

Non, la recherche ne dit pas aux praticiens ce qu'ils doivent faire.

Benoît Galand

Par dans la revue diversité n°192, mai-août 2018, pp. 107-112.

Partant du point de vue d'un praticien de la recherche, ce texte tente de faciliter le dialogue avec d'autres catégories d'acteurs éducatifs en explicitant quelques sources potentielles de malentendus autour des résultats scientifiques. Il vise également à montrer en quoi les caractéristiques mêmes de l'activité scientifique s'opposent à un usage de ses résultats à des fins de prescription.

Dans ma pratique de chercheur en éducation, je suis – comme tant d'autres – interpellé par le gouffre qui existe entre les connaissances scientifiques accumulées depuis des décennies et le fonctionnement actuel de nos systèmes éducatifs, confrontés à de nombreux défis. Sans prétendre traiter ces questions de manière exhaustive, ce texte tente d'explicitier ce qui constitue, d'après moi, quelques sources courantes de malentendus entre les chercheurs et d'autres acteurs éducatifs. Il cherche également à montrer en quoi les caractéristiques mêmes de l'activité scientifique vont à l'encontre d'une utilisation de ses résultats pour prescrire le travail des professionnels de l'éducation. Mon point de vue, forcément situé, est celui d'un universitaire, formé à la psychologie, réalisant des recherches en éducation, et ayant un rôle d'enseignant et de formateur. Mon espoir est qu'une meilleure compréhension du métier de chercheur et de ses codes pourra faciliter le dialogue entre acteurs éducatifs.

Qu'est-ce que la science ?

Une « révolution de l'ignorance » : telle est la manière dont l'historien Yuval Noah Harari (2015) qualifie l'importance grandissante accordée à l'approche scientifique dans nos sociétés, depuis la Renaissance. Selon lui, cette approche rompt avec l'idée selon laquelle traditions, religions et idéologies offrent toutes les réponses nécessaires et traitent de toutes les questions dignes d'intérêt. Le philosophe des sciences Karl Popper met quant à lui en avant la démarche de falsification comme un critère fondamental de scientificité. Selon lui, le caractère scientifique d'une recherche tient à la démarche intellectuelle consistant à soumettre ses idées à des tentatives de falsification pour en éprouver la robustesse (afin d'améliorer ses hypothèses, puis de les resoumettre à l'épreuve des faits), plutôt qu'à l'accumulation d'arguments en faveur de ses idées ou à l'utilisation de telle ou telle méthode de recherche (Gorman, 2005). Le recours à une méthode particulière ne garantit donc pas le caractère scientifique d'un travail de recherche, et de nombreux chercheurs en éducation défendent la nécessité d'une diversité d'approches méthodologiques (Cohen *et al.*, 2011). Ils soulignent cependant que certaines méthodes sont plus adaptées à certaines questions, et qu'il est important de ne pas prétendre répondre à des questions auxquelles la méthode choisie ne permet pas de répondre de manière adéquate.

Éviter les malentendus

Au-delà des débats méthodologiques, on peut identifier certaines pratiques du métier de chercheur qui peuvent être sources d'incompréhension vis-à-vis d'autres acteurs éducatifs.

Un mode de raisonnement inhabituel

S'appuyant sur la confrontation d'hypothèses, nombre de résultats de recherche peuvent paraître contre-intuitifs par rapport à l'expérience des acteurs éducatifs (notamment parce que ces résultats se basent sur des informations qui ne sont pas directement accessibles dans le quotidien de ces derniers) (Dubet, 2002). Par exemple, pour différentes raisons (rémission spontanée, régression vers la moyenne, « maturation »...), les situations très difficiles ou très problématiques ont souvent tendance à s'améliorer avec le temps. Si l'on réévalue la situation après un certain temps, il y a donc de fortes chances que l'on constate un progrès, « quoi que l'on ait fait » (Lilienfeld *et al.*, 2014). De même, la plupart des élèves progressent dans leurs apprentissages, peu importe le type d'enseignement qui leur est dispensé. Ainsi, mesurer le niveau d'acquis des élèves – par exemple avant et après la mise en place d'un dispositif didactique – et utiliser les progrès observés comme arguments en faveur de l'efficacité du dispositif est un raisonnement erroné d'un point de vue scientifique. Pour soutenir l'efficacité du dispositif, il faut entre autres montrer que les élèves qui l'ont reçu progressent davantage que d'autres, qui n'en ont pas bénéficié (ce que les chercheurs appellent un groupe témoin ou contrôle). Le même raisonnement peut s'appliquer au redoublement. Une enseignante peut observer qu'un élève qui recommence son année s'en sort mieux que l'année précédente, et peut-être un peu mieux que certains de ses camarades de classe. Est-ce un argument en faveur du redoublement ? D'une part, l'élève a déjà parcouru la matière une fois, il n'est par conséquent pas très surprenant qu'il la maîtrise davantage (Crahay, 2007). D'autre part, et surtout, un élève de même niveau que le précédent, que l'on a promu plutôt que de faire redoubler, a de forte chance de progresser davantage que son camarade qui redouble (et risque d'ailleurs moins de redoubler ou de décrocher par la suite). Un mode de questionnement typique de la démarche scientifique, en lien avec le critère de falsification évoqué ci-dessus, est ainsi de se demander : « Que se passe-t-il si l'on fait autre chose, si la situation est différente ? » Apporter des réponses rigoureuses à ce genre de questions n'a rien d'intuitif, d'autant plus quand on est engagé dans une prise de décision ou dans le cours d'une action (Baillargeon, 2006). C'est pourquoi les scientifiques recourent à des méthodes de recueil d'information parfois complexes, notamment pour se prémunir autant que possible de notre tendance à porter attention aux événements qui confirment nos croyances initiales (biais de confirmation d'hypothèse).

Isoler un effet ou lire un réel complexe

Une bonne partie de la démarche scientifique est analytique et cherche à découper la réalité et à isoler les effets de tel ou tel paramètre. L'expérience en laboratoire ou l'essai aléatoire contrôlé représente peut-être la forme ultime de ce type de démarche. Ces plans de recherche sont les seuls qui assurent un contrôle élevé des sources de variance et une bonne validité interne à la démarche, permettant notamment d'identifier les relations de causalité. Indispensable à l'avancée des connaissances scientifiques, cette démarche n'épuise pas le réel et peut paraître détachée des situations scolaires complexes et multiformes dans lesquelles les acteurs sociaux sont immergés, où une multiplicité de facteurs se combinent de manière plus ou moins simultanée – et peuvent donc annuler ou transformer l'effet identifié de manière isolée (Mitchell, 2012).

Effet moyen et cas particulier

En lien avec le point précédent, un grand nombre de travaux de recherche s'efforcent d'identifier des effets moyens « toute chose étant égale par ailleurs ». Cet effort est notamment orienté par le souhait de dégager des résultats qui dépassent les cas individuels étudiés, et parlent de ce fait à davantage de gens. Cependant, parents et professionnels, eux, sont face à des personnes particulières, singulières, « toute chose n'étant pas égale par ailleurs », et pour lesquelles il est souvent difficile de savoir dans quelle mesure les effets moyens s'appliquent. Un nombre croissant de recherches tentent par conséquent de mieux comprendre ce qui est efficace pour qui, dans quelles circonstances, et pourquoi (Galand, 2017-1).

Local et global

De même, si l'expérience des enseignants, des élèves ou des parents s'enracine dans une réalité souvent très locale, les études en éducation, elles, font fréquemment appel à des échantillons diversifiés ou les plus représentatifs possibles de la population visée. Cette différence de focale peut susciter des malentendus. Ainsi par exemple, l'évolution globale des violences scolaires parmi l'ensemble des élèves peut être différente de l'évolution constatée à l'échelle d'un établissement. De même, l'effet d'une innovation pédagogique peut avoir des effets divergents d'une classe à l'autre, par exemple selon le niveau initial ou l'origine sociale des élèves.

Pertinence théorique et pratique

Les concepts mobilisés par les chercheurs en éducation sont très souvent des construits latents, c'est-à-dire des notions théoriques non directement observables (motivation, processus cognitif, leadership, discrimination, émotion, activation de réseaux de neurones, etc.), pour lesquels des indicateurs vont devoir être élaborés. Impossibles à « toucher du doigt », ces concepts sont parfois difficiles à appréhender pour des non-initiés et peuvent être sources de confusions (y compris entre chercheurs). En outre, la pertinence concrète de ces concepts peut parfois être difficile à percevoir en dehors des cercles de spécialistes.

Du bon usage des probabilités

Ce point et les deux suivants concernent particulièrement les résultats quantitatifs. Les tests statistiques utilisés par les chercheurs portent sur des probabilités, pas sur des liens déterministes. Plus précisément, ces tests nous renseignent sur les chances/risques d'observer *par hasard* une association ou un écart différent de zéro dans les informations récoltées. Cette association ou cet écart sont jugés significatifs (c'est-à-dire réellement différents d'une absence d'association ou d'écart) quand les chances/risques sont considérés comme suffisamment faibles. Ainsi, on observe par exemple régulièrement que la présence de troubles d'apprentissage est associée à une fréquence accrue de redoublements. Comme on a une très faible probabilité d'observer cette association par hasard, elle est considérée comme significative. Attention, ce résultat ne signifie nullement que tout enfant présentant des troubles d'apprentissage va redoubler (nombre d'entre eux réussissent malgré leurs troubles), ni que l'on peut d'office expliquer le redoublement d'un enfant présentant des troubles d'apprentissage par ces mêmes troubles (cela peut être lié à de nombreux autres facteurs). La simple présence d'une association ou d'un écart, même significatif, ne permet pas d'établir une relation de causalité.

Significativité et taille d'effet

Le fait qu'un résultat soit ou non significatif ne nous renseigne pas sur la taille ou l'ampleur de cet effet. En d'autres mots, la significativité d'un effet ne nous dit pas si cet effet est fort ou faible, et s'il est important à prendre en compte. Un effet peut être très significatif, mais tellement faible qu'il n'a aucun intérêt pratique (certains chercheurs semblent d'ailleurs eux-mêmes, hélas, parfois confondre les deux). Il peut donc être intéressant de comparer les tailles de différents effets pour avoir une idée de leur poids relatif. C'est par exemple ce que propose John Hattie (2009), qui a mis en évidence que la plupart des facteurs étudiés en lien avec l'apprentissage ont un effet (au sens statistique) significatif sur celui-ci, et qui suggère donc de s'intéresser en particulier aux facteurs qui dépassent l'effet moyen. Mais avoir une idée du poids relatif d'un facteur ne permet pas encore de savoir à partir de quand on peut considérer qu'une taille d'effet est suffisamment importante pour avoir des répercussions concrètes dans la pratique, pour faire la différence dans la vie de quelqu'un.

Continuum et seuil

Cette question complexe de l'importance pratique n'est pas tranchée, mais prend un sens particulier quand on intègre le fait que nombre d'études recourent à des mesures continues alors que la plupart des décisions institutionnelles dans le cadre scolaire se réfèrent à des seuils (réussite ou échec,

exclusion ou non, orientation, admission, etc.). Une variable étudiée peut donc par exemple avoir un effet significatif et de taille modérée sur l'apprentissage, sans que l'on puisse dire à partir des résultats si cet effet sera suffisant pour permettre à un élève en difficulté d'atteindre le seuil de réussite demandé par les acteurs institutionnels.

Le métier de chercheur

De par sa nature, la recherche scientifique est une entreprise d'accumulation patiente plutôt que de percée fulgurante, qui tient plus du laborantin que du génie. Fondé sur une quête de vérité jamais atteinte, car toujours susceptible d'être remise en question, le travail du chercheur fait face à la tentation perpétuelle de prendre ses idées pour la réalité. D'où l'importance du regard critique des pairs, et des échanges au sein d'une communauté scientifique porteuse de débat sur la qualité de l'articulation entre propositions théoriques et résultats empiriques (Oreskes, Conway, 2012). Dépositaire d'un savoir spécialisé, intrinsèquement collectif (et qui tient pour une bonne part sa légitimité de sa reconnaissance par une communauté de pairs), le métier de chercheur s'accorde mal avec la figure de l'intellectuel médiatique et le vedettariat qui l'accompagne. Il s'accorde également mal avec les débats d'opinions où chaque point de vue doit être considéré de manière équilibrée. Si aucune idée n'est exclue a priori du débat scientifique, les arguments rhétoriques n'y ont que peu de valeur, et seule compte la qualité des résultats empiriques et de leur articulation à des explications théoriques cohérentes. Qui veut contester un consensus majoritaire doit venir avec des données solides étayant ses explications alternatives, pour avoir la moindre chance d'être écouté par ses pairs. La communauté scientifique n'accordera par ailleurs pas le même crédit à différents types de recherche. Un résultat répliqué dans différentes études réalisées par plusieurs équipes et dans divers contextes sera jugé plus convaincant qu'un résultat isolé ou porté par les études d'une même équipe (Makel, Plucker, 2014).

Des résultats aux préconisations ?

Les recherches qui nous aident à comprendre un phénomène nous permettent-elles de savoir comment agir ? Le lien n'est hélas pas direct, et un travail conséquent de recherche et développement est nécessaire pour passer d'une compréhension théorique valide d'un phénomène à des dispositifs ou des pratiques pouvant influencer de manière fiable sur celui-ci (Cèbe, Goigoux, 2009). Former des enseignants aux connaissances scientifiques actuelles concernant, par exemple, les processus de compréhension en lecture ou la dynamique du harcèlement à l'école risque donc de ne pas suffire pour aider ces enseignants à développer d'autres pratiques de classe, en partie pour les raisons de malentendus potentiels évoquées plus haut. Même s'ils le souhaitent, le temps et les ressources pour imaginer comment traduire ces connaissances théoriques en pratiques pertinentes pour leur contexte de travail a de forte chance de leur manquer, sans tenir compte des autres facteurs qui influencent leurs pratiques (Dupriez, 2015).

Un nombre croissant de travaux de recherche soulignent le rôle crucial que jouent les conditions de mise en œuvre ou d'implantation dans l'efficacité d'un dispositif (Durlak, DuPre, 2008). L'idée la plus brillante, déployée dans la précipitation, sans préparation, sans accompagnement, a très peu de chance de produire le moindre bénéfice observable. Ces travaux de recherche indiquent également que même les interventions solidement étayées d'un point de vue théorique et bénéficiant de conditions de mise en place optimales ont un effet moyen assez faible (Galand, 2017-2). La variabilité et l'adaptation sont la règle plutôt que l'exception dans le déploiement de « bonnes pratiques » à grande échelle, parfois au bénéfice et parfois au détriment de l'intervention originale.

Face à cet état des choses, poussant jusqu'au bout une vision « applicationniste », la Société pour la recherche en prévention a publié une liste de critères de qualité auxquels devraient répondre les études pouvant prétendre à l'efficacité et à la diffusion à large échelle de leurs résultats (Gottfredson *et al.*, 2015). Dans les faits, peu de recherches peuvent se targuer de répondre à ces multiples critères.

En fait, les « données probantes » sur lesquelles s'appuyer sont le plus souvent incertaines, dynamiques, contestables et incomplètes (Lomas *et al.*, 2005). Il semble difficile de se passer du jugement des professionnels, d'une analyse des contextes locaux, et d'une diversité d'outils et d'approches pour pouvoir s'y ajuster. Un nombre croissant d'études scientifiques s'inscrivent donc en faux contre la croyance selon laquelle il suffirait de diffuser largement des outils ou des dispositifs qui ont montré une certaine efficacité et de les reproduire ailleurs pour espérer obtenir les mêmes résultats. De son côté, l'Association américaine des psychologues (APA) a adopté en 2006 la définition des « pratiques fondées sur des preuves » suivante : « l'intégration de la meilleure recherche disponible et de l'expertise [professionnelle] en tenant compte des caractéristiques, de la culture et des préférences du [bénéficiaire] ».

Plus fondamentalement, si les recherches peuvent nous informer sur la probabilité d'aboutir à tel ou tel résultat suite à telle ou telle action, elles ne permettent pas d'établir la nature du résultat souhaitable. La démarche scientifique peut nous aider à lire le réel, à pointer les contradictions ou les incohérences entre les intentions annoncées et les effets probables, mais elle n'implique pas de prescription. Définir ce qu'est une bonne éducation et ses finalités, est et reste une question de valeurs, éminemment politique, ouverte au débat démocratique, et pas une affaire de science ou de technique. Ni transformer les enseignants en chercheurs, ni rajouter aux chercheurs la charge de devenir des experts en vulgarisation, ne paraissent des options viables. Si l'on veut limiter les risques d'instrumentalisation des résultats de la recherche et favoriser un dialogue fructueux entre acteurs éducatifs, il est dès lors crucial de créer des interfaces, des espaces intermédiaires, où une pluralité d'acteurs peuvent échanger et coconstruire (Rey, 2014). Des structures comme l'inspection, l'IFÉ ou le Cnesco par exemple, constituent peut-être des viviers pour des animateurs potentiels de ces espaces, maîtrisant à la fois les « réalités du terrain », les codes de la recherche, et les enjeux politiques.

BENOÎT GALAND

GIRSEF, Université catholique de Louvain, Belgique

<https://uclouvain.be/fr/repertoires/benoit.galand>

Références bibliographiques

- American Psychological Association (APA)(2006). Policy Statement on Evidence-Based Practice in Psychology. <http://www.apa.org/practice/guidelines/evidence-based-statement.aspx>
- Baillargeon N. (2006), *Petit Cours d'autodéfense intellectuelle*, Montréal, Lux Éditeur.
- Cèbe S., Goigoux R. (2009), *Lector et Lectrix. Apprendre à comprendre les textes narratifs*, Paris, Retz.
- Cohen L., Manion L., Morisson K. (2011), *Research methods in education* (7th edition), London, Routledge Falmer.
- Crahay M. (2007), *Peut-on lutter contre l'échec scolaire ?*, Bruxelles, De Boeck (3^e éd.).
- Dubet F. (2002), « Pourquoi ne croit-on pas les sociologues ? », *Éducation et Sociétés*, n° 9, p. 13-25.
- Dupriez V. (2015), *Peut-on réformer l'école ? Approches organisationnelle et institutionnelle du changement pédagogique*, Louvain-la-Neuve (Belgique), De Boeck supérieur.
- Durlak J.A., DuPre E.P. (2008), « Implementation Matters: A Review of Research on the Influence of Implementation on Program Outcomes and the Factors Affecting Implementation », *American Journal of Community Psychology*, vol. 41, issue 3-4, p. 327-350.
- Galand B. (2017-1), « Quels sont les effets de la différenciation pédagogique sur les dimensions cognitives et socio-affectives ? », conférence de consensus « Différenciation pédagogique : comment adapter l'enseignement pour la réussite de tous les élèves ? », Paris, 7-8 mars 2017, Notes des experts (p. 177-187), Cnesco, France. Disponible sur : http://www.cnesco.fr/wp-content/uploads/2017/04/170331_Notes_experts.pdf

- Galand B. (coord.), (2017-2), *Prévenir le harcèlement à l'école. Oui, mais comment ?*, Louvain-la-Neuve (Belgique), Presses universitaires de Louvain.
- Gorman D.M. (2005), « Drug and violence prevention: Rediscovering the critical rational dimension of evaluation research », *Journal of Experimental Criminology*, vol. 1, issue 1, p. 39-62.
- Gottfredson D.C. et al. (2015), « Standards of Evidence for Efficacy, Effectiveness, and Scale-up Research in Prevention Science: Next Generation », *Prevention Science*, vol. 16, issue 7, p. 893-926.
- Harari Y.N. (2015), *Sapiens. Une brève histoire de l'humanité*, Paris, Albin Michel.
- Hattie J. (2009), *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*, Taylor & Francis.
- Lilienfeld S.O. et al. (2014), « Why Ineffective Psychotherapies Appear to Work », *Perspectives on Psychological Science*, vol. 9, issue 4, p. 355-387.
- Lomas J., Culyer T., McCutcheon C. (2005), *Conceptualizing and combining evidence for health system guidance*, Canadian Health Services Research Foundation.
- Makel M.C., Plucker J.A. (2014), « Facts Are More Important Than Novelty: Replication in the Education Sciences », *Educational Researcher*, vol. 43, issue 6, p. 304-316.
- Mitchell G. (2012), « Revisiting Truth or Triviality: The external validity of research in the psychological laboratory », *Perspectives on Psychological Science*, vol. 7, issue 2, p. 109-117.
- Oreskes N., Conway E.M. (2012), *Les Marchands de doute*, Paris, Éditions Le Pommier (2^e éd. 2014).
- Rey O. (2014), « Entre laboratoire et terrain : comment la recherche fait ses preuves en éducation », *Dossier de veille de l'IFÉ*, n° 89, Lyon, ENS de Lyon.