

Concepts fondamentaux de l'évaluation quantitative de l'impact : Module 3

Aneta Bonikowska, PhD

Division de l'analyse sociale et de la modélisation

19 mars 2024



Statistique
Canada

Statistics
Canada

Canada

Module 3

Expériences contrôlées randomisées : l'étalon-or de l'EQI

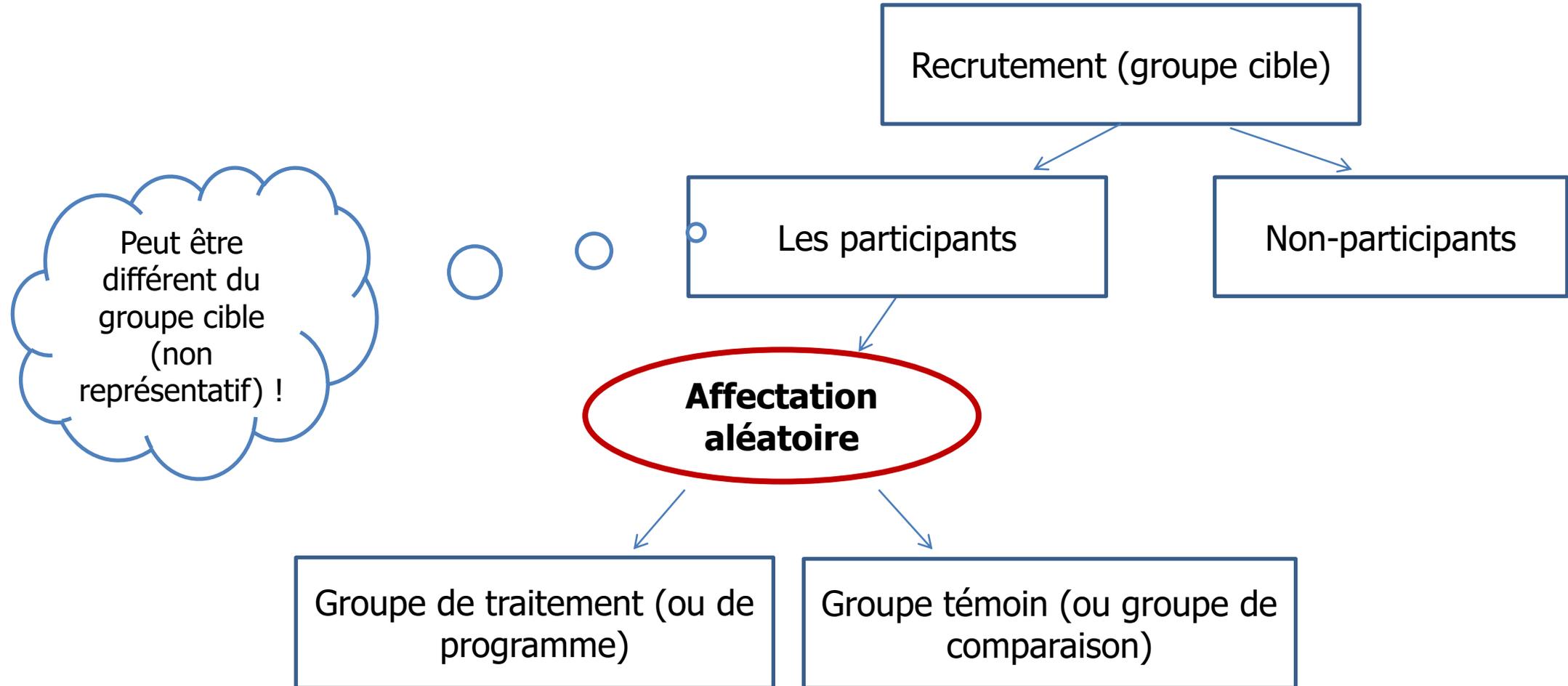
Contenu du module 3

- Deux types d'expériences
- Un plan expérimental de base
- Pourquoi expérimenter ?
- Limites des expériences
- Principaux points à retenir

Deux types d'expériences

- Expériences en laboratoire
 - La politique est mise à l'essai dans des conditions de laboratoire
 - Environnement (très) contrôlé
 - Critique :
 - peut être difficile de généraliser certaines découvertes à la vie réelle
 - Dans le contexte des sciences sociales, les caractéristiques individuelles telles que la capacité, l'effort, etc. sont importantes pour les résultats, mais ne peuvent pas être facilement mesurées et donc contrôlées
- Expériences sur le terrain
 - La politique est mise en œuvre dans la vie réelle (sur le terrain) au sein d'un petit groupe, qui peut ou non savoir qu'il fait partie d'une expérience
 - Aussi appelé un « projet de démonstration »
 - Nous nous concentrerons là-dessus

Un plan expérimental de base



Pourquoi expérimenter ?

- Nous aimerions diriger un groupe de personnes à travers la vie deux fois : une fois en passant par un programme, une fois sans - mais nous ne pouvons pas, au lieu de cela ...
- L'affectation aléatoire résout essentiellement le problème fondamental de l'évaluation de programme :
 - Le groupe de traitement et le groupe témoin sont les mêmes (en moyenne) au début de l'expérience (la chance leur a été attribuée, pas le choix personnel)
 - Enquête de base (avant l'expérience) nécessaire pour confirmer la similitude des caractéristiques entre le traitement et le groupe témoin
- **Impact = Résultat du groupe de traitement – Résultat du groupe témoin**

Pourquoi expérimenter ?

- L'affectation aléatoire supprime le choix de choisir le traitement (politiques ou programmes)
- Par conséquent, le groupe témoin sert de bon contrefactuel au groupe de traitement (ce qui se serait produit si le groupe de traitement n'avait pas été offert le traitement)
- L'affectation aléatoire est souvent appelée « l'étalon-or » dans l'EQI

Tout ce qui brille... Limites des expériences

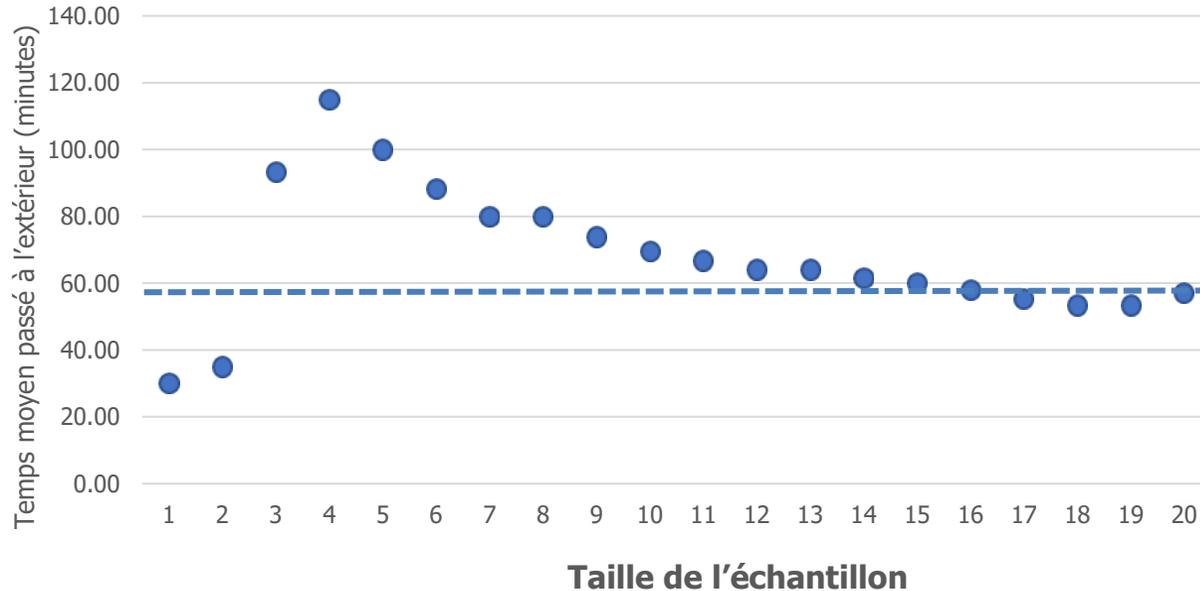
- Taille de l'échantillon
- Validité interne/externe
- Intention de traiter
- Attrition au sein du groupe témoin
- Biais d'abandon
- Biais de substitution
- Effets d'entraînement
- Préoccupations éthiques
- Coûts

Taille de l'échantillon

- On s'attend à ce que l'affectation aléatoire crée des groupes témoins et de traitement identiques
- Mais que se passe-t-il si la taille de l'échantillon est petite, comme c'est le cas dans certaines expériences ?
- Avec de petits échantillons, tout peut arriver (c.-à-d. que le groupe de traitement et le groupe témoin peuvent être différents)
- La solution consiste à prendre un échantillon plus grand

Taille de l'échantillon

Temps moyen passé à l'extérieur en fonction de la taille de l'échantillon



- Les petits échantillons produisent des estimations peu fiables, quelle que soit la taille de la population
- Peut faire correspondre le traitement et les groupes témoins en fonction de certaines caractéristiques ; p. ex., s'assurer que les personnes de même sexe, de même âge et de même niveau de scolarité sont réparties dans les deux groupes (par stratification)

Validité interne/externe

- **Validité interne** : Pour les participants qui ont accepté l'expérience, l'affectation aléatoire générera des résultats valides (tant que l'échantillon est suffisamment grand et que l'expérience a été bien conçue et mise en œuvre)
- **Validité externe** : les résultats s'appliquent également au groupe cible ou à d'autres groupes

Validité interne/externe

- Les expériences sur le terrain (sociales) ont été décrites comme valides en interne, mais pas nécessairement valides à l'externe
 - c.-à-d. que les expériences sociales sont valides pour ceux qui ont accepté (la possibilité de recevoir) un traitement (mais qui ne l'ont peut-être pas suivi)
- Pour fournir une preuve de validité externe :
 - Bien décrire les participants (par rapport au groupe cible ou à d'autres groupes)
 - Analyse des sous-groupes (les résultats sont-ils les mêmes pour chaque type de participant au programme ?)

Intention de traiter

- Nous ne pouvons pas forcer le groupe de traitement à prendre son « médicament »
- En comparant les résultats du traitement et du groupe témoin, nous estimons l'impact de « l'intention de traiter » (ou l'impact de l'offre de traitement)
- Nous ne pouvons pas estimer l'impact de la prise du traitement (effet de traitement sur le traitement)

Intention de traiter

- **Avertissement** : tentant de simplement regarder ceux dans le groupe de traitement qui ont réellement pris le traitement
 - La prise du traitement n'a pas été attribuée au hasard
 - Seule l'offre du traitement a été attribuée au hasard, de sorte que nous ne pouvons comparer que le traitement et les groupes témoins

L'intention de traiter est-elle si mauvaise ?

- Dans le monde réel, les programmes ne peuvent généralement être offerts qu'aux individus
- Certains prendront le traitement, d'autres ne le feront pas
- Tout comme une expérience !

Attrition au sein du groupe témoin

- Les expériences comportent généralement des enquêtes de suivi
- Dans toute enquête longitudinale, l'attrition est un problème
- Il est problématique avec les expériences si le taux d'attrition est différent pour le groupe de traitement et le groupe témoin :
 - Le groupe témoin peut avoir moins de raisons de continuer à participer (il n'a pas reçu le traitement)
- La solution consiste à dépenser beaucoup pour les relations avec les répondants et les appels de suivi, ou (si possible) à utiliser des données administratives (p. ex. les fichiers d'impôt sur le revenu) pour mesurer les résultats

Biais d'abandon

- Le groupe de traitement peut arrêter le traitement expérimental, peut-être en faveur d'un traitement non expérimental
 - Par exemple, le programme de formation professionnelle ne fonctionnait pas, alors j'ai demandé de l'aide ailleurs
- Le biais d'abandon réduit le traitement expérimental, mais l'intention de traiter peut encore être estimée
- En bref, pas un problème car ils le feraient dans la vraie vie
 - Nous estimons l'effet de l'offre du traitement

Biais de substitution

- Le groupe de traitement peut réduire la participation au traitement non expérimental
 - Par exemple, la formation professionnelle fonctionnait si bien qu'ils ont arrêté un autre cours qu'ils prenaient déjà
 - Pas un problème (ils feraient cela est la vraie vie)
- Le groupe témoin peut chercher d'autres formes de traitement (réaction à la cause de l'iniquité par randomisation – « Je vais vous montrer »)
 - Cela ne se produirait pas dans la vie réelle (à moins qu'il n'y ait randomisation)
 - L'impact estimé est biaisé vers zéro
 - Important à surveiller (enquêtes de suivi ou données administratives)

Effets d'entraînement

- Un traitement peut déborder sur le groupe témoin si les deux sont en contact étroit (p. ex. une séance d'information offerte à certains élèves d'une école)
 - Peu probable dans la vraie vie – des groupes similaires en contact étroit sont généralement tous les deux admissibles ou non à un programme gouvernemental
- Les biais ont un impact vers zéro
- La solution consiste à minimiser le contact entre le groupe de traitement et le groupe témoin

Préoccupations éthiques

- Les expériences peuvent être considérées comme refusant des services de programme au groupe témoin. Toutefois:
 - Les programmes ne sont pas toujours bénéfiques. A priori, les bénéfices sont inconnus, d'où l'expérience
 - Argument de l'amélioration de Pareto : personne n'est moins bien loti, certains sont (potentiellement) mieux lotis
 - Certains programmes de la vie réelle limitent les inscriptions, p. ex. garderies publiques du Québec, programmes d'immersion française en Colombie-Britannique (est-ce plus juste que l'affectation aléatoire ? Probablement pas, alors autant utiliser la meilleure méthode à des fins d'évaluation)
 - Dans les projets de démonstration, le consentement éclairé est fourni par les participants au programme (ils sont d'accord avec l'expérience et comprennent leurs chances de se voir offrir le traitement ou non)
 - Considérations relatives aux coûts du point de vue des contribuables : possibilité d'une facture coûteuse pour des programmes inefficaces qui n'ont pas été correctement testés

Coûts

- Les expériences sociales ont été critiquées pour être trop coûteuses
- Bien que les coûts soient plus élevés que dans les études non expérimentales, il est souvent impossible de tester une nouvelle politique avant des années plus tard, lorsque les données provenant d'enquêtes et de sources administratives sont disponibles
- Ce n'est qu'en effectuant un test pilote que nous pouvons savoir si la politique fonctionne
 - Si c'est le cas, autant le faire scientifiquement (affectation aléatoire)
- Ne pas évaluer l'impact des programmes peut s'avérer plus coûteux en fin de compte que de mener des expériences
 - Les projets pilotes sont beaucoup plus petits que le déploiement complet du programme
 - Si l'impact du programme est faible ou nul (ou négatif, c'est-à-dire préjudiciable), un projet pilote pourrait permettre d'économiser beaucoup plus que le coût de l'expérience

Exemple : Faire en sorte que le travail soit mieux rémunéré que l'aide sociale

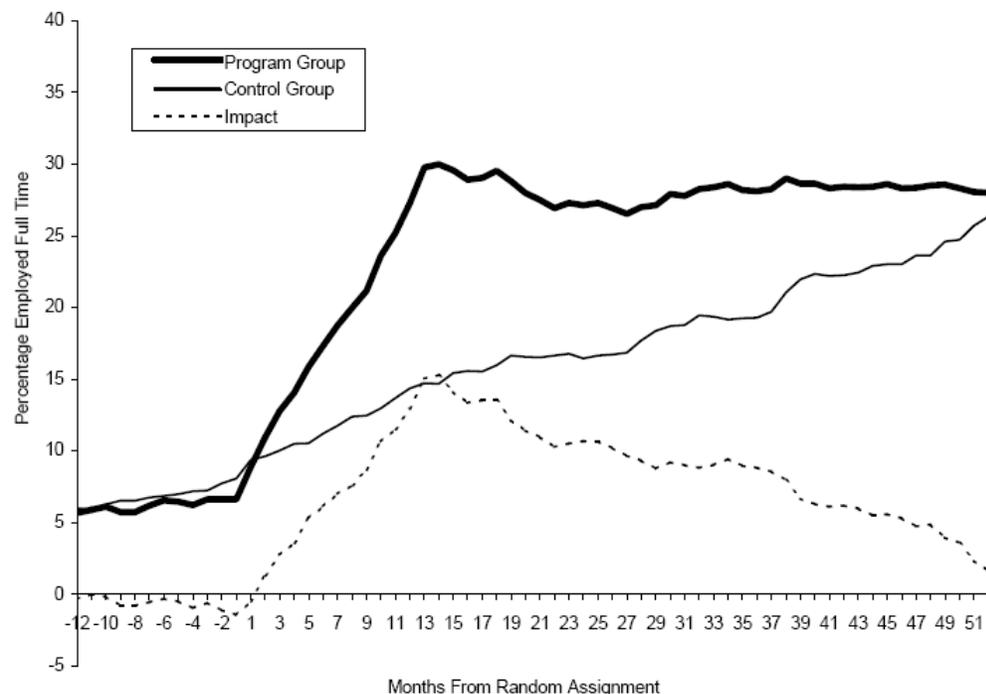
- Dans les années 1980 et 1990, l'aide sociale (AS) payait à peu près autant que le salaire minimum
- De nombreux bénéficiaires de l'AS ne pouvaient pas s'attendre à gagner plus que le salaire minimum s'ils quittaient l'AS, alors pourquoi chercher un emploi ?
- Le Projet d'autosuffisance (PA) a rendu le travail mieux rémunéré que l'AS, dans l'espoir de rendre les gens moins dépendants de l'AS
- Mise en œuvre par la Société de recherche sociale appliquée (SRSA)
- Les bénéficiaires de l'AS monoparentale (principalement les mères célibataires) pourraient être admissibles à un supplément de revenu pour rendre le travail plus attrayant
- Étude des bénéficiaires : les participants doivent avoir été sous l'AS pendant au moins 1 an (bénéficiaires à long terme de l'AS)

Exemple : Faire en sorte que le travail soit mieux rémunéré que l'aide sociale

- Les participants ont été assignés au hasard au groupe de traitement ou au groupe témoin
- Le groupe de traitement s'est vu offrir un supplément de revenu s'il trouvait un emploi d'au moins 30 heures par semaine dans un délai d'un an
- Supplément versé pour une période pouvant aller jusqu'à 3 ans, seulement en mois où le participant a travaillé à temps plein et n'a pas perçu d'AS
- Le supplément de revenus a essentiellement doublé le revenu avant impôt des salariés au salaire minimum
- Mise en œuvre au Nouveau-Brunswick et en Colombie-Britannique dans les années 1990, avec plusieurs entrevues de suivi

Exemple : Faire en sorte que le travail soit mieux rémunéré que l'aide sociale

Figure ES.1: Percentage Employed Full Time, by Months From Random Assignment



Sources: Calculations from baseline survey data and 18-month, 36-month, and 54-month follow-up survey data.

Note: "Employed full time" is defined as working 30 hours or more in at least one week during the month.

Source: Michalopoulos et al. 2002. Making Work Pay: Final Report on the Self-Sufficiency Project for Long-Term Welfare Recipients. Ottawa: Social Research and Demonstration Corporation.

1. La randomisation a fonctionné - le groupe de traitement et le groupe témoin avaient des taux d'emploi très similaires avant le début de l'expérience (d'autres caractéristiques, aussi)
2. Au cours de la période visée par le programme, l'emploi est plus élevé dans le groupe de traitement que dans le groupe témoin
3. Environ 1/3 du groupe de traitement a pris le supplément
4. Le pourcentage de personnes occupées augmentait de façon constante avec le temps passé au sein du groupe témoin, même s'ils ne recevaient aucune subvention
5. Les résultats du traitement et du groupe témoin ont convergé peu de temps après la fin du programme

Points à emporter

- Dans l'EQI, l'incidence d'une politique ou d'un programme est définie comme la différence entre le résultat réel (avec le programme) et un résultat contrefactuel (le résultat auquel nous pourrions nous attendre sans le programme)
- L'affectation aléatoire réduit les différences entre le groupe de traitement et le groupe témoin découlant du choix, de l'autosélection ; en ce sens, il fournit le meilleur groupe témoin et le meilleur résultat contrefactuel
- Les expériences randomisées ne sont pas sans problèmes
- Dans la plupart des contextes EQI, l'affectation aléatoire n'est pas présente, et la construction d'un groupe témoin approprié et d'un résultat contrefactuel nécessite plus de réflexion - heureusement, il existe des méthodes et des outils qui peuvent aider ...

Merci!



Aneta.Bonikowska@statcan.gc.ca



Restez branchés!

Application StatsCAN

Balado Hé-coutez bien!

StatsCAN Plus

Le Quotidien

Site Web

Enquêtes et programmes statistiques

Centre de service de données

Mon StatCan



Vous avez des questions? Contactez-nous : infostats@statcan.gc.ca

Statistique Canada – Votre organisme national de statistique



Éclairer grâce aux données, pour bâtir un Canada meilleur

