

Zeitschrift: Horizons : le magazine suisse de la recherche scientifique
Herausgeber: Fonds National Suisse de la Recherche Scientifique
Band: 33 (2021)
Heft: 131: Publier à tout prix!

Artikel: Mêmes informations conclusions différentes
Autor: Cartlidge, Edwin
DOI: <https://doi.org/10.5169/seals-1089013>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. [Mehr erfahren](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. [En savoir plus](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. [Find out more](#)

Download PDF: 11.02.2026

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

Mêmes informations, conclusions différentes

Le statut, le sexe, la couleur de la peau influencent-ils le temps de parole, la participation, les sanctions? Les études sur ce sujet sont parfois contradictoires, même lorsqu'elles ont analysé exactement les mêmes données. Pourquoi c'est ainsi et comment on pourrait y remédier.

Texte Edwin Cartlidge

Soumettez le même ensemble de données à plusieurs groupes de scientifiques et demandez-leur d'étudier un certain nombre d'hypothèses. Si ces chercheurs sont scrupuleux et vigilants, tous les groupes sont censés parvenir à des conclusions très similaires. C'est du moins ce qu'on attend d'une méthode scientifique quand elle fonctionne correctement.

Pourtant ce n'est pas le cas, selon une étude réalisée dans le cadre d'une large collaboration entre spécialistes en sciences sociales, en informatique et statisticiens publiée en juin 2021 dans la revue *Organizational Behavior and Human Decision Processes*. Des analystes indépendants ont reçu près de 4 millions de mots tirés d'un forum universitaire en ligne avec pour mandat d'évaluer deux hypothèses apparemment simples sur l'influence du genre et du statut universitaire sur les contributions aux discussions du forum. Les résultats ont été saisissants, avec de très fortes différences tant dans l'approche des analystes pour évaluer les données que dans leurs conclusions – parfois diamétralement opposées.

Ce résultat n'est pas le fruit du hasard. Ces dernières années, de nombreuses études ont montré qu'il est souvent impossible de reproduire les résultats d'une recherche dans des domaines allant de la sociologie à la médecine. Mais la discussion qui s'est ensuivie a surtout porté sur les pièges que représente la recherche de résultats statistiquement significatifs dans des données non homogènes et sur les biais dans la publication des résultats. La nouvelle étude révèle que les variations dans les méthodes de recherche représentent aussi un important problème.

Pour Abraham Bernstein, un informaticien de l'Université de Zurich, membre de la collaboration, cela signifie que les scientifiques doivent non seulement publier les données qui sous-tendent leurs recherches, mais aussi les étapes analytiques précises qu'ils ont suivies. «La marge d'interprétation d'une expression telle que 'nous avons réalisé x tests' est bien trop grande, dit-il. Tous les choix doivent être exposés très clairement.»

Pourquoi pas la règle du hors-jeu au lieu des cartons rouges?

La nouvelle étude est un exemple de la méthode de production participative – le recrutement, en général en ligne, de nombreux groupes de recherche indépendants pour analyser le même jeu de données. Un premier exemple de cette méthode, publié en 2018, avait demandé à des analystes d'étudier des données provenant des ligues de football pour déterminer s'il existait une corrélation entre la couleur de peau des joueurs et le nombre de cartons rouges reçus. Elle a débouché sur un large éventail de conclusions différentes – la plupart des analystes ayant trouvé un effet statistiquement significatif mais peu marqué, alors que d'autres n'ont constaté aucune corrélation.

Ce projet se concentrait toutefois sur une corrélation spécifique – impliquant les cartons rouges, plutôt que les hors-jeu par exemple. La dernière tâche, en revanche, laissait aux analystes la liberté de définir eux-mêmes les variables pertinentes. Elles devaient évaluer deux hypothèses concernant des commentaires postés pendant presque vingt ans sur le site *edge.org* par plus de 700 contributeurs, dont 128 femmes. La première hypothèse avançait que «la propension d'une femme à participer à une conversation est liée positivement au nombre de femmes dans la discussion». Et la seconde que «les participants de statut supérieur sont plus loquaces que ceux de statut inférieur».

L'étude a été menée par une collaboration internationale et coordonnée par Martin Schweinsberg, un psychologue de l'European School of Management and Technology à Berlin. Elle a pris en considération 19 analystes, issus d'un pool de 49, qui ont utilisé DataExplained – un site Internet conçu spécialement pour consigner et expliquer les étapes de leurs analyses – tant celles qu'ils ont finalement utilisées que celles qu'ils ont rejetées.

Les résultats de l'exercice montrent simplement à quel point les approches analytiques peuvent varier. En termes de méthode, les analystes ont employé un large éventail de techniques statistiques et un éventail encore plus large de variables. Pour saisir par exemple le «statut» individuel, ils ont, entre autres paramètres, pris en considération le titre du poste universitaire, la possession d'un doctorat, le nombre de citations et un chiffre – l'indice *h* – reflétant la quantité d'articles publiés qui sont cités souvent.

Avec préenregistrement contre les préjugés

La palette des résultats reflète la diversité des méthodes. Près de deux tiers des analystes ont conclu que les femmes tendaient à participer davantage en présence d'autres femmes, mais plus d'un cinquième est arrivé à la conclusion inverse. L'écart était encore plus grand pour la question du statut: 27% étaient en faveur de l'hypothèse et 20% y étaient opposés, tandis que les autres analystes arrivaient à des résultats statistiquement non significatifs.

Pour certains observateurs, ces résultats sont à interpréter avec circonspection. Leonhard Held de l'Université de Zurich remarque que, aussi libres qu'ils aient été dans le choix des variables et des méthodes statistiques, ce ne sont pas les analystes qui ont décidé d'étudier ce forum-là. Cette exclusion a peut-être exercé une influence sur les résultats puisque la taille de l'échantillon peut affecter la pertinence statistique. Il se demande aussi si ces tentatives de généraliser des dynamiques de groupe sont vraiment réalistes compte tenu du recours à ce seul forum.

Néanmoins, Leonard Held salue l'étude qui «illustre clairement les problèmes inhérents aux degrés de liberté des chercheurs dans les analyses exploratoires». Anna Dreber Almenberg, de l'Ecole d'écono-



Denis Zakaria est-il ici victime de discrimination en raison de sa couleur de peau? Il existe de nombreuses données sur le football, mais savoir lesquelles exactement doivent être analysées n'est pas clair. Au fait: aviez-vous des préjugés? Le carton rouge du match Suisse-Espagne était destiné à Remo Freuler, qui ne figure pas sur cette photo. Photo: Anton Vaganov/Pool/AFP/Keystone

mie de Stockholm, est aussi enthousiaste, qualifiant la recherche de «super importante» pour tenter d'améliorer encore la reproductibilité. Elle estime qu'elle met en évidence une limite du préenregistrement – l'obligation pour les scientifiques de définir leurs méthodes et leurs tests statistiques avant de collecter et d'analyser leurs données. Elle relève que si le préenregistrement peut améliorer la fiabilité des résultats, il ne peut pas dicter quelle analyse spécifique il faut utiliser.

Et, en effet, Martin Schweinsberg et ses collègues estiment que les différences dans les analyses représentent un «défi plus fondamental pour le savoir» que le p-hacking ou le fait de jeter un œil sur les données avant de procéder au test d'évaluation. Selon eux, ces problèmes peuvent être réglés soit par le préenregistrement, soit par l'analyse en aveugle – celle-ci, relèvent-ils, empêchant le chercheur de choisir, consciemment ou non, une analyse qui produira le résultat souhaité.

Toutefois, remarquent-ils, les connaissances, les croyances et les interprétations des analystes présentent des variations naturelles qui conduiront à des résultats différents, même s'ils agissent de manière transparente et en toute bonne foi. «Résultant souvent de présupposés théoriques, ces choix subjectifs et leurs conséquences pourraient bien être une composante inextricable du processus scientifique», écrivent-ils.

Présenter l'ensemble des choix explicites

Le projet «Many Paths» veut rendre ces variations explicites. Développé par un groupe de cinq universitaires d'Allemagne, des Pays-Bas et de Suisse, il entend mettre à nu «la confusion au milieu» du processus de recherche en amenant des praticiens de diverses disciplines à collaborer sur des projets communs. Pour commencer, il a demandé à des

politologues, des philosophes, des psychologues et d'autres spécialistes de discuter de la question séculaire: «Le pouvoir corrompt-il?» Jusqu'à présent, il a recruté une quarantaine d'expertes.

Many Paths utilise un instrument nommé Hypergraph pour documenter les recherches selon une procédure modulaire par étapes. Membre du projet, Martin Götz, de l'Université de Zurich, explique que l'idée est de «disséquer les publications scientifiques classiques» pour que chacun puisse travailler sur les éléments qui l'intéressent le plus, par exemple la théorie, la collecte de données ou encore la méta-analyse. Le but, explique-t-il, est de «remplacer le système actuel de publication» et ce qu'il considère comme des incitations à faire les grands titres plutôt que de la recherche solide.

Bien que de nombreux experts soient d'accord sur la nécessité d'améliorer la reproductibilité, certains mettent en garde contre un excès de pessimisme. Chercheur en médecine et épidémiologiste de l'Université Stanford, John Ioannidis relève que, même s'il existe des recherches où tous les résultats sont possibles, dans la plupart des cas, des analyses multiples montreront que certains d'entre eux sont plus plausibles que d'autres. «Nous ne devons pas tomber dans le nihilisme», dit-il.

Abraham Bernstein reste lui aussi optimiste, soutenant que le processus de recherche restera solide tant que les scientifiques seront clairs sur les choix qu'ils font au cours de leurs travaux. L'important, dit-il, «est que ces choix puissent être explicités et donc soumis à un examen minutieux».

Edwin Cartlidge est journaliste scientifique à Rome.