



Practice-based research : une aporie et des espoirs

Une revue critique de l'article d'Anthony S. Bryk

Pascal Bressoux



Édition électronique

URL : <http://journals.openedition.org/educationdidactique/2870>

DOI : 10.4000/educationdidactique.2870

ISSN : 2111-4838

Éditeur

Presses universitaires de Rennes

Édition imprimée

Date de publication : 31 décembre 2017

Pagination : 123-134

ISBN : 978-2-7535-7319-2

ISSN : 1956-3485

Référence électronique

Pascal Bressoux, « *Practice-based research* : une aporie et des espoirs », *Éducation et didactique* [En ligne], 11-3 | 2017, mis en ligne le 31 décembre 2017, consulté le 15 septembre 2020. URL : <http://journals.openedition.org/educationdidactique/2870>

PRACTICE-BASED RESEARCH : UNE APORIE ET DES ESPOIRS
UNE REVUE CRITIQUE DE L'ARTICLE D'ANTHONY S. BRYK

Pascal Bressoux
Université Grenoble Alpes
Laboratoire de recherche sur les Apprentissages en contexte – LaRAC (ex LSE)

LE CONSTAT

L'article de Bryk (2015) est établi sur un constat : il y a une demande et une exigence croissantes vis-à-vis de l'École ; le fonctionnement des écoles, de même que le métier d'enseignant, sont devenus de plus en plus complexes ; or, en dépit du premier point et pour partie en raison du second, nous ne savons pas passer des connaissances de la recherche à une amélioration de la pratique. Il nous faut donc changer de paradigme ou, pour le moins, l'amender et, pour cela, faire en sorte que le monde des enseignants et le monde des chercheurs soient moins étanches l'un à l'autre. Il ne faut plus considérer, d'un côté, ceux qui savent, et de l'autre, ceux qui font et sont censés mettre en place ce que les premiers leur disent de mettre en place. L'amélioration générale du système éducatif ne peut se faire sans une collaboration de l'ensemble. Mais de quelle collaboration est-il question ?

UNE TERMINOLOGIE BRUMEUSE

Il s'agit de promouvoir un paradigme de l'amélioration, fondé sur des communautés structurées en réseaux, et qui va permettre de réaliser de la *practice-based research*. Il s'agit là, bien sûr, d'un pas de côté par rapport à la notion d'*evidence-based practice*. Remarquons toutefois que la première braque la focale sur la recherche tandis que la seconde la braquait sur la pratique. C'est, à mon avis, une nuance importante : il s'agit de réfléchir à une façon de faire de la recherche, en la fondant sur la pratique (du terrain), plutôt que de savoir comment mettre en place (sur le terrain éducatif) une pratique fondée sur les « preuves »¹ scientifiques, ou données probantes. En toute rigueur, la terminologie *practice-based research* paraît cependant aporétique : comment la recherche pourrait-elle être fondée sur une pratique qui lui est externe ? Il y a là une difficulté qui me semble insoluble. La recherche se fait par une pratique scientifique qui est bien évidemment fondée sur des règles qui lui sont propres : explicitation, rationalité, vérifiabilité, cohérence, etc. Gageons que la similitude avec la terminologie *evidence-based practice* a été voulue par les concepteurs de ce nouveau « paradigme » et qu'il convient de ne pas la prendre trop au pied de la lettre, mais plutôt en un sens plus allusif, de recherche *qui prend en considération* la

pratique (de terrain). Il y a là toutefois un flou initial qui pourra se révéler gênant.

Il me semble que ce que Bryk critique dans la notion d'*evidence-based practice*, c'est l'idée que le terrain pourrait appliquer fidèlement, et à large échelle, ce qui a été montré comme « marchant » expérimentalement. On notera au passage que cela ne conduit en rien Bryk à dévaluer la recherche fondamentale ou encore expérimentale ; il présente à plusieurs endroits du papier la nécessité et l'intérêt de ce type de recherche. Il ne s'agit donc pas de mettre en doute sa validité scientifique mais, d'une part, d'étudier la variabilité des effets dégagés (j'y reviendrai) et, d'autre part, de questionner la difficulté à implanter à large échelle, du fait de la complexité même du fonctionnement du système scolaire, des pratiques qui semblent favorables. Pour le dire simplement : comment passer d'expériences « sous cloche » à des pratiques généralisées.

MAINTENIR DES STANDARDS SCIENTIFIQUES ÉLEVÉS

Balayons tout d'abord une conclusion que d'aucuns pourraient être tentés de tirer du constat précédent : le repli sur des communautés de pratiques. Selon Bryk, si entrevoir les enseignants comme des agents qui recevraient passivement la connaissance des chercheurs – et l'utiliseraient à bon escient dans leur classe – est erronée, l'idée qui consiste à faire reposer nos espoirs sur les communautés de pratiques l'est tout autant. En effet, celles-ci, bien que concevant les enseignants comme plus actifs et brisant le caractère solitaire de la pratique individuelle, posent le problème de la cumulativité des expériences et de leur évaluation. Les deux visions passent à côté d'une véritable collaboration dont toute la collectivité bénéficierait.

Plus fondamentalement, ce que Bryk veut dénoncer c'est cette idée qu'il y aurait quelque part des gens qui savent – les chercheurs –, d'autres qui font – les enseignants – et la séparation que cela engendre. La conséquence en est l'impossibilité d'une vraie collaboration. Il ne s'agit pas de nier la connaissance scientifique ; Bryk n'a pas à cet égard une vision relativiste. Il existe toutefois aussi une forme de connaissance, sans doute largement pratique, que les chercheurs n'ont pas, qui vient du métier, du travail enseignant et c'est à condition d'unir ces connaissances dans

une collaboration commune et de long cours qu'on pourra progresser dans un type de connaissance dont personne ne dispose actuellement : savoir comment améliorer le fonctionnement et les résultats d'un système éducatif. Remarquons bien qu'il ne s'agit pas seulement d'améliorer – même s'il s'agit là sans doute d'un but ultime – : on peut avoir réussi à améliorer dans un domaine, ou à un niveau particulier du système, mais on voudrait aussi avoir des règles générales d'amélioration, valables pour d'autres domaines et d'autres temps. Il s'agit donc surtout de *savoir comment améliorer*, ce qui représente un enjeu de connaissance et pas seulement d'action.

Les doutes de Bryk sur un modèle *top-down* allant de la recherche au terrain ne sont pas infondés. Marc Gurgand (2015) rappelle par exemple un constat édifiant établi par l'économiste Brian Jacob. Relatant la généralisation de 77 interventions menées depuis 2002 et jugées efficaces par l'*Institute of Education Sciences* (USA) selon des critères scientifiques (randomisation, etc.), Jacob a relevé que, une fois mises en place et évaluées en contexte écologique, seules 7 d'entre elles ont obtenu des effets positifs. Evidemment, on ne peut savoir à ce stade si ce nombre décevant est dû à un défaut d'implémentation (mauvaise compréhension des programmes, réticence à leur mise en place, défaut d'accompagnement, etc.), ou bien au fait que, implémentés correctement, beaucoup de programmes étaient en fait mal adaptés à certaines des conditions rencontrées dans la phase de généralisation. Voir aussi à ce sujet James-Burdumy *et al.* (2012) pour ce qui concerne des programmes d'intervention en lecture.

À l'évidence, tout n'est pas identique ; les contextes d'enseignement varient, et toute tentative d'implanter partout des solutions préfabriquées, sans se soucier des conditions de réception, de diffusion et d'exercice dans des contextes différents, a toutes les chances de se solder par un échec cuisant.

À ce niveau, un risque apparaît toutefois : celui de penser que rien n'est généralisable, que tout n'est qu'affaire de « rencontre », que l'on ne peut en aucun cas envisager de pratique menant à de meilleurs résultats et que, comme le récuse Bryk, on soit amené à « croire que chaque situation est unique et que, par conséquent, chaque éducateur, école, district doit inventer lui-même sa propre pratique » (p. 474).

J'ai pour ma part toujours été très agacé par les discours qui, sous prétexte de l'inextricable complexité du processus d'enseignement-appren-

tissage, prétendent qu'une connaissance générale sur l'efficacité des pratiques n'est pas possible. Que, par conséquent, cette efficacité ne peut être que le produit inattendu, imprédictible, d'un processus émergent, qui ne se transpose à rien et à personne, le résultat d'une rencontre unique entre des enseignants singuliers et des élèves singuliers.

Qui a jamais prétendu que l'acte d'enseignement n'avait aucune composante singulière ? Mais n'y a-t-il pas aussi une composante générique, commune à un grand nombre d'actes d'enseignement ? Ainsi, les travaux sur l'efficacité des enseignants montrent qu'un enseignant efficace dans une matière (p. ex. maths) a tendance à l'être aussi dans d'autres (p. ex. français) (Mingat, 1991), qu'un enseignant efficace dans une classe a tendance à l'être dans toutes celles où il enseigne (Felouzis, 1997), qu'un enseignant efficace une année a tendance à l'être les années suivantes (Chetty *et al.*, 2014). Autant de faits qui vont à l'encontre de l'hypothèse d'un phénomène émergent, imprédictible.

Je suis en grand accord avec Bryk quand il énonce la difficulté à généraliser des expériences scientifiques clairement contrôlées et localisées. Mais difficulté ne signifie pas impossibilité, et on connaît déjà quelques cas de généralisation réussie. Ainsi, le programme *Success for all* produit des effets constamment positifs depuis sa mise en place et sa généralisation (50 000 enseignants dans plus de 1 000 écoles dans 48 États et quatre autres pays touchant plus de 500 000 élèves d'après le site de la *Success for all Foundation*). Nul doute que les concepteurs et développeurs de ce programme ont acquis un savoir-faire en matière de diffusion de « bonnes pratiques ». Ce développement ne s'est toutefois pas fait du jour au lendemain.

Bryk signale que, souvent, l'urgence de trouver des solutions conduit à des montages précipités, avant même que le problème ait été correctement analysé. Il faut combattre la précipitation dans le changement. La « solutionite » nous guette tous. L'idée de Bryk selon laquelle il faut commencer petit, puis développer et amplifier au fur et à mesure que notre connaissance s'améliore, est sans doute très bonne.

Là où, selon moi, le propos de Bryk prend des tournants particulièrement opérationnels, c'est quand il propose de porter un nouveau regard sur les résultats des expérimentations. Selon lui, il faut s'intéresser, non pas seulement à la taille d'effet, c'est-

à-dire à l'effet moyen, mais aussi à la *variation* autour de cette moyenne. Si cela peut sembler anodin (tout statisticien sait qu'une distribution se décrit selon au moins deux paramètres, l'un de tendance centrale et l'autre de dispersion), cela renvoie à un renversement de conception où la variabilité n'est plus vue comme de l'erreur de mesure autour d'une tendance centrale « vraie », mais comme un phénomène à étudier *per se*. Le postulat de Bryk est que, dans cette variation, résident les causes mêmes de la difficulté à généraliser des expériences de recherche. La « science de l'amélioration » que réclame Bryk s'intéresserait à la variation davantage qu'à la tendance centrale et aurait ses propres méthodes d'investigation, favorisant la répétition des tests, la révision, la correction, et le recalibrage permanent des outils.

Cela conduit à porter l'accent sur l'étude de la répliquabilité qui est, selon moi, un autre point important de l'article. Il ne s'agit pas seulement de porter la focale sur « qu'est-ce qui marche », mais aussi et surtout sur « comment faire en sorte que ça marche dans différents contextes ». Et pour accomplir cela, il faut être capable de répliquer les résultats obtenus dans un contexte donné à d'autres contextes. Bryk érige ainsi la répliquabilité au rang de *Gold Standard*. Il s'agit bien sûr d'une manière de traiter la question de la généralisabilité des résultats et ce, non pas de façon abstraite, mais concrètement comme un problème posé d'emblée à la pratique. Car c'est par la pratique, parce qu'on apprend en faisant, que le système pourra s'améliorer. C'est par ce cycle d'actions et retours sur l'action que nous pourrions progresser dans la connaissance de ce que nous ne savons pas encore faire : améliorer le système complexe qu'est le système éducatif.

Il s'agit donc de mieux comprendre le travail et l'environnement de travail des enseignants. Pour ce faire, enseignants, chercheurs, techniciens, etc. doivent s'engager mutuellement dans le test, l'amélioration de ce qui se passe concrètement dans les écoles, par un travail collaboratif au long cours, qui permet de prendre en compte les spécificités des contextes d'implantation des programmes, des expérimentations, et en éprouve la généralisabilité.

LA QUESTION DE LA GÉNÉRALISABILITÉ AU CŒUR DES SCIENCES SOCIALES

Cette question de la généralisabilité est redoutable ; un détour par les travaux de Jean-Claude Passeron va nous en montrer le caractère inéluctable en sciences sociales. Passeron est probablement le chercheur qui a le mieux posé l'impossibilité pour les sciences sociales de produire un savoir nomologique, c'est-à-dire de produire des lois universelles. Il est dans le fondement même des sciences sociales de se poser la question de la généralisabilité de leurs résultats (Passeron, 2006). En effet, étant contextualisées, se pose toujours la question du transfert des résultats à d'autres contextes. Les sciences de la nature peuvent avoir des modèles « clos » (Passeron parle de « modèles universels ») au sens où tous les facteurs supposés influents sur le phénomène étudié sont contrôlés, ou inclus dans le modèle. Au contraire, les sciences sociales élaborent des modèles qui ne peuvent jamais être totalement clos car une multitude de facteurs potentiellement influents n'ont pu y être intégrés (Passeron parle de « modèles à déictiques »). Les phénomènes sociaux sont en effet marqués par leur historicité, c'est-à-dire leur insertion dans un contexte historique, social, géographique, économique, etc. déterminés.

Ainsi, les résultats produits par un modèle scientifique doivent-ils être recontextualisés (« re-historicisés »). Prenons l'exemple de la taille de la classe : si mon modèle me dit qu'il existe une relation entre celle-ci et les acquisitions des élèves, encore devrais-je préciser de quel type d'acquisitions il s'agit (français ? maths ? etc. ; s'il s'agit de maths, parle-t-on de géométrie ?, de numération ?), auprès de quels élèves (quel âge ? quel milieu social ?), dans des classes dont la taille varie de combien à combien, etc. Il y a peu de chances, par exemple, que mon résultat, s'il est obtenu en France auprès d'élèves de CP dans des classes variant de 15 à 27 élèves, soit généralisable à des contextes très différents : il ne dit sans doute pas grand-chose de ce qui se passe au Tchad dans des classes de lycée dont la taille varie de 50 à 200 élèves... Les coordonnées spatio-temporelles² de ces deux contextes, relatives à l'objet d'étude, sont bien trop différentes pour qu'une indifférenciation des contextes soit envisageable. En revanche, et à charge de vérification, il y a des chances raisonnables que les résultats comme ceux de la taille des classes se généralisent à des contextes aux coordon-

nées proches : ces résultats sont en effet très voisins dans les pays occidentaux (USA, France, Danemark, Suède, Pays-Bas, etc.) et ce, y compris sur des configurations relativement précises : effets bénéfiques des petites classes, plus marqués chez les élèves défavorisés et en début de scolarité.

À l'inverse, plusieurs travaux ont montré que, parmi les facteurs qui semblent définir une direction efficace d'établissement aux États-Unis, certains voyagent assez mal en Europe (France, Pays-Bas...) (Bressoux, 1994). C'est le cas, par exemple, d'un *leadership* pédagogique affirmé. Cela renvoie aux prérogatives très différentes qui sont données aux directeurs d'école aux États-Unis et dans plusieurs pays d'Europe : là où, aux États-Unis, le directeur a des rôles étendus qui s'assimilent à celui d'un chef d'entreprise, ailleurs, il est plutôt vu comme *primus inter pares*, avec des pouvoirs administratifs et pédagogiques beaucoup plus limités. Dans ce cas, les « coordonnées administratives » séparent de façon pertinente les deux types de contextes.

Dans les sciences sociales, la question de la généralisabilité (appelée aussi validité externe) fait donc encore partie de la question de recherche. Cela ne signifie pas qu'aucune connaissance générale n'existe ou ne soit possible, mais que ces connaissances sont nécessairement indexées à un contexte plus ou moins étendu. Aucun précepte pédagogique un tant soit peu opérationnalisé pour la pratique ne peut être *solus et omni definito* :

« Il n'y a pas de didactique en soi, des mathématiques par exemple, valable pour tous publics y compris les extra-terrestres (même s'ils avaient la même mathématique que nous). "Didactique" est donc une recherche qui soumet un savoir donné à un traitement particulier déterminé par une fin particulière, à savoir la transmission de ce savoir dans des conditions historiques, sociales et psychologiques qui sont toujours particulières » (Passeron, 2006, p. 532-533).

Pour autant, cela ne conduit nullement Passeron à un défaitisme scientifique, à un abandon de toute recherche d'optimisation d'efficacité, bien au contraire :

« Il n'est nul besoin d'argumenter pour établir que l'apprentissage des savoirs scientifiques ne saurait être, par une exception épistémologique, le seul objet, psychologique ou sociologique, soustrait à la recherche

fondamentale en sciences humaines, pas plus que le processus pédagogique ne saurait être le seul processus qui échapperait à la possibilité de rationalisation, c'est-à-dire à l'application de technologies de l'efficacité elles-mêmes issues de la recherche scientifique » (Passeron, 2006, p. 529, souligné par nous).

La nature contextualisée de l'acte pédagogique n'exclut donc en rien la question de la recherche scientifique d'une plus grande efficacité. Si la « contrainte déictique » (Passeron, 2006, p. 573) empêche de postuler la constance des contextes, elle n'empêche évidemment en rien l'existence d'une parenté entre contextes. Reste à savoir quelles sont les « coordonnées » qui vont permettre de définir la parenté – ou l'étrangeté – des contextes relativement à l'objet en question : en l'occurrence, les pratiques qui doivent conduire à améliorer les apprentissages. Ces coordonnées doivent-elles ressortir à des caractéristiques des élèves (lesquelles?), à celles des enseignants (lesquelles?), à celles de l'environnement socio-économique (lesquelles?), à tout cela à la fois, à d'autres choses encore (lesquelles?) ?

GÉNÉRALISATION : UNE QUESTION DE « VALIDITÉ EXTERNE PURE » ET « D'EXTENSION »

Il me semble qu'une première distinction, importante, est à faire entre les recherches en laboratoire et les recherches écologiques ; je veux dire par là une distinction entre les recherches qui se font en dehors de la classe, de son fonctionnement habituel, et les recherches qui s'effectuent en conditions de classe.

Il faut en effet séparer ce qui relève de la *généralisation au contexte de classe* de ce qui relève de la *généralisation aux autres classes*. Ce sont là, me semble-t-il, deux problèmes tout à fait différents même si leur dénomination sous le terme de « généralisation » est commune. Dans le premier cas, se pose ce que j'appellerais une question de « *validité externe pure* » tandis que, dans le second, se pose une question « *d'extension* ».

Concernant la validité externe pure, il me suffira de prendre un exemple : celui des recherches sur le travail coopératif. S'il est assez facile, en condition de laboratoire, de s'assurer que les élèves participent effectivement à la tâche proposée, ce n'est plus du tout vrai en condition réelle de classe, où l'enseignant va devoir gérer plusieurs groupes d'élèves. Il y a là un

art pédagogique, qui consiste à faire fonctionner des groupes simultanément dans une situation ordinaire de classe, qui n'entre pas du tout en jeu dans la situation de laboratoire où, en général, un seul groupe est constitué et observé de près et ce, dans des conditions tout à fait exceptionnelles pour les sujets. Ainsi, les recherches en laboratoire ont clairement montré que certaines compositions de groupes étaient plus favorables que d'autres aux acquisitions des élèves, en particulier pour générer un conflit sociocognitif, (Darnon, Butera & Mugny, 2008). Il faut s'assurer que les sujets se lancent dans une relation épistémique, c'est-à-dire une véritable discussion sur la manière de réaliser la tâche prescrite, et non dans une relation sociale où l'enjeu est le statut que l'on a au sein du groupe. Il a été montré par exemple que de trop fortes asymétries de compétences dans la composition des groupes pouvaient conduire les forts à asseoir leur domination et les autres à suivre leur opinion sans aucune mise en jeu de leur propre conception, sans conflit sociocognitif donc.

Même si ces éléments sont très importants à prendre en compte, y compris en situation de classe, les recherches qui ont porté sur le travail collaboratif en situation de classe ont, quant à elles, assez vite pointé qu'une des conditions était que les élèves... travaillent, ce qui n'est pas si évident car les travaux de groupes supposent généralement une dose d'autonomie. Aussi, les recherches écologiques ont-elles montré qu'il était efficace, voire nécessaire, que deux types d'objectifs soient assignés lors d'un travail coopératif : un objectif de réussite du groupe et un objectif de réussite individuelle (Slavin, 1980). Ainsi, même si la collaboration entre les membres du groupe demeure fondamentale, chacun est responsabilisé et personne ne doit se laisser aller à de la « paresse sociale » (*social loafing*), phénomène parasite bien connu des travaux à plusieurs.

Ainsi, construire un programme d'enseignement coopératif où la consigne d'organisation donnée aux enseignants consiste seulement à leur demander de s'assurer d'une bonne composition de groupe risquerait de produire des résultats bien décevants même si, du point de vue expérimental, le programme semblait prometteur.

Que peut-on tirer de cela ? Assurément pas que la recherche en laboratoire serait non pertinente. Elle peut être au contraire d'une formidable utilité, parce qu'elle permet justement d'isoler les phénomènes : dans l'exemple précédent, ces recherches permettent

à l'évidence de mieux comprendre les processus de groupe susceptibles de favoriser les apprentissages. Toutefois, sa force heuristique fait aussi sa faiblesse écologique, et il y a probablement très peu de résultats obtenus en laboratoire qui ont une traduction immédiate en classe. On peut en tirer que, pour passer de « efficace en laboratoire » à « efficace en classe », il faut nécessairement une phase de test à petite échelle dans les classes. Il s'agit alors de vérifier que les conditions d'implantation en contexte écologique sont valides, que cela ne va pas entraîner de perturbations négatives qui annihileraient l'effet observé en laboratoire.

Vient alors le problème de l'extension. Dans ce cas, la question peut être celle de l'échelle : comment passer d'une petite échelle à une grande échelle ? Pourra-t-on assurer une aussi bonne formation aux enseignants impliqués, un aussi bon suivi, une proximité qui fournissait des *feed-back* fréquents et pertinents quant à la mise en place des activités, une relation qui assurait également de la confiance, etc. ? Il s'agit là d'une question importante.

Cependant, le problème de l'extension peut aussi recouvrir une autre question, qui risque d'être plus redoutable encore : en étendant, ne risque-t-on pas de changer certains contextes ? Et on en revient alors au problème fondamental soulevé par Passeron : l'indexation sur les contextes, dont les « coordonnées » peuvent varier, et qui peuvent rendre non pertinents les résultats obtenus ailleurs.

Pour simplifier, supposons qu'une pratique soit efficace, ce qui est toujours problématique parce que 1) une pratique n'est pas « bonne » en elle-même mais « meilleure » ou « moins bonne » que telle autre, car l'efficacité est un concept relatif et 2) il est difficile de définir une pratique en dehors de son actualisation – parce que c'est une pratique justement. En ce sens, un programme expérimental qu'on souhaiterait implanter dans les écoles ne constitue pas une pratique mais un *discours* qui prescrit une pratique. En tant que programme, il ne peut pas définir tous les éléments de la pratique qui sera effectivement mise en place, mais seulement certains axes, des structures fondamentales, des modes d'organisation, de gestion, des contenus, etc. C'est pourquoi le programme est, *per se*, toujours débordé par la pratique. Même dans le cas d'une bonne volonté d'application, le programme est toujours tordu, fait à la main de l'enseignant. Mais, par commodité, restons sur cette supposition qu'une pratique soit

efficace et que, dans sa généralisation, seule se pose une question de mise en place. Pour quelle raison cette généralisation pourrait-elle conduire à des résultats décevants ?

On peut placer une série d'hypothèses :

- l'enseignant *ne veut pas* mettre en place tout ou partie du programme. N'oublions pas que nombre de programmes expérimentaux sont testés sur des *enseignants volontaires*, ce qui n'est plus le cas en phase de généralisation. Non seulement on n'a plus affaire (seulement) à des volontaires, mais certaines personnes peuvent être réticentes initialement ou le devenir en cours de mise en place. On voit qu'un point important est de gagner l'adhésion des enseignants et de garder cette adhésion. Or, on peut ne pas gagner l'adhésion d'un enseignant car :

- cela va à l'encontre de ses connaissances théoriques ;

- cela va à l'encontre de son expérience ;

- c'est trop dur, trop coûteux à mettre en place. La pratique quotidienne, quelle qu'elle soit, va toujours au plus simple : si le rapport coût / efficacité perçu est défavorable, la pratique est systématiquement biffée (et il faut insister sur le fait que ce n'est pas une question de paresse, mais d'économie cognitive) ;
- l'enseignant *ne peut pas* mettre en place tout ou partie du programme. Malgré une adhésion initiale de l'enseignant, des éléments extérieurs empêchent cette bonne mise en place :

- élèves, parents, collègues, sont réticents. L'enseignant n'est en effet pas le seul agent concerné : il a de nombreux « partenaires » qui, eux, peuvent ne pas vouloir du nouveau programme. Le cas des évaluations sans notes en est une illustration typique : si les parents veulent des notes pour leur enfant, il sera très difficile de tenir cette pratique ;

- sans réticence *a priori* des élèves ou autres personnes impliquées, la nouvelle pratique ne « prend » pas ; elle conduit au chahut ou inversement à la passivité ; élèves et parents n'en comprennent pas l'utilité, etc. ;

- l'enseignant *ne sait pas* mettre en place tout ou partie du programme :

- il y a toujours des coins aveugles dans un programme, qui présupposent l'existence de savoirs théoriques, pédagogiques ou didactiques, ou simplement de savoir-faire d'expérience : or, la formation initiale peut être insuffisante ou inadaptée à la mise en place de

la nouvelle pratique, ou bien encore l'enseignant peut manquer d'expérience ;

- il y a un déficit de prescription (le « comment faire » et le « quoi faire »). Par exemple, un enseignant peut être convaincu du bien-fondé d'un programme d'enseignement explicite sans pour autant savoir comment le mettre en place concrètement pour l'apprentissage de certaines notions ;

- il y a un défaut d'accompagnement : on touche là, à mon sens, un aspect très crucial. Je suis persuadé que, au-delà de toute prescription, formation, information, l'accompagnement dans la mise en place est une clé pour l'efficacité finale. Il ne faut pas seulement informer, mais aussi conforter, rassurer, encourager, justifier, etc.

On voit que même dans un cas simplifié à l'extrême (nous avons que supposé que le problème n'était que celui de la mise en place d'une pratique efficace), les raisons qui pourraient conduire à une faible efficacité dans sa généralisation sont nombreuses.

D'autant que l'efficacité d'une pratique n'a pas la limpidité qu'on serait tenté de lui accorder naïvement. Il est en effet très difficile, du point de vue des acteurs, d'avoir une vue claire sur l'efficacité des pratiques, y compris pour les enseignants. Comme on dit, ça ne saute pas aux yeux. Il est très difficile pour des enseignants de savoir si leur pratique est efficace. D'une part, parce que les élèves progressent, quelle que soit la pratique. D'autre part parce que, je le répète, l'efficacité est relative ; une pratique est efficace relativement à une autre. Elle est donc, selon certains critères d'efficacité établis, meilleure, équivalente, ou moins bonne qu'une autre. La conséquence est qu'il faudrait pouvoir comparer les résultats avec ceux issus d'une autre pratique, ce que les enseignants ne peuvent évidemment pas faire : ils sont dans l'impossibilité de mettre en place simultanément deux pratiques. En conséquence, ils ne disposent pas du contrefactuel. Enfin, les apprentissages des élèves se construisent dans le long terme, ils n'ont pas l'immédiateté de résultat d'une pièce sortant d'une machine. Un enseignant n'est pas dans la position d'un opérateur qui voit de suite, au calibrage des pièces sorties de sa machine, si celle-ci est bien réglée ou non.

DES COMMUNAUTÉS D'ENSEIGNANTS ET CHERCHEURS STRUCTURÉES EN RÉSEAUX?

Que peut-on tirer de ces divers éléments? Tout d'abord qu'il ne suffit pas de faire des appels incantatoires à la relation recherche-terrain, tant que les acteurs de ces deux mondes ont des contraintes qui les repoussent les uns des autres, des intérêts modestes à collaborer, ou que l'attraction subie ne peut durer qu'un temps fort limité.

Le temps constitue ici, à mon sens, la ressource et la contrainte majeures. Si la répliquabilité des résultats et le développement, au sens de recherche et développement, sont les maîtres-mots (le nouveau *Gold standard*), alors la durée est au cœur du nouveau paradigme. Comment éviter qu'une collaboration enseignant-chercheur ne soit éphémère? Une fois la collaboration réalisée, que les enseignants ont œuvré, appris, apporté leur expertise professionnelle au projet et au raffinement du protocole de recherche, que les chercheurs ont eux-mêmes appris, mis en place la recherche et publié les résultats, qu'est-ce qui peut les inciter à maintenir leur collaboration? La bonne volonté ne suffit pas : il faut que les deux communautés aient un intérêt à collaborer et, pour cela, la première condition est qu'elles en perçoivent l'une et l'autre l'utilité. La deuxième condition est que cette collaboration soit soutenue, facilitée par des infrastructures adaptées. Le temps de la recherche n'étant pas le temps pédagogique, ces nouvelles infrastructures doivent aider à coordonner ces deux temporalités. Je rejoins en cela l'analyse de Bryk. Il faut de plus que cette communauté soit vaste : là où deux groupes de chercheurs et enseignants s'épuiseront vite, la multiplicité des acteurs assure un renouvellement permanent.

Reste que Bryk ne dit ni ce que doivent être ces communautés ni la manière de les organiser, ni même ce que peut être une « théorie en marche partagée » (*shared working theory*) (p. 473) qu'il appelle de ses vœux. On peut toutefois, modestement, relater quelques expériences et avancer certaines propositions.

Tout d'abord, le lien partenarial avec les enseignants peut être assuré dans les recherches elles-mêmes quand celles-ci sont conduites sur le terrain. C'est ce qui a été fait par exemple dans l'étude PARLER. Celle-ci a consisté à mettre en place, de la grande section de maternelle jusqu'en fin de CE1, un dispositif pédagogique visant à favoriser l'acquisition

du langage : enseignement explicite et systématique, dans des « groupes de besoins », des habiletés liées à l'acquisition du code écrit et des habiletés liées à la compréhension pratiquée à l'oral (cf. Zorman *et al.*, 2015). Les résultats positifs engendrés par le programme PARLER sont, à notre sens, inséparables des conditions de mise en place et de suivi du dispositif. En ce sens, PARLER n'est pas seulement un ensemble de séquences pédagogiques, éventuellement présentables dans un cadre programmatique, mais c'est aussi un dispositif de mobilisation et d'accompagnement des équipes pédagogiques sachant que, au départ, les enseignants n'étaient pas volontaires pour participer à l'expérimentation. L'orientation, la préfiguration des outils, les habiletés à travailler *viennent bien des acquis de la recherche* (voir par exemple Bianco, 2015), mais *la manière de les mettre en œuvre a été spécifique*, adaptée aux conditions de réalisation pratique de l'expérimentation et au type d'écoles (toutes étaient en REP). Il suffit de prendre des résultats de recherche, aussi bien établis soient-ils, et de tenter de les mettre en place dans les écoles, pour se rendre compte que ces résultats ne disent pas tout – et ne peuvent pas tout dire – de leur implantation concrète. On est alors confronté à une multitude de décisions à prendre, d'arbitrages à faire. C'est là que la participation des enseignants est primordiale. Ainsi, les enseignants de PARLER ont participé à l'adaptation des outils et au montage des séquences, de même qu'à leur programmation dans le temps. Ils ont bénéficié d'un temps de formation initiale et continue, ils ont aussi reçu un soutien constant des équipes de recherche et de circonscription. Si bien que PARLER n'est pas un programme clé en mains, mais un programme co-élaboré entre équipes de recherche, équipes d'enseignants et équipes de circonscription. Implanté ailleurs, il n'aurait pas eu exactement la même configuration, *bien qu'il eût conservé les mêmes principes*.

Tous ces éléments font que la généralisation d'une expérience comme celle de PARLER est problématique et que cela constitue un enjeu tant pour les chercheurs que pour les enseignants. Il est difficile d'envisager que, développée à grande échelle, cette expérience puisse se dérouler avec le même degré d'articulation recherche-terrain. Sans doute, des phases intermédiaires d'élargissement seraient-elles nécessaires pour apprendre ce qui marche et marche moins bien. Mais tout cela réclame du temps, et surtout un mode d'organisation qui soutienne conti-

nûment, et soit lui-même pérenne. C'est pourquoi se pose inévitablement la question des institutions ou, dans les termes de Bryk, des infrastructures, qui pourraient favoriser cette mise en place.

Il semble évident que, pour que la recherche alimente la pratique et que la pratique éclaire la recherche sur ses contraintes, il faut des contacts fréquents et prolongés entre les chercheurs et les praticiens.

La formation des enseignants est évidemment au cœur de cette question. Le couplage avec la recherche doit débiter avec la formation initiale et se prolonger en formation continue. C'est parce qu'ils auront été initiés à des résultats et à des méthodes de recherche que les enseignants seront persuadés que celle-ci peut leur être utile et qu'ils sauront ce qu'ils peuvent en attendre. Les ESPE sont normalement chargées d'assurer ce lien entre recherche et formation. Même si beaucoup de choses restent à améliorer, leur création a sans doute contribué à un meilleur partenariat entre formateurs, chercheurs et cadres de l'Éducation nationale. Cette formation initiale est importante car, couplée aux avancées de la recherche, elle amène à une interrogation perpétuelle qui est difficilement compatible avec la simple transmission de l'expérience de génération à génération.

La formation continue des enseignants est cruciale elle aussi. Les connaissances scientifiques sur l'enseignement-apprentissage sont en progrès permanent. Comment pourrait-on justifier que les enseignants en soient coupés au long de leur carrière ? Il faut qu'ils y aient un accès direct, via les chercheurs eux-mêmes, et qu'ils ne soient pas contraints à aller chercher eux-mêmes une information qui n'est pas toujours facilement accessible. Et quand bien même elle le serait, il est bien souvent difficile de faire la part de ce qui est scientifiquement validé de ce qui ne l'est pas ou peu. D'ailleurs, cette distinction est d'autant plus difficile qu'on n'a pas reçu initialement de formation à la recherche, ce qui rejoint la question de la formation initiale. Des enseignants formés sont aussi des enseignants peuvent exercer un regard critique sur les études et résultats qui leur sont présentés.

Aujourd'hui, la formation continue des enseignants se fait largement en vase clos, via les DAF. Pourquoi n'y aurait-il pas des modules universitaires que les enseignants pourraient suivre et créditer, ce qui leur permettrait de progresser dans l'obtention d'un diplôme universitaire. Les formations pour devenir conseiller pédagogique ou formateur, le

CAFIPEMF au premier degré et le CAFFA récemment créé pour le niveau secondaire, ne bénéficieraient-elles pas d'une collaboration – et non pas seulement d'interventions ponctuelles – avec des chercheurs. Les masters MEEF, et en particulier la mention Pratiques et Ingénierie de Formation (PIF) ne seraient-ils pas un bon support pour ces formations, sous des formes qui restent à inventer mais, toujours, s'appuieraient sur une solide recherche universitaire ?

Je suis persuadé que les enseignants auront d'autant plus envie de collaborer avec les chercheurs qu'ils sauront ce qu'ils peuvent en attendre et qu'ils sauront mieux formuler leur demande. Car la demande de recherche peut venir aussi du terrain. Quelles institutions peuvent supporter une telle manière de fonctionner ?

Les Instituts Carnot Éducation (ICE), dont le premier a été mis en place à titre expérimental dans la Région Auvergne-Rhône-Alpes, pourraient être des structures qui favorisent la création de ces communautés d'enseignants et chercheurs structurées en réseaux que Bryk réclame. En effet, l'ICE fonctionne dans deux sens : dans un sens, les chercheurs proposent des projets de recherche et, dans l'autre, les équipes enseignantes formulent des « projets d'action éducative » qui doivent être adossés à la recherche. Il y a un nécessaire couplage à trouver entre les deux. Dans certains cas, les projets de recherche et les projets d'action éducative se sont déjà trouvés et les demandes d'affiliation à l'ICE reflètent ce couplage. Dans d'autres cas, les deux types de projets se font indépendamment ; le rôle de l'ICE est alors de mettre en correspondance ceux qui peuvent l'être. Les projets de recherche sont l'objet d'une expertise scientifique rigoureuse de façon à assurer une qualité scientifique irréprochable. De l'autre côté, les projets d'action éducative sont examinés du point de vue de leur intérêt pédagogique et de leur potentiel de développement de compétences dans les équipes.

L'ICE réalise des séminaires où chercheurs et enseignants viennent exposer conjointement l'avancée de leur travail commun. Par ailleurs, l'ICE a recruté des « passeurs », personnes qui sont à la fois dans le monde enseignant et ont un pied dans la recherche (p. ex. études doctorales.). Ces passeurs assurent le lien entre recherche et terrain, favorisent l'avancée du travail ou la formulation de nouvelles demandes, etc.

L'ICE est donc une tentative intéressante de structurer les relations entre recherche et terrain. La présence de différentes instances dans ses conseils (rectorats, ESPEs³, universités...) en est le reflet institutionnel. La question qui demeure est celle de la pérennité, non seulement sur un plan politique (quel est l'avenir des ICE ?) mais aussi du point de vue du suivi des projets. Les ICE devraient être, à mon sens, des institutions qui permettent de faire de la recherche et développement. Ils sont un cadre favorable à la mise en place de dispositifs à petite échelle dans un premier temps, puis une échelle progressivement élargie, et avec un temps suffisamment long qui permet de raffiner les méthodes et les dispositifs.

CONCLUSION

L'article de Bryk constitue un argumentaire stimulant pour la recherche en éducation. Ce que j'en retiens pour ma part, c'est que la recherche en éducation a besoin de faire et d'apprendre à faire du développement, au sens que prend ce mot dans « recherche et développement » et qui recouvre à la fois un aspect diffusion de plus en plus large et un aspect amélioration continue. Progresser dans nos connaissances sur la base de l'étude de la variabilité des effets, avec un souci de répliquabilité des effets de nos dispositifs me semble primordial. Je suis convaincu en effet que nous ne savons pas encore bien faire de développement dans la recherche en éducation. L'une des raisons en est que le temps de la pratique et le temps de la recherche académique ne coïncident guère. L'urgence à publier, la nécessité de trouver des fonds de recherche, de même que la difficulté à nouer des liens stables et durables avec les établissements scolaires font que les chercheurs ont une propension plus forte à proposer de nouveaux protocoles qu'à améliorer lentement, progressivement des dispositifs déjà en partie éprouvés. D'un autre côté, les innovations qui viennent du terrain sont souvent disparates, non évaluées, non cumulatives et elles ne peuvent donc que très rarement rencontrer la recherche académique. De plus, il n'existe guère de support institutionnel à ce qu'enseignants et chercheurs puissent engager et maintenir leur collaboration dans le long terme. C'est pourquoi, couplée à la question de la formation initiale et continue des enseignants, la question des infrastructures me semble essentielle et que j'ai esquissé que les ICE

pouvaient constituer une initiative intéressante dans cette direction.

Si la notion de *practice-based research* signifie faire de la R&D en éducation, alors cette injonction me semble pertinente (bien que je conserve les réserves que j'ai données à cette terminologie). Elle ne peut toutefois, selon moi, remplacer la notion d'*evidence-based practice* parce que les deux ne se situent pas au même niveau. Il me semble utile de distinguer la *production* des données probantes, de celle de leur *diffusion et amélioration*. Bien entendu, ces deux ensembles ne sont pas disjoints et mon recours aux travaux de Passeron avait pour but de rappeler qu'il ne peut y avoir en sciences sociales de vérité « hors-sol ». Toute donnée probante suppose son indexation à un contexte qui en donne à la fois la pertinence et les limites : en ce sens, il ne peut y avoir de donnée probante en éducation qui ne contienne en elle-même la question de sa généralisation à d'autres contextes. Toutefois, il me semble que la production de connaissances sur l'efficacité de dispositifs éducatifs et la production de connaissances sur la diffusion / extension de ces dispositifs sont des moments heuristiques différents.

Il y a certes une vision radicale de l'*evidence-based practice* qui consiste à montrer ce qui marche et à dire aux enseignants qu'ils n'ont plus qu'à l'implanter dans leur classe. C'est cette vision que semble critiquer Bryk et l'expérience montre que les choses ne fonctionnent pas comme cela. Il existe toutefois aussi une vision plus mesurée qui consiste à dire que les connaissances issues de la recherche doivent parvenir jusqu'aux enseignants, de façon régulière et continue, et que même si cela ne conduit pas à une implantation littérale, directe et automatique, ces connaissances constituent un repère permanent pour les enseignants par rapport auquel ils peuvent juger et ajuster leur pratique. En somme, *evidence-based practice* et *practice-based research* seraient deux paradigmes complémentaires plutôt que substituables.

Gageons que l'article de Bryk va alimenter les discussions. Il serait souhaitable que celles-ci ne restent pas confinées à l'espace des chercheurs mais qu'elles touchent aussi celui des enseignants ainsi que le cercle des décideurs. Car l'amélioration du système éducatif est un enjeu majeur qui ne pourra se réaliser sans la participation de tous les acteurs concernés.

NOTES

1. Le terme anglais « *evidence* » est un faux-ami. Il ne signifie pas ce qui est évident, mais ce qui a à voir avec la preuve. Encore ce terme de « preuve » doit-il être précisé : il ne s'agit pas d'une preuve au sens abstrait et démonstratif du terme, mais d'une preuve empirique (dans le contexte d'une enquête policière, le terme « *evidence* » signifie « pièce à conviction »), c'est-à-dire d'un résultat empirique scientifiquement validé. Le terme « donnée probante » est sans doute la meilleure approximation qu'on puisse donner au terme « *evidence* ».
2. Nous reprenons ici, pour commodité le terme « coordonnées spatio-temporelles » employé par Passeron, tout en indiquant qu'il s'agit d'un raccourci pour signifier « coordonnées spatio-socio-économico-culturo- etc.-temporelles ». Bref, toute coordonnée qui permet de caractériser le contexte de façon pertinente au regard de l'objet d'étude.
3. Les ESPEs font certes partie des universités, mais elles ont aussi une représentation spécifique dans l'ICE, en tant que centres de formation de personnels de l'Éducation nationale.

RÉFÉRENCES

- Bianco, M. (2015). *Du langage oral à la compréhension de l'écrit*. Grenoble : Presses universitaires de Grenoble.
- Bressoux, P. (1994). Les recherches sur les effets-écoles et les effets-maîtres. *Revue Française de Pédagogie*, 108, 91-137.
- Bryk, A. S. (2015). 2014 AERA distinguished lecture. Accelerating how we learn to improve. *Educational Researcher*, 44(9), 467-477.
- Chetty, R., Friedman, J. N., & Rockoff, J. E. (2013). *Measuring the impact of teachers I : evaluating bias in teacher value-added estimates*. NBER working paper series, Working Paper 19423, National Bureau of Economic Research. Repéré à : [<http://www.nber.org/papers/w19423>]
- Darnon, C., Butera, F., & Mugny, G. (2008). *Des conflits pour apprendre*. Grenoble : Presses universitaires de Grenoble.
- Felouzis, G. (1997). *L'efficacité des enseignants*. Paris : Presses universitaires de France.
- Gurgand, M. (2015, décembre). Comment interpréter les effets maîtres? Colloque *La cognition du maître*, organisé par le Groupe COMPAS-École normale supérieure. Paris : 16-18 décembre.
- James-Burdumy, S., Mansfield, W., Deke, J., Carey, N., Lugo-Gil, J., Hershey, A., Douglas, A., Gersten, R., Newman-Gonchar, R., Dimino, J., Faddis, B., & Pendleton, A. (2009). *Effectiveness of selected supplemental reading comprehension interventions : Impacts on a first cohort of Fifth-Grade students*. Washington, DC : Institute of Education Sciences, US Department of Education.
- Mingat, A. (1991). Expliquer la variété des acquisitions au cours préparatoire : les rôles de l'enfant, la famille et l'école. *Revue Française de Pédagogie*, 95, 47-63.
- Passeron, J.-C. (2006). *Le raisonnement sociologique* (édition revue et augmentée). Paris : Albin Michel.
- Slavin, R. E. (1980). Cooperative learning. *Review of Educational Research*, 50(2), 315-342
- Success for all Foundation. *Results*. Consulté le 3/10/17, repéré à : [<http://www.successforall.org/results/>]
- Zorman, M., Bressoux, P., Bianco, M., Lequette, C., Pouget, G., & Pourchet, M. (2015). « PARLER » : un dispositif pour prévenir les difficultés scolaires. *Revue Française de Pédagogie*, 193, 57-76.