

Quand les chercheurs jugent « la fiabilité réelle des études scientifiques »

il est encore trop facile de publier des études annonçant des effets expérimentaux qui, en réalité, n'existent pas ou sont plus faibles qu'on ne l'a prétendu, explique dans sa chronique l'économiste Paul Seabright.

LE MONDE | 08.09.2018 à 10h30 | Par Paul Seabright (Professeur à l'Institut d'études avancées de Toulouse)



« On peut donc conclure que la communauté des chercheurs dispose d'indices qui lui permettent de juger de la fiabilité réelle des études, même lorsqu'elles ont été acceptées et publiées dans les revues les plus prestigieuses. » F. Cirou & S. Leong Ho/PhotoAlto / Photononstop

Recherches. Quel crédit peut-on accorder aux études scientifiques publiées dans les revues les plus prestigieuses ? Même si leurs procédures de sélection sont les meilleures, même si le filtrage qu'elles ont mis en place est très efficace, il est toujours possible que des études soutiennent des hypothèses fausses et que personne ne les remette en question avant publication.

Depuis quelque temps, la communauté scientifique s'intéresse aux tentatives de reproduire des études déjà publiées. Le Reproducibility Project : Psychology a essayé de dupliquer 97 études en psychologie et a constaté, en 2015, que seules 36 % d'entre elles pouvaient être confirmées. L'Experimental Economics Replication Project a fait de même, en 2016, pour 18 études en économie expérimentale, avec un taux de confirmation de 61 % – plus élevé, mais loin d'être écrasant.

Une équipe de chercheurs vient de publier une duplication de 21 études phares en sciences sociales expérimentales, publiées entre 2010 et 2015 dans *Nature* et *Science*, les deux revues scientifiques les plus prestigieuses (« *Evaluating the replicability of social science experiments in "Nature" and "Science" between 2010 and 2015* » (<https://doi.org/10.1038/s41562-018-0399-z>), par Colin F. Camerer et al., « *Nature Human Behaviour Letters* », 27 août 2018). L'une des particularités de ce projet est que la taille des échantillons utilisés pour la duplication est environ cinq fois plus grande que celle des échantillons des études d'origine, ce qui augmente sensiblement la fiabilité de la reproductibilité.

Des progrès nécessaires

Treize d'entre elles confirment, de façon statistiquement significative, les effets constatés dans l'expérience originale, soit un taux de 62 %. Mais l'ampleur de l'effet confirmé pour ces 13 études est inférieure de moitié à l'ampleur constatée dans la première étude !

Certes, il peut y avoir des faiblesses dans la mise en œuvre des tentatives de duplication. Mais, globalement, la conclusion semble être que la réalisation et la publication d'études scientifiques, au moins en sciences sociales, a besoin de faire de grands progrès : il est encore trop facile de publier des études annonçant des effets expérimentaux qui, en réalité, n'existent pas ou sont plus faibles qu'on ne l'a prétendu...

Lire aussi : L'innovation, meilleure amie de la publication scientifique ?

(/idees/article/2016/04/28/l-innovation-meilleure-amie-de-la-publication-scientifique_4910330_3232.html)

Mais les auteurs ont complété leur démarche en recrutant 397 chercheurs pour mener une expérience destinée à mesurer le degré de confiance que la communauté scientifique accordait aux études originales. Avant que les résultats de la duplication soient connus, ces cobayes ont joué au jeu du *prediction market* (« marché de prédiction »), où ils pouvaient acheter ou vendre des « actions » pour chacune des études originales.

Ces titres ne donnaient une récompense que si et seulement si l'étude était validée. Le prix de chacune d'elles sur ce marché de prédiction donnait ainsi une mesure du degré de confiance implicite des chercheurs à l'égard des études originales ; mais ils ont également répondu à une enquête pour indiquer leur degré de confiance explicite.

Transparence nécessaire

En l'occurrence, la confiance moyenne accordée aux études, mesurée par le marché, fut de 63 %. Celle mesurée par l'enquête fut de 61 %. Soit deux chiffres très proches du véritable taux de confirmation, de 62 %. De plus – détail important –, il y avait une forte corrélation entre le degré de confiance mesuré par chacune des deux méthodes et l'ampleur de l'effet constaté par la duplication.

Lire aussi : Erreur ou fraude : la sanction des chercheurs (/idees/article/2015/05/21/erreur-ou-fraude-la-sanction-des-chercheurs_4638183_3232.html)

On peut donc conclure que la communauté des chercheurs dispose d'indices qui lui permettent de juger de la fiabilité réelle des études, même lorsqu'elles ont été acceptées et publiées dans les revues les plus prestigieuses. Cette information n'est rendue visible que grâce à l'agrégation réalisée par le jeu du marché de prédiction, mais elle ne l'est pas pour le lecteur profane d'une revue scientifique !

La publication dans une revue, si sérieuse soit-elle, n'est donc pas équivalente à une ouverture totale de la méthode scientifique au public non averti. La science peut faire mieux pour placer la transparence au cœur de ses activités.