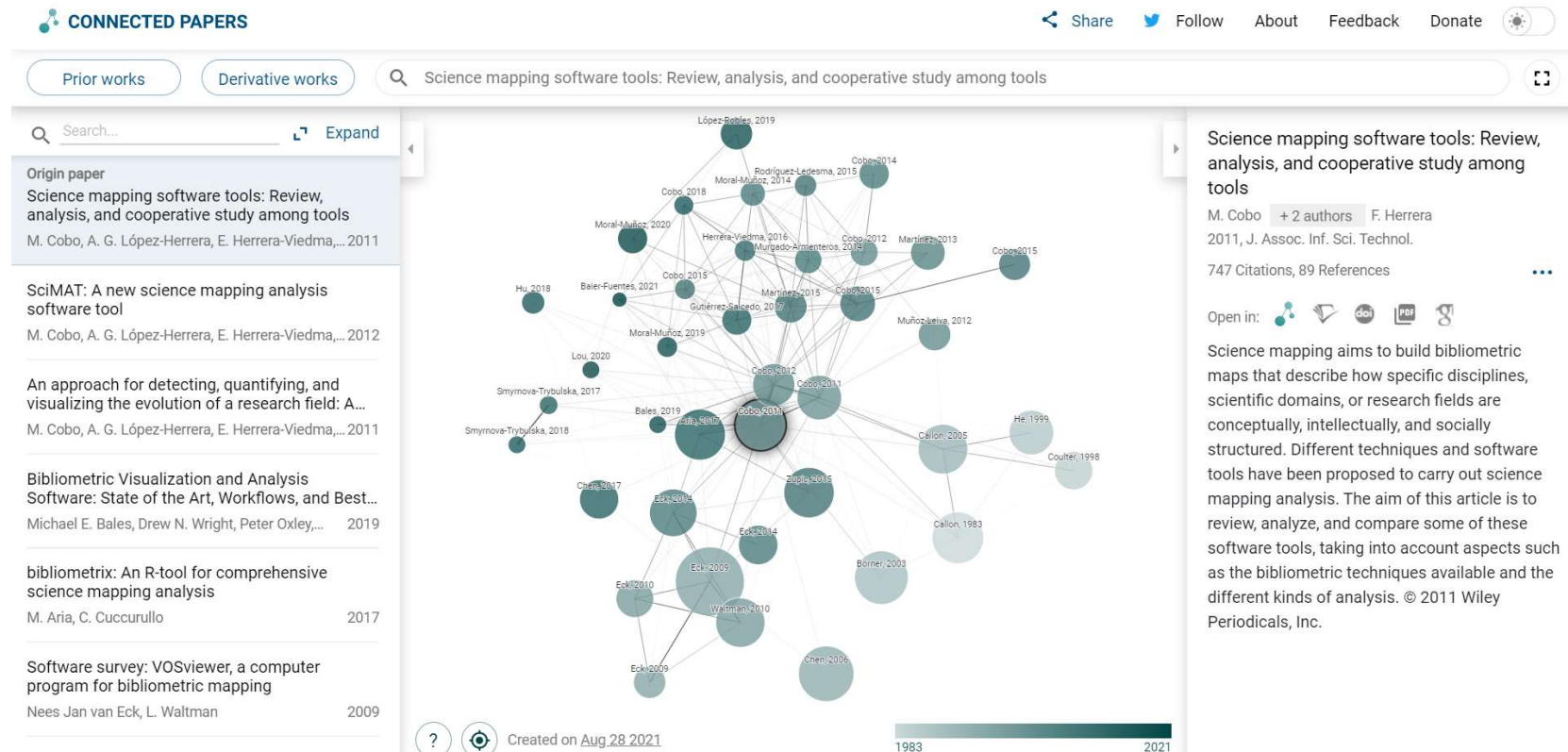
- Éclaire les principales **décisions** politiques (en l'absence de suffisamment d'éléments probants)
 - Adoption des preuves existantes: Rend les éléments probants **exploitables** et les met à l'échelle
 - Test de la “**robustesse**” des résultats et réduction des biais de publication

Lorsqu'il n'y a que **peu** de preuves scientifiques: Identifie les **manquements du corps de preuves scientifiques existantes**

- afin d'orienter les nouveaux efforts de recherche dans de nouveaux contextes/ avec d'autres groupes
- afin d'éliminer les **conflits** entre les preuves scientifiques

Quelques outils utiles...

Les Analyses documentaires de Connected Papers



[\(Link\)](#)

Les “Evidence Gap Maps” de 3ie

→ gapmaps.3ieimpact.org/evidence-maps/map-evidence-maps-relating-sustainable-development-lmics



A map of evidence maps relating to sustainable development in L&MICs

Last modified date: 17 April 2017

Evidence map About

? **HOVER OVER** a bubble to see details with links to studies. **CLICK ON** a link in the axes to see an explanation of the Intervention / Outcome. **SELECT** an area of the chart to zoom in. **TOGGLE** study categories on and off using the legend at the bottom of the chart. **EXPORT** the chart using the menu button at the top right of the chart.

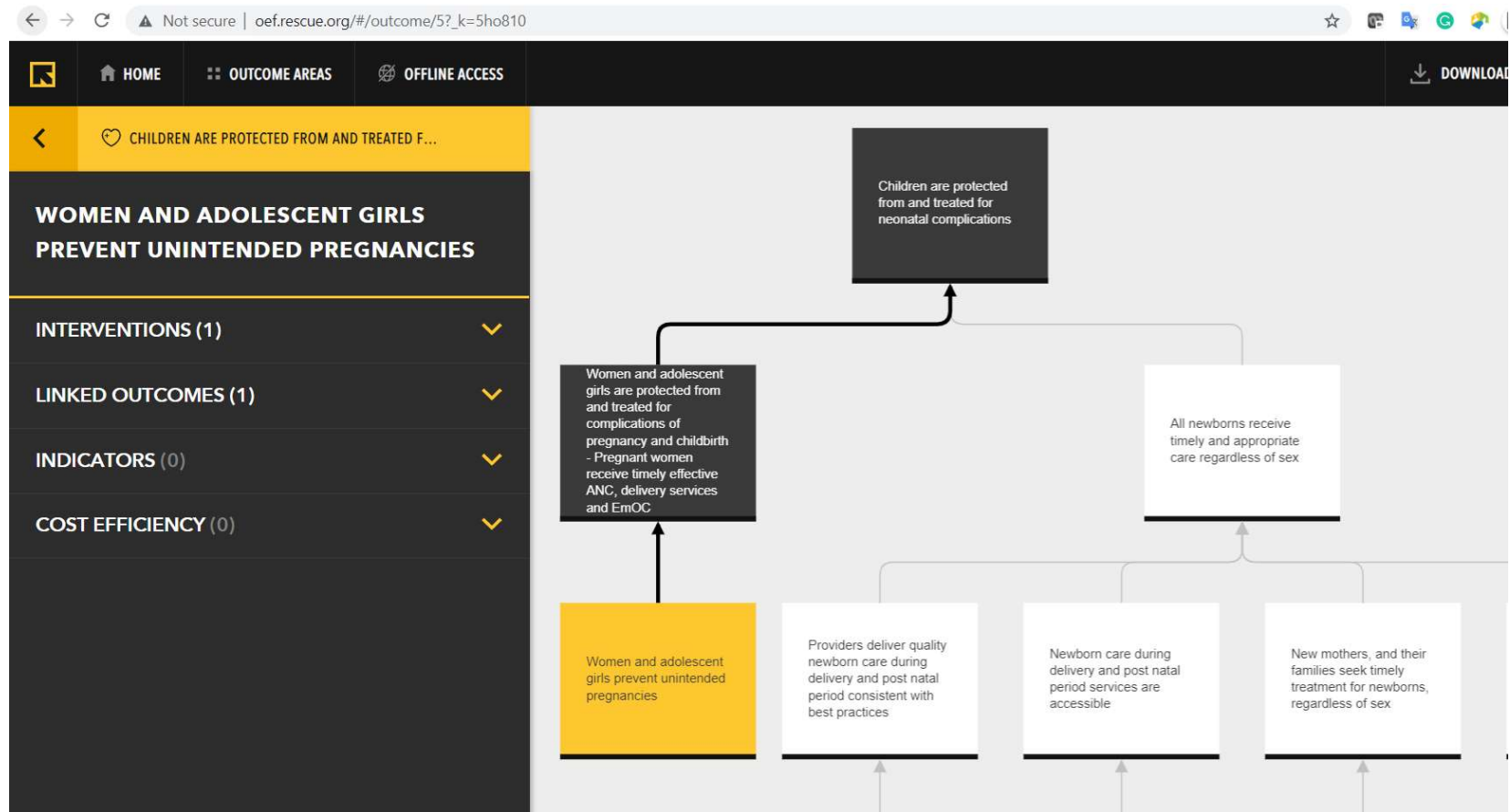
Region: All Study design: All Population: All [Update chart](#)

Sustainable Development Goals

Sectors	Sustainable Development Goals (SDGs)																
	SDG1 - Poverty	SDG2 - Agriculture nutrition and food security	SDG3 - Health	SDG4 - Education and Learning	SDG5 - Gender equality and empowerment	SDG6 - WASH	SDG7 - Energy	SDG8 - Economic growth and employment	SDG9 - Infrastructure	SDG10 - Inequality	SDG11 - Urban and rural development	SDG12 - Consumption	SDG13 - Climate change	SDG14 - Marine	SDG15 - Ecosystem related outcomes	SDG16 - Peaceful and inclusive societies	SDG17 - Global Partnership
Agriculture and rural development	●	●	●	●	●	●	●	●	●	●	●	●	●		●	●	
Climate change and environment	●	●	●	●		●		●	●	●	●		●		●	●	
Conflict management and post-conflict reconstruction	●	●	●	●	●	●		●		●		●			●	●	
Cross-sectoral	●		●	●	●	●		●		●	●	●				●	

(Link)

Le « Cadre de résultats et de données probantes » de l'International Rescue Committee



[\(Link\)](#)

La « Boîte à outils d'enseignement et d'apprentissage » de la fondation Education Endowment

educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/



The Big Picture

Evidence summaries

Practical Tools

Projects and Evaluation

Support for schools

School Improvement Planning

News

More ▾



Teaching and Learning Toolkit

An accessible summary of the international evidence on teaching 5-16 year-olds

Filter Toolkit

Toolkit Strand ^

Cost ▾

Evidence Strength ^

Impact (months) ▾

Filter results by keywords



Cost



Evidence



Months Impact

Reset

Arts participation

Low impact for low cost, based on moderate evidence.



+2

Aspiration interventions

Very low or no impact for moderate cost, based on very limited evidence.



0

Behaviour interventions

Moderate impact for moderate cost, based on extensive evidence.



+3

(Link)

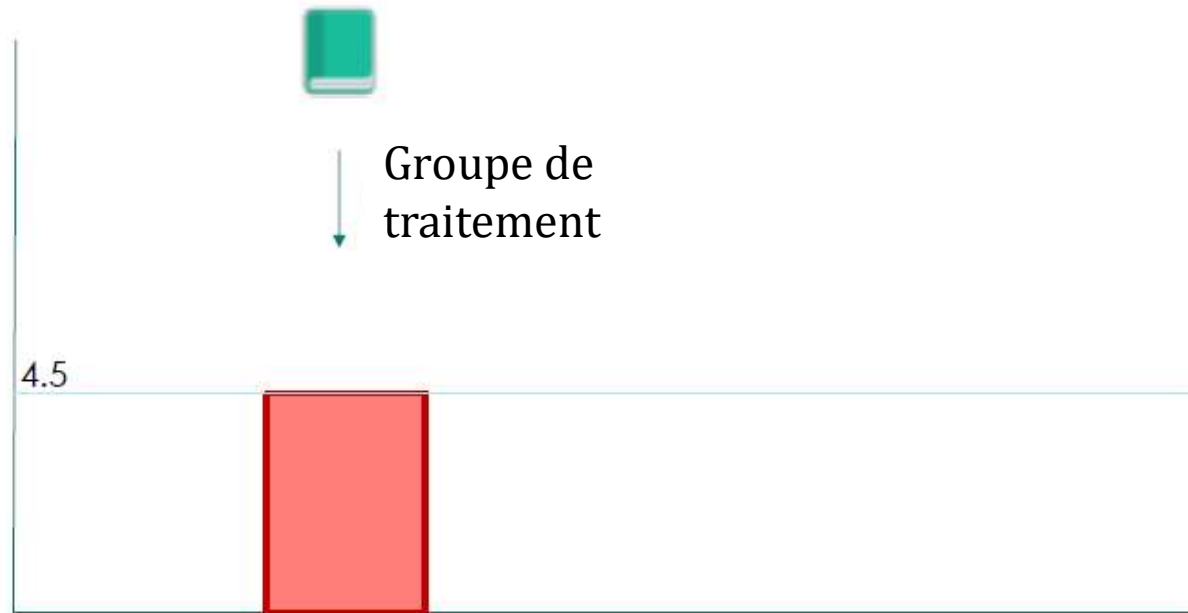
...Plus de détails sur les méthodes
d'évaluation d'impact contrefactuelle

Sélection des participants

- La sélection de participants très motivés pour intégrer le programme est un défi pour une EIC
- L'« auto-sélection » dans le programme rend plus difficile l'appréciation des impacts spécifiques du programme

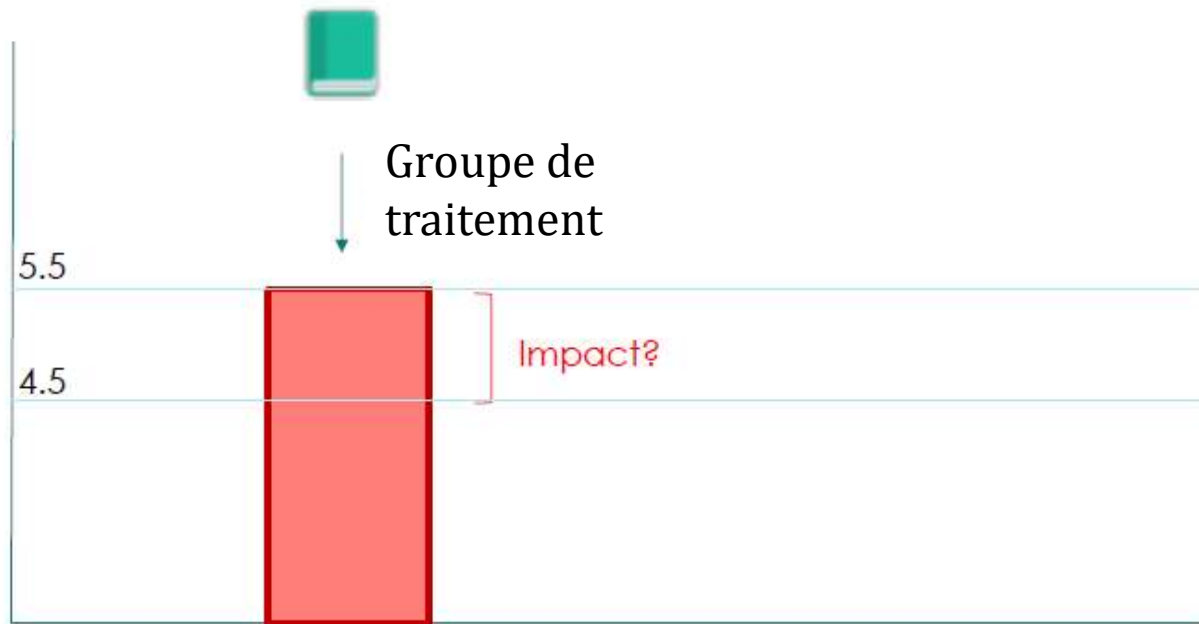
Biais de sélection

Variable de résultat:
Note moyenne en maths



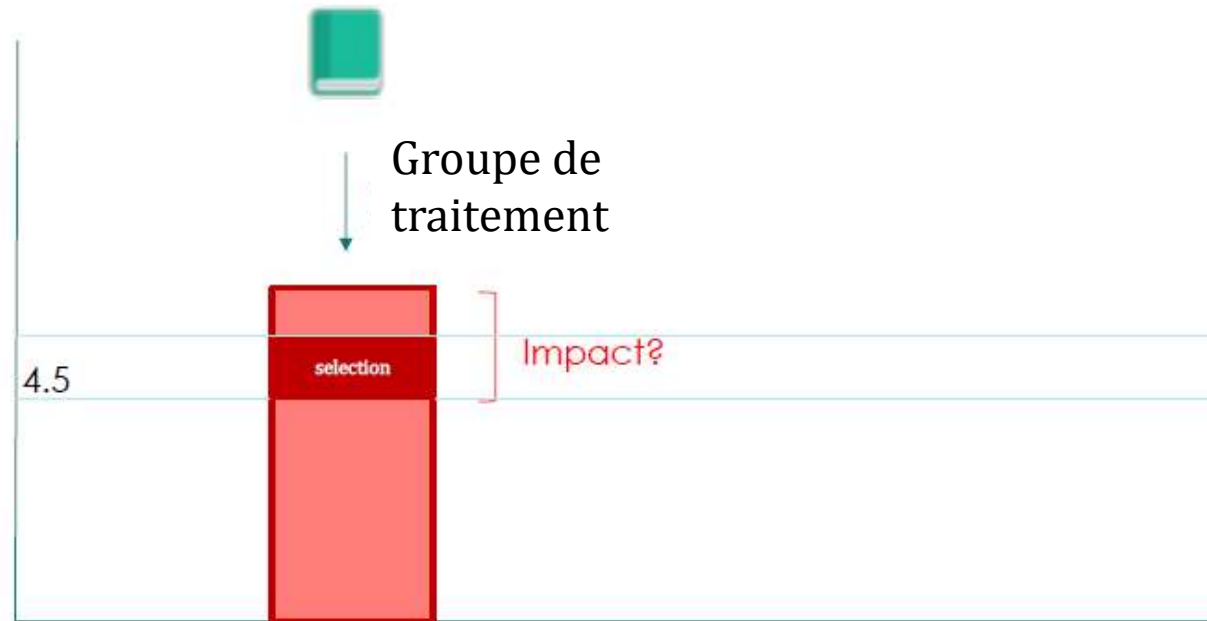
Biais de sélection

Variable de résultat:
Note moyenne en maths



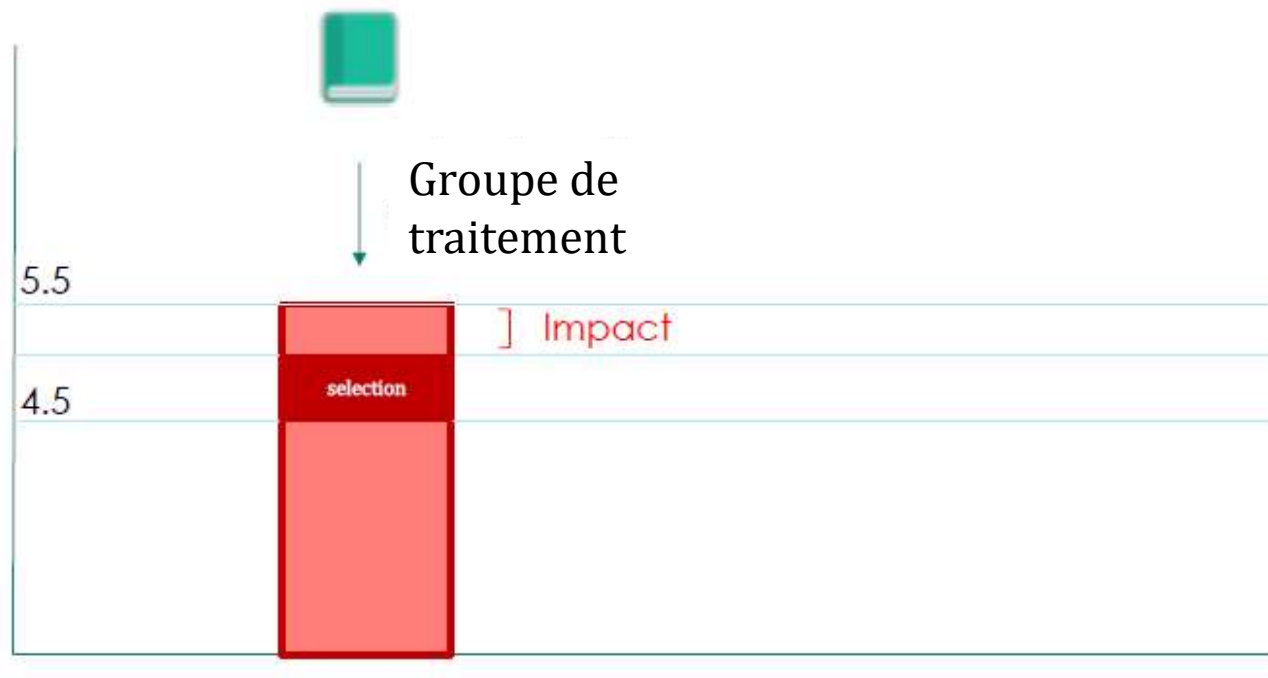
Biais de sélection

Variable de résultat:
Note moyenne en maths



Biais de sélection

Variable de résultat:
Note moyenne en maths



Y a-t-il besoin d'un groupe de contrôle/comparaison pour une EIC?

oui

non

Y a-t-il besoin d'un groupe de contrôle/comparaison pour une EIC?

oui

non

En quoi consiste un groupe de contrôle ou de comparaison dans le cadre d'une EIC ?

que se serait-il passé pour les NON-bénéficiaires, s'ils avaient participé au programme?

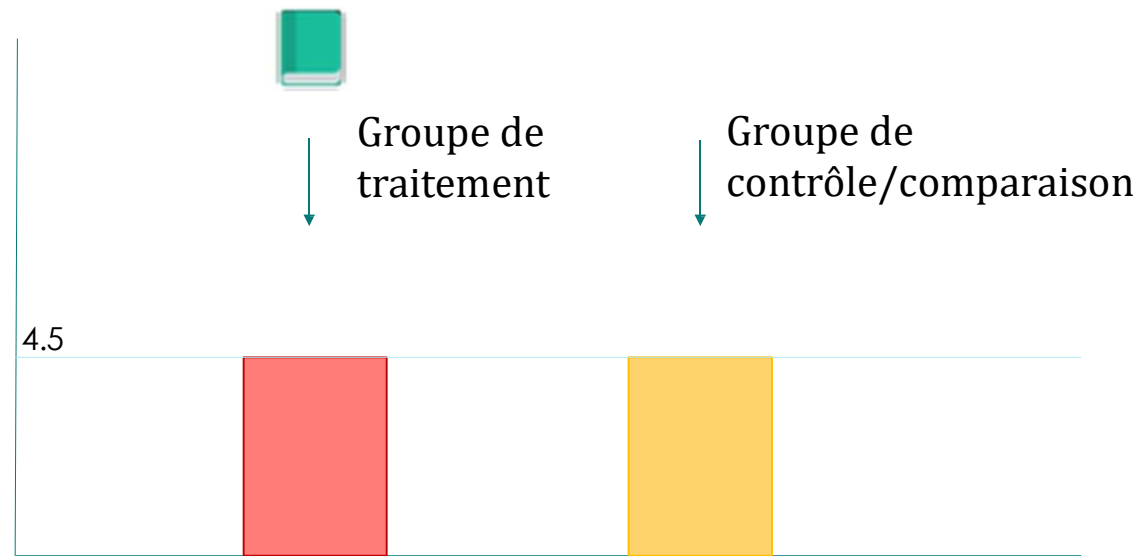
que se serait-il passé pour les bénéficiaires du programme, s'ils N'AVAIENT PAS participé au programme ?

En quoi consiste un groupe de contrôle ou de comparaison dans le cadre d'une EIC ?

que se serait-il passé pour les NON-bénéficiaires, s'ils avaient participé au programme?

que se serait-il passé pour les bénéficiaires du programme, s'ils N'AVAIENT PAS participé au programme ?

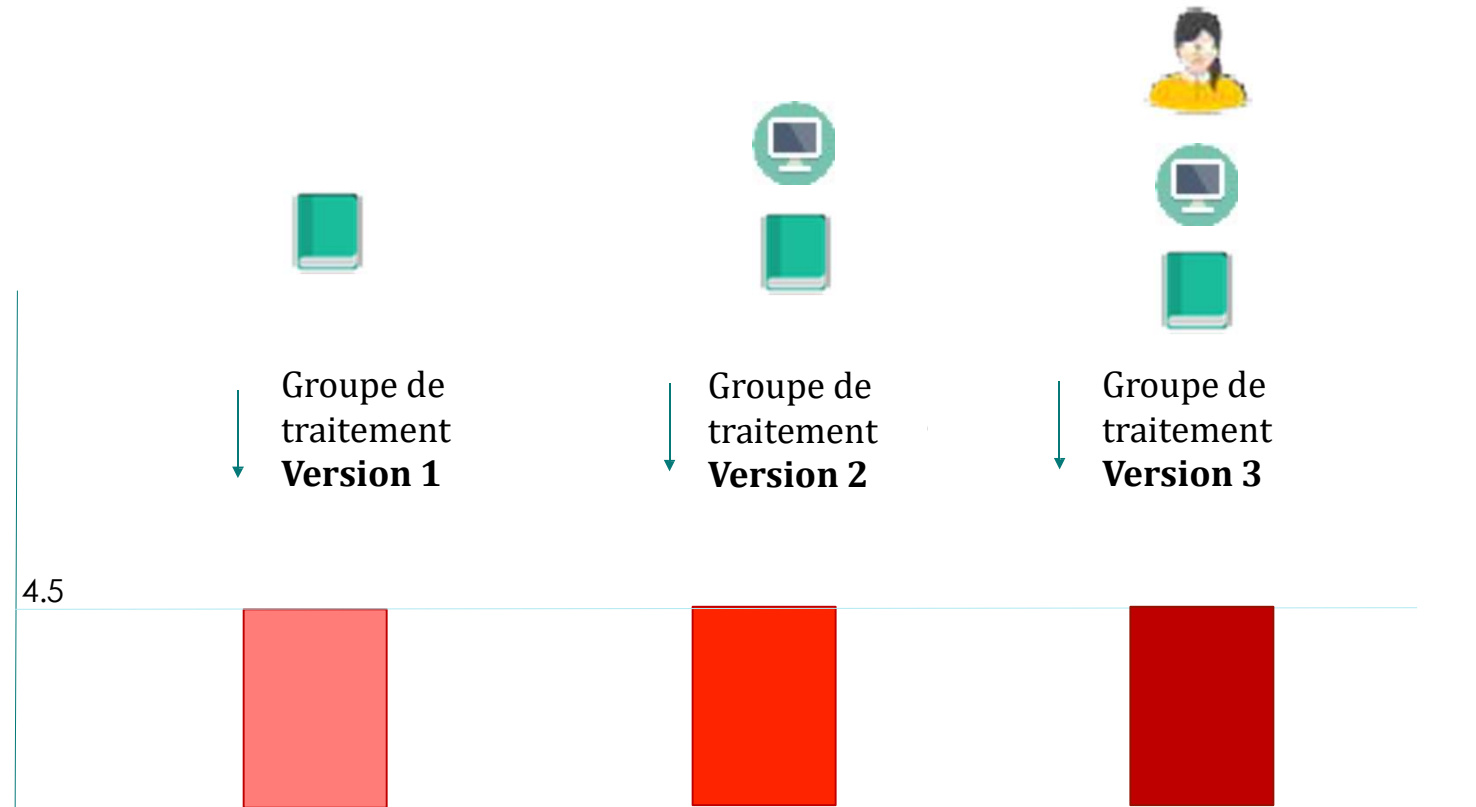
Variable de résultat:
Note moyenne en maths



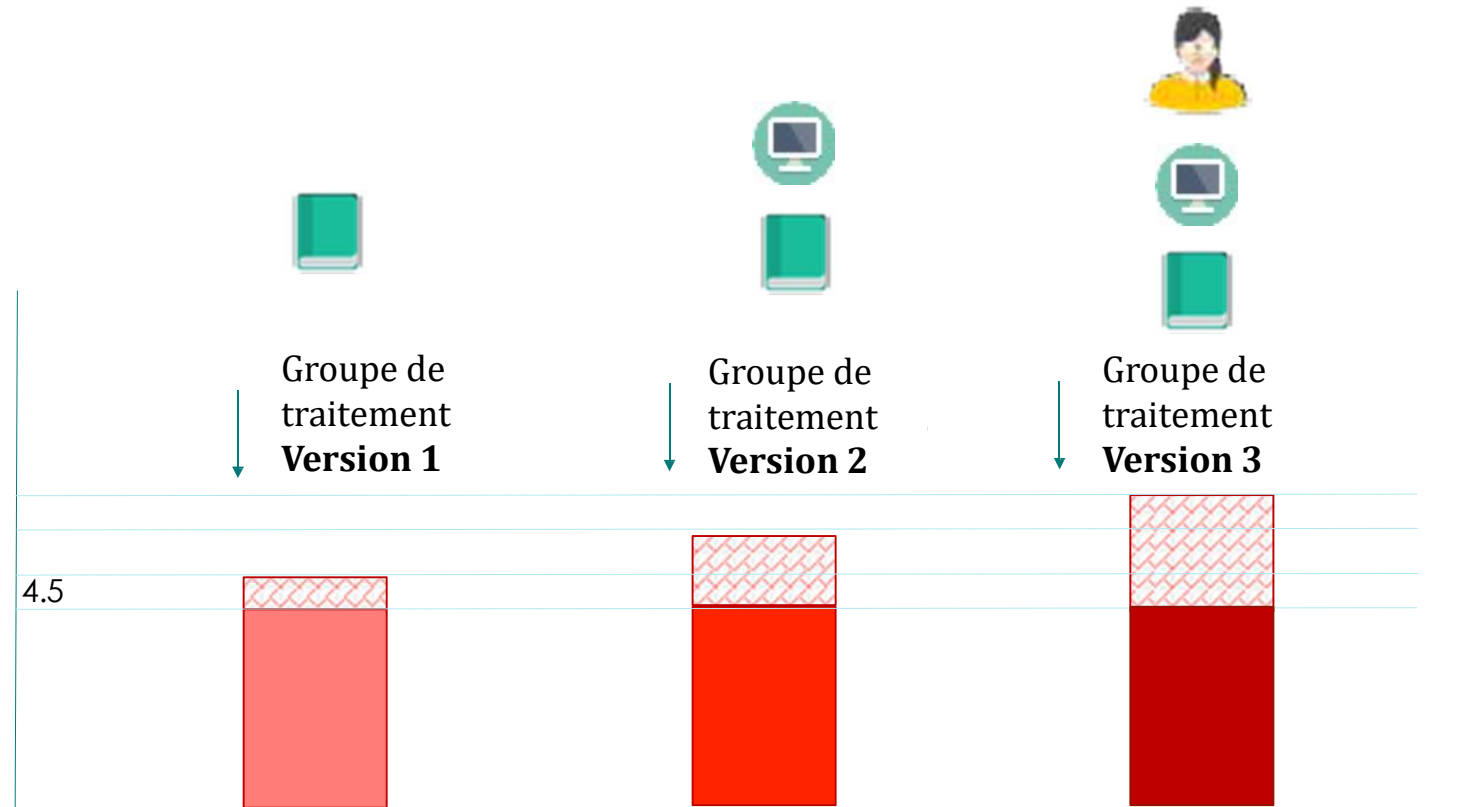
Groupe de contrôle/comparaison dans le cadre d'une EIC

- Le groupe de comparaison d'une EIC n'est pas forcément un groupe qui ne participe pas au programme

Variable de résultat:
Note moyenne en maths



Variable de résultat:
Note moyenne en maths



L'EIC peut capturer les impacts inattendus du programme

oui, si les outils (les
questionnaires) tiennent
compte d'autres résultats
potentiels

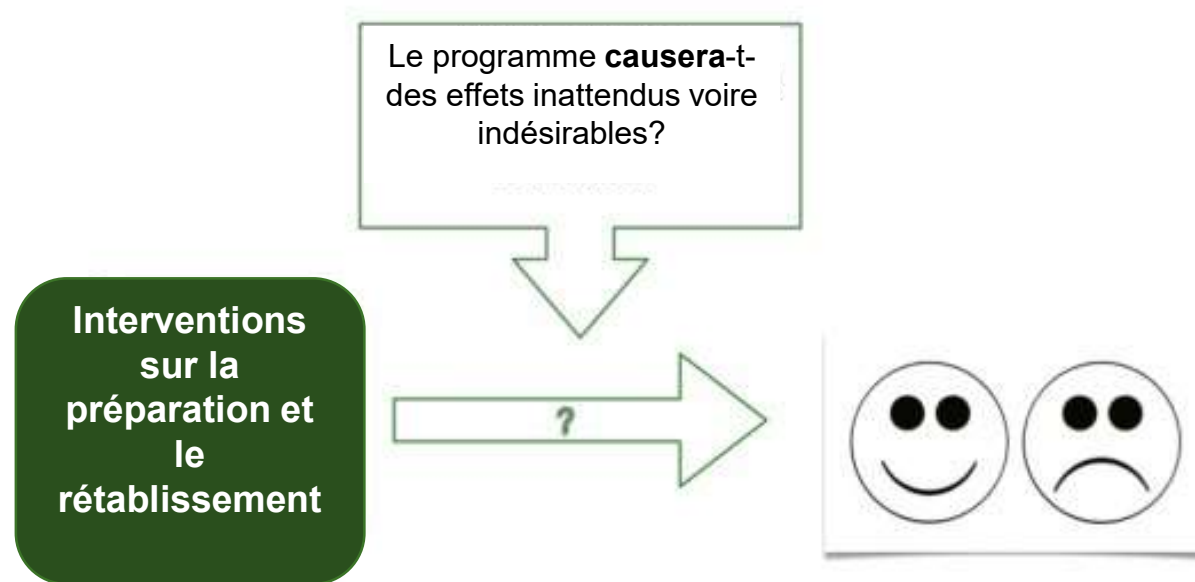
non

L'EIC peut capturer les impacts inattendus du programme

oui, si les outils (les
questionnaires) tiennent
compte d'autres résultats
potentiels

non

Effets inattendus



En résumé: l'EIC capture

**Effets
escomptés**

Attribution

**Effets
inattendus**

**... plus d'informations sur les conditions
requises pour les Evaluations d'Impact
Contrefactuelles**

Quand démarrer une EIC ?

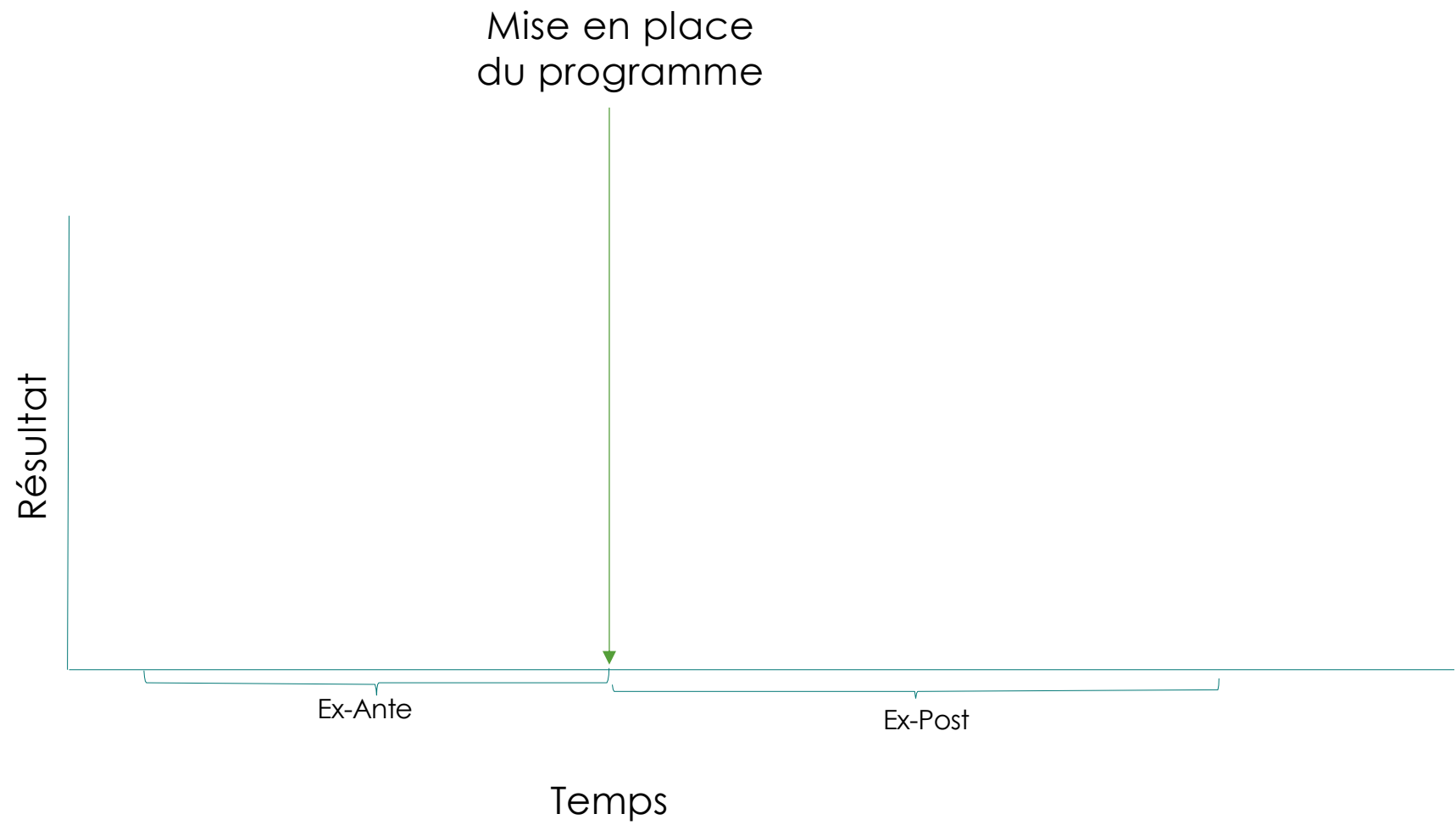
ex ante (avant le début du programme à évaluer)

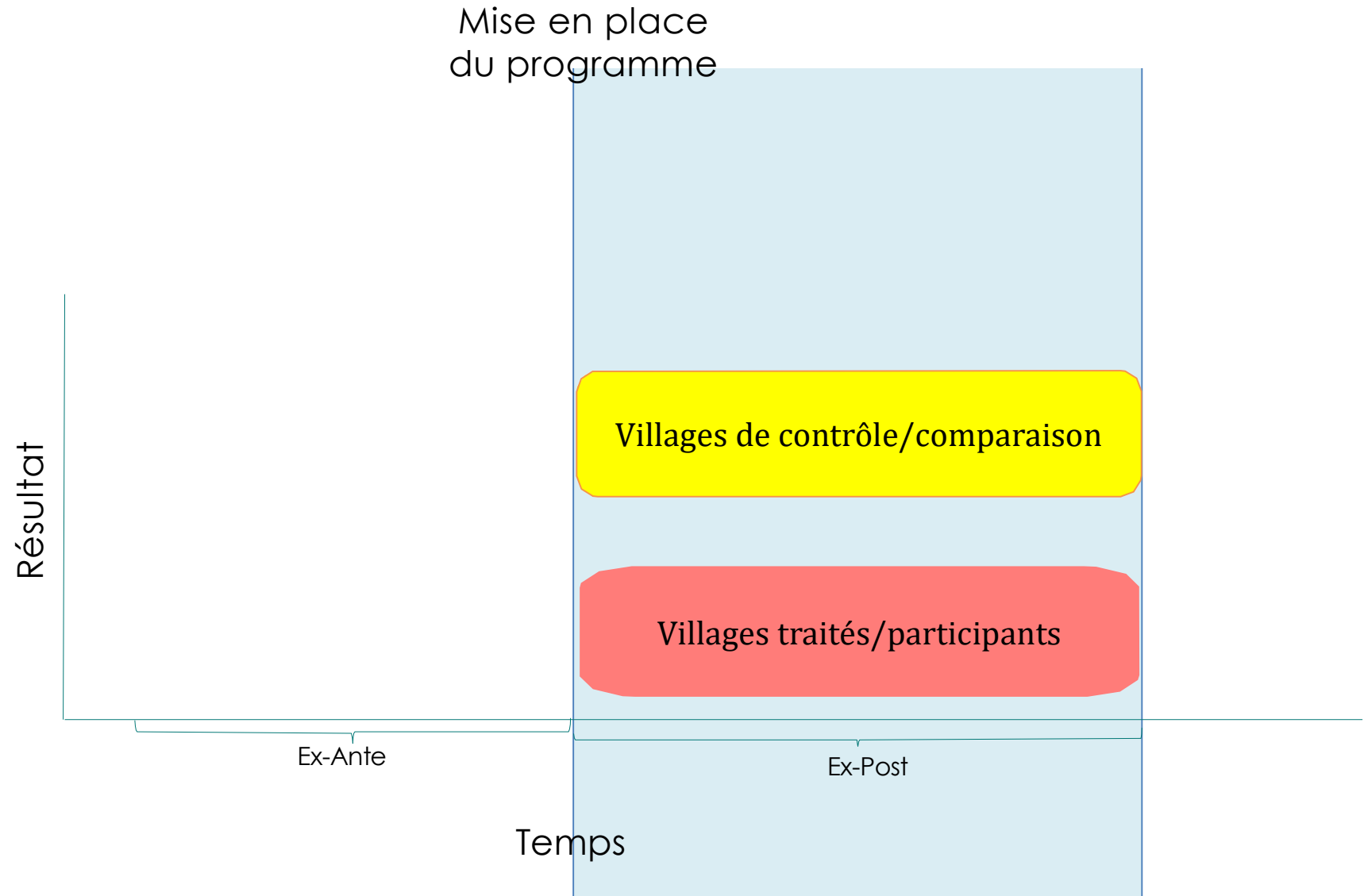
ex post (après la fin du programme à évaluer)

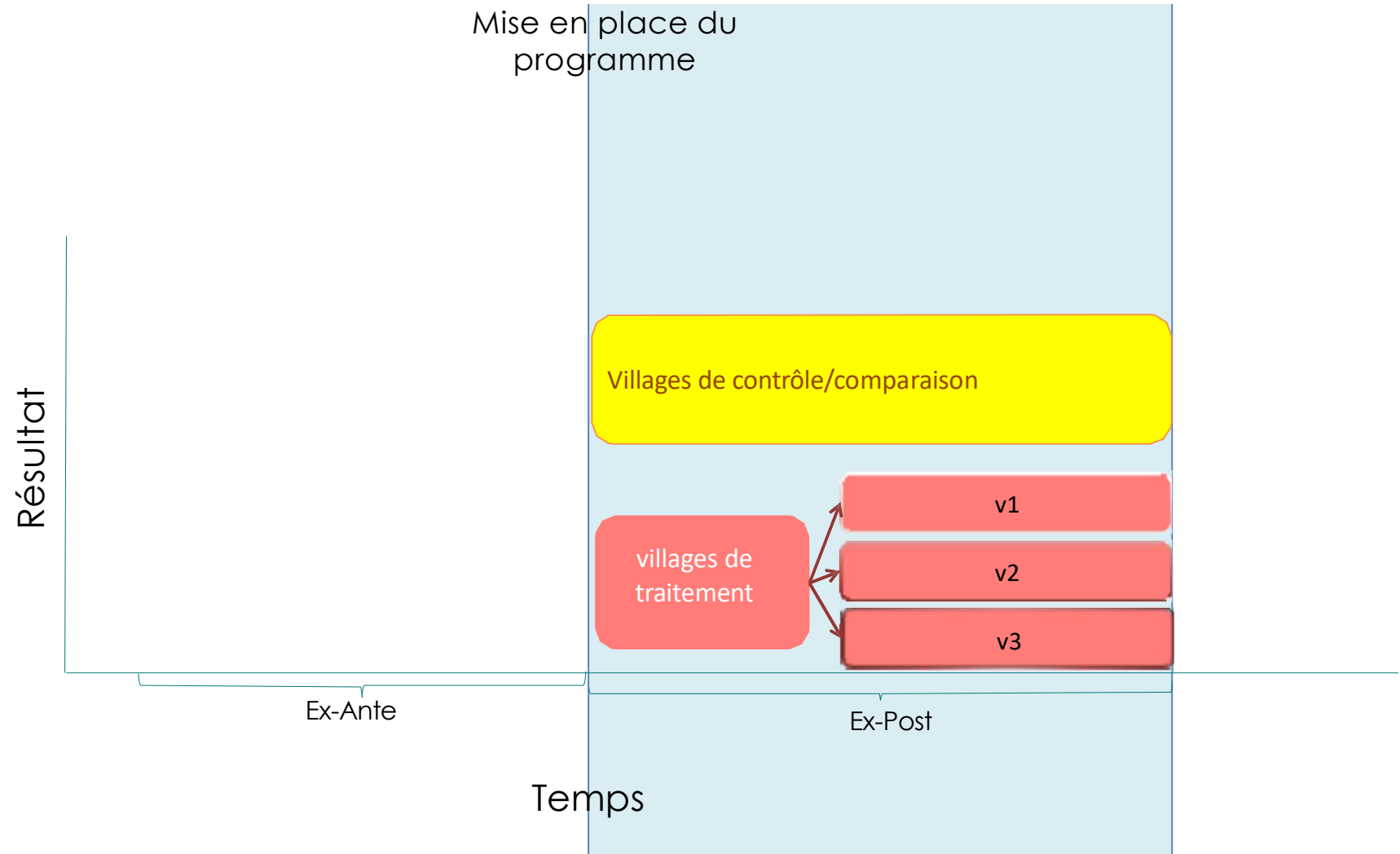
Quand démarrer une EIC ?

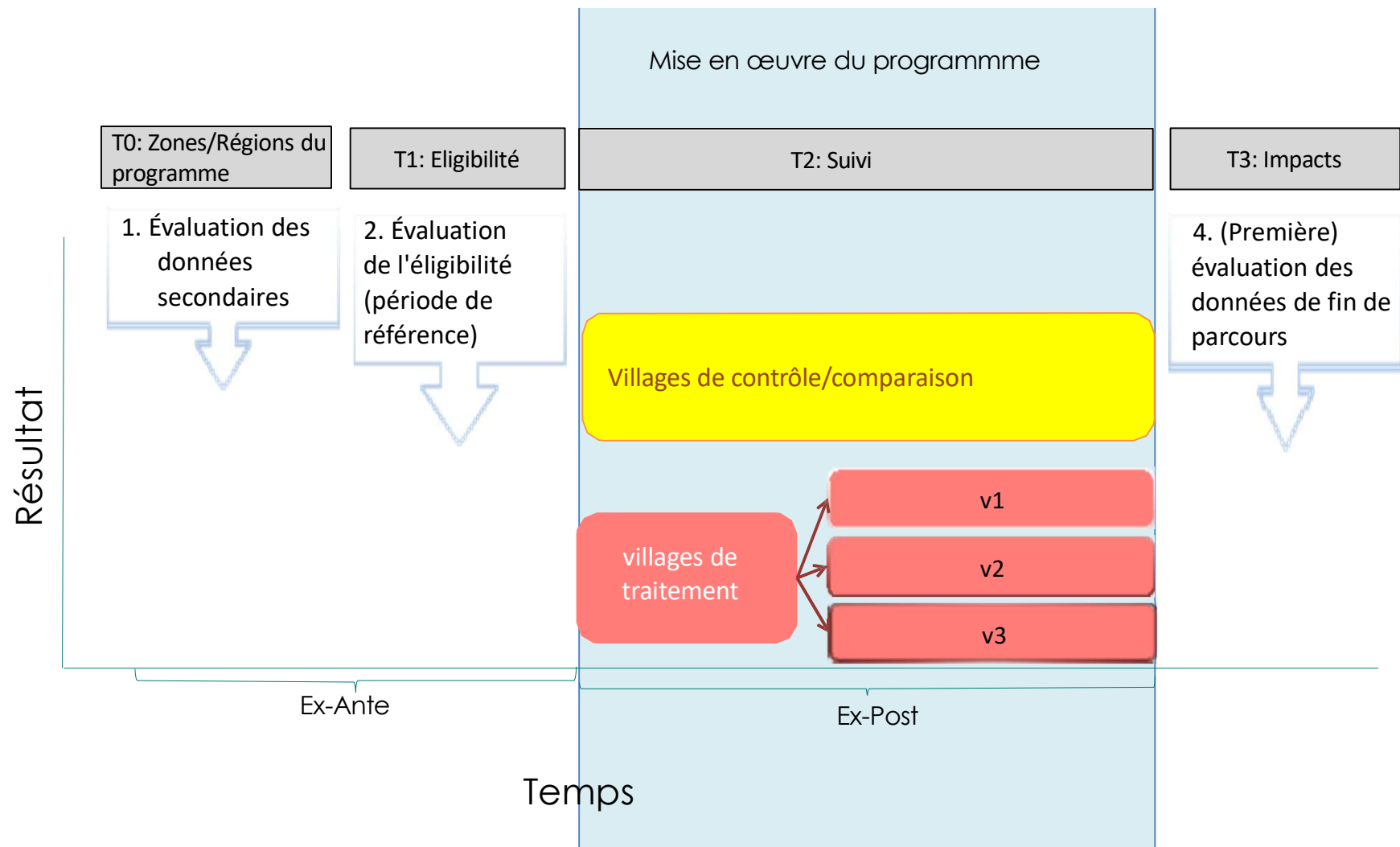
ex ante (avant le début du
programme à évaluer)

ex post (après la fin du
programme à évaluer)









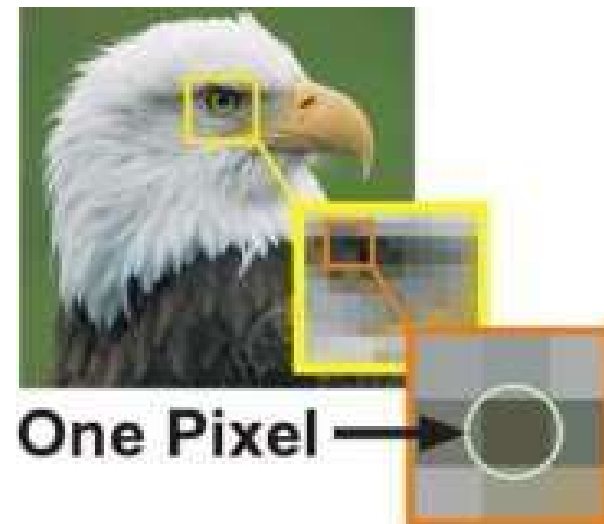
Données nécessaires pour une EIC

- Une EIC requiert des informations sur le groupe de traitement ainsi que sur le groupe de contrôle/comparaison
- Si l'allocation des participants au programme est aléatoire, on peut en théorie n'utiliser qu'une seule série d'observations (mesures prises à un seul point du temps)
- Toutefois, l'analyse est plus optimale si l'on dispose de mesures répétées au cours du temps

Taille d'échantillon pour une EIC

- Dans la mesure du possible, une EIC opte toujours pour un plus grand nombre d'unités d'observations plutôt que moins
- Unités d'observation: individus, ménages, écoles, entreprises, etc.
- La taille de l'échantillon détermine la **précision des mesures**.
Plus il y a d'observations :

- plus l'étude est précise
- plus vous aurez confiance en les conclusions de votre évaluation.



Les besoins d'une EIC: résumé

- Une EIC rigoureuse peut être planifiée **avant ou après** la mise en place du programme
(ATTENTION: pour certaines méthodes, l'EIC doit être planifiée en amont)
- Les résultats doivent être mesurés pour le groupe qui recoit le programme **ET** pour le groupe de contrôle/comparaison
- Les informations doivent être collectées au moins une fois dans chaque groupe, et **idéalement plusieurs fois au cours du temps**
- Une EIC requiert en général un **grand nombre d'observations**

**... un risque pour toutes les évaluations
d'impact**

L'erreur de type II



L'erreur de type II en EIC

Hypothèse de base : Aucun impact du programme de prévention de la COVID-19.

Erreur de type II : Le programme a eu un impact (les enfants portent des masques), mais nous ne pouvons pas le détecter.



L'erreur de type II en EIC

Admettons qu'un programme a un impact positif sur les bénéficiaires :

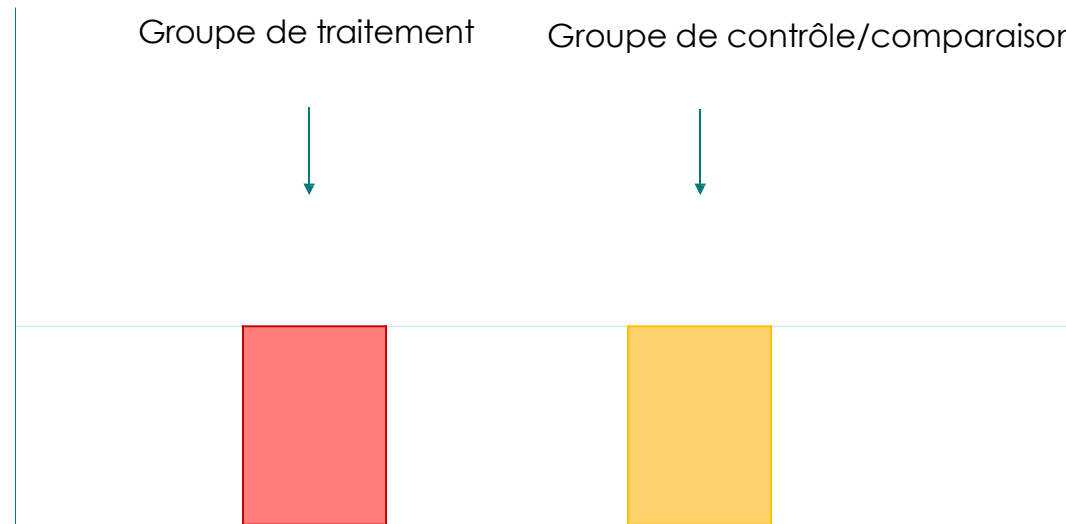
- Si l'échantillon utilisé pour l'évaluation est trop petit, on risque de ne pas pouvoir détecter cet impact positif
- « **L'erreur de type II** » : Le risque de **ne pas pouvoir** conclure que votre programme a un impact **même lorsqu'il en a effectivement un**
- Pourrait aboutir à des décisions politiques visant à éliminer le programme, ce qui serait préjudiciable aux bénéficiaires et à la société

Effets de contagion (“spillovers”)

- EIC peut être menacée lorsque le groupe de traitement affecte indirectement les personnes qui ne bénéficient pas du programme
- Des effets de contagion (« spillovers ») qui « polluent » le groupe de contrôle:
 - physiques (partage des ressources/environnement)
 - information (« enseignement social »)
 - à l'échelle du marché (par exemple, les effets de déplacement)
 - Pour en savoir plus, consultez "Running Randomized Evaluations" de Glennerster et Takavarasha
- La contagion peut entraîner une sous-estimation ou une surestimation de l'impact

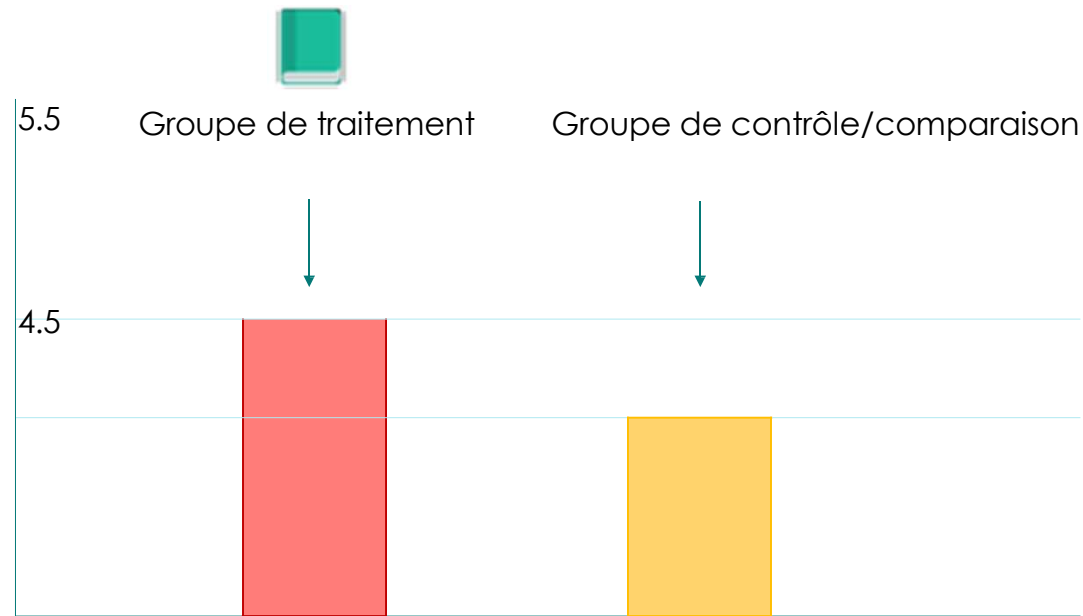
Effets de contagion (“spillovers”)

Variable de résultat:
Note moyenne en maths



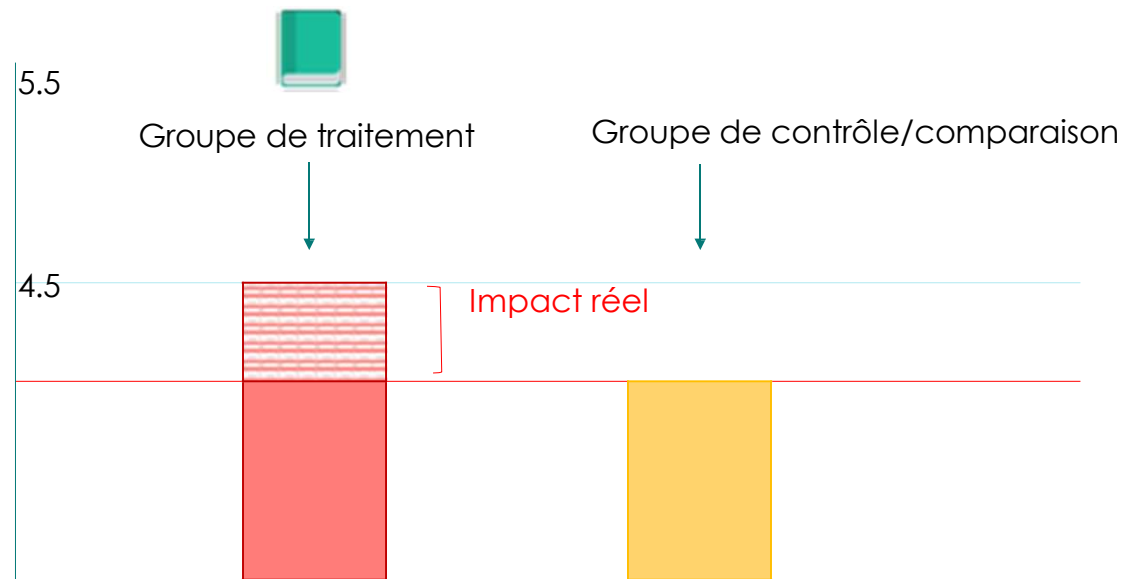
Effets de contagion (“spillovers”)

Variable de résultat:
Note moyenne en maths



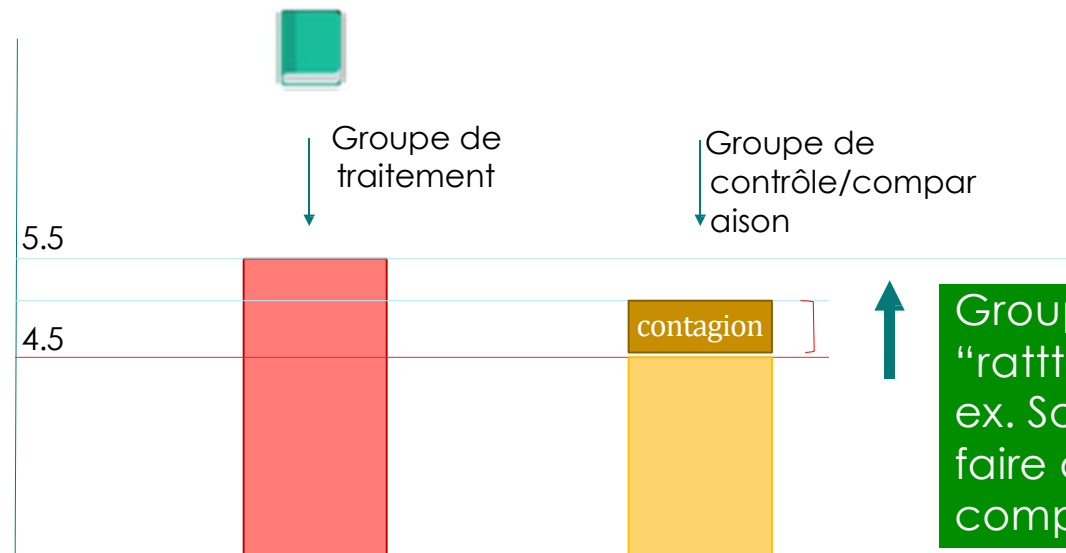
Effets de contagion (“spillovers”)

Variable de résultat:
Note moyenne en maths



Effets de contagion (“spillovers”)

Variable de résultat:
Note moyenne en maths

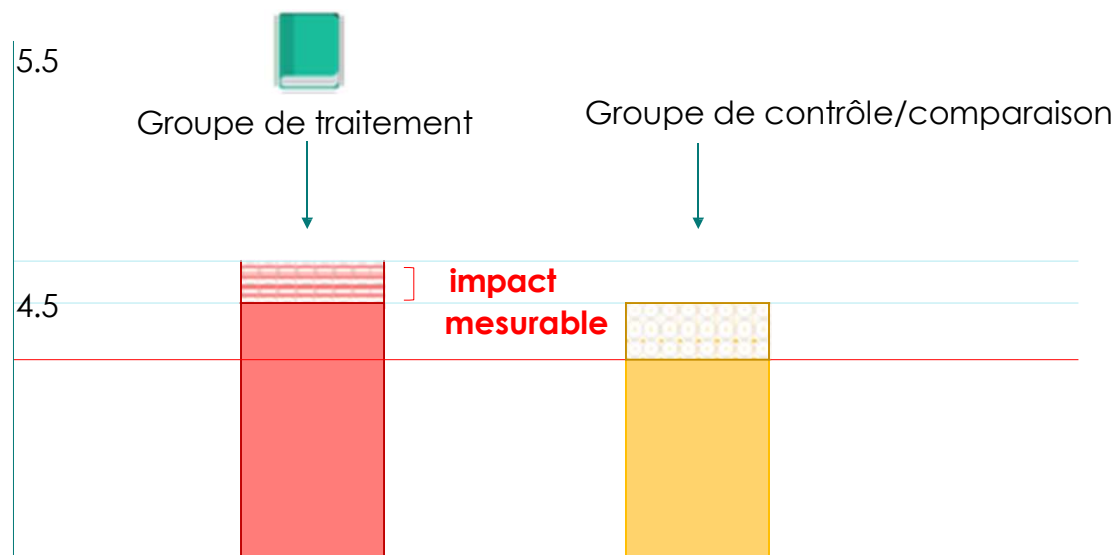


Group de contrôle
“ratttrappe” l’autre (par
ex. Sont **encouragés** de
faire aussi bien; **imitent** le
comportement)

Effets de contagion (“spillovers”)

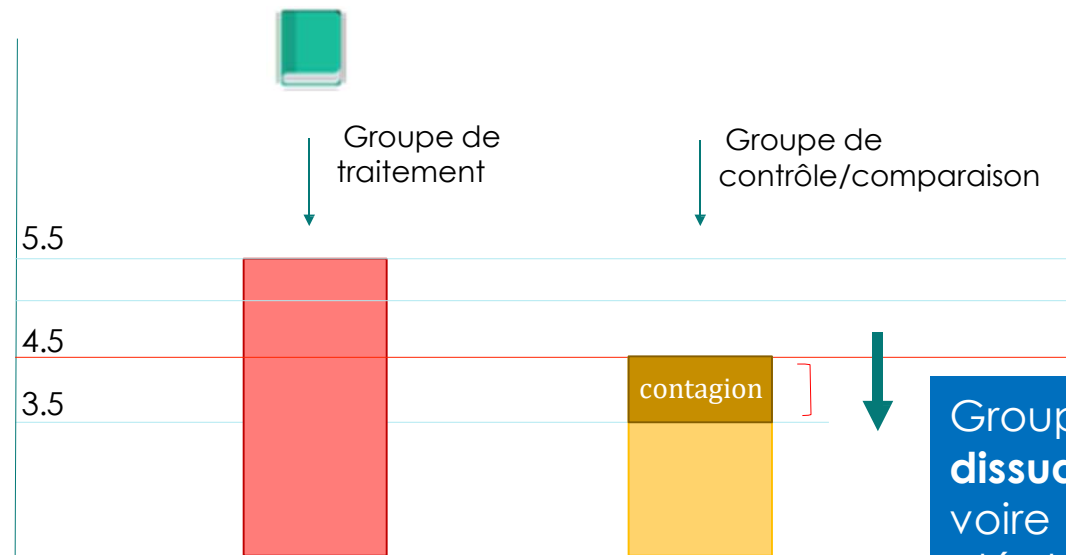
Variable de résultat:
Note moyenne en maths

→ **SOUS-ESTIMATION** de l'IMPACT
REEL du programme



Effets de contagion (“spillovers”)

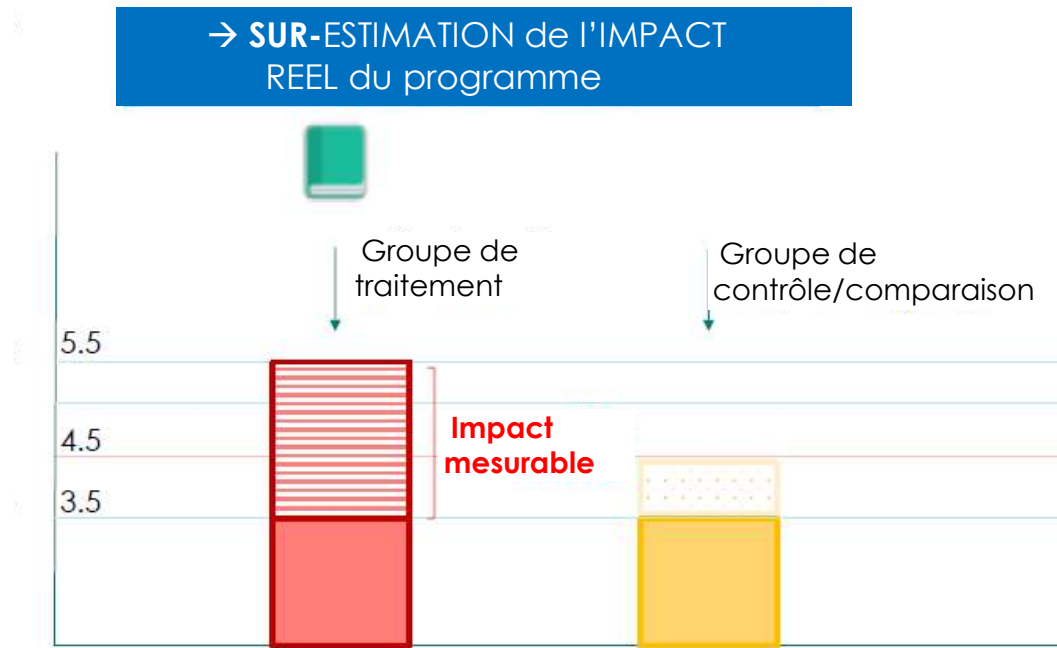
Variable de résultat:
Note moyenne en maths



Groupe de contrôle
dissuadé de bien faire,
voire (non
aléatoirement) **déplacé**

Effets de contagion (“spillovers”)

Variable de résultat:
Note moyenne en maths



Session 2 : Identification des besoins du programme, de la théorie du changement, des questions d'évaluation, et des indicateurs

C4ED – EUTF
Septembre 2021

La Théorie du Changement

- La Théorie du Changement est une approche structurée servant à la conception et à l'évaluation des programmes.
- Elle établit la chaîne logique selon laquelle les intrants du programme aboutissent à des changements dans les résultats.

La Théorie du Changement...

- ...décrit un programme et sa vision du changement
- ...est la feuille de route conceptuelle d'un programme pour atteindre son objectif
- ...articule le programme et ses hypothèse sous-jacentes
- ...soutient le suivi et l'évaluation

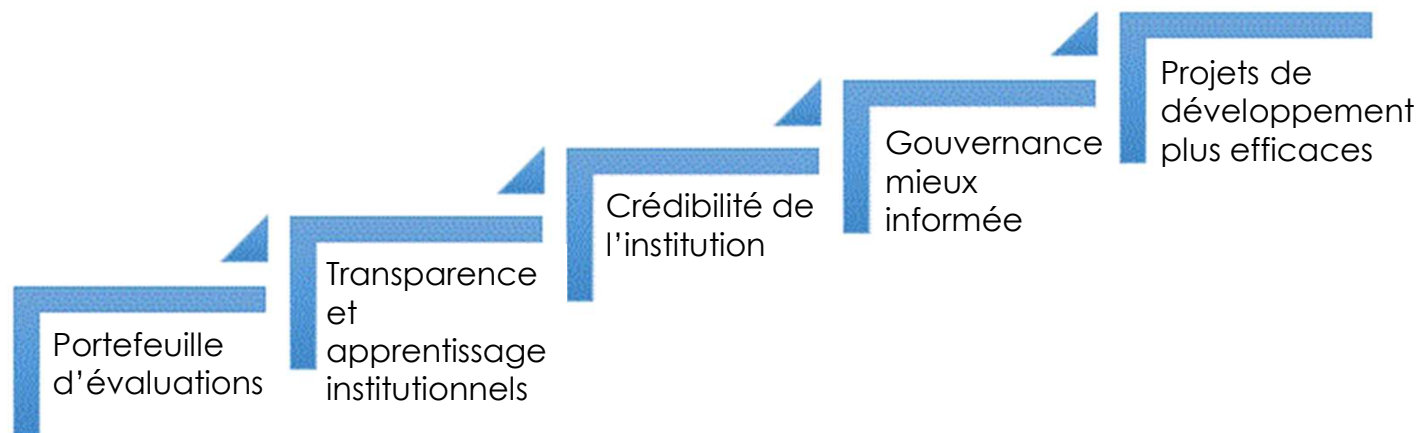
La Théorie du Changement

- Les théories du changement doivent être élaborées au moyen d'un processus **collaboratif** et **consultatif**
- **L'identification des hypothèses** est un élément clé de la théorie du changement. L'analyse et la discussion de ces postulats peuvent permettre d'identifier et **prévoir des risques potentiels**

Les objectifs des théories du changement

Les théories du changement permettent de déterminer les projets à évaluer en priorité (et ceux à ne pas évaluer)

Les TdC favorisent également le développement institutionnel et la transparence, tout en renforçant la crédibilité et la capacité de gouvernance d'une institution



La théorie du changement : Doubter du programme

- La théorie du changement décrit tous les éléments constitutifs nécessaires à la réalisation d'un objectif donné en posant des questions fondamentales :
 - Pourquoi le programme réussira-t-il ?
 - Quels sont les canaux ou mécanismes qui permettent de présumer que les résultats escomptés seront atteints ?
 - Quelles sont les hypothèses formulées ?
 - Quelles sont les incertitudes ?

Théorie du changement et indicateurs

- À chaque étape de la Théorie du Changement, on définit les indicateurs à mesurer afin de vérifier si le programme a réussi.
- Si le programme ne réussit pas, les indicateurs des étapes intermédiaires nous permettront de savoir à quelle étape de la chaîne le programme n'a pas produit les résultats escomptés.

Préparer une évaluation

- Quatre étapes :
 - 1) Élaborer une théorie du changement
 - 2) Mettre au point d'une chaîne de résultats
 - 3) Définir des questions d'évaluation (QE)
 - 4) Définir des indicateurs de résultat et de performance
- Les QE sont conçues comme des hypothèses vérifiables dérivées de la TdC
- Les QE doivent être inspirées par le principal intérêt politique en question, par exemple:
 - L'impact du programme sur les résultats ? Ou...
 - Certaines modalités du programme sont-elles plus avantageuse que d'autres ? Ou...
 - Quel est l'effet de l'introduction d'une innovation dans la conception du programme censée modifier les comportements ?

Source: Gertler et al. (2016)

Préparer une évaluation

- **Expérimentations sur les mécanismes** : des EI qui ne testent pas un programme, mais un mécanisme de causalité qui sous-tend le choix d'un programme

Exemple: l'octroi de subventions aux grossistes en fruits et légumes en vue de lutter contre l'obésité dans les quartiers pauvres

- ➡ Hypothèse sous-jacente : Un accès plus facile aux aliments sains entraînera une hausse de leur consommation
- ➡ Mécanisme: Les foyers qui reçoivent gratuitement des corbeilles de fruits et légumes accroissent-ils leur consommation de fruits et légumes?

Si la réponse est non, on ne peut pas espérer que le programme fonctionne !

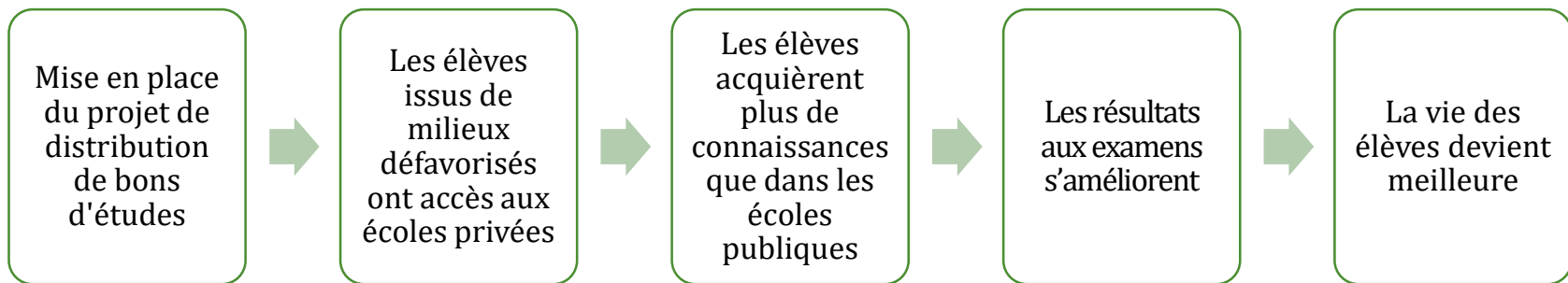
Source: Gertler et al. (2016)

Préparer une évaluation

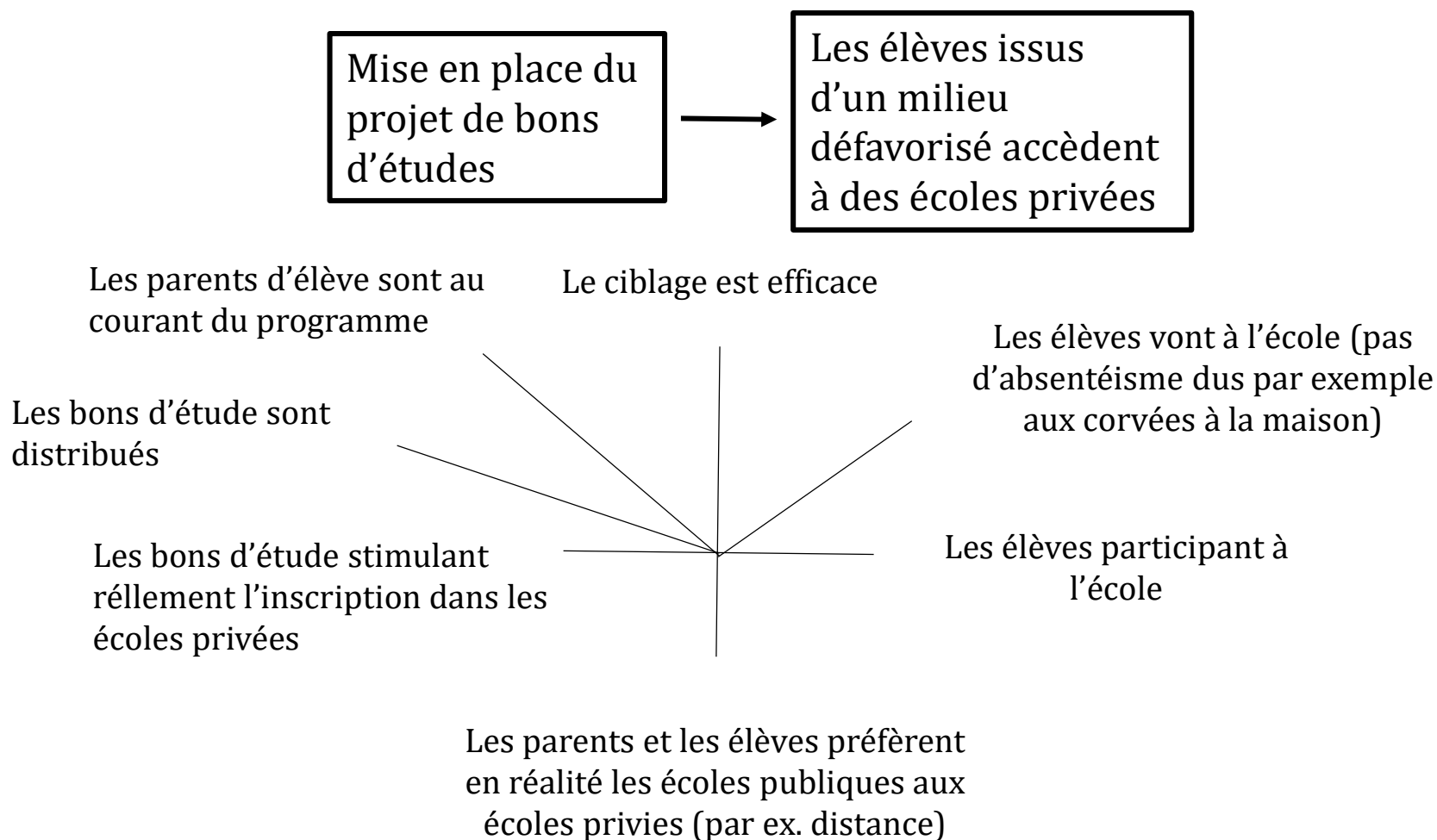
- Les principales parties prenantes du projet doivent s'entendre sur les principaux résultats d'intérêt et les changements attendus
- Les indicateurs doivent être Spécifiques, Mesurables, Attribuables, Réalistes et Ciblés
(en anglais « SMART - Specific, Measurable, Attributable, Realistic, Targeted »)
- Identifier des indicateurs tout au long de la chaîne de résultats : ils sont essentiels pour déterminer pourquoi une intervention réussit ou pas

Gertler et al. (2016)

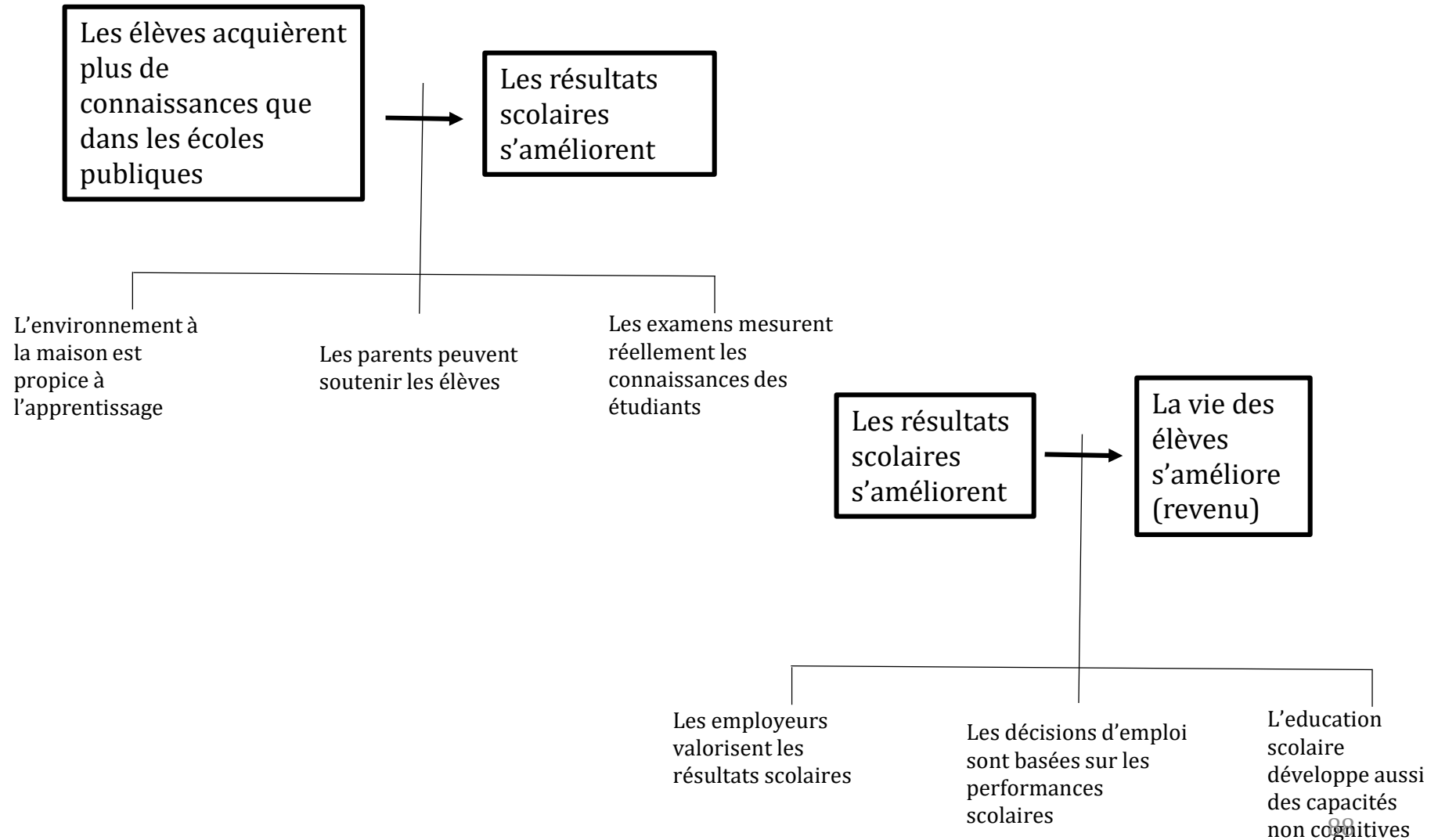
La théorie du changement d'un programme de distribution de « bons d'études »



La théorie du changement d'un programme de distribution de bons d'études



La théorie du changement d'un programme de distribution de bons d'études



Le choix des questions d'une évaluation d'impact

Pour le choix des questions d'une évaluation d'impact rigoureuse....

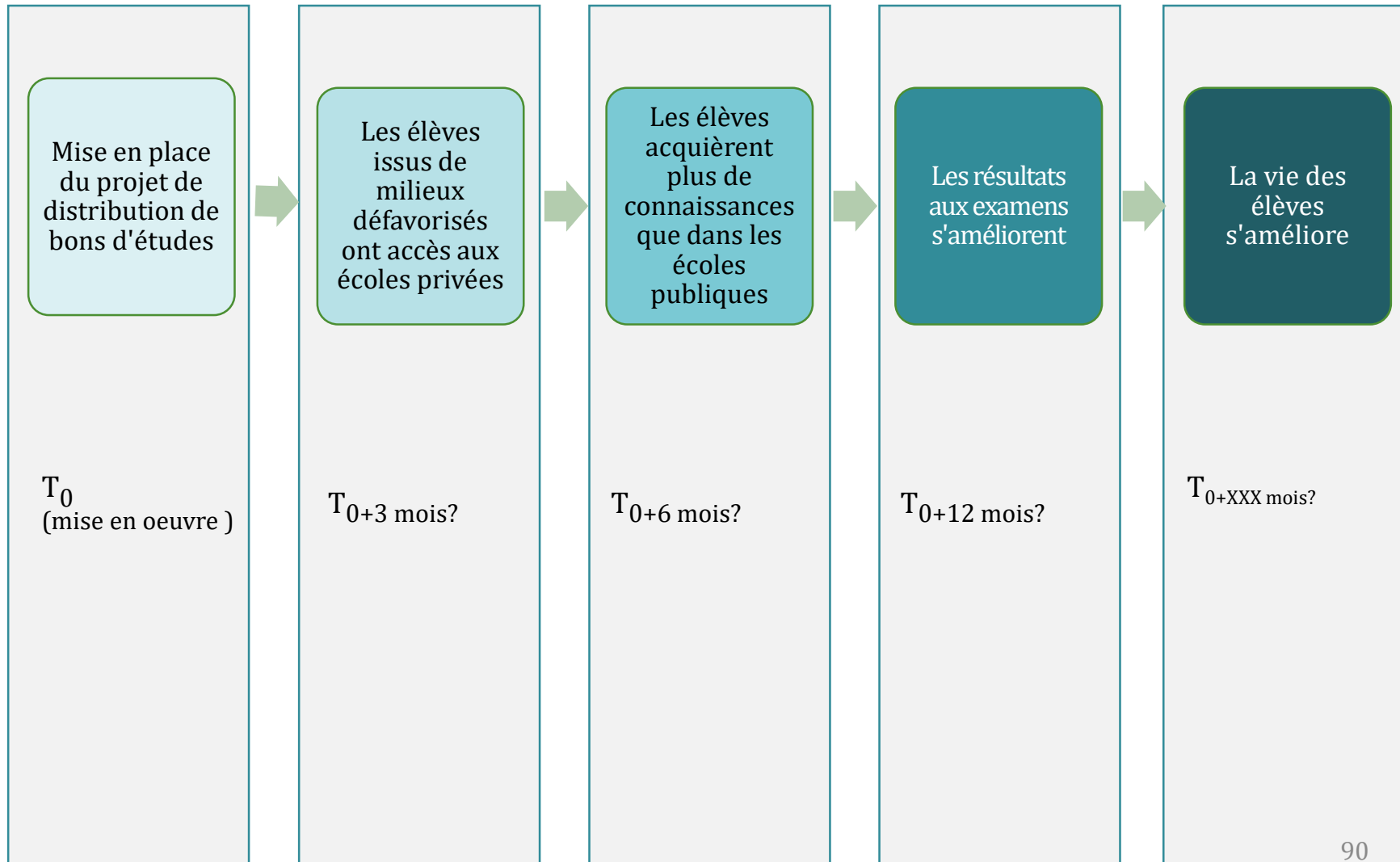
- QUE DOIT-ON MESURER?

- il faut suivre la théorie du changement et essayer de mesurer les changements à chaque étape

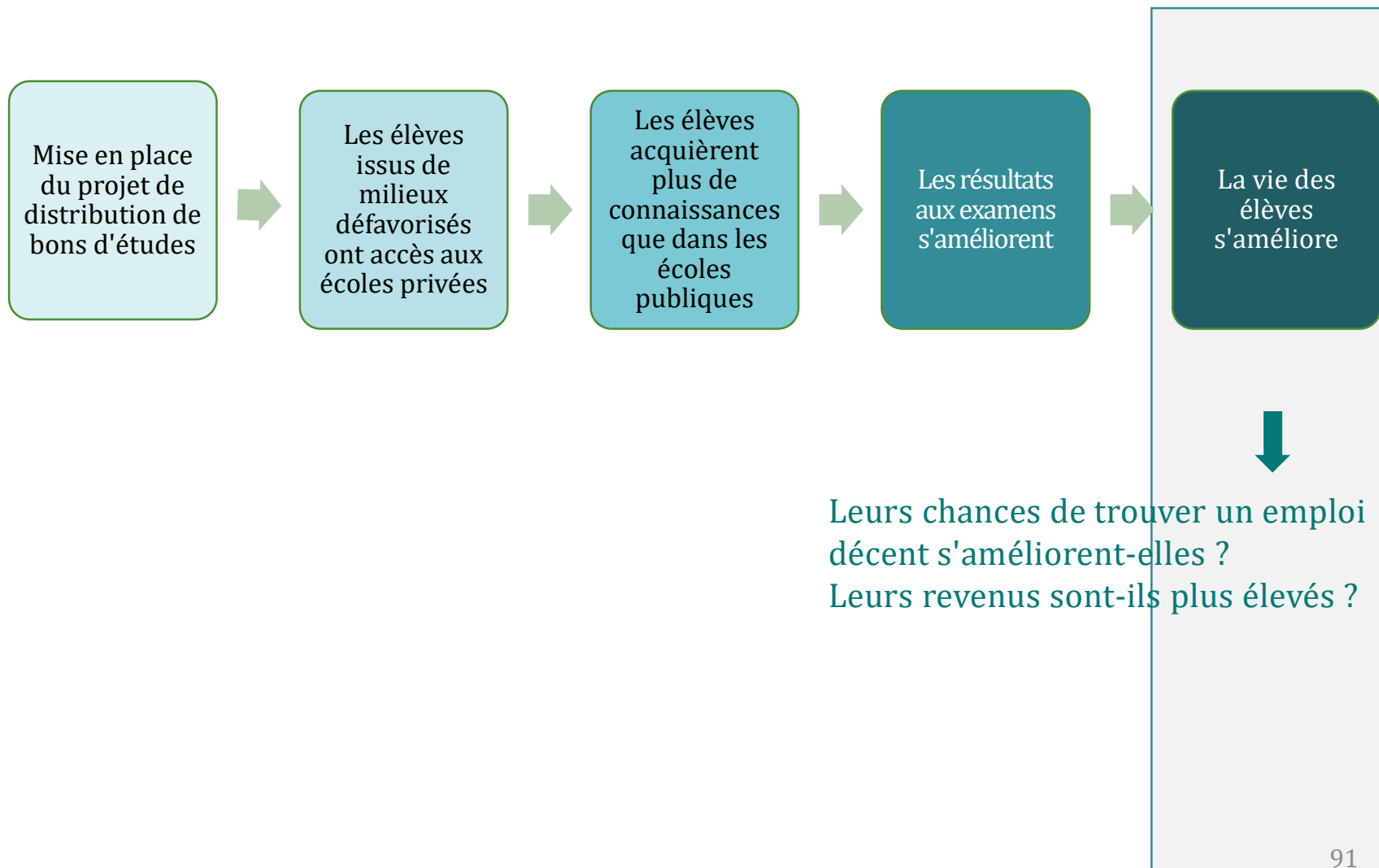
- QUAND DOIT-ON LE MESURER?

- soyez réaliste non seulement quant aux changements à mesurer, mais aussi quant à la possibilité de les capturer au moment de la mesure

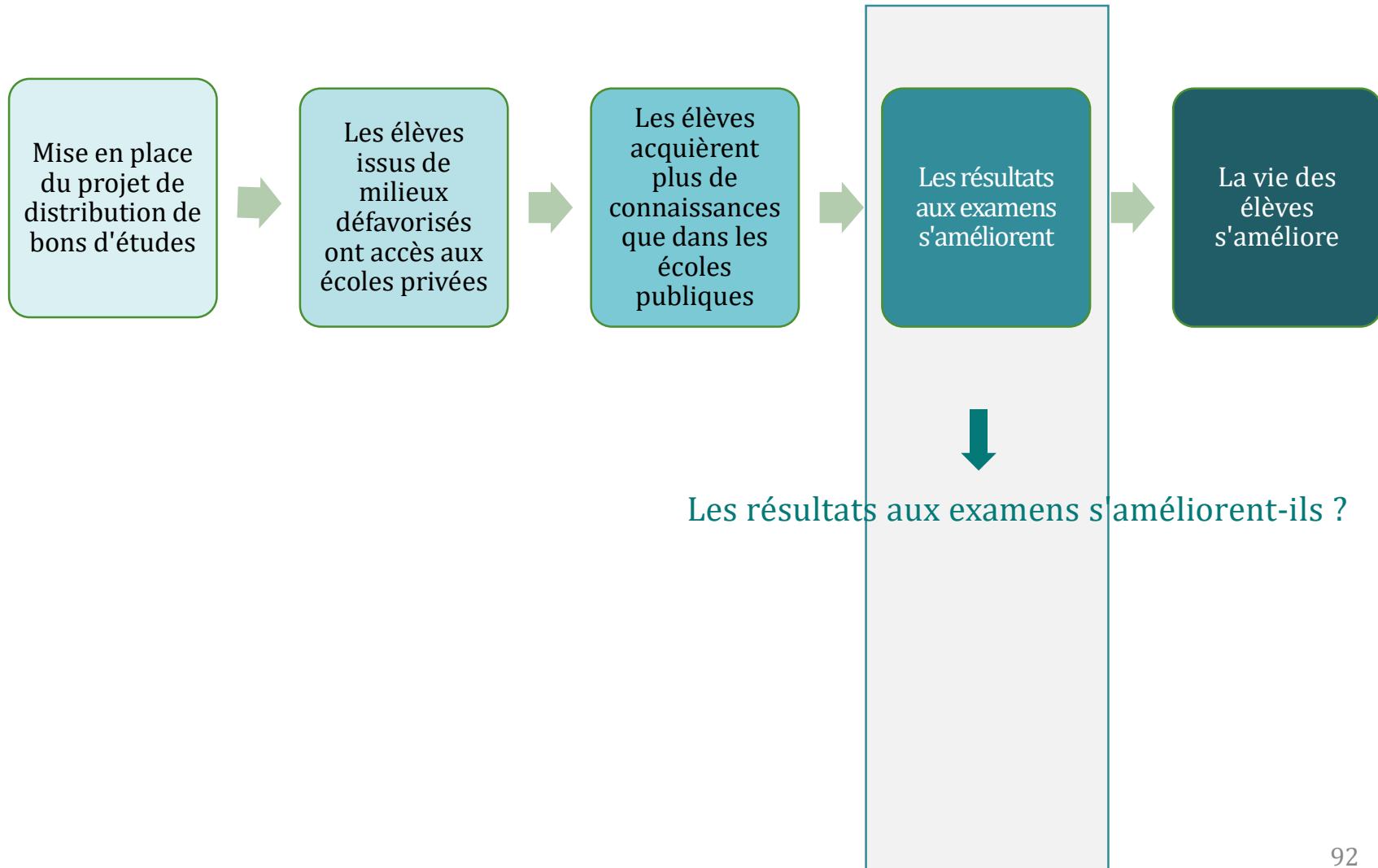
Le choix des questions d'une évaluation d'impact



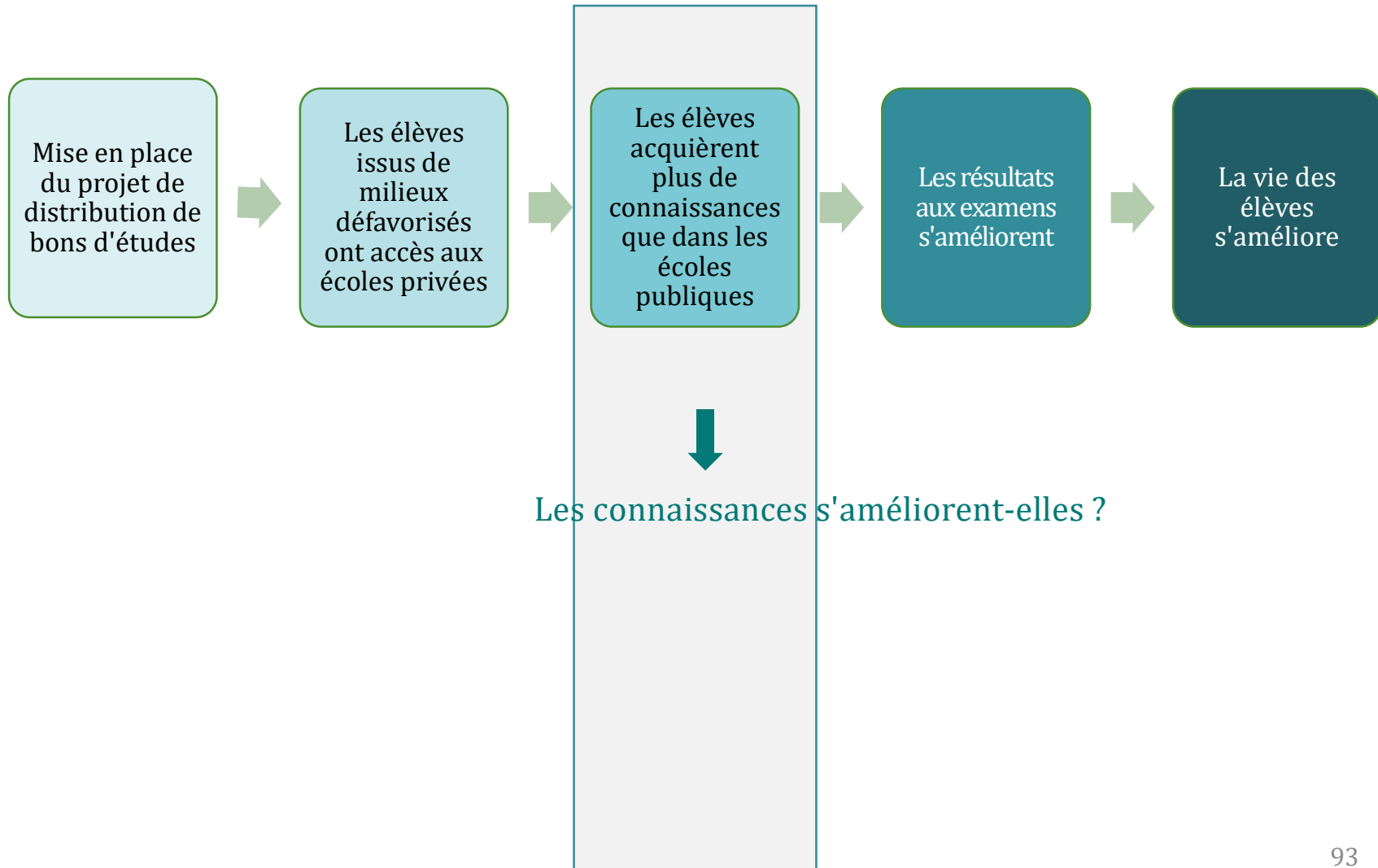
Le choix des questions d'une évaluation d'impact



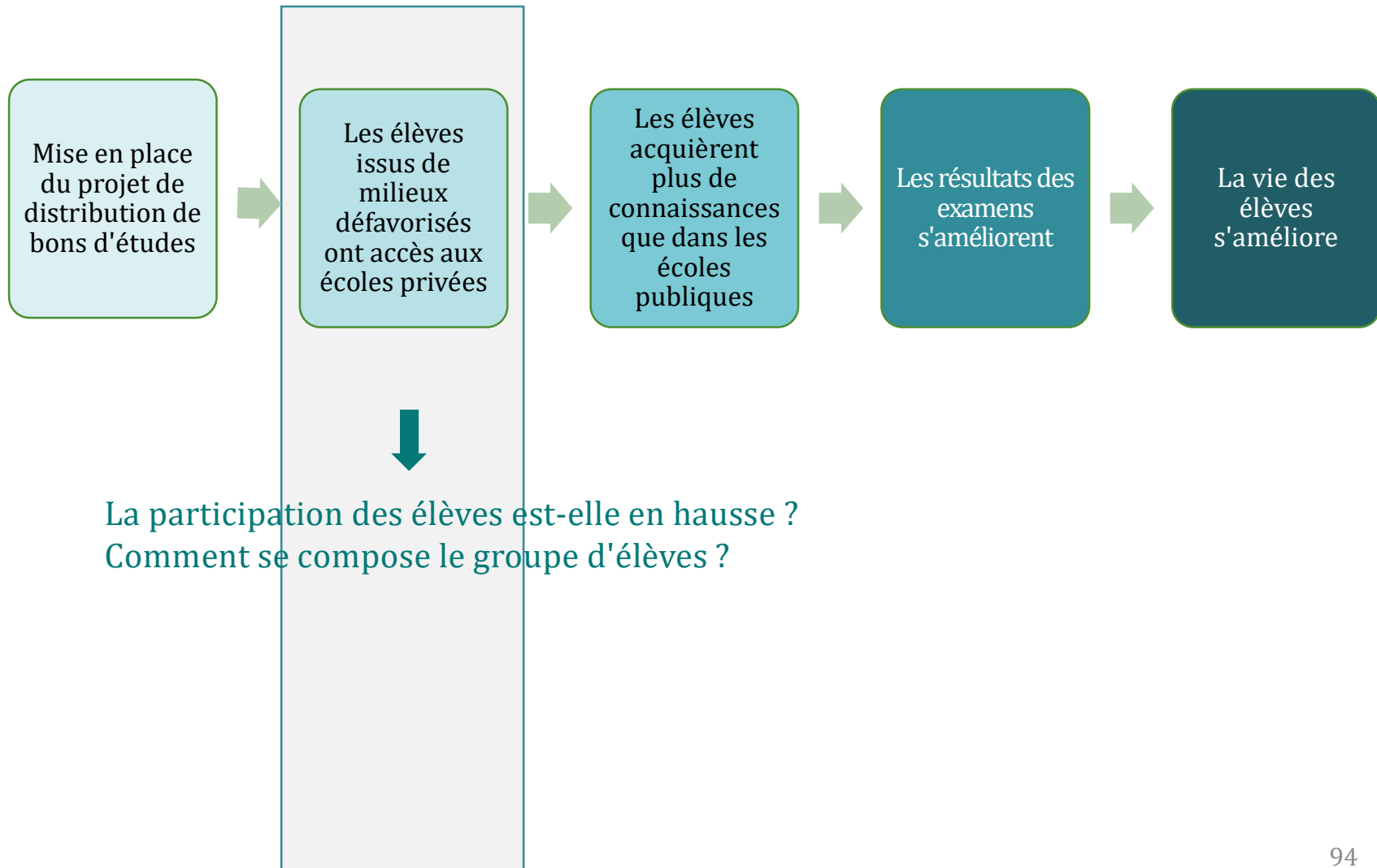
Le choix des questions d'une évaluation d'impact



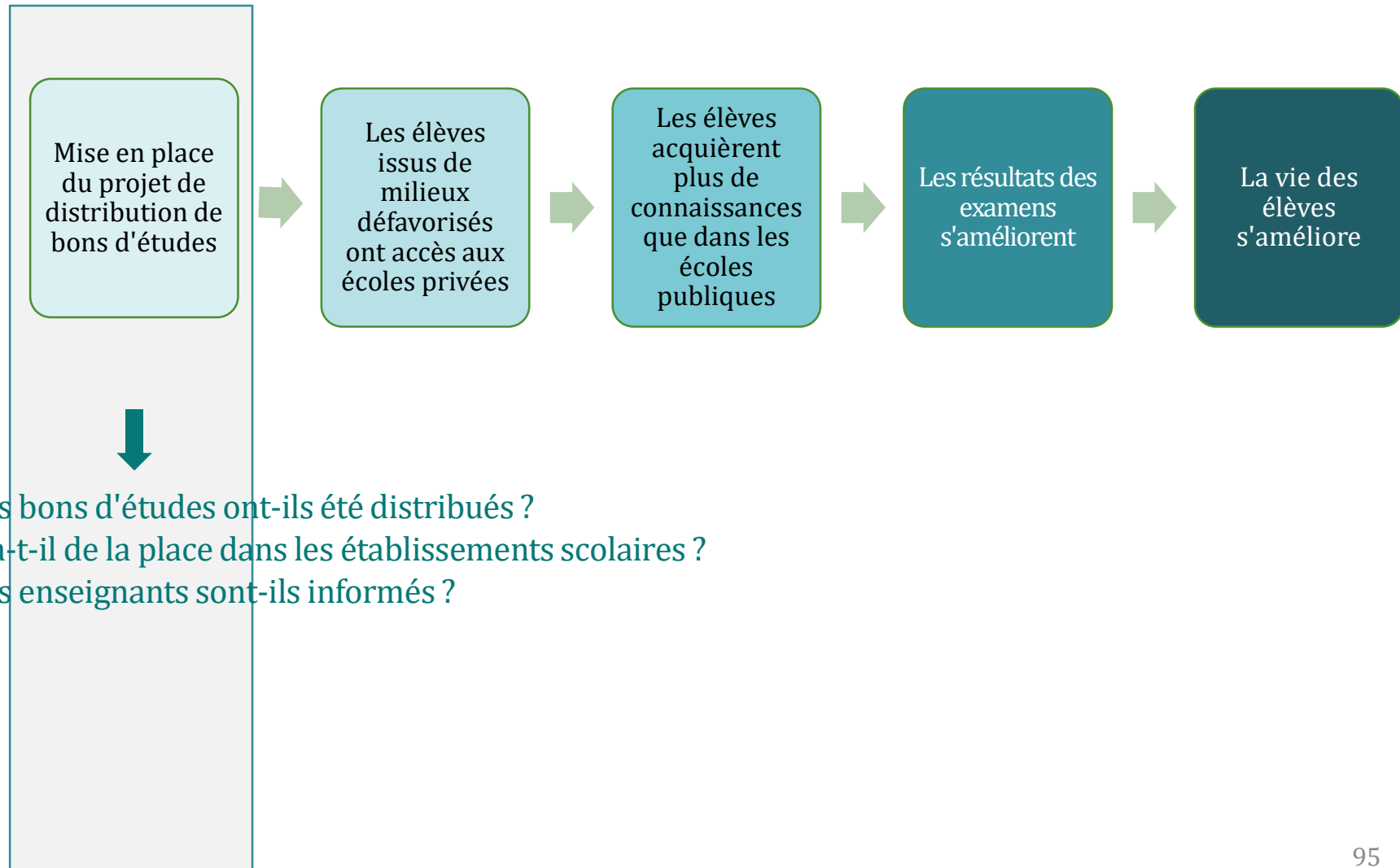
Le choix des questions d'une évaluation d'impact contrefactuelle



Le choix des questions d'une évaluation d'impact



Le choix des questions d'une évaluation d'impact



Définition des questions et indicateurs clés d'une évaluation

1. L'évaluation d'impact requiert l'identification des bonnes questions d'évaluation à examiner

2. Mesurer l'impact d'une intervention nécessite la définition d'indicateurs pertinents

- Objectifs didactiques :

- 1. élaborer les questions d'évaluation

- 2. élaborer et peaufiner des indicateurs

spécifiques, quantifiables, imputables, réalistes et ciblés

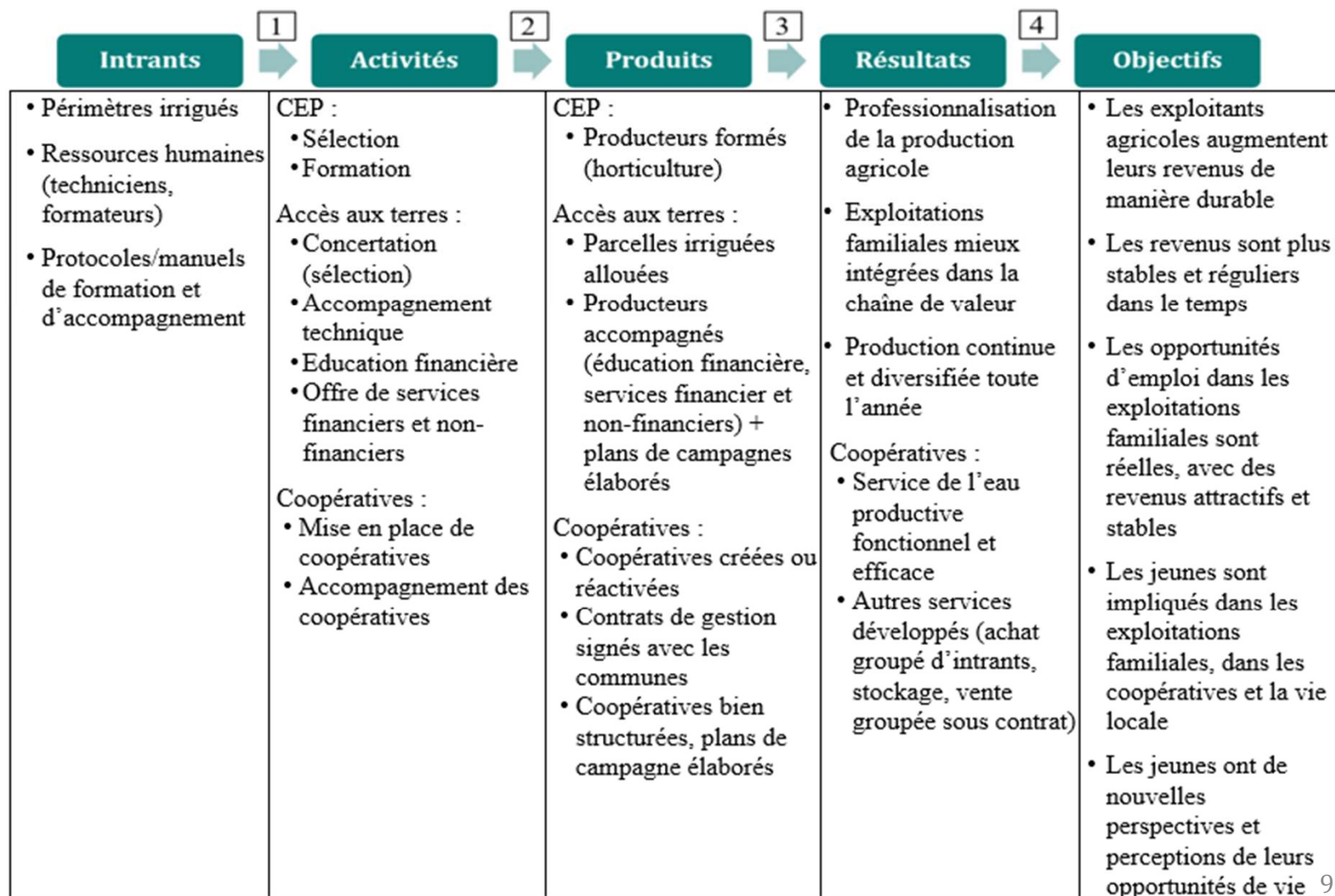
(« SMART »)



Définition des questions et indicateurs clés d'une évaluation

- *Phase 1: Identifier les interventions du projet et leurs objectifs*
- *Phase 2: Identifier les éléments de preuve existants relatifs à ces interventions*
- *Phase 3: Vérifier les hypothèses sous-jacentes de la TdC*
- *Phase 4: Définir les objectifs principaux de l'évaluation d'impact*
- *Phase 5: Valider et peaufiner les questions d'évaluation avec les parties prenantes*

Projet de Recherche sur l'Evaluation d'Impact Economique du FFU au Sénégal: TdC

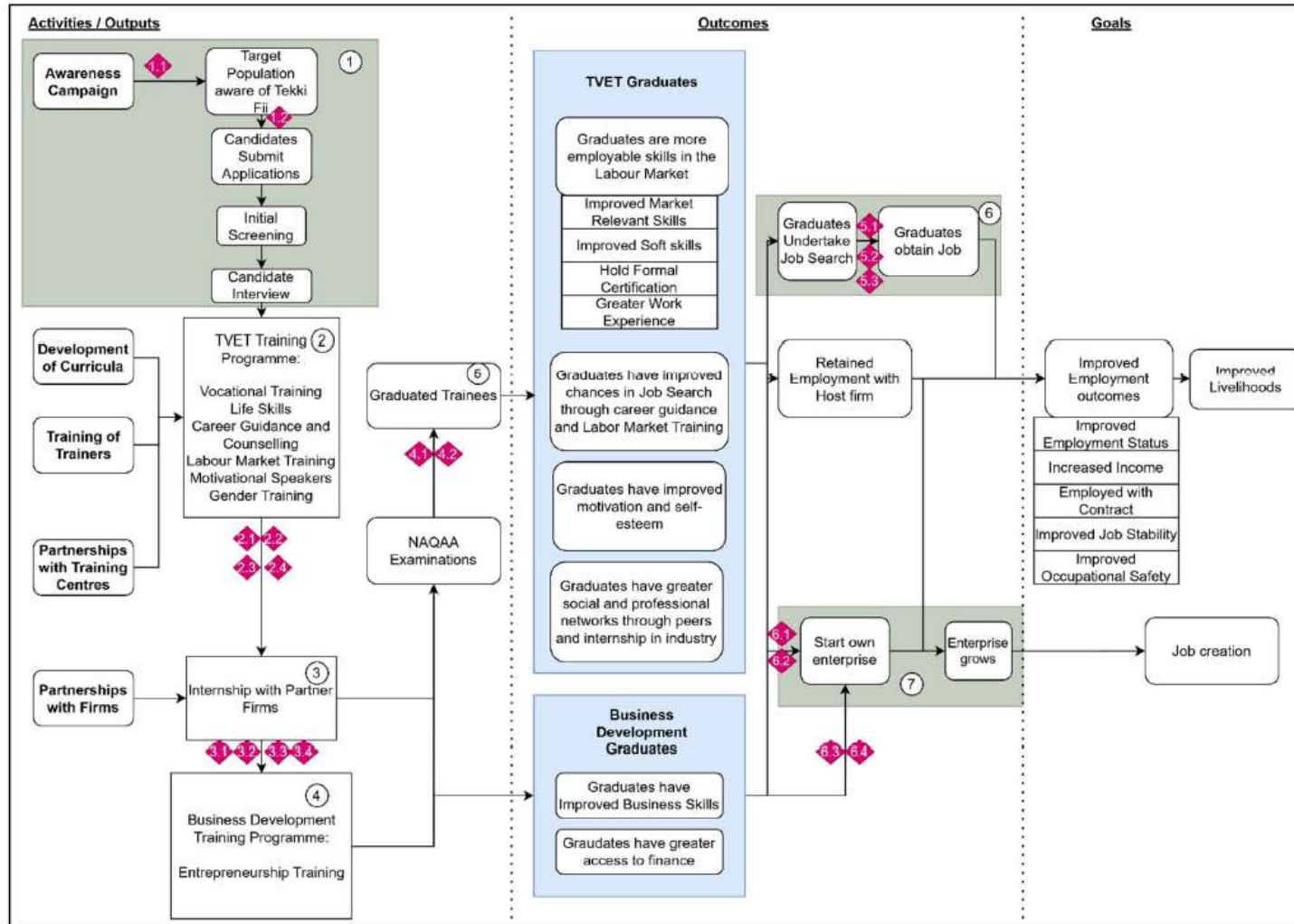


Projet de Recherche sur l'Evaluation d'Impact Economique du FFU au Sénégal: TdC

Hypothèse sous-jacentes

1	• Une promotion adéquate du projet est nécessaire pour garantir l'implication des communautés dans les processus de concertation/sélection et la participation des bénéficiaires aux dispositifs d'accompagnement non financier/financier • La promotion du projet peut aussi renforcer/créer l'intérêt des populations cibles pour une agriculture plus intensive
2	• Les concertations locales doivent être efficaces dans l'allocation des parcelles aux exploitants • Les exploitants sélectionnés ont un intérêt pour une agriculture plus intensive • Les exploitants sélectionnés participent aux formations • Les processus de création et de mise en place des coopératives sont efficaces • L'accompagnement offert aux coopératives est adapté à leurs besoins
3	• L'accompagnement offert aux bénéficiaires est adapté à leurs besoins • Les coopératives fonctionnent bien, gèrent correctement les périmètres irrigués et assurent efficacement le service de l'eau productive • Les dispositifs d'achat groupé d'intrants, de stockage et de vente groupée sous contrat développés par les coopératives sont efficaces et fiables pour permettre aux exploitants d'être mieux intégrés dans la chaîne de valeur • Les paysans se donnent les moyens d'entrer en campagne et de réaliser leurs campagnes selon les plans élaborés par les coopératives
4	• Les plans de campagne sont respectés • Les producteurs réalisent 2 voire 3 campagnes sur l'année • Les périmètres irrigués sont bien gérés par les coopératives de manière durable • Les concertations locales ont permis l'inclusion des jeunes dans le processus, y compris en tant que chefs d'exploitation • Les jeunes bénéficient effectivement des fruits de leur travail

Projet de Recherche sur l'Evaluation d'Impact Economique du FFU au Ghana: TdC



Projet de Recherche sur l'Evaluation d'Impact Economique du FFU au Sénégal: Questions et indicateurs d'évaluation

Questions d'évaluation d'ensemble du portefeuille	Sous-questions d'évaluation propres au pays	Dimensions	Indicateurs	Méthode d'évaluation
QE1. Dans quelle mesure les interventions du FFU contribuent-elles à l'emploi, la création d'emploi et aux compétences ?				
<i>Dans quelle mesure le PARERBA contribue-il à l'emploi, la création d'emploi et aux compétences ?</i>				
1.2. Quel sont les effets du soutien au MPME sur l'employabilité des bénéficiaires et l'accès à l'emploi (décent)?	1.2.SEN.a. L'accompagnement des exploitants agricoles aboutit-il à la création de nouveaux emplois, en particulier chez les jeunes ?	Emploi Emploi des jeunes	- Nombre de chefs d'exploitation et d'employés dans les exploitations agricoles - Nombre de jeunes chefs d'exploitation et employés dans les exploitations agricoles	Quantitative
	1.2.SEN.b. L'accompagnement des AGR/MPE aboutit-il à la création de nouveaux emplois, en particulier chez les jeunes ?		- Nombre de nouvelles entreprises créées et consolidées - Nombre de nouvelles entreprises créées et consolidées par des jeunes - Nombre de responsables et d'employés dans les MPE/AGR - Nombre de jeunes responsables et employés dans les MPE/AGR	Quantitative
	1.2.SEN.c. Quel est l'impact du projet sur la création d'emploi dans les PME accompagnées, notamment l'emploi des jeunes ?		- Nombre d'employés dans les PME - Nombre de jeunes employés dans les PME	Quantitative
	1.2.SEN.d. Quelle est la qualité des emplois créés ?	Mécanismes encourageant la création de conditions d'emploi décentes Motivation des bénéficiaires	- Perceptions sur la qualité de l'accompagnement vers un emploi de qualité - Présence de mécanismes qui encouragent la création de conditions d'emploi décentes - Motivation des bénéficiaires pour emplois de qualité et potentiels défis pour en trouver	Qualitative

Recherche sur l'évaluation de l'impact économique d'un projet EUTF en Gambie : Questions et indicateurs d'évaluation

Questions d'évaluation d'ensemble du portefeuille	Sous-questions d'évaluation propres au pays	Dimensions	Indicateurs	Méthode d'évaluation
QE2. Dans quelle mesure les interventions du FFU ont-elles affecté le niveau de vie et la résilience des bénéficiaires?				
<i>Dans quelle mesure le PARERBA a-t-il affecté le niveau de vie et la résilience des bénéficiaires?</i>				
2.2. Quels sont les effets du soutien au MPME sur le niveau de vie et la résilience ?	2.2.SEN.a. Les revenus des bénéficiaires du projet augmentent-ils et sont-ils plus stables ?	Niveau et stabilité du revenu	- Revenu mensuel - Dépenses mensuelles - Variabilité du revenu - Variabilité des dépenses	Quantitative
	2.2.SEN.b. Quel est l'impact du projet sur le niveau moyen et la stabilité du revenu des jeunes ?	Niveau et stabilité du revenu	- Revenu mensuel des jeunes - Dépenses mensuelles des jeunes - Variabilité du revenu des jeunes - Variabilité des dépenses des jeunes	Quantitative

Session 3 : Introduction aux méthodes expérimentales

Conception de l'évaluation, méthodes d'EIC

C4ED – EUTF

Septembre 2021

Lauréats du prix « Nobel » d'économie en 2019



Abhijit Banerjee,
Esther Duflo,
Michael Kremer



2019 Sveriges Riksbank Prize
in Economic Sciences in
Memory of Alfred Nobel



- « L'un des enjeux les plus pressants pour l'humanité est la **réduction de la pauvreté mondiale**, sous toutes ses formes. **Abhijit Banerjee, Esther Duflo, et Michael Kremer** ont introduit une **nouvelle approche** permettant d'obtenir des réponses fiables quant aux meilleurs moyens de lutter contre la pauvreté dans le monde. **Elle consiste à subdiviser cette problématique en questions plus simples, plus faciles à aborder.** Ils ont réussi, depuis le milieu des années 1990, à tester une série d'interventions dans différents domaines au moyen d'expériences de terrain, par exemple pour les résultats scolaires ou la santé des enfants.»

Source: The Royal Swedish Academy (2019), *The Prize in Economic Sciences 2019*, Press Release

Durant cette session :

- Comprendre toute la richesse des méthodes expérimentales et explorer différentes méthodes
- Comprendre les limitations de ces méthodes
- Il vous faudra une pièce de monnaie ...

Essai Comparatif Randomisé (ECR)

Objectif : Simuler une situation « **contrefactuelle** »

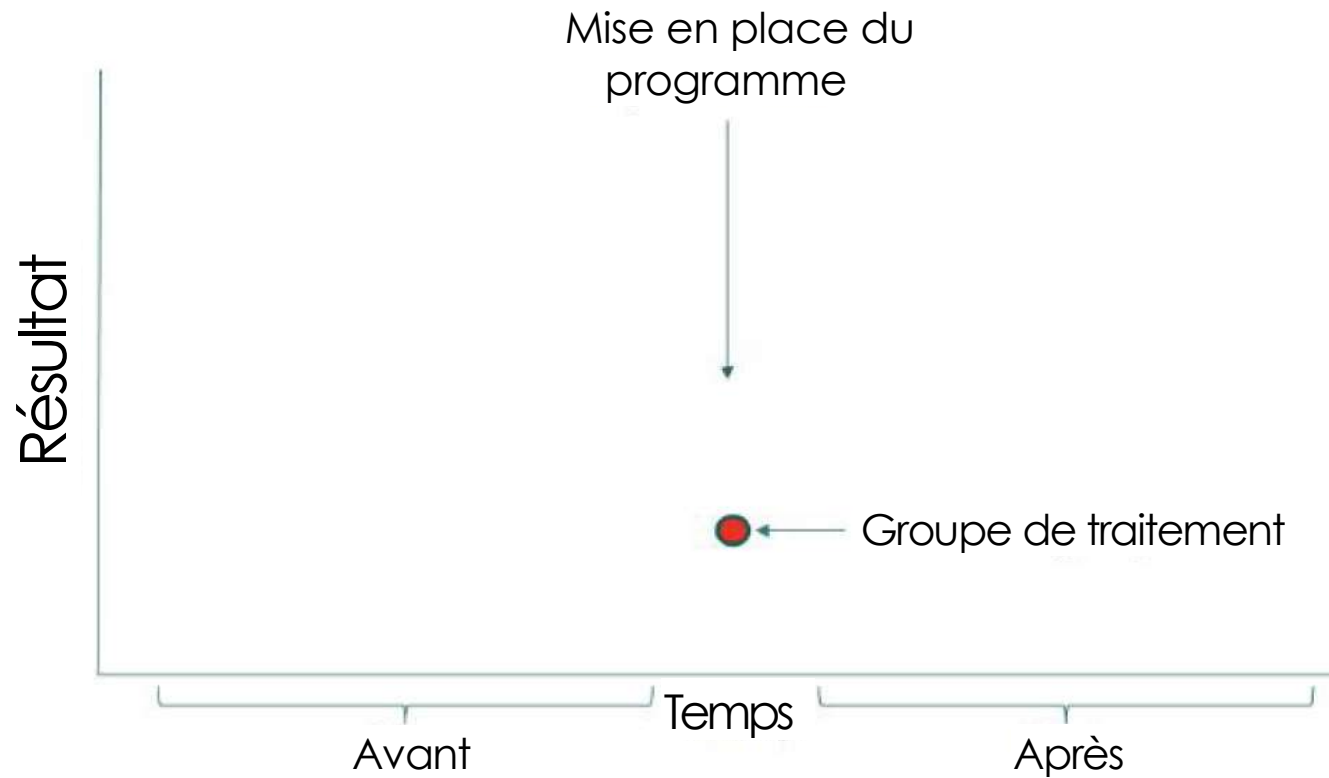
- Que se serait-il passé si le programme n'avait pas existé ?
- **Meilleur moyen de trouver un contrefactuel :**
Essai Comparatif Randomisé (ECR)
- **Idée :** Allouer aléatoirement le programme à un groupe de traitement et un groupe de contrôle/comparaison
→ le groupe de contrôle/comparaison « imite » le résultat contrefactuel du groupe de traitement

Essai Comparatif Randomisé (ECR)

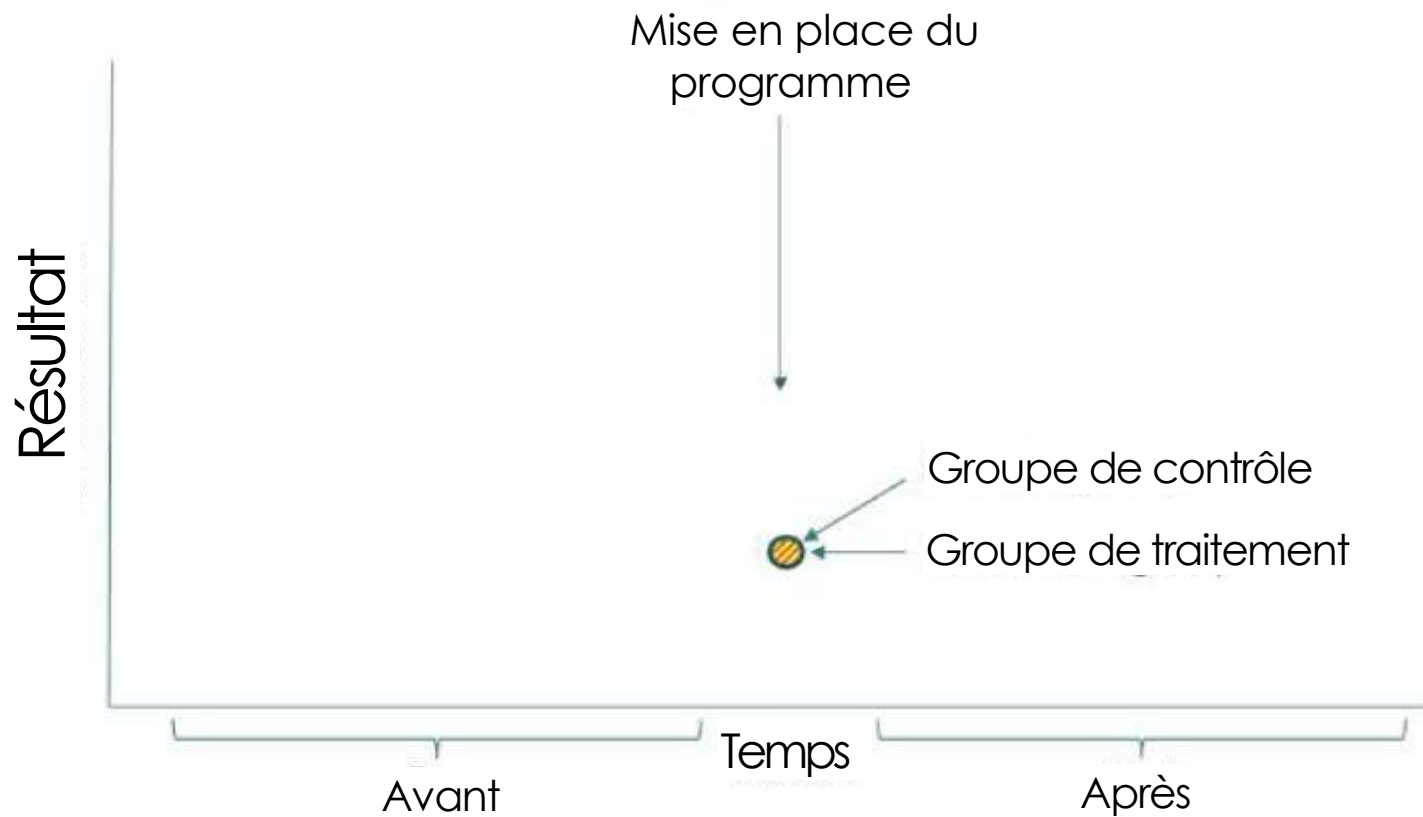
Pourquoi l'ECR est-il considéré comme **l'idéal** des évaluations d'impact quantitatives ?

- Pas nécessaire de formuler des hypothèses sur les niveaux de résultats ou les tendances entre les groupes si ceux-ci sont aléatoires au sein des groupes
- Fournit des preuves solides
- Facile à interpréter

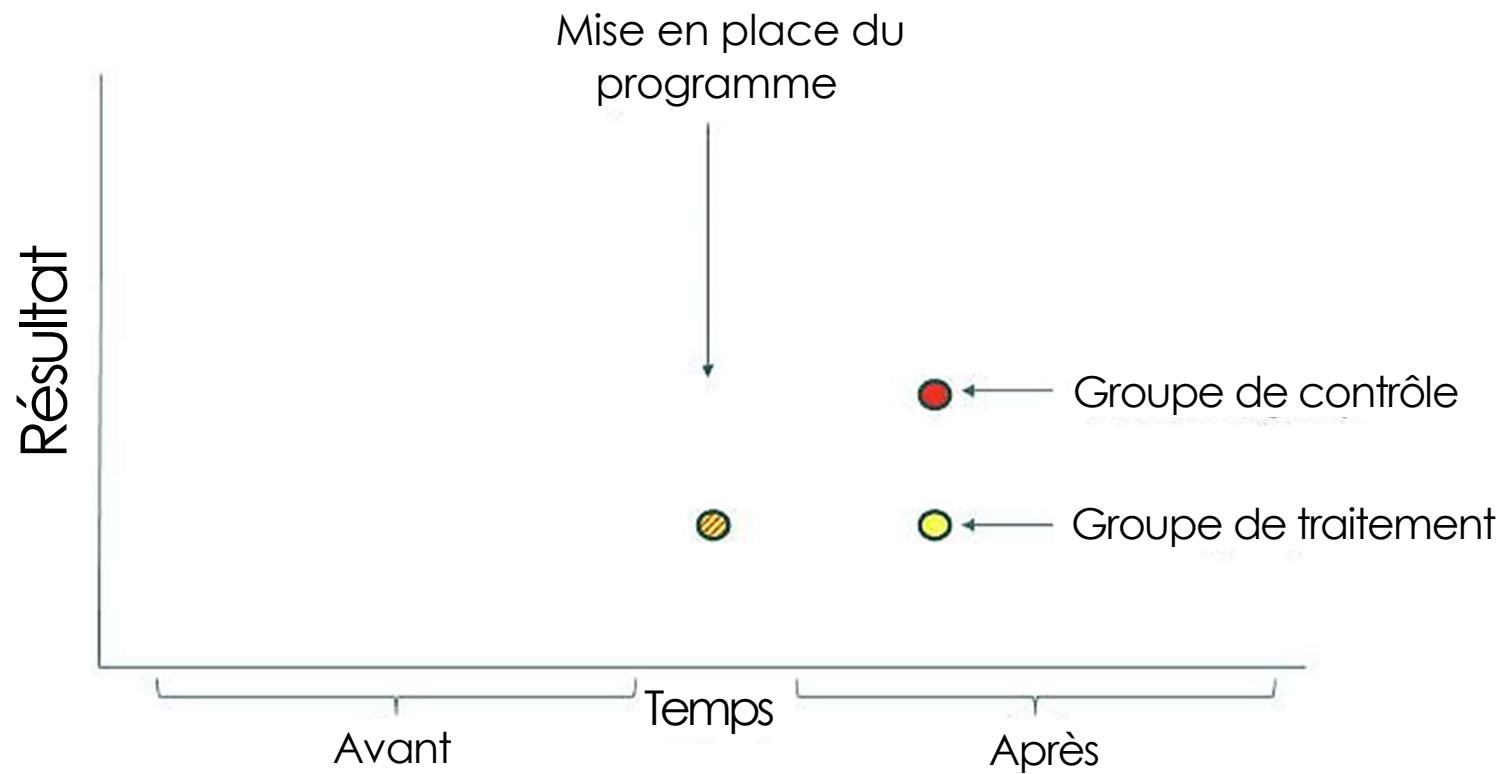
Essai Comparatif Randomisé (ECR)



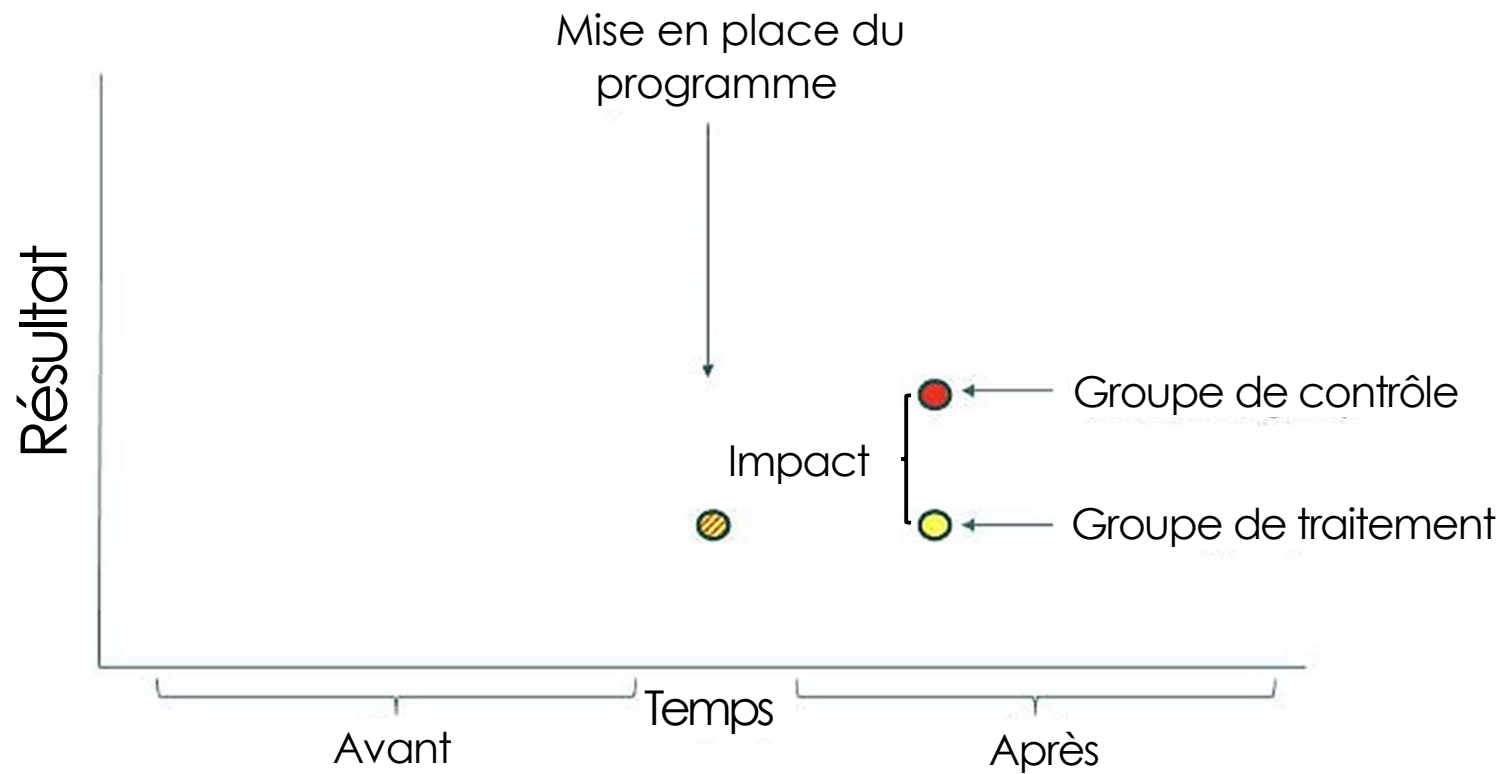
Essai Comparatif Randomisé (ECR)



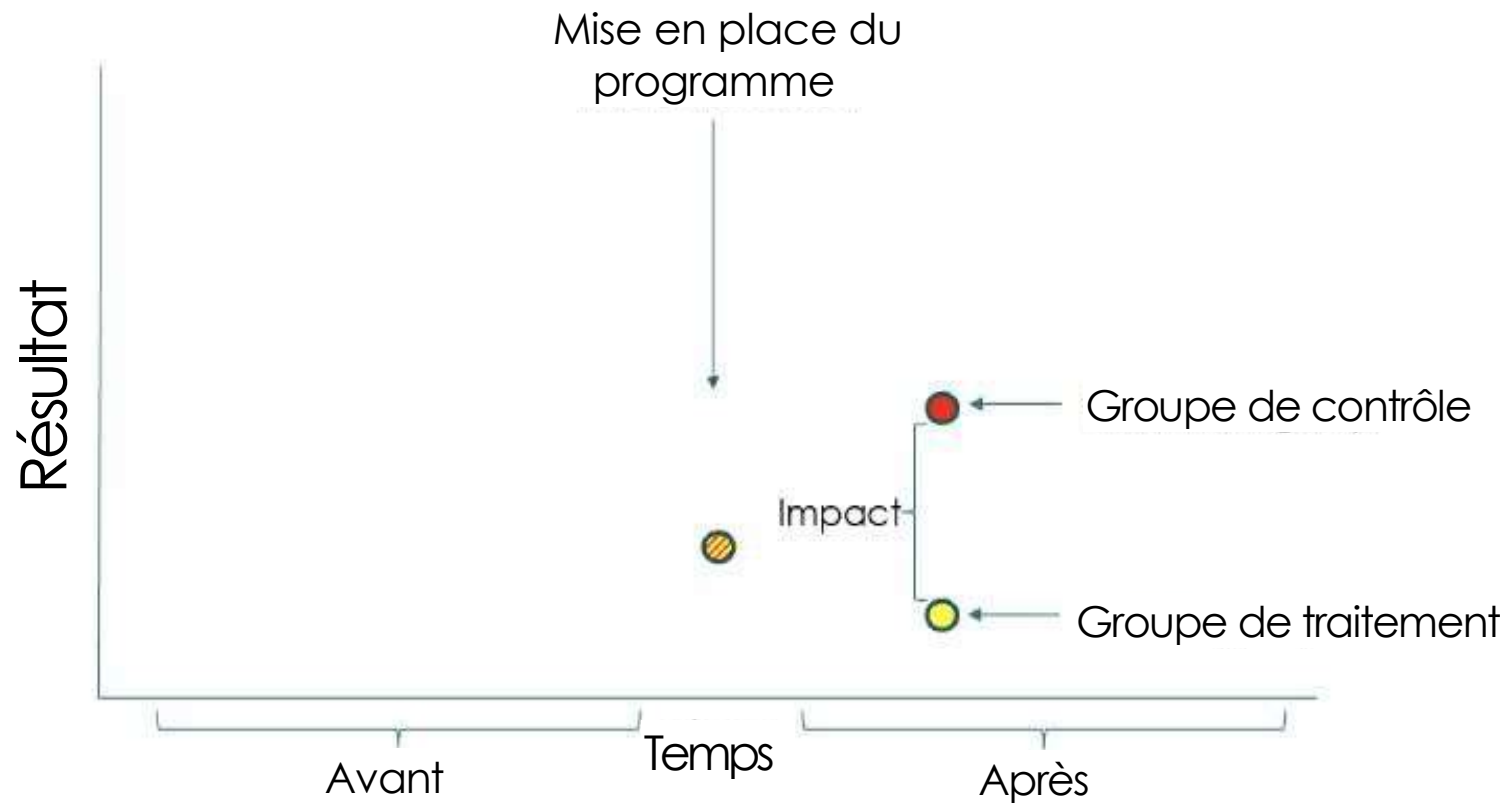
Essai Comparatif Randomisé (ECR)



Essai Comparatif Randomisé (ECR)



Essai Comparatif Randomisé (ECR)



Comment mettre en œuvre un ECR :

White, H., & Raitzer, D. A. (2017)

1. Identifier les questions d'évaluation
2. Isoler le(s) traitement(s) souhaité(s)
3. Considérer les effets de contagion
4. Déterminer le niveau de randomisation, de traitement et d'analyse
5. Décider du type de randomisation
6. Identifier le groupe de personnes éligibles au programme
7. Sélectionner l'échantillon à analyser
8. Randomiser
9. Collecter les données de référence et vérifier l'équilibre des groupes
10. Garantir l'intégrité du cadre d'évaluation et faire le suivi du programme

Pourquoi randomiser est une bonne idée

- Les ECR offrent **les meilleures données possibles sur la causalité**, dans la mesure où ils soient bien conçus
- Les ECR sont **faciles à analyser**
- La randomisation est un moyen **transparent et juste** de répartir les bénéfices d'une intervention

Toujours sceptique?

- Il existe des modèles sans groupe de contrôle « pur »
- La mise en place de votre projet peut déjà correspondre à la conception d'un ECR (par phases)
- La randomisation peut être appliquée seulement à un sous-échantillon du groupe ciblé

Niveau de randomisation

- Le niveau d'observation auquel on alloue le traitement de manière aléatoire
- Par exemple: individu, ménage, école, village, entreprise, etc.

Quelques situations propices aux essais randomisés

- Quand la demande excède l'offre
- Quand une innovation ne peut pas être proposée à tous les bénéficiaires en même temps
- Quand les unités d'observation expérimentales peuvent être isolées dans le temps
- Quand les unités d'observation expérimentales sont séparées dans l'espace ou que la communication entre unités est limitée
- Quand un changement est nécessaire et qu'on ne sait pas si les solutions marchent
- Quand certaines personnes n'ont pas de préférence entre les alternatives
- Quand on a le contrôle sur les unités d'observation expérimentales
- Quand il y a un tirage au sort

Source: SHADISH, W.R., COOK, T.D. ET CAMPBELL, D.T., (2002)

Quand un ECR doit-il être élaboré et mis en place?

Avant le début du programme

Une fois que le programme a démarré

Après la fin du programme

Quand un ECR doit-il être élaboré et mis en place?

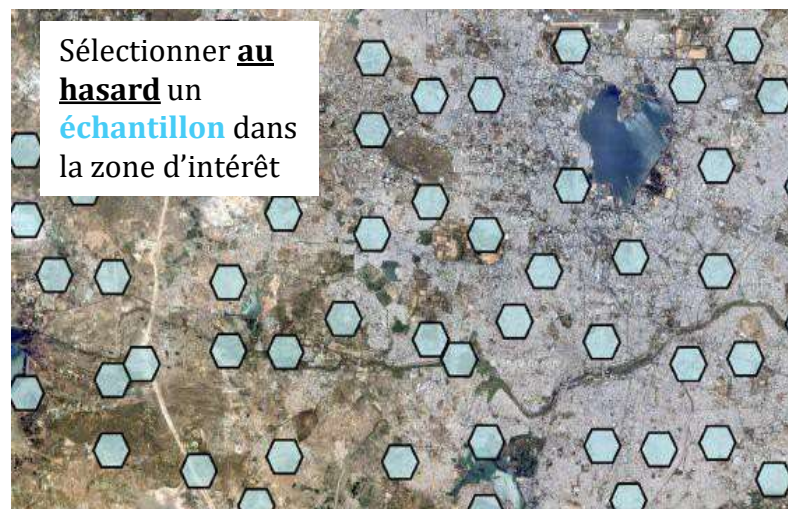
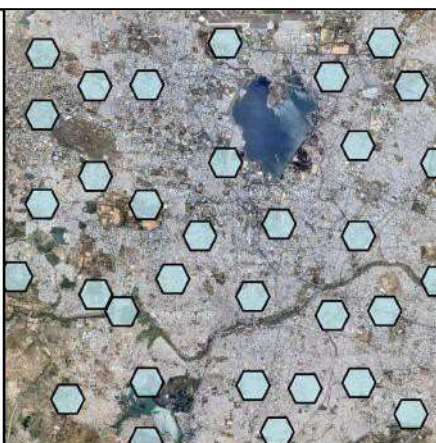
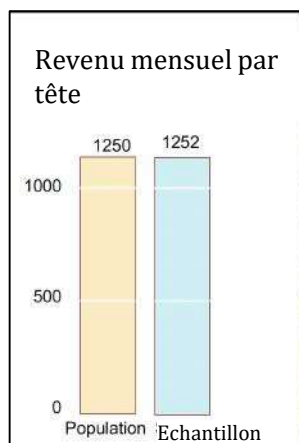
Avant le début du
programme

Une fois que le
programme a
démarqué

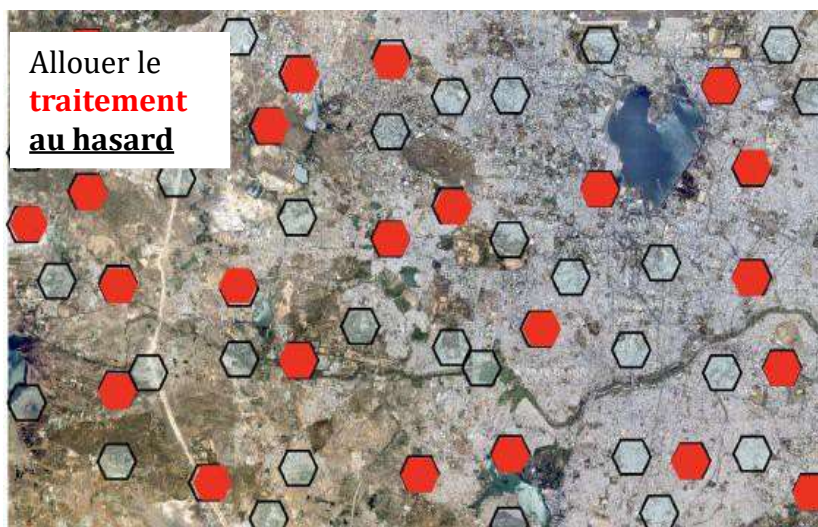
Après la fin du
programme

Les ECR en PRATIQUE

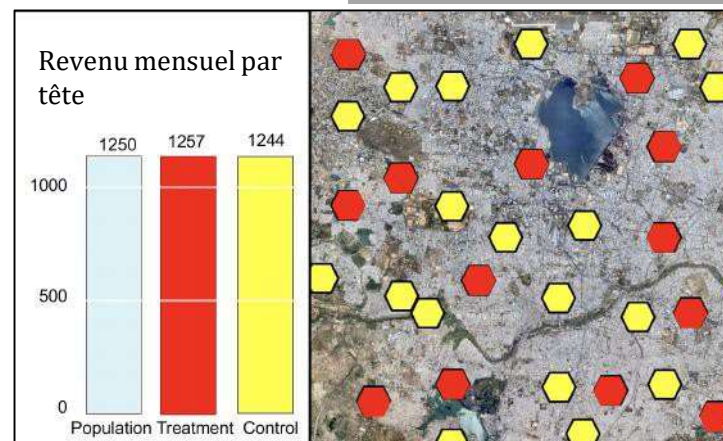
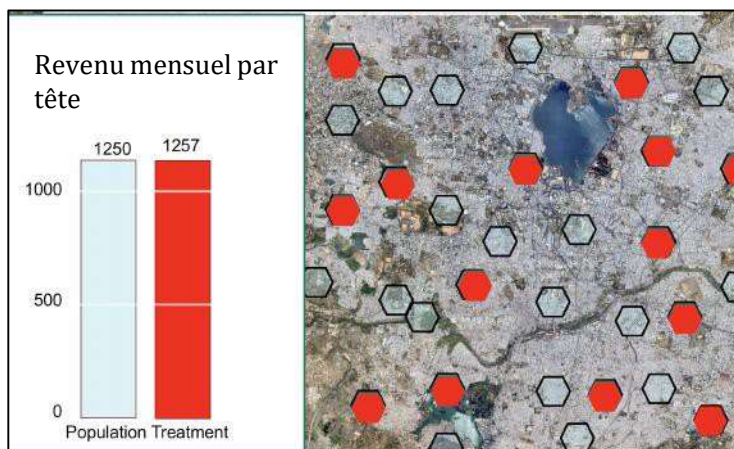
Echantillonnage ALÉATOIRE d'une sous-population pour l'étude



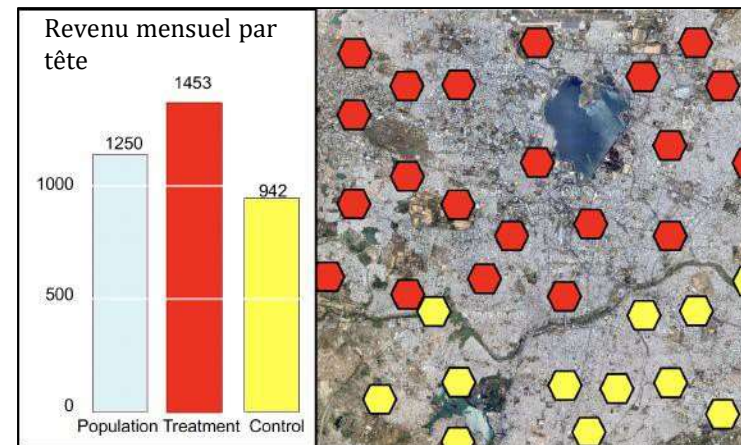
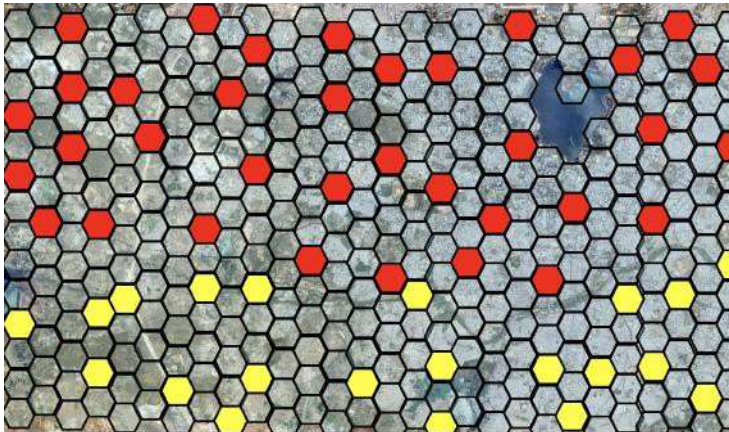
Sélection/allocation ALÉATOIRE des bénéficiaires et non bénéficiaires du programme



Le reste de l'échantillon forme le **groupe de comparaison**



Allocation non aléatoire (possiblement)



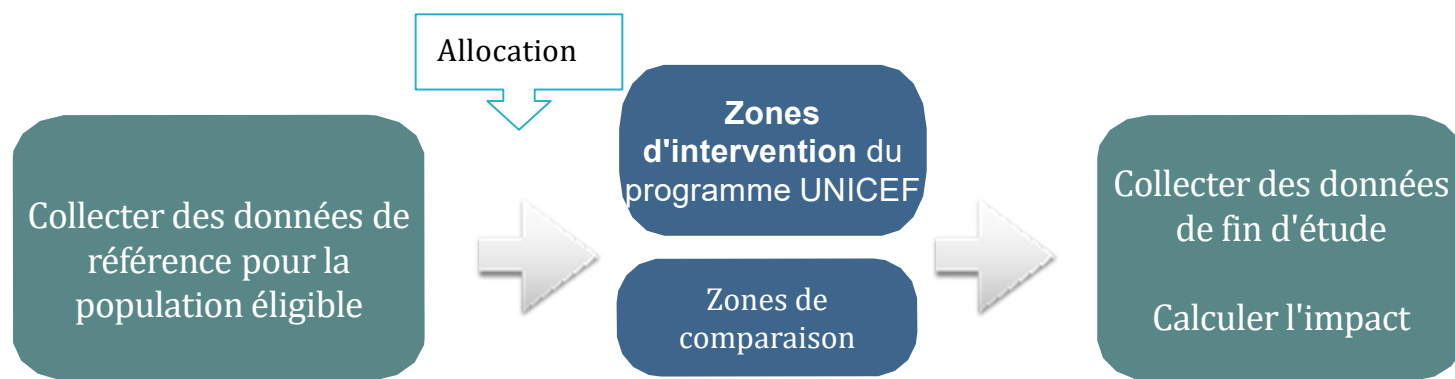
Un ECR en pratique

Collecter des données de
référence pour la
population éligible

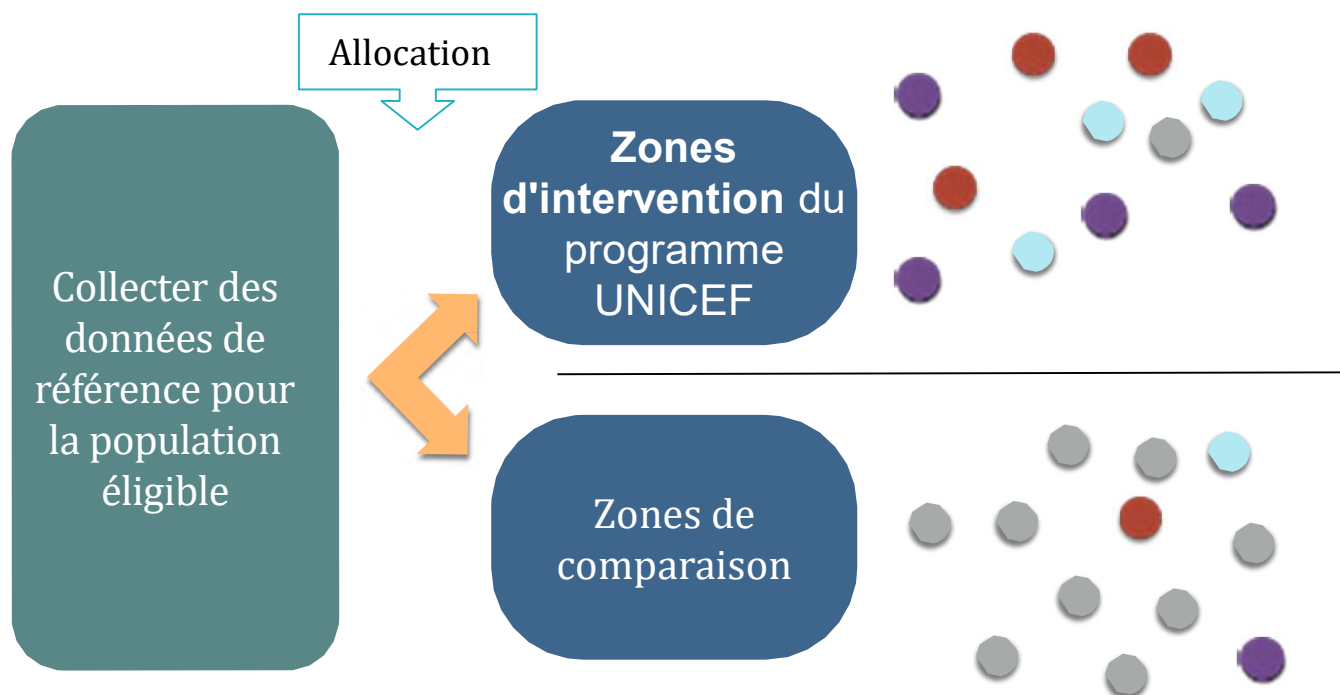
Programme
UNICEF-IKEAF

Collecter des données
de fin d'étude
Calculer l'impact

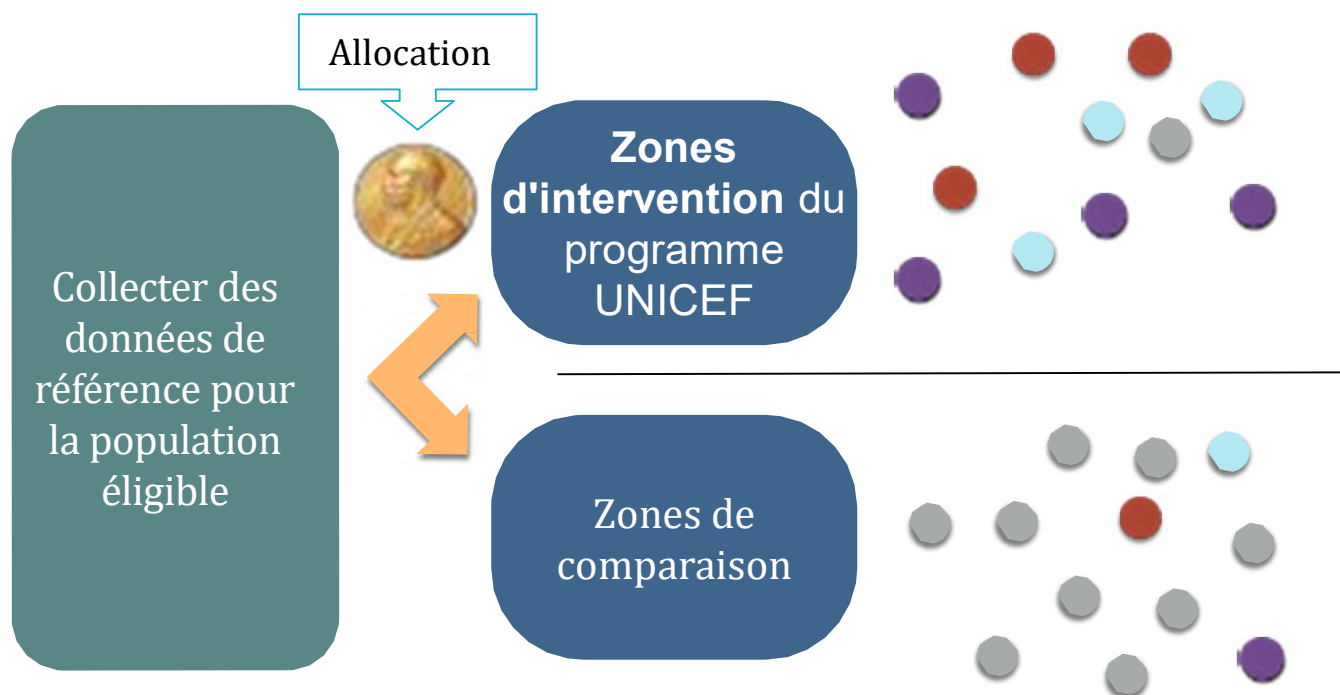
Un ECR en pratique



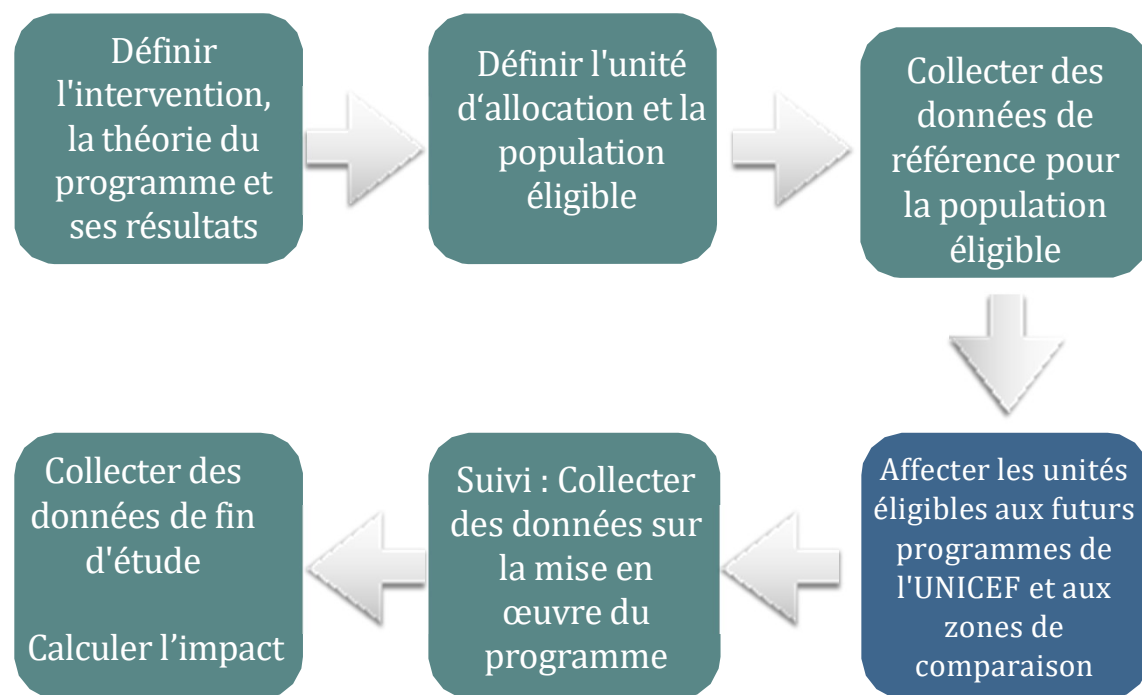
Un ECR en pratique



Un ECR en pratique



Un ECR en pratique



Source: White et al. (2014) Randomized Controlled Trials (RCTs), Methodological Briefs: Impact Evaluation n° 7, Office de la recherche de l'UNICEF, Florence. Modifié par les auteurs.

Les ECR : Diverses approches

Comment **RANDOMISER** le groupe de traitement et le groupe de contrôle/comparaison ?

1. Tirage au sort
2. Groupes de traitement multiples
3. Introduction progressive aléatoire (par phases)
4. Intensités de traitement aléatoires / ECR à étapes multiples
5. Approche par « encouragement »

1. Tirage au sort

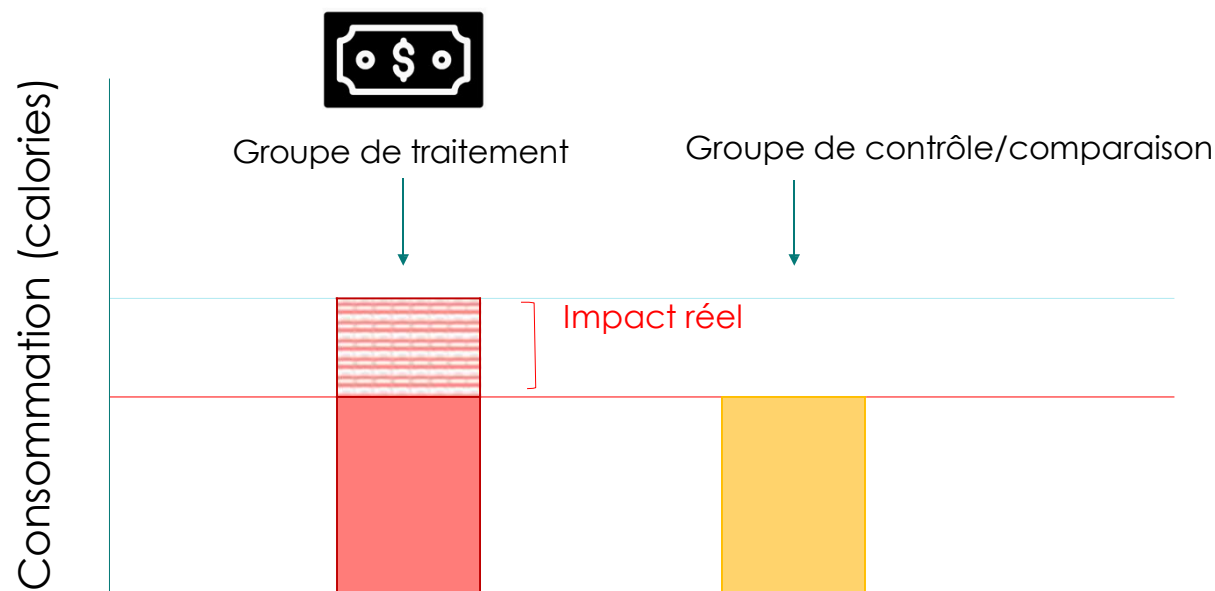
- Sélection aléatoire des unités de traitement (personnes, rues, communautés) par tirage au sort
- Deux formes de tirage au sort :
 - Public: tiré d'un chapeau ou d'un seau, équitable et transparent, mais susceptible de provoquer des changements comportementaux dans le groupe de contrôle/comparaison
 - Privé: Moins transparent, mais permet d'atténuer les réactions comportementales du groupe de contrôle

Tirage au sort: exemple IDinsight (2020)

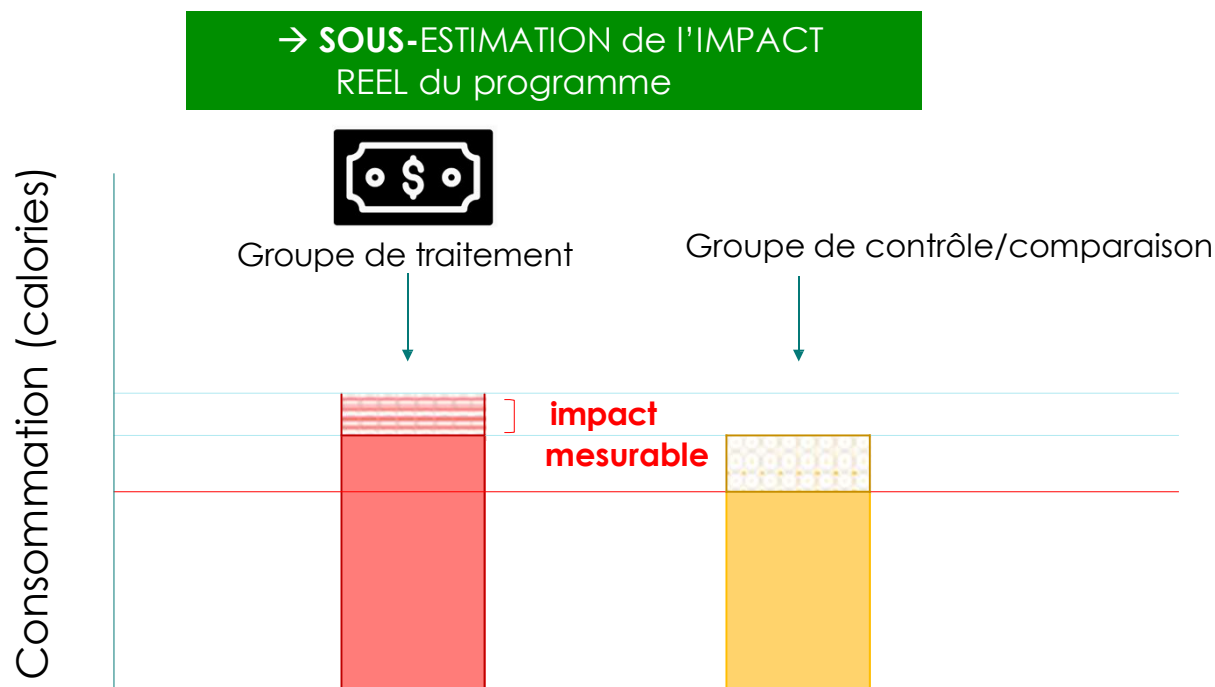
- Transfert monétaire unique d'une valeur de 1 000 USD dans un camp de réfugiés en milieu rural



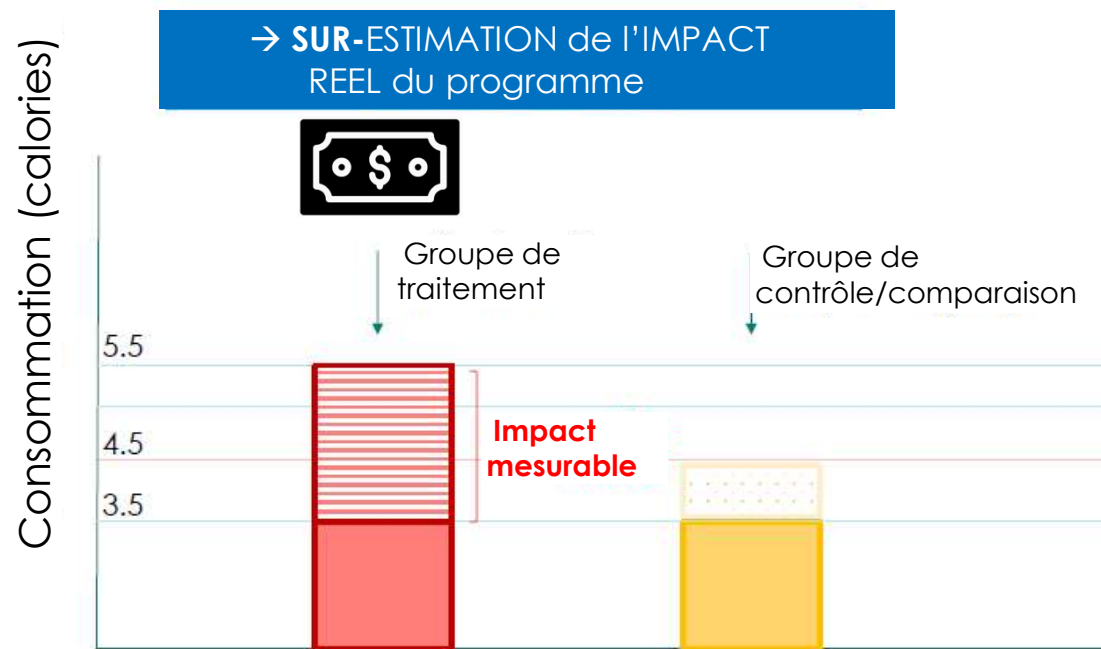
Pourquoi les réponses comportementales du groupe de contrôle ont-elles de l'importance ?



Pourquoi les réponses comportementales du groupe de contrôle ont-elles de l'importance ?



Pourquoi les réponses comportementales du groupe de contrôle ont-elles de l'importance ?

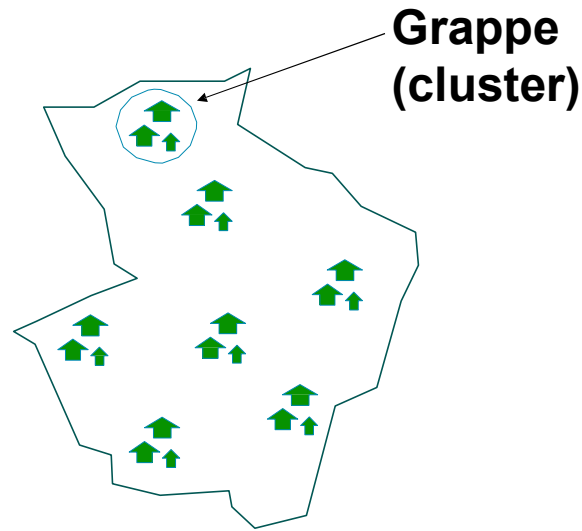


Tirage au sort

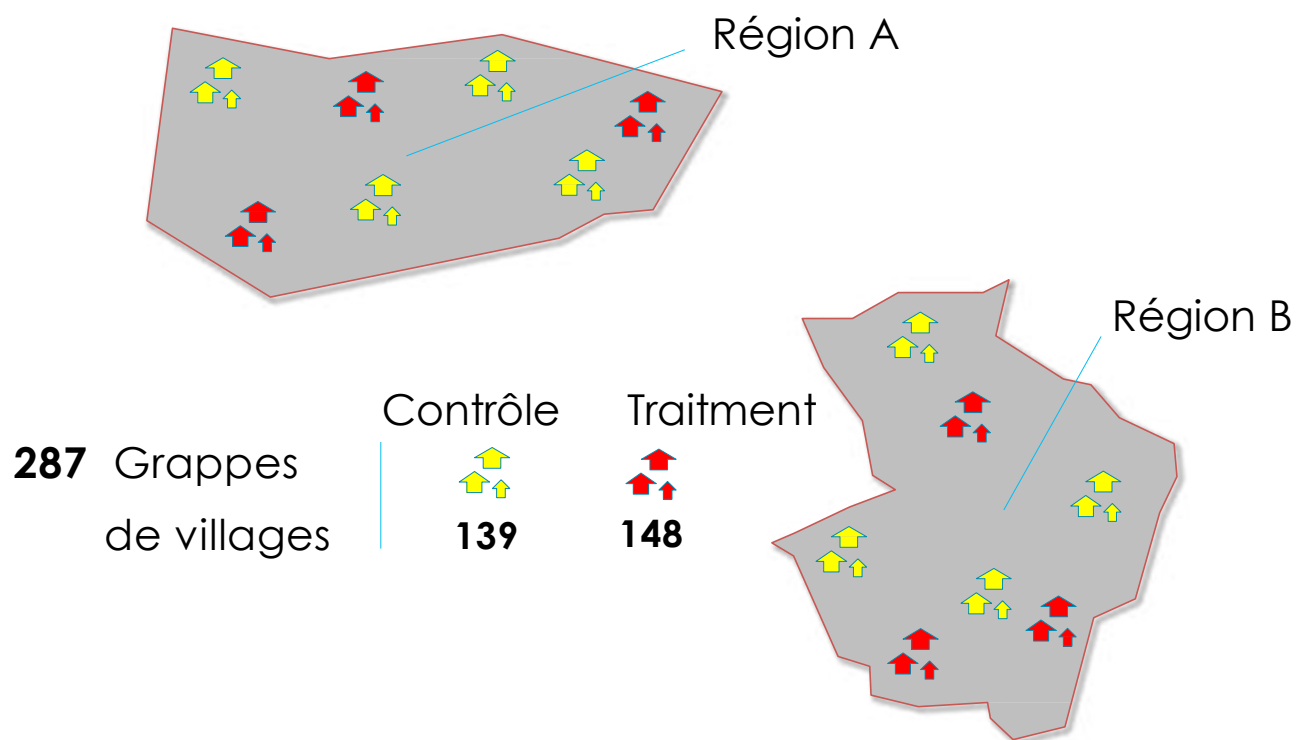
- **Avantage:** Transparent et équitable
- **Inconvénient:** Le groupe témoin ou de comparaison est susceptible de changer également de comportement dans le cas d'un tirage au sort public ; nécessité d'un échantillon de grande taille ; difficile à stratifier

Allocation aléatoire par grappes

- Considérer des groupes d'unités (grappes, ou “cluster”) pour une mise en place plus facile et pour prendre des marges de sécurité
- Allocation du programme à un niveau agrégé, même si les données sont collectées à un niveau plus fin



Allocation aléatoire par grappes



2. Groupes de traitement multiples

- Affecter aléatoirement des unités (personnes, communautés, etc.) à différents groupes de traitement (versions du programme ; types d'interventions) et au groupe de contrôle
 - Analyser les différentes versions et composantes d'un programme les unes par rapport aux autres et par rapport au groupe de contrôle ou de comparaison
- Possible également en l'absence de groupe de contrôle/de comparaison

Groupes de traitement multiples - Exemple

Intervention pédagogique dispensée à travers deux options:

- Visite en personne
- Appel téléphonique

Groupes de traitement multiples - Exemple

- Les combinaisons d'interventions possibles

	Appel téléphonique	Pas d'appel téléphonique
Visite	 	
Pas de visite		

Groupe de traitement multiples

Très utile en cas d'incertitude, notamment lorsque l'on veut déterminer quelle intervention est la meilleure pour atteindre ses objectifs

- **Avantage:** Donne une meilleure idée des canaux de transmission de l'impact plutôt que de simplement répondre à la question de savoir si l'intervention est efficace ou non
- **Inconvénient :** Nécessite un échantillon de plus grande taille

Groupes de traitement multiples

Exemple : Sensibilisation pendant la pandémie de la COVID-19

Les normes sociales, la motivation et les habitudes ont un rôle crucial à jouer dans l'adoption d'un comportement sain sur le plan de l'hygiène

- L'échange de SMS avec une grande figure publique : le destinataire et son entourage observent davantage les recommandations nationales (Banerjee et al. 2020 ; Bengale Occidental, Inde)

25 millions de
SMS envoyés au
Bengale
Occidental



Académique célèbre porte des messages liés au COVID-19

-  Encourage à signaler les symptômes aux autorités de santé locales
-  Met l'accent sur un geste barrière spécifique (2 mètres, masque, hygiène)
-  Un message sur le risque pour les autres
-  Expliquer que l'ostracisation des gens infectés est inacceptable

Banerjee et al. 2020



Le rôle essentiel des messages émis par des personnes crédibles

- Les comportements et les connaissances ne vont pas forcément de paire : d'importants changements dans la distanciation, l'hygiène et le port du masque malgré des changements minimales ou nuls dans les connaissances
- la teneur exacte du message peut être sans importance ; l'essentiel, c'est d'inciter à y prêter attention

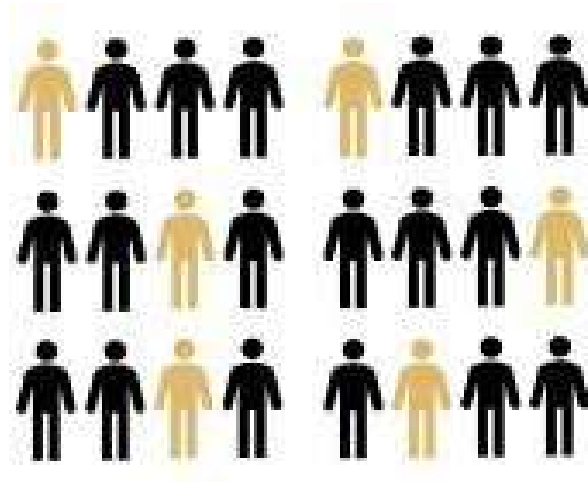
3. ECR par phases (introduction progressive)

Tout le monde reçoit le traitement, mais à différentes périodes

- Ceux qui reçoivent le traitement en dernier constituent le groupe de comparaison
- Il faut laisser un délai suffisant entre les différentes phases pour que les effets se développent

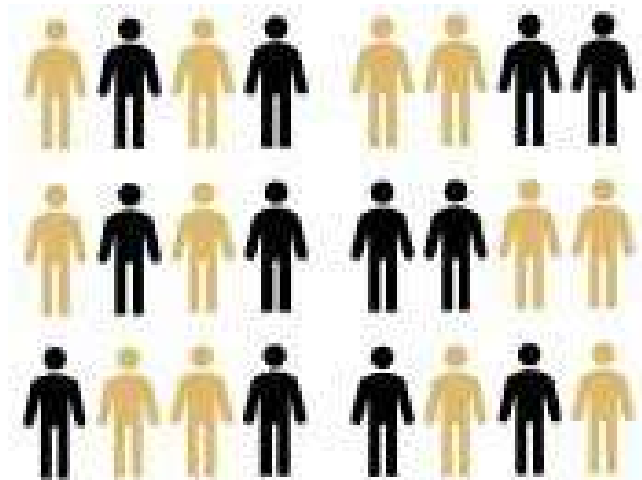
ECR par phases (introduction progressive)

- Phase 1: **SELECTION AU HASARD !!!**
 - Seulement $\frac{1}{4}$ reçoivent le traitement.
 - Les $\frac{3}{4}$ restant constituent le groupe témoin



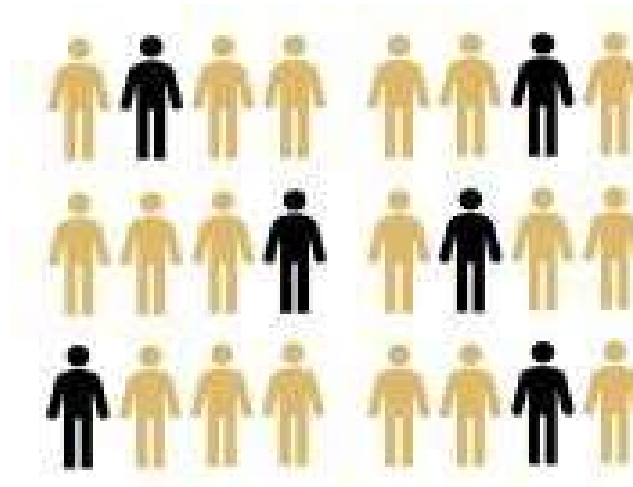
ECR par phases (introduction progressive)

- Phase 2:
 - Seulement $1/2$ reçoivent le traitement
 - La moitié restante constitue le groupe témoin



ECR par phases (introduction progressive)

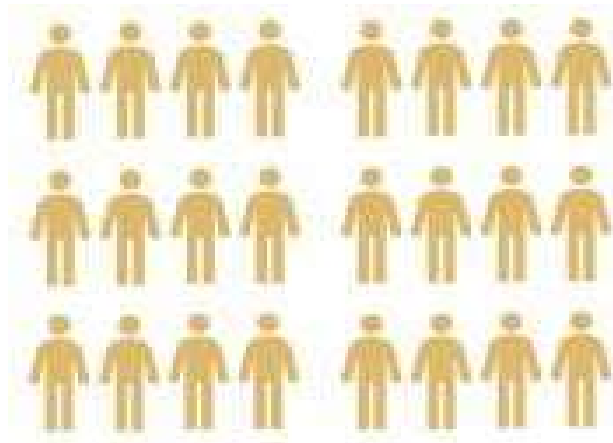
- Phase 3:
 - Seulement $\frac{3}{4}$ reçoivent le traitement.
 - Le $\frac{1}{4}$ restant constitue le groupe témoin



ECR par phases (introduction progressive)

- Phase 4:

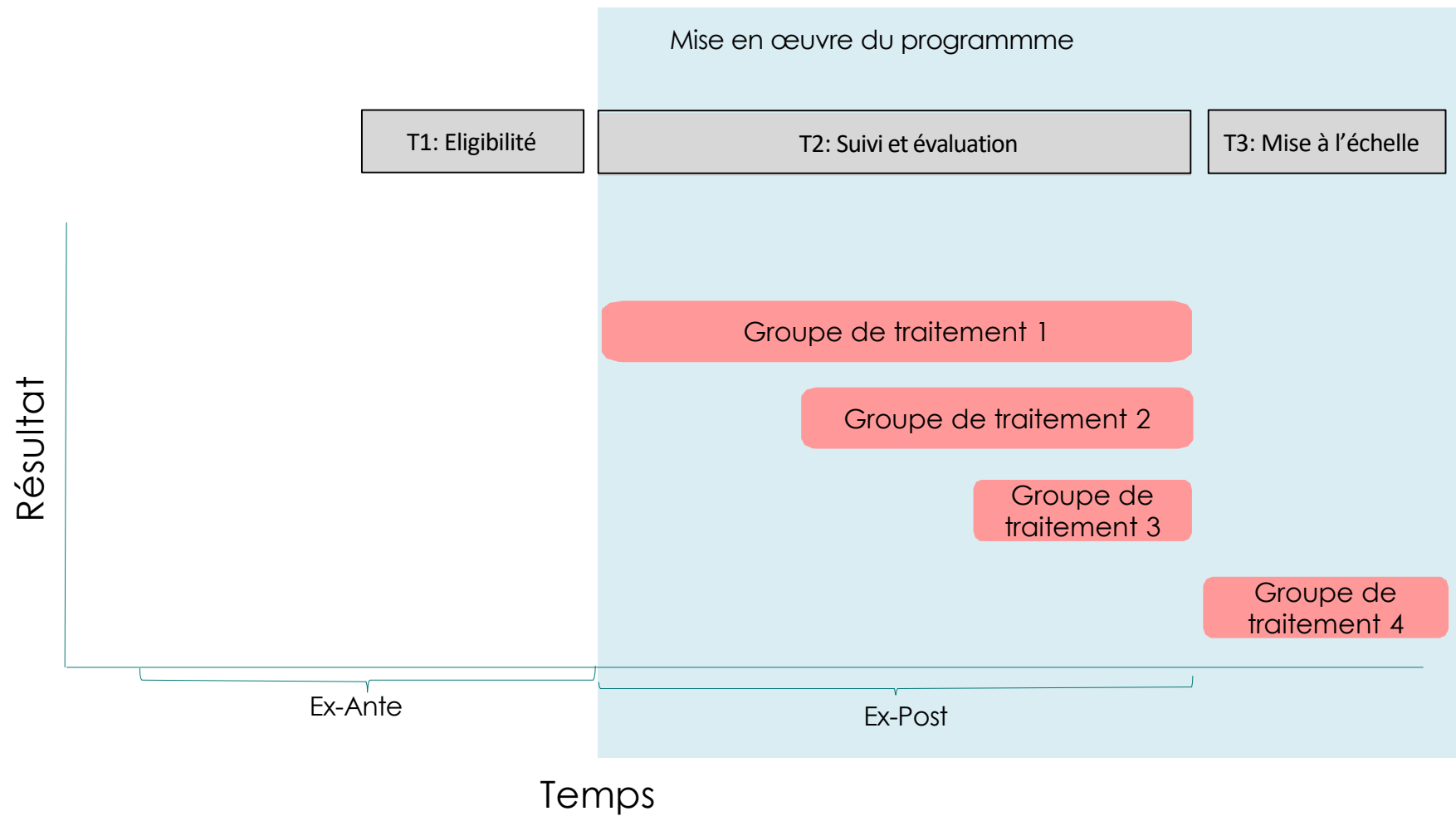
→ Tout le monde reçoit l'intervention. Plus de groupe témoin.



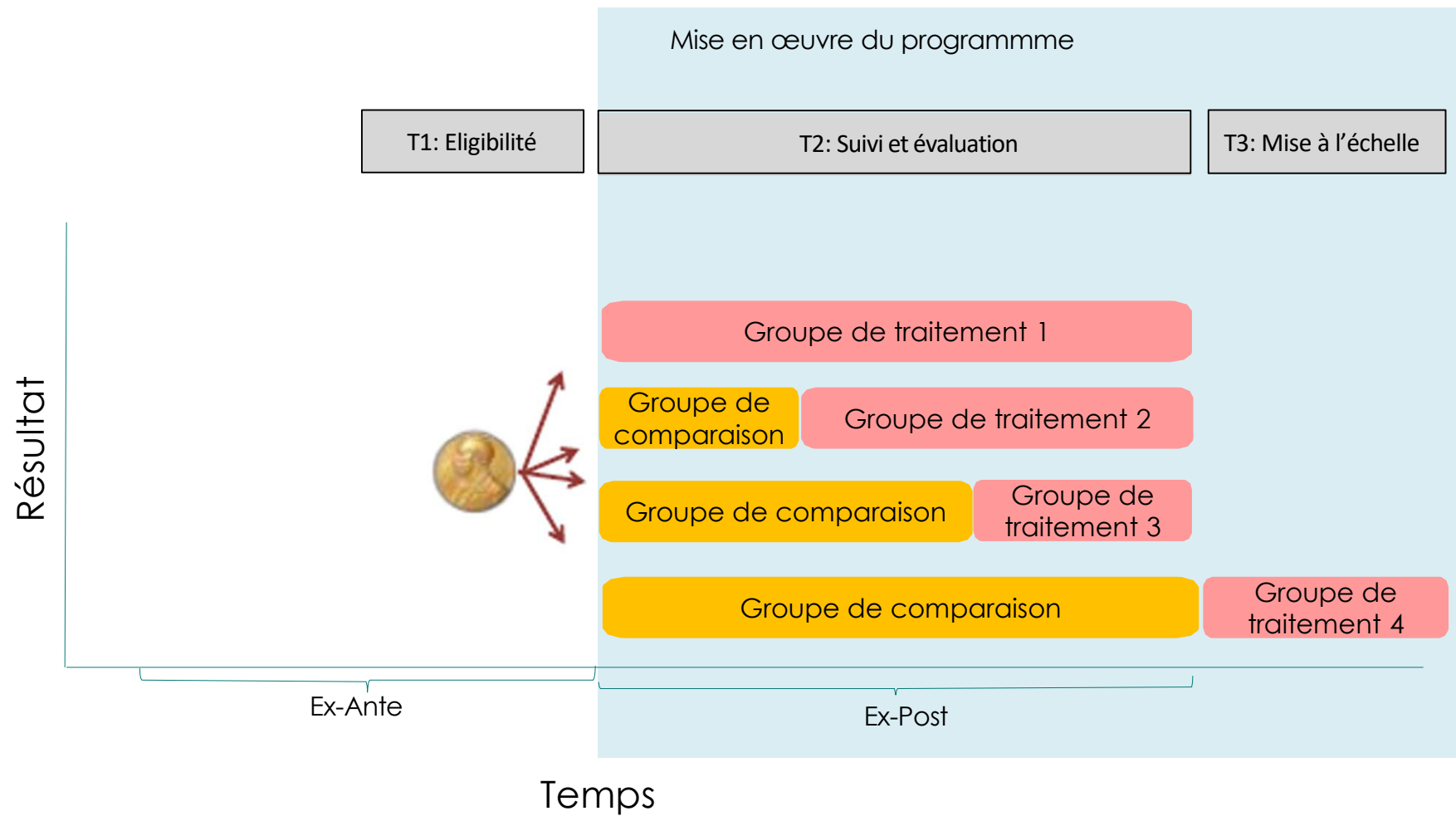
ECR par phases (introduction progressive)

- **Particulièrement utile lorsque** : Des contraintes de budget, de temps ou de capacités ne permettent pas de mettre en place un programme pour tout le monde en même temps
- **Avantages** : Tout le monde reçoit l'intervention. La prise de décision est transparente et équitable. Permet d'analyser les effets avec différentes intensités (durée) d'exposition à l'intervention
- **Inconvénients** : Il n'est pas possible (en général) de capturer les impacts à long terme

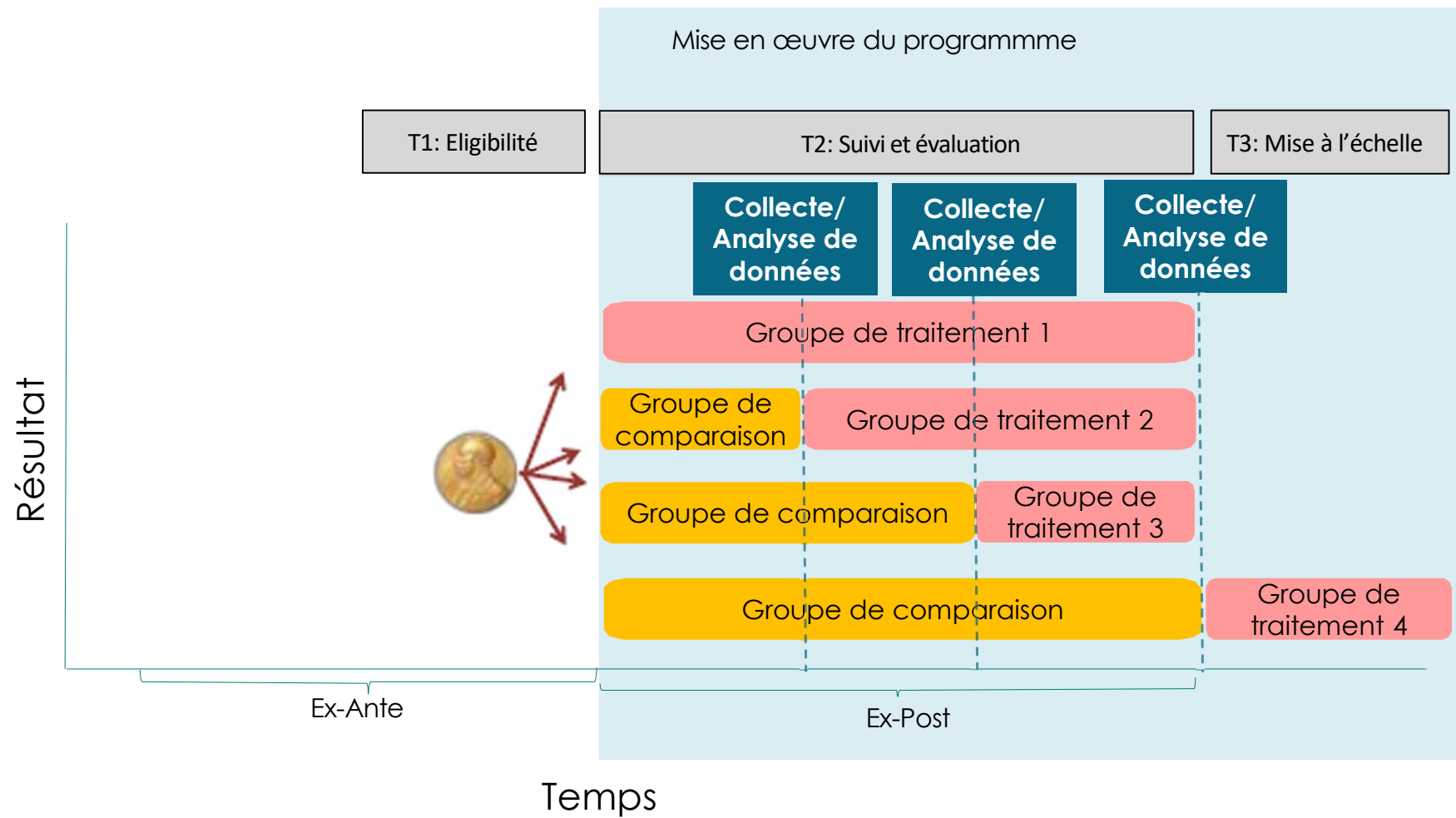
ECR par phases (introduction progressive)



ECR par phases (introduction progressive)



ECR par phases (introduction progressive)



4. Intensités aléatoires / ECR à étapes multiples

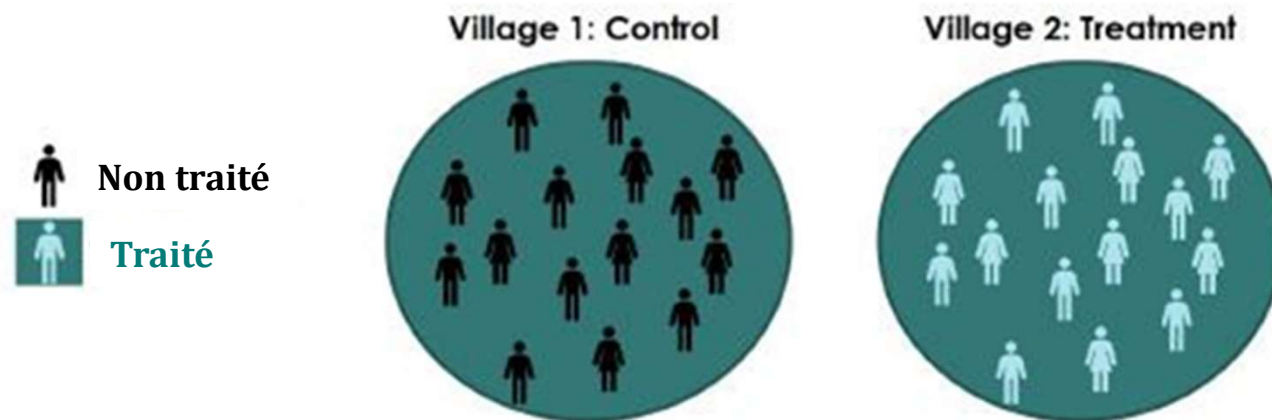
Allocation aléatoire de l'intervention:

- Dans certaines communautés 100% des personnes éligibles reçoivent l'intervention
- Dans d'autres communautés, une partie seulement (par exemple 50%) reçoit l'intervention
- Et enfin, dans un troisième groupe de communautés, personne ne reçoit l'intervention (groupe de contrôle ou de comparaison)

Intensités aléatoires / ECR à étapes multiples

Exemple : Intervention sur la sensibilisation à l'hygiène en vue de combattre les maladies infectieuses

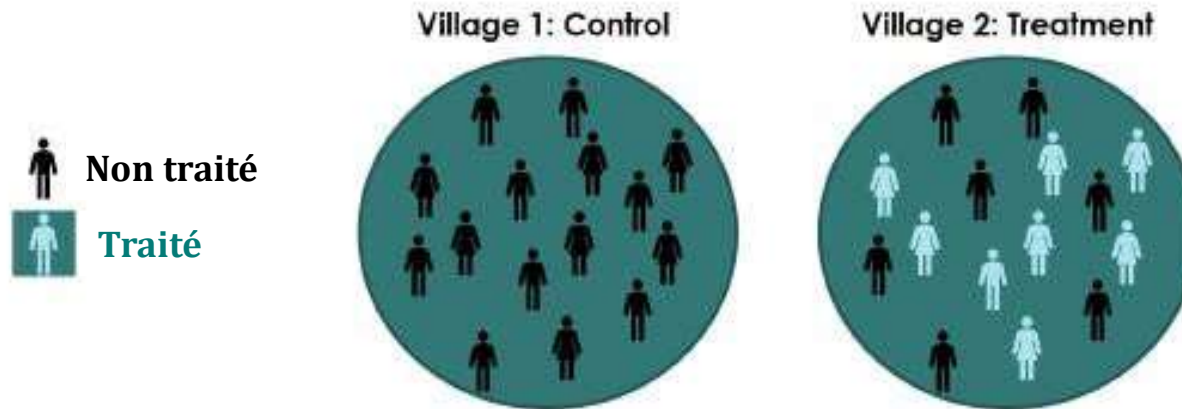
- Village 1 : Contrôle (pas de traitement)
- Village 2 : 100% si les individus reçoivent un traitement (information sur l'hygiène)



Intensités aléatoires / ECR à étapes multiples

Maintenant, 50% seulement des individus du village 2 reçoivent le traitement :

- Hypothèse : Les gens partagent l'information entre eux
- Comparez les résultats des personnes traitées ET non traitées du village 2 avec le groupe témoin ou de comparaison du village 1.



Intensités aléatoires / ECR à étapes multiples

- **Très utile lorsque** : On veut analyser un programme dans lequel des effets de contagion sont probables. Par exemple, la propagation de l'information dans un programme d'information
- **Avantage** : permet d'analyser les effets de contagion possibles
- **Inconvénient** : Nécessite un échantillon de plus grande taille

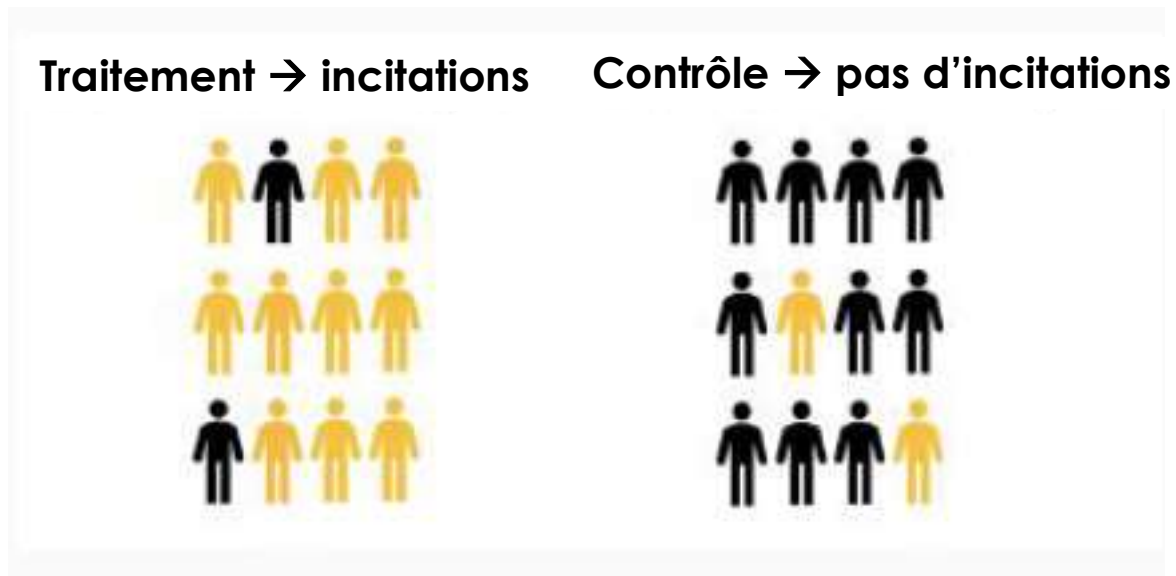
5. Approche par «encouragement»

Situation : Il est impossible de priver une personne d'un traitement (par ex. programme de distribution de nourriture gratuite)

- Solution possible : Randomisez les incitations → Faites la publicité et donnez des informations à certains, et rien à d'autres. Les participants peuvent toujours bénéficier de l'intervention, même sans incitation
- Les incitations agissent comme un instrument (VI)
- On compare les personnes qui ont reçu des incitations avec celles qui n'en ont pas reçu

5. Approche par «encouragement»

- Solution : Le « traitement » est simplement une **incitation** à utiliser les bons d'alimentation



Reçoit les bons d'alimentation

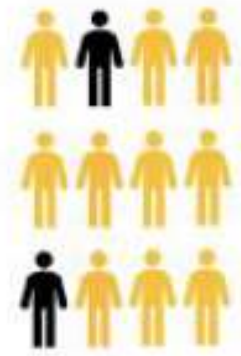


Ne reçoit pas les bons d'alimentation (mais en avait la possibilité)

5. Approche par «encouragement»

- Comparez l'ensemble du groupe de traitement avec l'ensemble du groupe de contrôle/de comparaison
- Analysez l'effet de l'incitation à l'usage des bons alimentaires (variable instrumentale) et indirectement l'impact des bons alimentaires sur la nutrition elle-même

Traitement → incitations



Contrôle → pas d'incitations



Reçoit les bons d'alimentation

Ne reçoit pas les bons d'alimentation (mais en avait la possibilité)

5. Approche par «encouragement»

- **Avantage** : Permet d'analyser des programmes accessibles à tous (du moins les personnes éligibles)
- **Inconvénient** : Seule une partie de l'impact est mesurée ; cela pourrait être problématique pour l'évaluation si le groupe de contrôle est fortement incités (par exemple si l'encouragement/l'incitation ou la mobilisation via la télévision ou la radio n'est pas ciblée)

Exercise

Le problème

- Un sondage réalisé auprès des ménages du pays de Fantasia a démontré que 60% de la population rurale cuisine encore au bois. Or, la cuisine au bois aggrave les problèmes actuels de déforestation du pays et accentue la pollution intérieure.
- Le gouvernement a par conséquent **décidé de distribuer des cuisinières écologiques**
- Sont éligibles les ménages des **zones rurales** possédant **moins d'un hectare de terres** dans **deux des dix départements** du pays
- La cuisinière peut être retirée dans les antennes locales d'agences gouvernementales, qui sont largement implantées et détiennent les registres de tous les ménages d'agriculteurs de leur secteur.
- La mise en œuvre s'étend sur la période **2022-2025**.

Le gouvernement vous demande de l'aide :

Ils veulent en savoir plus sur l'**impact de l'intervention** !

Ils sont ouverts aux approches **expérimentales**.

Exercice

1. Proposez une question d'évaluation et spécifiez les indicateurs à mesurer.
2. Suggérez une approche expérimentale permettant de répondre à la question donnée dans la situation décrite.
 - Décrivez comment vous identifierez les groupes de traitement et de contrôle.
 - Quelles sont les forces et les limites de la méthode proposée?
3. De quel type de données avez-vous besoin pour votre évaluation? A quel moment celles-ci doivent-elles être collectées?

Exercice Jour 1 - Discussion

Groupe cible: ménages ruraux, possédant moins d'1ha de terres, dans 2 des 10 départements du pays

Question(s) d'évaluation: Dans quelle mesure le **programme de distribution de cuisinières écologiques** aux **ménages possédant moins d'un hectare de terres** a contribué à diminuer la **déforestation** et la **pollution (intérieure)** dans les **deux départements ciblés** ? → pourrait séparer en deux questions (impact sur la déforestation et impact sur la pollution intérieure)

→ REMARQUE: pollution intérieure n'est pas l'objectif final → impact final est d'améliorer la santé des ménages

Variables de résultat: Nombre d'ha de forêt dans le département (déforestation); Nb. de ménages utilisant le bois de cuisine / les cuisinières éco. (mécanisme → si cuisinière éco pas utilisées → pas d'impact possible sur déforestation, pollution intérieure et santé!; peut être détecté par le système de contrôle/suivi du projet); Indice moyen de pollution de l'air dans un ménage qui utilise un foyer à bois; Nb. de consultations au centre de santé pour maladies respiratoires (santé); Nb. De ménages/individus rapportant des problèmes de santé chroniques, par ex. toux

→ REMARQUE: certains indicateurs de santé (par ex. infection pulmonaires) peuvent prendre plus de temps pour réagir, impact peut-être pas détectable durant courte vie du programme (3 ans). Possible d'utiliser d'autres indicateurs « prédicteurs » de la variable d'impact de long terme, par ex. toux ou bronchite chronique

Méthode de randomisation: Si randomise au sein d'un village → risque de contagion (partage de ressources, imitation des comportements). Solution: Randomisation **par grappes** au niveau du village → tous les ménages éligibles dans un village sont traités; risque très faible de contagion entre villages.

- ECR par phases → avantage: tout le monde sera traité / inconvénient: impact à court/moyen terme
- ECR par encouragement → toutes les personnes éligibles ont accès au programme; MAIS incitations plus fortes dans certains villages choisis au hasard (groupe de traitement)

Groupe de traitement: Population éligible dans les villages des départements ciblés alloués aléatoirement au traitement

Groupe de contrôle: Population éligible dans les villages des départements ciblés alloués aléatoirement au groupe de contrôle

Données: collectées au niveau des ménages (enquêtes), des centres de santé, données environnementales/climatiques (eaux et forêt), données forestières, qualité de l'air (maisons); données collectées au niveau communautaire; données sur l'indice de végétation; Cartographie par image satellitaire

A collecter avant, pendant et après l'intervention: Données de référence (AVANT le début du programme); données de mi-parcours (par ex. 18 mois) → possibilité d'adapter le programme; données de fin → impact à moyen terme; (IMPORTANT: Données de suivi !)

Questions et Réponses

Avez-vous des questions?

Vous pouvez aussi m'écrire à l'address email suivante:

m.gillaizeau@c4ed.org

Merci pour votre attention.
A demain!

- Gertler, P. J., Martinez, S., Premand, P., Rawlings L., Vermeersch, C. M. J. (2011) Impact Evaluation In Practice, First Edition, The World Bank.
- Gertler, P. J., Martinez, S., Premand, P., Rawlings L., Vermeersch, C. M. J. (2016) Impact Evaluation In Practice, Second Edition, The World Bank.
- White, H., & Raitzer, D. A. (2017). Impact evaluation of development interventions: A practical guide. Asian Development Bank.
- Correlations: <https://towardsdatascience.com/nicholas-cage-pool-saviour-9c13feaff6f>
- White, H. (2019). The twenty-first century experimenting society: the four waves of the evidence revolution. Palgrave Communications, 5(1), 1-7.
<https://www.nature.com/articles/s41599-019-0253-6>
- Connected Papers: <https://www.connectedpapers.com/>
- 3ie's "Evidence Gap Maps": <https://gapmaps.3ieimpact.org/evidence-maps/map-evidence-maps-relating-sustainable-development-lmics>
- International Rescue Committee's "Outcomes and Evidence Framework": <http://oef.rescue.org/#/home?k=zhw0nz>
- Education Endowment's Foundation "Teaching and Learning Toolkit": <https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/>
- Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. (2016). Impact evaluation in practice. World Bank Publications.
<https://www.nobelprize.org/prizes/economic-sciences/2019/press-release/>
- Shadish, W. R., & Cook, T. D. (86). Campbell. DT (2002). Experimental and quasi-experimental designs for generalized causal inference.
- J-PAL, Poverty Action Lab, L4, 2019: https://www.povertyactionlab.org/sites/default/files/L4_How%20to%20Randomize.pdf
- IDinsight (2020): <https://www.idinsight.org/article/implementing-a-10000-refugee-household-public-lottery-in-eight-days-lessons-from-kiryandongo-uganda/>
- Banerjee, A., Alsan, M., Breza, E., Chandrasekhar, A. G., Chowdhury, A., Duflo, E., ... & Olken, B. A. (2020). Messages on COVID-19 prevention in India increased symptoms reporting and adherence to preventive behaviors among 25 million recipients with similar effects on non-recipient members of their communities (No. w27496). National Bureau of Economic Research.
- Avdeenko et al. 2020: Evidence on Covid-19 Pandemic Control Interventions and their Impacts on Health-Related Outcomes
- World Health Organization. 2014. Indoor air quality guidelines: household fuel combustion.