

- Résumé
- 1. Une assertion causale est une déclaration sur ce qui ne s'est pas produit.
- 2. Pas de causalité sans manipulation.
- 3. Les causes sont non rivales.
- 4.  $X$  peut causer  $Y$  même si  $X$  n'est pas une condition nécessaire ou une condition suffisante pour  $Y$ .
- 5. Le problème fondamental de l'inférence causale.
- 6. Vous pouvez estimer l'effet causal moyen même si vous ne pouvez observer aucun effet causal individuel.
- 7. L'estimation de l'effet causal moyen ne nécessite pas que les groupes de traitement et de contrôle soient identiques.
- 8. La corrélation n'est pas la causalité.
- 9. Si vous savez qu'en moyenne  $A$  cause  $B$  et  $B$  cause  $C$ , cela ne veut pas dire qu'en moyenne  $A$  cause  $C$ .
- 10. Il est plus facile d'en apprendre davantage sur les "effets des causes" que sur les "causes des effets".

## Résumé

Le philosophe David Lewis a décrit la causalité comme "quelque chose qui fait une différence, et cette différence doit être une différence par rapport à ce qui se serait passé sans elle".<sup>1</sup> Ceci est l'interprétation de la causalité pour la plupart des expérimentalistes. Même si la définition semble simple, elle a de nombreuses implications subtiles. Voici dix idées impliquées par cette notion de causalité qui importent pour le design de recherche.<sup>2</sup>

# 1. Une assertion causale est une déclaration sur ce qui ne s'est pas produit.

Pour la plupart des expérimentalistes, la déclaration " $X$  a causé  $Y$ " signifie que  $Y$  est présent *et*  $Y$  n'aurait pas été présent si  $X$  n'avait pas été présent. Cette définition requiert une notion de ce qui aurait pu arriver, mais ne s'est pas produit.<sup>3</sup> De même, "l'effet" de  $X$  sur  $Y$  est la différence entre la valeur que  $Y$  aurait prise étant donné une valeur de  $X$  et la valeur que  $Y$  aurait prise étant donné une autre valeur de  $X$ . En raison de l'accent mis sur les différences entre les résultats, cette approche est parfois appelée approche "des différences" ou "contrefactuelle" de la causalité.

**Note technique:** Les statisticiens emploient le cadre des "résultats potentiels" pour décrire les relations contrefactuelles. Dans ce cadre,  $Y_i(1)$  désigne le résultat pour l'unité  $i$  qui serait observé sous une condition (par exemple, si l'unité  $i$  a reçu un traitement) et  $Y_i(0)$  désigne le résultat qui serait observé dans une autre condition (par exemple, si l'unité  $i$  n'a pas reçu le traitement). Un effet causal du traitement pour l'unité  $i$  pourrait être une simple différence des résultats potentiels  $\tau_i = Y_i(1) - Y_i(0)$ . Un traitement a un effet causal (positif ou négatif) sur  $Y$  pour l'unité  $i$  si  $Y_i(1) \neq Y_i(0)$ .

## 2. Pas de causalité sans manipulation.

La définition “contrefactuelle” de la causalité exige que l’on soit capable de réfléchir aux résultats qui peuvent entraîner des conditions différentes. Quelle serait la situation si un parti plutôt qu’un autre était élu ? Les déclarations causales de tous les jours ne répondent souvent pas à cette exigence de l’une des deux manières suivantes.

- Premièrement, certaines déclarations ne précisent pas de conditions contrefactuelles claires. Par exemple, l’affirmation selon laquelle “la récession a été causée par Wall Street” n’indique pas de contrefactuel évident — devons-nous examiner s’il y aurait eu une récession si Wall Street n’avait pas existé ? Ou est-ce une déclaration sur des décisions particulières que Wall Street aurait pu prendre mais n’a pas prises ? Si oui, quelles décisions ? La validité de telles déclarations est difficile à évaluer et peut dépendre des conditions contrefactuelles impliquées par une déclaration.
- Deuxièmement, certaines déclarations impliquent des conditions contrefactuelles qui ne peuvent être imaginées. Par exemple, l’affirmation selon laquelle Peter a obtenu le poste parce qu’il est Peter implique une considération de ce qui se serait passé si Peter n’était pas Peter. Alternativement, l’affirmation selon laquelle Peter a obtenu le poste parce qu’il est un homme nécessite de considérer Peter comme autre qu’un homme. Le problème est que les contrefactuels dans ces cas impliquent un changement non seulement de la condition à laquelle fait face un individu, mais de l’individu lui-même.

Pour éviter de tels problèmes, certains statisticiens recommandent de restreindre les assertions causales aux traitements qui peuvent en théorie (pas nécessairement en pratique) être manipulés.<sup>4</sup> Par exemple, alors que nous pourrions avoir des difficultés avec l’affirmation selon laquelle Peter a obtenu le poste parce qu’il était un homme, nous n’avons pas de telles difficultés avec l’affirmation selon laquelle Peter a obtenu le poste parce que l’agence de recrutement pensait qu’il était un homme.

### 3. Les causes sont non rivales.

Même si nous pouvons nous concentrer sur l’effet d’une seule cause  $X$  sur un résultat  $Y$ , nous ne nous attendons généralement pas à ce qu’il n’y ait qu’une seule cause de  $Y$ .<sup>5</sup> De plus, si vous additionnez les effets causaux de différentes causes, il n’y a aucune raison de s’attendre à ce qu’ils totalisent 100 %. Par conséquent, il ne sert à rien d’essayer de “répartir” les résultats entre différents facteurs de causalité. En d’autres termes, les causes sont non rivales. La National Rifle Association soutient, par exemple, que les armes à feu ne tuent pas les gens, les gens tuent les gens. Cette déclaration n’a pas beaucoup de sens dans le cadre contrefactuel. Enlevez les armes à feu et vous n’aurez pas de morts par balles. Les armes à feu sont donc une cause. Enlevez les gens et vous n’aurez pas non plus de décès par balle, donc les gens sont aussi une cause. En d’autres termes, ces deux facteurs sont simultanément les causes des mêmes résultats.

### 4. $X$ peut causer $Y$ même si $X$ n’est pas une condition nécessaire ou une condition suffisante pour $Y$ .

On parle souvent des relations causales en termes déterministes. Même la citation de Lewis en haut de cette page semble suggérer une relation déterministe entre les causes et les effets. On pense parfois que les relations causales impliquent des conditions nécessaires (pour que  $Y$  se produise,  $X$  doit se

produire); on pense parfois que de telles relations impliquent des conditions suffisantes (si  $X$  se produit, alors  $Y$  se produit). Mais une fois que nous parlons d'unités multiples, il y a au moins deux façons de penser que  $X$  cause  $Y$  même si  $X$  n'est ni une condition nécessaire ni une condition suffisante pour  $Y$ . La première consiste à tout réinterpréter en termes probabilistes : par  $X$  cause  $Y$ , on entend simplement que la probabilité de  $Y$  est plus élevée lorsque  $X$  est présent. Une autre consiste à tenir compte des contingences — par exemple,  $X$  peut causer  $Y$  si la condition  $Z$  est présente, mais pas dans le cas contraire.<sup>6</sup>

## 5. Le problème fondamental de l'inférence causale.

Si les effets causaux sont des déclarations sur la différence entre ce qui s'est produit et ce qui aurait pu se produire, alors les effets causaux ne peuvent pas être mesurés. Mauvaise nouvelle ! De manière prospective, vous pouvez organiser les choses de manière à observer ce qui se passe si une personne reçoit un traitement ou ce qui se passe si elle ne reçoit pas le traitement. Pourtant, pour la même personne, vous ne pourrez jamais observer ces deux résultats et leur différence. Cette incapacité à observer les effets causaux au niveau de l'unité est souvent appelée le “problème fondamental de l'inférence causale”.

## 6. Vous pouvez estimer l'effet causal moyen même si vous ne pouvez observer aucun effet causal individuel.

Même si vous ne pouvez pas observer si  $X$  cause  $Y$  pour une unité donnée, il est peut-être toujours possible de déterminer si  $X$  cause  $Y$  en moyenne. L'effet causal moyen est égal à la différence entre le résultat moyen pour toutes les unités si elles étaient toutes dans la condition de contrôle et le résultat moyen pour toutes les unités si elles étaient toutes dans la condition de traitement. De nombreuses stratégies d'identification causale (voir 10 stratégies pour déterminer si  $X$  a causé  $Y$  (<http://egap.org/resource/10-strategies-for-figuring-out-if-x-caused-y>)) se concentrent sur des façons d'en savoir plus sur ces résultats potentiels moyens.<sup>7</sup>

10 choses à savoir sur les tests d'hypothèse (<https://egap.org/resource/10-things-to-know-about-hypothesis-testing/>) décrit comment en savoir plus sur les effets causaux individuels plutôt que sur les effets causaux moyens étant donné le problème fondamental de l'inférence causale.

## 7. L'estimation de l'effet causal moyen ne nécessite pas que les groupes de traitement et de contrôle soient identiques.

Une stratégie que les gens utilisent pour en savoir plus sur l'effet causal moyen consiste à créer des groupes de traitement et de contrôle par randomisation (voir 10 Stratégies pour déterminer si  $X$  a causé  $Y$  (<http://egap.org/resource/10-strategies-for-figuring-out-if-x-caused-y>)). Ce faisant, les chercheurs s'inquiètent parfois s'ils constatent que les groupes de traitement et de contrôle qui en résultent ne sont pas comparables sur certaines dimensions importantes.

La bonne nouvelle est que l'argument expliquant pourquoi les différences dans les résultats moyens entre les groupes de traitement et de contrôle assignés de manière aléatoire capturent l'effet moyen du traitement (en espérance pour des randomisations répétées au sein du même groupe d'unités) ne repose *pas* sur le fait que les groupes de traitement et de contrôle ont des caractéristiques observées similaires. Il repose uniquement sur l'idée que, en moyenne, les résultats dans les groupes de traitement et de contrôle captureront les résultats moyens pour toutes les unités du groupe expérimental si elles étaient, respectivement, dans le traitement ou dans le contrôle. En pratique, les groupes de traitement et de contrôle réels ne seront pas identiques.<sup>8</sup>

## 8. La corrélation n'est pas la causalité.

Une corrélation entre  $X$  et  $Y$  est une déclaration sur les relations entre les résultats réels, et non sur la relation entre les résultats réels et les résultats contrefactuels. Ainsi, les déclarations sur les causes et les corrélations n'ont pas grand-chose à voir les unes avec les autres. Des corrélations positives peuvent être cohérentes avec des effets causaux positifs, aucun effet causal ou même des effets causaux négatifs. Par exemple, la prise de médicaments contre la toux est positivement corrélée à la toux mais a, espérons-le, un effet causal négatif sur la toux.<sup>9</sup>

## 9. Si vous savez qu'en moyenne $A$ cause $B$ et $B$ cause $C$ , cela ne veut pas dire qu'en moyenne $A$ cause $C$ .

Vous pourriez vous attendre à ce que si  $A$  cause  $B$  et  $B$  cause  $C$ , alors  $A$  cause  $C$ .<sup>10</sup> Mais il n'y a aucune raison que les relations causales moyennes soient transitives. Imaginez que  $A$  cause  $B$  pour les hommes mais pas les femmes et  $B$  cause  $C$  pour les femmes mais pas les hommes. Ensuite, en moyenne,  $A$  cause  $B$  et  $B$  cause  $C$ , mais  $A$  ne cause pas  $C$  à travers  $B$ .

## 10. Il est plus facile d'en apprendre davantage sur les "effets des causes" que sur les "causes des effets".

Bien que cela puisse sembler être deux façons de dire la même chose, il y a une différence entre comprendre quel est l'effet de  $X$  sur  $Y$  (les "effets d'une cause") et si un résultat  $Y$  était dû à une cause  $X$  (la "cause d'un effet").<sup>11</sup> Considérez l'exemple suivant. Supposons que nous menions une expérience avec un échantillon qui contient un nombre égal d'hommes et de femmes. L'expérience assigne de manière aléatoire des hommes et des femmes à un traitement binaire  $X$  et mesure un résultat binaire  $Y$ . De plus, supposons que  $X$  ait un effet positif de 1 pour tous les hommes, c'est-à-dire le résultat potentiel de contrôle des hommes est de zéro,  $Y_i(0) = 0$ , et leur résultat potentiel traité est de un,  $Y_i(1) = 1$ . Pour toutes les femmes,  $X$  a un effet négatif de  $-1$ , c'est-à-dire que le résultat potentiel de contrôle des femmes est de un,  $Y_i(0) = 1$ , et leur résultat potentiel traité est de zéro,  $Y_i(1) = 0$ . Dans cet exemple, l'effet moyen de  $X$  sur  $Y$  est nul. Mais pour tous les participants du groupe de traitement avec  $Y = 1$ , il est vrai que  $Y = 1$  car  $X = 1$ . De même, pour tous les participants du groupe de

traitement avec  $Y = 0$ , il est vrai que  $Y = 0$  car  $X = 1$ . L'expérimentation peut obtenir une réponse exacte à la question sur les "effets d'une cause", mais il n'est généralement pas possible d'obtenir une réponse exacte à la question sur la "cause d'un effet".<sup>12</sup>

1. Lewis, David. "Causation." The journal of philosophy (1973): 556-567.↵
2. Auteur d'origine : Macartan Humphreys. Révisions mineures : Winston Lin et Donald P. Green, 24 juin 2016. Révisions MH 6 janvier 2020. Révisions Anna Wilke mai 2021. Le guide est un document vivant et peut être mis à jour par les membres de EGAP à tout moment ; les contributeurs répertoriés ne sont pas responsables des modifications ultérieures.↵
3. Holland, Paul W. "Statistics and causal inference." Journal of the American Statistical Association 81.396 (1986): 945-960.↵
4. Holland, Paul W. "Statistics and causal inference." Journal of the American Statistical Association 81.396 (1986): 945-960.↵
5. Certains appellent cela le "problème des causes de prodigalité".↵
6. Mackie a présenté l'idée de conditions dites "INSS" ("INUS" en anglais) pour capturer la dépendance des causes sur d'autres causes. Une cause peut être une partie *Insuffisante* mais *Nécessaire* d'une condition qui est elle-même *Superflue* mais *Suffisante*. Par exemple, composer un numéro de téléphone est une cause de "contacter quelqu'un" car avoir une connexion et composer un numéro est suffisant (S) pour passer un appel téléphonique, alors que composer seul sans connexion ne suffirait pas (I), ni avoir une connexion (N). Il existe bien sûr d'autres moyens de contacter quelqu'un sans passer d'appels téléphoniques (S). Mackie, John L. "The cement of the universe." London: Oxford Uni (1974).↵
7. **Note technique :** La principale idée technique est que la différence des moyennes est la même que la moyenne des différences. C'est-à-dire, en utilisant "l'opérateur d'espérance",  $\mathbb{E}(\tau_i) = \mathbb{E}(Y_i(1) - Y_i(0)) = \mathbb{E}(Y_i(1)) - \mathbb{E}(Y_i(0))$ . Les termes à l'intérieur de l'opérateur d'espérance dans la deuxième quantité ne peuvent pas être estimés, mais les termes à l'intérieur de l'opérateur d'espérance dans la troisième quantité peuvent l'être.<sup>7</sup>Voir l'illustration ici (<https://raw.githubusercontent.com/egap/20methodes-guides/master/causal-inference/PO.jpg>).↵
8. Pour cette raison, les tests- $t$  pour vérifier si "la randomisation a fonctionné" n'ont pas beaucoup de sens, du moins si vous savez qu'une procédure randomisée a été suivie — simplement par hasard, 1 test sur 20 montrera des différences statistiquement détectables entre les groupes de traitement et de contrôle. En cas de doute sur la mise en œuvre correcte d'une procédure randomisée, ces tests peuvent être utilisés pour tester l'hypothèse selon laquelle les données ont bien été générées par une procédure randomisée. Ces tests peuvent alors être particulièrement importants pour des expériences de terrain où les chaînes de communication entre la personne randomisant et la personne mettant en œuvre l'assignation du traitement peuvent être longues et complexes.↵
9. **Note technique:** soit  $D_i$  un indicateur pour savoir si l'unité  $i$  a reçu un traitement ou non. Alors, la différence des résultats moyens entre ceux qui reçoivent le traitement et ceux qui ne le reçoivent pas peut s'écrire  $\frac{\sum_i D_i \times Y_i(1)}{\sum_i D_i} - \frac{\sum_i (1-D_i) \times Y_i(0)}{\sum_i (1-D_i)}$ . En l'absence d'informations sur la manière dont le traitement a été assigné, nous ne pouvons pas dire si cette différence est un bon estimateur de l'effet moyen du traitement, c'est-à-dire de la différence entre les résultats

potentiels moyens pour les groupes de traitement et de contrôle pour toutes les unités. Ce qui importe est de savoir si  $\frac{\sum_i D_i \times Y_i(1)}{\sum_i D_i}$  est une bonne estimation de  $\frac{\sum_i 1 \times Y_i(1)}{\sum_i 1}$  et si  $\frac{\sum_i (1-D_i) \times Y_i(0)}{\sum_i (1-D_i)}$  est une bonne estimation de  $\frac{\sum_i 1 \times Y_i(0)}{\sum_i 1}$ . Cela pourrait être le cas si ceux qui ont reçu un traitement sont un échantillon représentatif de toutes les unités, mais sinon il n'y a aucune raison de s'attendre à ce qu'il le soit.↵

10. Interprétez “ $A$  cause  $B$ , en moyenne” comme “l’effet moyen de  $A$  sur  $B$  est positif”.↵
11. Certains réinterprètent la question des “causes des effets” comme suit : quelles sont les causes qui ont des effets sur les résultats. Voir Andrew Gelman and Guido Imbens, “Why ask why? Forward causal inference and reverse causal questions”, NBER Working Paper No. 19614 (Nov. 2013).↵
12. Voir, par exemple, Tian, J., Pearl, J. 2000. “Probabilities of Causation: Bounds and Identification.” *Annals of Mathematics and Artificial Intelligence* 28:287–313.↵