

Les questions causales en évaluation : une introduction

Il est beaucoup question dans l'actualité de l'impact des politiques publiques, sans toujours s'accorder sur ce que ce terme recouvre. Un préalable est de s'interroger sur les questions de causalité, c'est-à-dire de relation de cause à effet (I) ; puis d'en comprendre les enjeux pour l'action publique (II). Nous verrons ensuite les différentes stratégies qui existent pour faire la preuve d'une relation causale (III).

DE QUOI PARLE-T-ON ?

Le terme de *causalité* désigne la relation entre deux phénomènes, l'un étant désigné comme une cause, l'autre comme un effet. **Chacun d'entre nous, dans notre vie de tous les jours, fait à chaque instant des inférences causales** : cela fait partie de notre rapport au monde. L'inférence nous permet de réduire l'incertitude sur nos actions et celle des autres, elle permet de faire des choix. Elle est aussi souvent implicite : lorsque vos parents vous disent : « il fait froid, mets ton écharpe », ils infèrent que le froid donne le rhume.

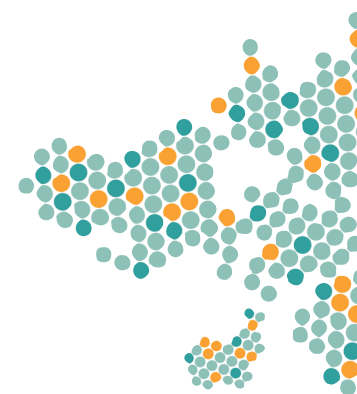
Contrairement aux phénomènes associés par la relation causale, la causalité elle-même n'est pas observable, c'est une construction, une analyse que nous faisons, une *inférence causale*.

La recherche, en sciences naturelles ou sociales, commence souvent par des questions qui sont causales par nature (Pourquoi observe-t-on ce phénomène ? *Parce que...*) auxquelles elle répond en posant puis en testant des hypothèses causales. La causalité est ainsi un des modes possibles pour chercher à connaître et comprendre les réalités qui nous entourent.

UNE NOTION SITUÉE


La notion de causalité telle que nous l'utilisons est située dans le temps et dans l'espace.

C'est une notion occidentale, qui est liée à l'évolution des sciences modernes. Il y a d'autres approches, en [Inde](#) ou en [Chine](#) par exemple, qui sont assez largement différentes dans leurs fondements, ce qui peut amener à des incompréhensions culturelles.



💡 -D'ailleurs c'est faux : c'est la promiscuité, plus forte l'hiver, qui favorise la transmission du virus. Mettez une écharpe quand même pour faire plaisir à vos parents.

Pour aller plus loin sur les questions de causalité :

-  Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R., & Befani, B. (2012). Broadening the range of designs and methods for impact evaluations. Report of a study commissioned by Dfid.

Ensuite, c'est une notion qui a évolué dans le temps, depuis la causalité aristotélicienne, puis scolastique qui s'attache à comprendre la nature des choses ; puis le passage à la causalité moderne^{HULSWIT, 2004}, qui cherche des lois exactes pour expliquer les phénomènes du monde. Galilée peut ainsi écrire :

« La cause est ce qui est tel que, lorsqu'elle est posée, l'effet s'ensuit ; lorsqu'elle est ôtée, l'effet est ôté »

Il y a donc un transfert d'une compréhension des phénomènes à celui du rapport, inobservable, mais essentiel, entre plusieurs phénomènes jugés indépendants.

Dans cette vision, toute cause a un effet, et les causes produisent (dans les mêmes conditions) les mêmes effets. Cette approche déterministe fait l'objet de débats importants. Elle est à la fois raffinée et progressivement déconstruite aux 19^e et 20^e siècles, tant avec l'avancée des sciences que dans le champ de la philosophie.

Les principales approches contemporaines de la causalité vont alors être :

- ▶ la causalité **contrefactuelle** qui est une façon pratique de définir la causalité par la différence entre une situation avec une situation sans. Dire que B est un effet de A, c'est dire que s'il n'y avait pas eu A, il n'y aurait pas eu B. Il y a une chaîne logique d'évènements qui mène de A à B.
- ▶ la causalité comme une **probabilité** : A est une cause de B si B apparaît plus souvent quand A est présent. Par exemple, fumer augmente la chance d'avoir un cancer (par 10 ou 15), mais je peux avoir un cancer sans fumer et je peux fumer sans avoir le cancer (quoique dans des probabilités bien moindres) ;
- ▶ la causalité **circonstanciée** : dans cette logique, une cause A peut certes mener à un effet B, mais 1/elle n'est pas nécessaire (une cause Y pourrait produire cet effet), 2/elle ne peut être une cause que parmi d'autres conditions (un champ causal), qui peuvent évoluer selon les circonstances. J.L. Mackie va plus loin en considérant que ce que nous appelons des causes sont souvent des « conditions INUS », c'est-à-dire un élément *insuffisant*, mais *nécessaire* d'une condition qui est elle-même *non nécessaire*, mais *suffisante* pour le résultat. L'exemple qu'il donne est le suivant :

Lorsque nous disons qu'un court-circuit est la cause de l'incendie d'une maison, nous savons bien que l'explication est incomplète. En l'absence de rideau à proximité, il n'y aurait pas eu de feu, et celui-ci aurait pu être causé par autre chose. Dire donc qu'un court-circuit a causé l'incendie, c'est dire que cet évènement insuffisant en tant que tel était nécessaire à un paquet de causes (le court-circuit produisant une étincelle qui enflamme le rideau), qui en lui-même est suffisant pour produire l'incendie, mais pas nécessaire (il y avait d'autres façons de produire cet incendie).

Plusieurs romans ont montré le basculement dans la science et la causalité moderne, notamment *Le nom de la rose*, d'Umberto Eco, et *L'Œuvre au noir*, de Marguerite Yourcenar.

L'IMPACT, UNE DÉFINITION

Le mot « impact » décrit une vive, importante conséquence d'une action, par métaphore de l'impact d'un projectile. En évaluation, il est venu à décrire le lien entre une cause (généralement l'intervention évaluée) et des effets désirés ou non.


La volonté d'estimer les effets d'une intervention est ancienne et a donné lieu à des débats sur les méthodes, importées de la recherche, qui pouvaient être employées à cette fin. Donald Campbell notamment importe la **méthode expérimentale** dans l'action publique^{CAMPBELL, 1979}. De son point de vue, le monde réel est trop complexe : en testant les assertions causales dans un cadre contrôlé, il est possible d'obtenir des preuves causales qui peuvent guider les choix politiques. La preuve ainsi obtenue doit guider les choix politiques, en matière de politiques sociales notamment.


Cette vision alors largement partagée par les évaluateurs (« quantitavistes » ou « qualitatistes ») de l'importance de la méthode en général et de la preuve de causalité en particulier va être largement remise en cause dans les années 1980^{SHADISH ET AL, 1991}. Quoi qu'ils en disent, les décideurs n'accordent pas la même importance que les chercheurs à la rigueur méthodologique. Pour rendre les évaluations plus utiles, les évaluateurs se concentrent sur le fait de savoir *pourquoi* et *comment* les interventions obtiennent ou non leurs effets (plutôt que de savoir *si*), notamment à travers les registres de la pertinence et de la cohérence, et insistent sur la dimension participative des évaluations.

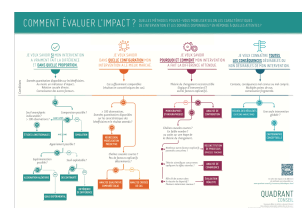
Au début des années 2000, cependant, et dans le mouvement de la politique basée sur la preuve (*evidence-based policy*), la méthode expérimentale revient sous une forme nouvelle. Le *randomised controlled trial* (RCT) ou méthode expérimentale par assignation aléatoire, s'impose dans l'aide au développement à travers des figures telles que [Esther Duflo](#), puis se diffuse plus largement dans les politiques publiques, notamment en France^{DEVAUX-SPATARAKIS, 2014}. C'est dans ce cadre que le terme impact a gagné en popularité, dans un sens très restreint : celui de l'effet attribué à une intervention à travers la méthode expérimentale ou des approches quasi-expérimentales.

Ce « coup de fouet » a suscité une controverse méthodologique au sein de la communauté évaluative^{DONALDSON, 2009}, mais il a également stimulé le développement de nouvelles approches évaluatives, également issues de la recherche, et permettant de traiter de façon plus rigoureuse et systématique les questions d'impact dans des approches causales différentes de l'approche contrefactuelle : l'analyse de contribution, l'évaluation réaliste, l'analyse qualitative comparée, et d'autres approches se construisent toutes au cours des années 1990 et 2000 et ont pris leur pleine mesure dans les années 2010.

Ce que nous appelons ici « évaluation d'impact » désigne **une approche évaluative centrée sur la question des effets de l'intervention, et déployant une méthode spécifique d'analyse de la causalité pour en rendre compte**. Elle n'est jamais une simple mesure du changement, et s'appuie systématiquement sur la collecte et l'analyse de données (et pas un simple raisonnement ou des théories existantes, par exemple) pour établir sa démonstration.

 Cette méthode compare deux groupes de « jumeaux » tirés au sort, l'un bénéficiant de l'intervention, l'autre non, observés au regard d'un indicateur mesurant le changement escompté. Le résultat de l'expérimentation ou « impact » est la différence entre la valeur constatée pour le groupe avec intervention et un contrefactuel, c'est-à-dire le groupe sans intervention.

 Pour une liste de méthodes d'évaluation d'impact, voir notre arbre des méthodes d'évaluation d'impact.



DES DIFFICULTÉS PROPRES À L'ACTION PUBLIQUE

Nous avons parlé jusqu'ici de causalité de façon générique, mais il se trouve que le sujet des évaluations, c'est-à-dire des questions publiques appliquées aux actions publiques ou d'intérêt général, a ses difficultés propres.

De quelle type d'action publique parlons-nous et quelle action publique est évaluée ? Selon la focale employée, on peut la juger particulièrement complexe, ce qui en rend les conséquences particulièrement incertaines. Pourtant, nous cherchons tous à établir des relations causales évidentes qui permettent une relation plus simple au monde^{ARENDET, 1972}. Dans les politiques publiques, les « baguettes magiques » et les « mythes » abondent et sont pris pour argent comptant (par exemple : « le problème c'est les banlieues, et la solution c'est la fermeté »)^{LACASSE, 1995}, à un moment où paradoxalement, la légitimité de l'action publique tient de plus en plus à la preuve de son efficacité.

Dans ce contexte, que faire ? Les évaluateurs peuvent reconnaître la complexité de nos sociétés, mais chercher à s'en abstraire en agissant dans un cadre plus simple, contrôlé. C'est ainsi qu'ils pourront affirmer un lien de causalité, voire des lois générales et des propriétés à valeur intrinsèque (des « baguettes magiques »), permettant au décideur de décider plus facilement ce qu'il choisit de faire ou pas. La méthode expérimentale établie par Donald Campbell s'inscrit dans ce sens.

Ou au contraire, on peut vouloir évaluer les actions publiques dans leur contexte, malgré leur complexité, quitte à établir des causalités moins certaines, et à accepter que les résultats de l'évaluation ont une valeur relative à l'environnement dans lequel ils ont été établis^{MONNIER ET AL., 1992}. Mais la reconnaissance de l'incertitude des résultats peut être à double tranchant : trop peu conclusifs pour décider, suffisamment malléables pour ne pas remettre en cause profondément les politiques menées.

Enfin, on peut estimer que dans un cadre complexe, l'établissement de relations causales est impossible, ou estimer que les évaluations sont peu utiles, ou tendront quoiqu'il arrive à consolider les politiques engagées, même si elles sont non fondées, ce qui fonde parmi un certain nombre de personnes le refus de l'évaluation d'impact, voire de la pratique évaluative.


UNE DÉMARCHE D'INFÉRENCE CAUSALE ?


Dans l'évaluation d'impact, il faut donc se mettre d'accord sur la question causale, c'est-à-dire le lien entre l'intervention évaluée et l'effet étudié, qu'il soit désiré ou non. L'hypothèse sur ce lien est une inférence causale.


Construire une hypothèse causale

Ce lien initial peut être conçu dans une approche :

- ▶ **inductive.** Il s'agit de partir de l'observation des faits, sans hypothèse préalable, pour établir progressivement une relation causale sur une base probabiliste.

 Qu'est-ce qui rend une intervention complexe ? La multiplicité de ces acteurs, des autres politiques, des contextes et des échelles d'intervention ; la durée de l'intervention ; le fait que sa mise en œuvre dépende de la façon dont elle est réceptionnée par les acteurs, qui s'en saisissent ou non, et des rétroactions entre eux ; et que l'intervention évolue au gré de ce processus. SUS, ROBERT, RIDDE, 2013

 Quitte à filer la métaphore, ce que dit E. Monnier c'est qu'un même projectile n'aura pas le même impact sur un mur de brique, de terre ou de béton armé.

 Un exemple de démarche inductive: En mesurant systématiquement le poids des enfants à la naissance et en le comparant aux caractéristiques de leurs parents, on peut faire l'hypothèse que, par exemple, les enfants de fumeurs ont un poids plus faible que la moyenne.

- ▶ **déductif**. Une hypothèse initiale est établie, issue de la recherche ou de la logique par exemple. L'observation sert à valider ou à réfuter les prémisses de l'hypothèse.
- ▶ **abductif**. Il s'agit de définir quelle est la cause la plus vraisemblable, ou la plus partagée, à un événement. On affirme alors cette relation à titre d'hypothèse, qui sera testée ultérieurement.

Ces approches ne sont pas irréductiblement opposées et peuvent se combiner ou se compléter dans le temps.

Par exemple, pour évaluer une politique polémique, on peut partir des théories de la recherche, ou au contraire des croyances des acteurs publics pour les tester. Ce lien peut être confirmé, circonstancié, ou réfuté, et dans ce cas des théories issues de la recherche utilisées pour mieux rendre compte des liens causaux observés.

De même, l'approche inductive peut être employée pour rechercher les conséquences potentielles d'une intervention et ainsi tourner les projecteurs vers une relation causale inattendue, qui pourra ensuite être théorisée et testée.

Dans tous les cas, l'évaluation commence et/ou se termine par l'élaboration de théories sur le changement, c'est-à-dire une explication sur la façon dont une intervention a des effets. Cette théorie est progressivement raffinée au sein d'une même évaluation ou au fur et à mesure que des évaluations sont menées sur un même sujet.

Penser la relation entre la cause et l'effet

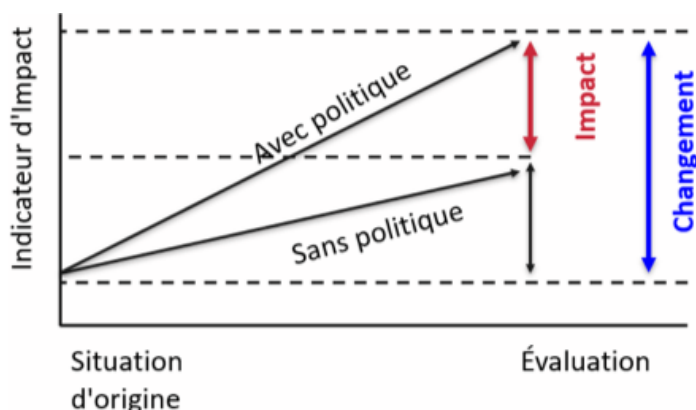
La façon dont nous construisons l'hypothèse causale est indissociablement liée à la façon dont nous concevons la relation entre la cause et l'effet.

On peut distinguer quatre approches prégnantes en évaluation :

- ▶ une **approche successionniste** cherche à établir des relations directes entre un certain nombre de variables : des variables indépendantes qui représentent la ou les causes testées, et une variable dépendante qui traduit l'effet des variables indépendantes. Bien sûr, corrélation n'est pas causalité, et dans cette approche on cherchera à multiplier les tests, par exemple en manipulant dans les expérimentations diverses variables indépendantes pour les observer séparément. L'association causale augmente ainsi avec le nombre de cas où il y a une conjonction entre la cause et l'effet, et baisse avec le nombre de cas où seul un des termes de l'équation, cause ou effet est présent : c'est la **régularité** qui assure l'impact.
- ▶ l'**approche contrefactuelle** s'inscrit aussi dans l'approche successionniste. Ce qui varie, c'est surtout la façon dont la preuve de l'impact est apportée, non plus par l'agrégation de cas multiples, mais par différence entre une situation « avec » et une situation « sans » (schéma à droite). Pour prouver au mieux cette différence, il faut alors envisager de constituer une expérience, aux conditions contrôlées, per-


💡 - En psychiatrie, on parle d'apophénie pour les personnes qui font des associations fortes entre des événements qui n'ont pas de lien entre eux.

💡 - C'est notamment une des promesses du *Big Data* que d'identifier plus clairement des relations inattendues entre des événements divers MONROE, 2015.



mettant de répartir les variables indépendantes entre deux groupes, ou de simuler cette expérience à partir de données ex-post (approche quasi-experimentale). C'est ce qui permettra d'identifier les véritables causes d'un phénomène, et d'estimer la taille et la significativité de leur effet.

- ▶ l'**approche configurationnelle** rejette cette vision successionniste dans laquelle on se focalise sur des causalités simples et immédiates, alors que plus probablement, la plupart des changements sociaux traités dans les évaluations sont l'objet de causes multiples, certaines pouvant être immédiates, mais d'autres préparant le terrain nécessaire pour que ces causes immédiates agissent^{BEFANI, 2012}. Comment attribuer à une nouvelle méthode éducative les résultats scolaires des enfants quand ceux-ci dépendent tout autant du travail des enseignants, de l'entourage familial, etc ? Dans ce cadre, on ne parle donc plus en termes de variables à manipuler, mais de conditions. À partir de l'analyse d'un certain nombre de cas présentant des similitudes et des différences, on cherche à identifier la conjonction ou la configuration de conditions associées à un changement, qu'il faut donc pouvoir mesurer ou a minima qualifier systématiquement. Ces « grappes de conditions » peuvent être suffisantes et/ou nécessaires pour aboutir à un changement donné ou pour expliquer différents effets, ce qui tient lieu d'explication causale.
- ▶ dans une **approche générative** ou explicative, on commence également par une analyse des changements et des régularités qui peuvent être observées. Mais là où les approches successionnistes ou configurationnelles portent leurs efforts sur l'existence du lien causal, l'approche générative cherche avant tout à l'expliquer par des chemins d'impact (quelles séries d'évènements mènent aux changements observés ?) et par des mécanismes sous-jacents. Ces derniers ne sont généralement pas ou mal mesurables, puisqu'ils décrivent des processus humains, qui relèvent de la psychologie intime ou des interactions avec l'environnement. Le déclenchement de ces mécanismes par les interventions évaluées dépend de contextes, propres à chacun, qui devront également être étudiés. Les systèmes étudiés ne sont pas stables : les processus causaux transforment les contextes et se reconfigurent sans cesse. La causalité tient alors à la force explicative de la théorie qui est produite, et donc de la profondeur de l'analyse, qui en contrepartie porte sur des objets restreints.

 Pour le dire autrement, l'approche configurationnelle donne les ingrédients de la recette, là où l'approche générative cherche surtout à en connaître l'ordre, les proportions, et pourquoi et comment elle fonctionne.

Donnons l'exemple d'un accompagnement individualisé vers l'emploi :

- ▶ dans l'approche par **régularité**, je peux dans un premier temps identifier si les personnes accompagnées retournent plus vite vers l'emploi que la moyenne.
- ▶ Mais je peux aussi constituer un contrefactuel, c'est-à-dire un groupe de personnes non accompagnées, partageant les mêmes caractéristiques que les personnes accompagnées, pour vérifier si à caractéristiques mesurables égales, les personnes accompagnées retrouvent plus vite un emploi.

.....
« les personnes accompagnées ont 32 % de chances supplémentaires de retrouver un emploi en comparaison des personnes non accompagnées »
.....

.....
« Toutes choses égales par ailleurs, l'accompagnement se traduit par un effet net de +12 points sur le retour à l'emploi des personnes accompagnées »
.....

- ▶ dans l'approche configurationnelle, je m'intéresse aux personnes accompagnées uniquement : certaines ont retrouvé un emploi, d'autres non. Peut-on identifier parmi les caractéristiques des personnes, ou de leur accompagnement, ou de leur environnement, des conditions qui seraient associées au succès, et d'autres qui seraient associées à l'échec ?
- ▶ dans l'approche générative, je cherche à comprendre pourquoi, comment et dans quels cas l'accompagnement permet ou non le retour vers l'emploi.

Conclure à la relation causale

Chacune des quatre grandes approches a ses propres critères permettant de conclure à la relation causale. Nous proposons ici quelques grands principes transversaux.

Un premier principe tient à la **gradation de l'affirmation causale**. Les éléments qui sont recueillis dans une évaluation d'impact peuvent être de nature différente. Nous postulons ici que l'on peut définir trois grandes étapes :

Une bonne théorie

▶ **une bonne théorie**, appuyée sur des recoupements, des premières observations, un travail de réfutation logique, lorsque c'est possible l'intégration de théories existantes issues de la recherche, est le premier stade de l'affirmation causale. Cette théorie est alimentée en permanence par les deux étapes à venir ;

Des données sur les changements escomptés

▶ **appuyer cette théorie par des données** sur les changements escomptés, les caractéristiques des bénéficiaires, vérifier comment elles évoluent dans le temps, apprécier les corrélations apparentes ou leur absence, et comparer ces données avec la théorie initiale, devrait constituer une seconde étape ;

Tester la théorie par inférence causale

▶ **tester systématiquement** la théorie selon une démarche d'inférence causale, dans une ou plusieurs des approches présentées plus haut, est une dernière étape.

Les commandes d'évaluation d'impact veulent souvent aller directement à la dernière étape, sans avoir mené à bien les deux premières, qui sont pourtant essentielles et dans bien des cas suffiraient à mener des politiques informées. C'est vrai pour les programmes sociaux comme pour la crème solaire^{OSTER, 2014}.

Un deuxième principe tient à la qualité des résultats. La **validité interne** est une mesure de cette qualité. Dans une approche successionniste, elle va être pensée en termes de réalité de la causalité (n'y a-t-il pas d'autres causes possibles qui n'ont pas été testées ?) et (d'absence) de biais (l'évaluation a-t-elle été faite dans les règles de l'art ?) ; les termes sont différents, mais l'idée globalement la même, pour les autres approches : a-t-on fait pour le mieux pour se permettre de conclure ? a-t-on collecté suffisamment d'information ? a-t-elle été traitée de la bonne façon, sans omettre les résultats qui ne cadraient pas avec la conclusion ? etc.

« les chômeurs de longue durée, issus du secteur industriel, dans des zones de déprise industrielle, sont peu susceptibles de retrouver un emploi à l'issue du dispositif lorsque l'accompagnement ne comprend aucun module de formation »

« les accompagnateurs, par leur discours, font porter la faute du chômage sur les ouvriers chômeurs de longue durée, sans leur donner d'échappatoire. Cela renforce un processus d'autodépréciation qui les pousse à ne plus entreprendre de démarches, qui sont synonymes pour eux d'humiliations supplémentaires. En revanche, ce discours fonctionne très bien avec des chômeurs plus jeunes qui y voient un appel à compter sur leurs propres forces pour retrouver un travail et sont ainsi remotivés »



Personne n'a jamais décidé d'exposer 20 ans une population en plein soleil, la moitié badigeonnée de crème solaire pendant que l'autre serait laissée sans protection pour estimer la différence entre le taux de cancer de chaque groupe. Pour autant, nous disposons d'une théorie solide pour expliquer pourquoi la crème solaire réduit le risque de cancer, et d'études longitudinales qui tendent à valider cette théorie.

Un point important à retenir à ce stade est qu'**une parfaite validité, pour l'évaluation de l'action publique, est pratiquement inaccessible.**

« Ne pas reconnaître l'incertitude et la complexité n'est pas seulement une erreur technique, mais aussi éthique. » CILLIERS, 2005


Les individus ne sont pas des rats de laboratoire aux variables manipulables, il y a des facteurs que nous n'envisageons, ou que nous ne pouvons pas mesurer, des influences externes inconnues. Ensuite, il faut se représenter ce qu'est une évaluation d'impact dans les faits : l'approche d'ensemble est déclinée en méthodes, elles-mêmes appuyées sur des outils de collecte et d'analyse, dont la mise en œuvre dépend in fine du savoir-faire des personnes qui conduisent cette évaluation, voire du hasard. Chacun des maillons compte. Ce qu'il est possible de faire, c'est donc de faire œuvre de transparence pour permettre à d'autres de critiquer et de compléter au mieux.


Un dernier principe tient à la **répétition des approches**. Ce principe tient du bon sens : 1 évaluation peut se tromper 1 fois, mais 1 000 études peuvent-elles se tromper 1 000 fois ? Des méthodes ont été développées pour distinguer, dans les différentes approches exposées ci-dessus, des enseignements de valeur issus de différentes évaluations portant sur les mêmes sujets. Les méta-analyses permettent de monter en généralité là où des évaluations ponctuelles ne le pourraient pas. La synthèse réaliste (approche générative) vise à distinguer systématiquement quels mécanismes mènent à quels résultats dans quels contextes. Pour autant, il est probablement nécessaire de garder une certaine prudence : si 1 000 études ont été menées de la même manière, ne risque-t-on pas de garder les mêmes zones d'ombre qu'avec une seule ? Peut-être plus ambitieux, l'idée d'avoir une triangulation permanente, c'est-à-dire répondre à la même question de recherche, mais avec des démarches d'inférence causale et des approches et méthodes différentes, est sans doute une bonne façon de consolider progressivement les résultats obtenus (MUNAFÒ & SMITH, 2019).

CONCLUSION

La causalité est un sujet central, à la fois dans notre quotidien, mais aussi dans la recherche et dans l'évaluation. Il existe plusieurs façons de comprendre cette notion, qui mènent à des approches d'évaluation d'impact fondamentalement différentes. Les attentes que nous avons vis-à-vis de l'évaluation doivent permettre de guider les choix de stratégie évaluative, selon que l'évaluation serve plutôt une démarche de redevabilité (est-ce que ça marche ?) ou d'amélioration (dans quels cas ça marche, pourquoi, comment ?).

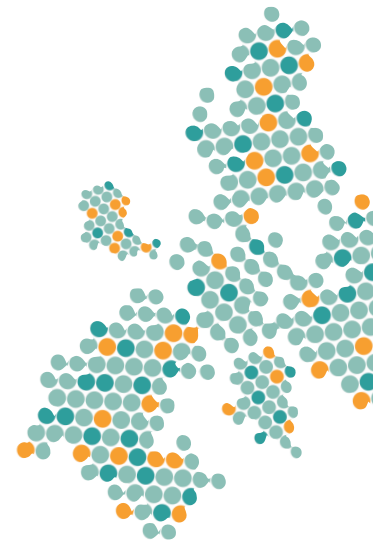
Reste la question : comment avoir une vraie discussion sur les démarches d'inférence causale et ce que l'on peut en attendre, pour faire des choix informés ?

 - Voir à ce sujet la récente reproduction d'une expérimentation par assignation aléatoire, qui en reproduisant les calculs en remet largement en cause la validité. BÉDÉCARRATS ET AL., 2018

 - Une méta-analyse récente répond par exemple à la question : la taille de la classe a-t-elle un effet sur les résultats des élèves ? CAMPBELL, COLLABO' RATION, 2018

BIBLIOGRAPHIE

- Bédécarrats, F., Guérin, I., Morvant-Roux, S., & Roubaud, F. (2018). Verifying the internal validity of a flagship RCT: A review of Crépon, Devoto, Duflo and Pariente (American Economic Journal: Applied Economics, 2015.)
- Befani, B. (2012). Models of Causality and Causal Inference.
- Campbell, D.T. (1979). Assessing the impact of planned social change. *Evaluation and Program Planning*, 2(1), 67–90.
- Campbell Collaboration (2018). Small class sizes for improving student achievement in primary and secondary schools. [lien](#)
- Cilliers, P. (2005), Complexity, deconstruction and relativism, *Theory, Culture & Society* 22 (5) 255–267.
- Devaux-Spatarakis, A. (2014). La méthode expérimentale par assignation aléatoire: Un instrument de recomposition de l'interaction entre sciences sociales et action publique en France?
- Donaldson, S. I., Christie C., & Mark M. Éd. (2009), What counts as credible evidence in applied research and evaluation practice.
- Hulswit, M. (2004). A Short History of 'Causation'. *SEED*, 4(3), 16–42.
- Hulswit, M. (2004). Causality and Causation: The Inadequacy of the Received View. *SEED*, 4(2), 3–23.
- Lacasse, F. (1995). Mythes, savoirs et décisions politiques.
- Mill, J.S. (1843). A system of logic
- Monnier, E., Conan, M., Allen, B., Duran, P., Spenlehauer, V., & Toulemonde, J. (1992). Évaluations de l'action des pouvoirs publics.
- Monroe, B. L., Pan, J., Roberts, M. E., Sen, M., & Sinclair, B. (2015). No! Formal Theory, Causal Inference, and Big Data Are Not Contradictory Trends in Political Science. *Political Science & Politics*, 48(01), 71–74.
- Munafò, M. R., & Smith, G. D. (2018). Robust research needs many lines of evidence. *Nature*, 553(7689), 399–401.
- Shadish, W. R., Cook, T. D., & Leviton, L. C. (1991). *Foundations of Program Evaluation: Theories of Practice*. SAGE.
- Oster, E. (2014, août 12). Is Sunscreen A Life-saver Or A Poison? [lien](#)



Pour en savoir plus

Retrouvez des guides, des outils sur l'évaluation sur notre site <http://quadrant-conseil.fr/ressources.php>

Impliquez-vous dans l'évaluation en vous rendant sur le site de la SFE : www.sfe-asso.fr



Cette fiche a été rédigée dans le cadre d'une commande de Pôle Emploi. Réutilisation possible sans permission en citant la source: « Quadrant Conseil, 2019 »