

Fondements empiriques de la méthode expérimentale

Mattia A. Fritz

27/03/2022

Introduction

Ce document offre un survol des fondements empiriques de la méthode expérimentale.

Document en construction, pour l’instant se référer aux slides disponibles dans l’espace Moodle du cours

1 Sources de la connaissance empirique

Les connaissances empiriques sont basées sur trois sources :

1. Les observations

La « réalité » est observée de manière plutôt passive, en exploitant des phénomènes qui se manifestent spontanément

2. Les simulations

La « réalité » est (re)construite à travers des mécanismes computationnels qui illustrent l’évolution d’un phénomène

3. Les expériences

La « réalité » est sollicitée par une intervention active, pondérée et planifiée, dont l’effet est évalué sur un phénomène d’intérêt

Il y a eu dans le passé la tendance à séparer nettement les trois sources, notamment au niveau de la distinction entre observations et expériences selon le principe “*correlation is not causation*”. Dans ce sens, les expériences sont identifiées comme le *golden standard* pour établir des mécanismes causales. Plus récemment, il y a la tendance à voir la causalité comme un effet émergent qui peut être établi également lorsque l’intervention n’est pas effectuée de

manière active, mais inférer depuis des situations qui répondent à des critères spécifiques (Pearl, 2000 ; Pearl et al., 2016 ; Pearl & Mackenzie, 2018 ; Rohrer, 2018).

De plus, il existe des recouvrements entre les trois sources qui sont souvent utilisées de manière complémentaires pour maximiser la connaissance qu'on peut avoir sur un sujet. Le reste de cette section s'intéresse néanmoins aux expériences de manière plutôt *classique* (Maxwell et al., 2017), afin de fournir un aperçu conceptuel qui peut être ensuite décliné dans des *formats hybrides* plus complexes.

2 Les étapes d'une expérience

On peut identifier 4 étapes principales dans une expérience :

1. La **justification** qui consiste à établir la *raison d'être* d'une intervention active pour en étudier l'effet sur un ou plusieurs phénomènes d'intérêt
2. L'**explication causale potentielle** qui consiste à émettre des hypothèses –avec différents niveaux de précision possibles – sur le mécanisme causale qui détermine/explique l'effet de l'intervention
3. Le **processus génératif de données** qui consiste à créer un micro-monde dans lequel l'intervention puisse être menée dans des conditions qui visent à exclure, idéalement, toute autre source de variabilité potentielle
4. L'**inférence** qui combine des outils statistiques et des connaissances dans le domaine pour déterminer à quel point le processus génératif de données (le micro-monde) fourni des éléments utiles et fiables qu'on peut généraliser au contexte plus large (le macro-monde/la réalité)

3 Justification

Les expériences impliquent une forme d'action – l'intervention – qui est susceptible d'avoir un effet sur la réalité à travers au moins un phénomène identifiable :

- Par **intervention**, on se réfère de manière très flexible à une action qui modifie la « réalité », y compris de manière épistémique (i.e. ce que l'on sait sur un sujet). L'action

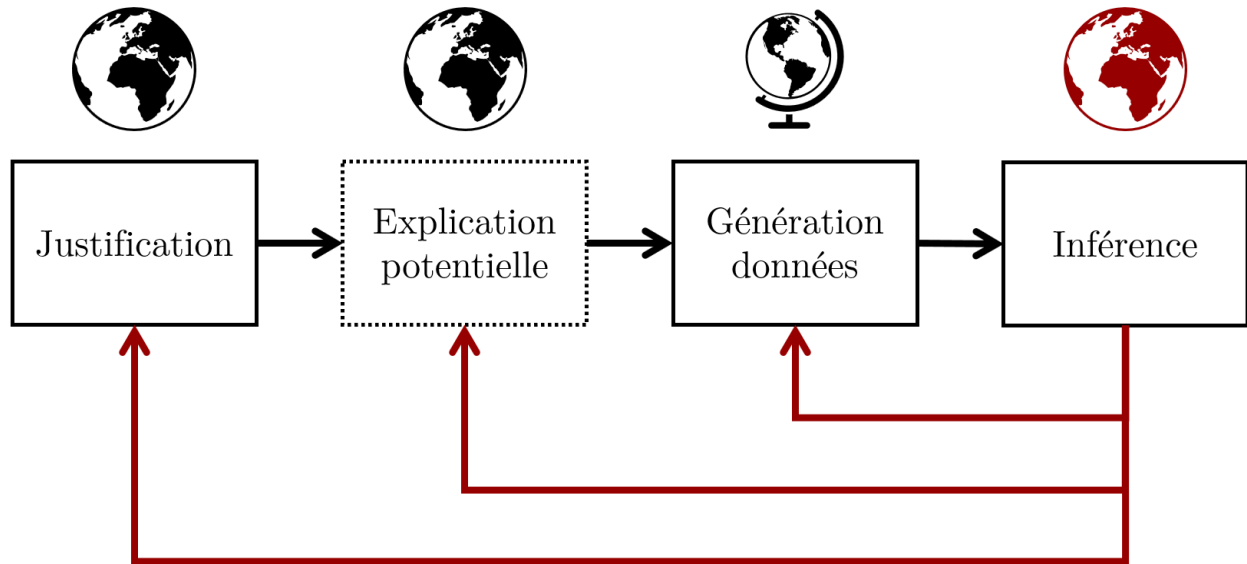


FIG. 1 : Les 4 étapes d'une recherche expérimentale

peut être concrète et appliquée (e.g. introduire une technologie en salle de classe, proposer un scénario pédagogique innovant, mettre à jour une interface utilisateur, ...) ou plus abstraite et théorique (e.g. un principe ou une théorie pédagogique, un référentiel de compétences, un plan d'études, ...).

- Par **phénomène**, on se réfère de manière très flexible à un pattern d'événements qui se répète de manière congruente, par exemple dans des conditions similaires. Le pattern peut être concret et appliqué (e.g. la rétention mnésique d'éléments dans une liste, la perception d'utilité d'un artefact technologique, ...) ou plus abstrait et théorique (e.g. la compréhension de texte, la collaboration, le bien-être des étudiant-es, ...).

De ce fait, les expériences doivent être planifiées et pondérées attentivement selon au moins deux principes phares :

- Quel serait le bénéfice à appliquer l'intervention, comparé à ne pas l'appliquer
- Quel serait la perte potentielle à appliquer l'intervention, comparé à ne pas l'appliquer

Si on prend l'exemple de l'introduction d'une technologie éducative pendant un cours, les deux questions à se poser seront donc (1) qu'est-ce que les étudiant-es obtiennent *en plus* par l'introduction de la technologie ; et (2) qu'est-ce que les étudiant-es *risquent de perdre* à cause de l'introduction de la technologie. Les étudiant-es pourraient gagner en interactivité,

collaboration, maîtrise technique, etc. ; et en même temps risquer de perdre en termes de temps nécessaire pour apprendre à utiliser la technologie, en charge cognitive pour articuler plusieurs environnements de travail, etc.

Dans le domaine de la recherche, en plus, l'introduction d'une intervention doit aussi respecter une double justification :

1. **Scientifique** : est-ce que l'intervention est pertinente par rapport aux connaissances disponibles dans le domaine de référence ?
2. **Éthique** : est-ce que l'intervention considère attentivement le rapport entre bénéfices et risques potentiels, notamment au niveau des participant-es ?

La combinaison entre les deux justifications jettent les bases pour une question de recherche à laquelle on peut essayer de répondre à travers une expérience.

3.1 Justification scientifique

La justification scientifique est représentée dans un article ou contributions scientifiques en général par l'introduction et le cadre théorique.

L'introduction illustre la problématique, sa pertinence par rapport aux connaissances actuelles ou phénomènes émergents. La problématique peut être abordée idéalement dans un espace défini par deux continuum :

1. Exploratoire vs. Confirmatoire

Dans une expérience exploratoire, les connaissances actuelles sont jugées encore incomplètes ou trop fragiles/contradictoires pour formuler une explication causale potentielle précise (voir point suivant). L'objectif de l'expérience est principalement celui de mettre en relief des effets potentiels qui pourraient justifier l'intérêt pour des recherches ultérieurs. Au contraire, les expériences confirmatoires cibles de manière précise une intervention, qui peut être aussi une théorie, avec l'objectif d'en corroborer la validité ou d'en remettre en question la pertinence (e.g. la rejeter). Ce mécanisme en technologie éducative – et plus en général dans les sciences sociales – est très complexe, car les phénomènes sont souvent articulés (Meehl, 1990 ; Scheel et al., 2020).

2. Fondamentale vs. Appliquée

La recherche fondamentale s'intéresse à la compréhension de phénomènes de manière transversale et indépendante à des applications concrètes, tandis que la recherche appliquée vise généralement un contexte et un cadre d'intervention plus spécifique. La

distinction est néanmoins délicate et avec plusieurs recouvrements possibles. On peut retenir généralement l'intention de la recherche : est-ce qu'elle vise à apporter des contributions sur un large domaine d'application ou plutôt sur un cadre plus restreint et spécifique ?

Le cadre théorique articule ce que la communauté scientifique connaît (ou ne connaît pas encore) à propos d'un sujet et comment ces connaissances ont été acquises : observations, simulations, expériences, mais aussi raisonnement, argumentation, synthèse de connaissances actuelles avec des revues de la littérature ou méta-analyses. Le cadre théorique met en général en relief des manques ou évidences ambivalentes qui justifient la nécessité de proposer une nouvelle expérience ou de répliquer une expérience déjà menée afin de *contrôler* la fiabilité et robustesse des résultats.

La justification scientifique dépend des connaissances du domaine spécifique duquel l'intervention s'inspire et/ou dans lequel elle s'applique.

3.2 Justification éthique

La justification éthique considère l'ensemble des implications internes et externes à l'expérience, surtout d'un point de vue des entités impliquées (e.g. les participant-es, les écoles, ...). Outre à la sensibilité de données recueillies qui est partagée aussi avec les études de type observation, les expériences sont caractérisées par une intervention active, décidée en amont par les chercheurs, qui peut par exemple être appliquée seulement à une partie des participant-es. Il faut donc pondérer à quel point l'intervention peut produire des effets *objectivement* ou *subjectivement* négatifs pour les participant-es.

De plus, les implications éthiques ne concernent pas seulement ce qui est fait, mais également ce qui n'est pas fait (et aurait pu être fait). Si on imagine une intervention technopédagogique qui est censée apporter des avantages énormes au groupe qui peut en bénéficier, est-il acceptable d'un point de vue éthique d'avoir un *groupe contrôle* sous forme d'une classe qui n'aura pas accès à cette technologie ?

Ce type de dilemmes étant très délicats, les expériences doivent être acceptées par une commission éthique, à laquelle les chercheurs doivent fournir toute information utile, comme par exemple le public cible, le type d'intervention envisagée, la présence d'éventuels mécanismes cachés aux participant-es, etc.

3.3 Question de recherche expérimentale

La combinaison entre la justification scientifique (qu'est-ce qui est pertinent par rapport aux connaissances actuelles) et la justification éthique (qu'est-ce qui est raisonnable tester en respectant l'intégrité de toutes les parties prenantes) détermine la question de recherche expérimentale.

Comparé à une question de recherche scientifique au sens plus large, une question de recherche expérimentale se caractérise généralement par la prédisposition à identifier précisément l'intervention et le(s) phénomène(s) d'intérêt souvent déjà à partir du titre de la contribution ou de l'article scientifique. Ce mécanisme est d'ailleurs illustré par exemple dans ces sujets de mémoire du Master MALTT appliquant la méthode expérimentale :

- *The influence of background music on learning from text* (Adam, 2019)
- Collaboration en environnement médiatisé par ordinateur : Des usages et de l'impact d'un outil de feedback émotionnel (Perrier, 2017)
- Étude de l'effet de l'esthétique sur l'utilisabilité d'une interface lors d'une tâche de recherche d'informations sur un site Internet (Venni, 2017)

La recherche expérimentale se traduit souvent par des formulations qui suggèrent – avec différents degrés de *précaution* selon le caractère exploratoire ou confirmatoire vu plus haut – des mécanismes causaux potentiels entre l'intervention et l'évolution du phénomène d'intérêt.

4 Explication causale potentielle

L'explication causale potentielle est la seule étape du processus qui n'est pas formellement indispensable à la méthode expérimentale. On peut tout à fait imaginer de mener des expériences dans lesquelles les chercheurs n'émettent pas d'hypothèses sur les mécanismes qui déterminent l'effet de l'intervention sur le phénomène d'intérêt. Cette utilisation non-explicative ou a-théorique s'applique notamment dans certains type de test A/B utilisés en expérience utilisateur (est-ce que le bouton vert génère plus d'inscriptions du bouton bleu?). Cependant, même lorsqu'une hypothèse causale n'est pas formellement testée dans une expérience, l'expérience en elle-même est soutenue au moins par une réflexion de causalité potentielle (j'imagine que le bouton vert peut attirer d'avantage le regard, ou avoir une meilleure lisibilité, ou les deux en même temps). Si ce n'était pas le cas, alors on devrait tester indistinctement toutes les combinaisons potentielles entre éléments, avec des coûts et pertes de temps exorbitants.

4.1 L'effet de X sur Y

La plupart des recherches qui adoptent la méthode expérimentale s'intéressent donc à des mécanismes causaux potentiels qui pourraient expliquer/prédire l'effet de l'intervention X sur un phénomène Y .

$$X \rightarrow Y$$

D'un point de vue logique, on essaie d'établir une *fonction causale* selon laquelle le phénomène Y écoute l'intervention X pour déterminer son état. Nous avons donc une sorte de *fonction* similaire au principe de Input-Output en programmation :

$$Y = f(X)$$

La fonction causale $f(X)$ correspond en général à un modèle mathématique/statistique (Maxwell et al., 2017 ; Pearl, 2000 ; Pearl et al., 2016 ; Rodgers, 2010), sujet qui sera traité dans les fondements statistiques de la méthode expérimentale. Pour l'instant, on peut retenir en termes logiques que l'état du phénomène Y (e.g. le résultat à un examen) est une *fonction* de l'intervention X (e.g. est-ce que l'étudiant-e a résumé ses notes sous forme de carte conceptuelle ou de résumé textuel?). En d'autres termes, on peut s'attendre à un *pattern* dans la *distribution* du phénomène Y qui dépend/est fonction de l'intervention X . Par exemple : on s'attend à un effet de la modalité avec laquelle les notes d'un cours sont résumées (carte conceptuelle vs. résumé textuel) sur les notes obtenus à l'examen. Le *pattern* attendu consiste dans une distribution des notes obtenues par les étudiant-es ayant utilisé les cartes conceptuelles autour d'une moyenne plus élevée par rapport à la distribution des notes obtenus par les étudiant-es ayant utilisé le résumé textuel.

D'un point de vue logique, la fonction causale est simple à déclarer en termes formels. En revanche, dans la pratique, déterminer un effet causal dans un environnement complexe est très difficile.

4.2 Causalité en technologie éducative

Le contexte de la technologie éducative est particulièrement sensible à cette complexité, car il regroupe plusieurs types d'interaction entre les technologies éducatives, les processus d'apprentissage, et des facteurs qui peuvent influencer les unes, les autres ou les deux en même temps. On peut néanmoins essayer de synthétiser la recherche en technologie éducative d'un point de vue causal comme l'investigation des effets des technologies éducatives sur

les processus d'apprentissage, avec des facteurs qui peuvent intervenir dans cette relation causale, comme illustré dans l'image ci-dessous.

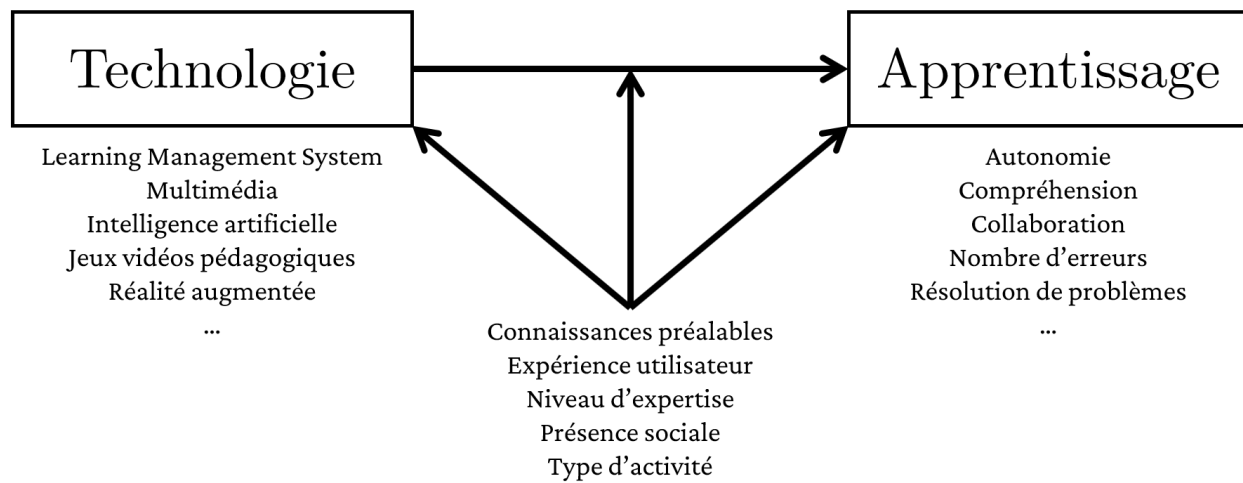


FIG. 2 : Relation causale entre technologies éducatives et processus d'apprentissage, avec des facteurs qui peuvent intervenir dans la relation.

Du côté des technologies éducatives on peut trouver les *Learning Management Systems*, les principes multimédia, les jeux vidéos pédagogiques, etc. L'apprentissage peut être à son tour divisé dans des processus tels que la compréhension, la capacité à collaborer, la résolution de problèmes, etc. Parmi les facteurs qui peuvent intervenir dans cette relation figurent les connaissances préalables, l'expérience utilisateur dans l'appropriation et utilisation des technologies, le niveau d'expertise, et ainsi de suite.

La recherche expérimentale en technologie éducative consiste donc principalement à formuler des hypothèses sur les relations causales entre les technologies éducatives et les processus d'apprentissage, ainsi que sur les effets de facteurs qui peuvent faciliter, entraver ou modérer ces effets.

4.3 Difficulté à établir des relations causales

À l'intérieur d'un système complexe, dans lequel les relations entre concepts peuvent s'influencer mutuellement, réussir à détecter un mécanisme causal qui détermine la direction d'un effet (i.e. c'est l'intervention X qui détermine le phénomène Y) s'avère très difficile (Mackie, 1965 ; Pearl, 2000 ; Pearl et al., 2016 ; Pearl & Mackenzie, 2018). Il existe différentes mécanismes qui peuvent rendre la détection d'un mécanisme de cause à effet compliquée. Cette liste, adaptée des travaux de Pearl (Pearl, 2000 ; Pearl et al., 2016 ; Pearl & Mackenzie, 2018) – voir aussi McElreath (2020) – en propose 4 : un mécanisme de relation inverse,

la présence d'une cause commune, la présence d'un médiateur, et la présence d'un effet de *collision*.

4.3.1 Un mécanisme de relation inverse

Ce phénomène s'avère si on attribut un effet de X sur Y, lorsqu'en vérité c'est Y qui a un effet sur X.

$$X \leftarrow Y$$

On peut imaginer que l'utilisation de technologies de conception et fabrication assistées par ordinateur (CFAO) – imprimantes 3D, broderie numérique, ... – ont un effet sur la créativité des étudiant-es, mais il peut également être le cas que des étudiant-es avec à la base une créativité plus prononcée s'intéressent davantage à ce type de technologies.

4.3.2 La présence d'une cause commune

Ce phénomène s'avère si on attribut un effet de X sur Y, lorsqu'en vérité X et Y n'est sont pas en relation causale, mais l'illusion d'un effet émerge à cause de la présence d'un troisième facteur qui influence à la fois X et Y.

$$X \leftarrow Z \rightarrow Y$$

Par exemple, le fait qu'un-e étudiant-e ait ou pas des stickers sur son ordinateur portable (X) n'a en soi aucun effet sur ses capacités en programmation (Y). Cependant, les programmeurs, qui ont donc des connaissances en programmation, ont souvent la tendance à mettre sur leurs propres ordinateurs des stickers des technologies ou langages de programmation adoptés (Z).

$$\textit{Stickersordinateur} \leftarrow \textit{Personneprogramme} \rightarrow \textit{Compétencesprogrammation}$$

Le fait d'avoir des stickers et d'avoir des compétences en programmation dérivent de la même cause : être programmeur. En revanche, si on ne considère pas cette cause commune et on se limite à observer que les étudiant-es avec les stickers ont des bonnes compétences en programmation, on pourrait faire l'hypothèse que les stickers ont un effet sur les connaissances en programmation (e.g. motivation, rappel, ...).

4.3.3 La présence d'un médiateur

Ce phénomène s'avère quand X et Y sont vraiment dans une relation causale, mais le mécanisme est transmis à travers un troisième facteur Z. À ce moment, le risque est que l'effet causale entre X et Y peut disparaître si le facteur Z n'est pas représenté correctement.

$$X \rightarrow Z \rightarrow Y$$

Imaginons que les responsables d'une formation en technologie éducative se questionnent sur l'opportunité d'assigner aux étudiant-es des livrables intensifs d'un point de vue computationnel comme par exemple le montage vidéo, le rendering 3D, ou la modélisation de phénomène complexe avec des simulations Markov Chain Monte-Carlo (X). Ces livrables risquent d'augmenter le temps de latence que les étudiant-es passe en attendant la fin des processus computationnels, ce qui les obligent à diminuer leur temps libre (Y). Cette relation est néanmoins médiée par la puissance des ordinateurs des étudiant-es (Z).

$$\text{Livable computationnel} \rightarrow \text{Puissance ordinateur} \rightarrow \text{Temps libre}$$

Or, si les responsables de la formation prenaient en considération dans un test/sondage/entretien seulement des étudiant-es avec des ordinateurs très performants, cette médiation pourrait donner l'illusion qu'il n'y a pas de mécanisme de cause à effet entre les livrables intensifs d'un point de vue computationnel et le temps libre des étudiant-es.

4.3.4 La présence d'un effet de *collision*

Ce phénomène s'avère lorsque X et Y ne sont pas en relation causale, mais les deux influencent le même facteur. Il peut s'avérer que sous certaines conditions spécifiques de Z, X et Y résultent associés.

$$X \rightarrow Z \leftarrow Y$$

Imaginons que pour être accepté-es dans un Master (Z) il faut soit avoir suivi un bachelor (X) ou avoir une idée de projet pertinente avec le type de formation dispensée (Y). Les deux variables ne sont pas liées : on peut avoir une bonne idée avec ou sans bachelor, et on peut avoir un bachelor avec ou sans une bonne idée de projet.

$$\text{Bachelor} \rightarrow \text{Admission Master} \leftarrow \text{Projet intr}$$

Imaginons que les responsables du Master émettent l'hypothèse que l'obtention du bachelor détermine la qualité du projet proposé. Pour cela ils ne prennent en considération que les étudiant-es qui ont été admis-es. À ce moment, les responsables vont croire qu'il y a en réalité un effet. Lorsque deux phénomènes sont indépendants, la probabilité que l'un **et** l'autre se manifestent est plus faible par rapport à la manifestation de l'un **ou** l'autre individuellement. Si l'étudiant-e a un Bachelor, il/elle est admis-e même sans un projet intéressant. Il y a donc plus de chances que les étudiant-es avec bachelor aient proposé un projet pas très intéressant. La même chose s'applique par reflet aux étudiant-es admis-es avec un projet intéressant : ils ont moins de chances d'avoir un bachelor. En combinant les deux choses, il semble émerger un effet négatif du bachelor sur la qualité du projet, mais qui est dû à la sélection seulement des étudiant-es admis-es.

4.4 L'échelle de causalité

Dans ces travaux à propos de la causalité, Pearl (Pearl, 2000 ; Pearl et al., 2016 ; Pearl & Mackenzie, 2018) propose d'imaginer une échelle avec trois marches. Chaque marche qui est montée permet de s'approcher davantage à l'explication d'un mécanisme causale.

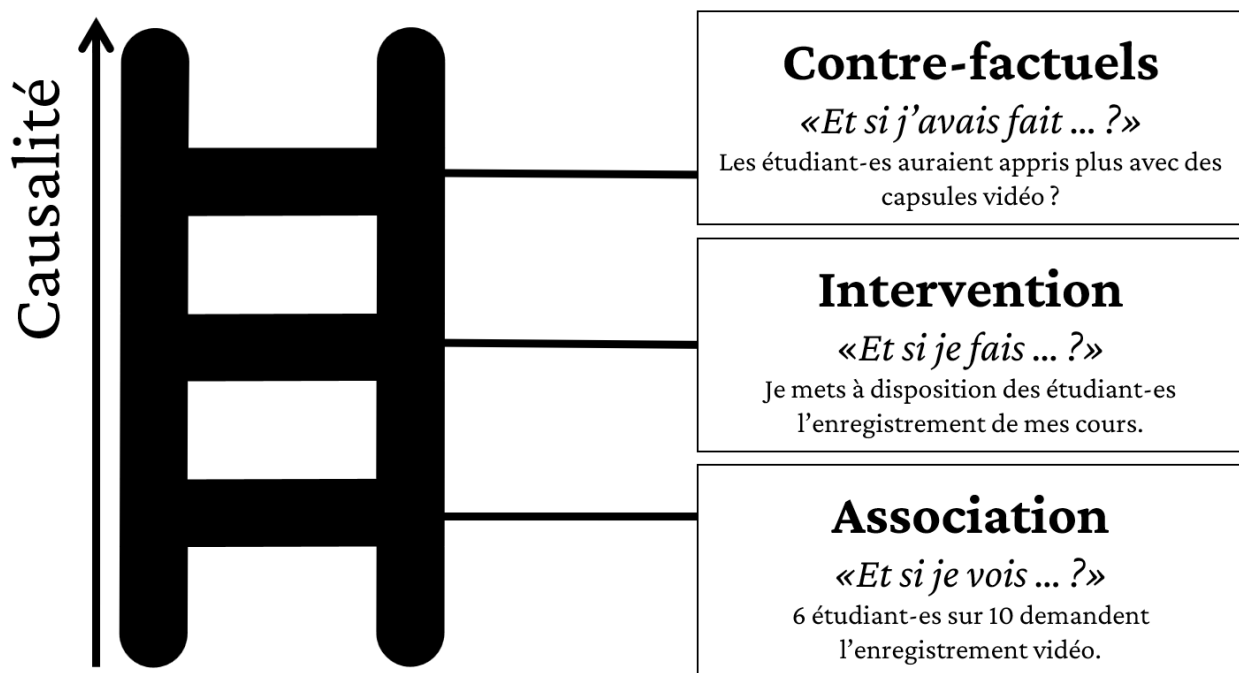


FIG. 3 : L'échelle de causalité avec trois marches. Adapté de (Pearl & Mackenzie, 2018)

4.4.1 Association

La première marche en bas de l'échelle consiste dans l'association de deux (ou plusieurs) phénomènes qui ont la tendance à se manifester selon des patterns récurrents. Comme indiqué plus haut dans la section consacrée aux difficultés dans l'établissement d'un mécanisme causale, une *simple* association ne permet cependant pas ni de déterminer si un phénomène est la cause de l'autre ou vice-versa, ni d'exclure d'autres facteurs potentiels qui pourraient expliquer l'association (e.g. cause commune, médiateur ou effet de *collision*).

Une association peut être déterminée de manière systématique par exemple à travers un sondage, comme dans l'exemple de l'image : 6 étudiant-es sur 10 demandent l'enregistrement vidéo des cours. On peut inférer depuis cette information que les étudiant-es associent l'enregistrement vidéo du cours à quelques formes de bénéfices pour leur apprentissage.

Une autre manière de déterminer systématiquement une association consiste dans les études dites *corrélationnelles*, dans lesquelles le terme *association* est remplacé justement par le terme *corrélation* qui présente une connotation statistique précise. On peut en effet calculer la corrélation entre deux phénomènes mesurés quantitativement sur une échelle dans laquelle :

- Un score s'approchant de -1 témoigne d'une corrélation *négative*, caractérisée par le fait que A et B ont la tendance à varier de manière opposée : une valeur élevée de A est associée à une faible valeur de B (ou vice-versa) ; ou une valeur faible de A est associée à une valeur élevée de B (ou vice-versa).
- Un score s'approchant de 0 témoigne l'absence d'une corrélation, c'est-à-dire que les deux phénomènes ont la tendance à varier sans un pattern identifiable (e.g. si l'un est élevé, l'autre peut être parfois élevé aussi, parfois faible, et parfois moyen).
- Un score s'approchant de 1 témoigne d'une corrélation *positive*, caractérisée par le fait que A et B ont la tendance à varier de manière spéculaire : une valeur élevée de A est associée à une valeur élevée de B (ou vice-versa) et une faible valeur de A est associée à une faible valeur de B (ou vice-versa).

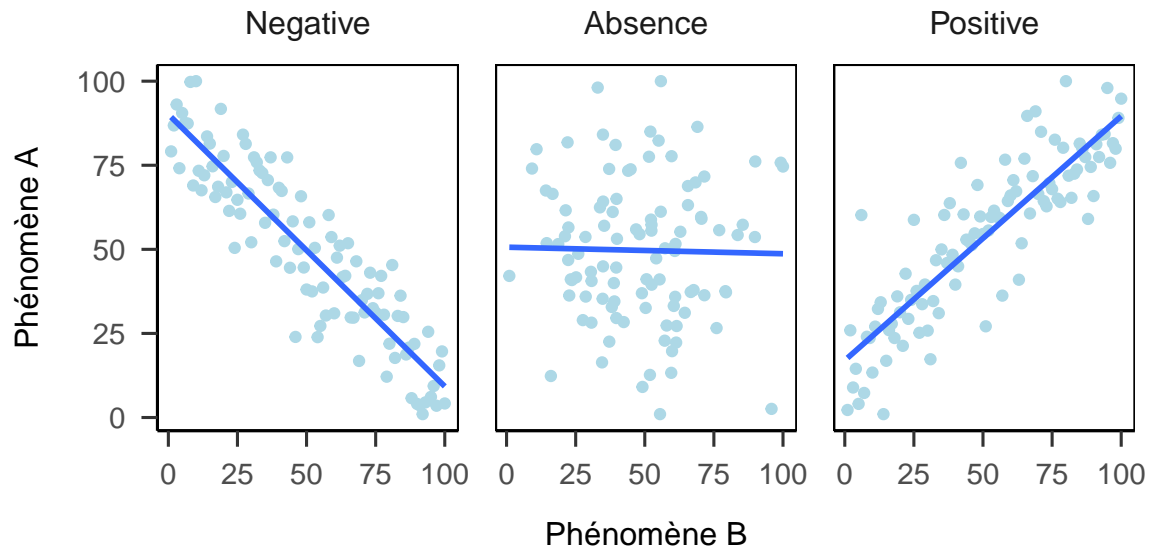


FIG. 4 : Exemple de corrélations

4.4.2 Intervention

4.4.3 Contre-Factuels

5 Processus génératif des données

6 Inférence

Références

- Mackie, J. L. (1965). Causes and Conditions. *American Philosophical Quarterly*, 2(4), 245-264.
- Maxwell, S. E., Delaney, H. D., & Kelley, K. (2017). *Designing experiments and analyzing data : a model comparison perspective* (Third edition). Routledge.
- McElreath, R. (2020). *Statistical rethinking : A Bayesian course with examples in R and Stan* (2 éd.). Taylor ; Francis, CRC Press.
- Meehl, P. E. (1990). Appraising and Amending Theories : The Strategy of Lakatosian Defense and Two Principles That Warrant It. *Psychological Inquiry*, 1(2), 108-141.
- Pearl, J. (2000). *Causality : models, reasoning, and inference*. Cambridge University Press.
- Pearl, J., Glymour, M., & Jewell, N. P. (2016). *Causal inference in statistics : a primer*. Wiley.
- Pearl, J., & Mackenzie, D. (2018). *The book of why : the new science of cause and effect*.

Basic Books.

- Rodgers, J. L. (2010). The epistemology of mathematical and statistical modeling : A quiet methodological revolution. *American Psychologist*, 65(1), 112. <https://doi.org/10.1037/a0018326>
- Rohrer, J. M. (2018). Thinking Clearly About Correlations and Causation : Graphical Causal Models for Observational Data. *Advances in Methods and Practices in Psychological Science*, 1(1), 27-42. <https://doi.org/10.1177/2515245917745629>
- Scheel, A. M., Tiokhin, L., Isager, P. M., & Lakens, D. (2020). *Why Hypothesis Testers Should Spend Less Time Testing Hypotheses*. 1-12.