



Accélérer la manière dont nous apprenons à améliorer

Accelerating How We Learn to Improve

Anthony S. Bryk



Édition électronique

URL : <http://journals.openedition.org/educationdidactique/2796>

DOI : 10.4000/educationdidactique.2796

ISSN : 2111-4838

Éditeur

Presses universitaires de Rennes

Édition imprimée

Date de publication : 6 décembre 2017

Pagination : 11-29

ISBN : 978-2-7535-7318-5

ISSN : 1956-3485

Référence électronique

Anthony S. Bryk, « Accélérer la manière dont nous apprenons à améliorer », *Éducation et didactique* [En ligne], 11-2 | 2017, mis en ligne le 06 décembre 2017, consulté le 19 avril 2019. URL : <http://journals.openedition.org/educationdidactique/2796> ; DOI : 10.4000/educationdidactique.2796

ACCÉLÉRER LA MANIÈRE DONT NOUS APPRENONS À AMÉLIORER

Anthony S. Bryk

Canergie Foundation for the Advancement of Teaching, Stanford, CA

Un¹ gouffre croît, entre nos aspirations pour nos systèmes éducatifs, qui augmentent rapidement, et ce que les écoles peuvent accomplir au quotidien. L'éducation a besoin d'un paradigme de l'amélioration qui reconnaisse la complexité du travail de l'éducation et la grande variabilité des résultats (*outcomes*) que nos systèmes produisent actuellement. Cet article décrit ce paradigme. Il réunit la discipline de la « science de l'amélioration » avec le pouvoir de communautés structurées en réseau pour accélérer la manière dont nous apprenons à améliorer les capacités à s'améliorer. Ces communautés d'amélioration en réseau (en anglais NIC : *Networked Improvement Community*) combinent la pensée analytique et des méthodes systématiques pour développer et tester, de manière plus fiable, des changements qui peuvent améliorer les résultats. Les NIC sont inclusives dans la mesure où elles rassemblent l'expertise des praticiens, des chercheurs, des concepteurs, des technologues et bien d'autres encore. La manière dont elles organisent leurs activités est proche de celle d'une communauté scientifique. Elles développent des preuves fondées sur la pratique (*practice-based evidence*) comme un complément essentiel aux résultats (*findings*) d'autres formes de recherche en éducation. Il ne s'agit pas seulement de savoir ce qui peut améliorer ou faire empirer les choses; il s'agit de développer le savoir-comment (know how) nécessaire pour améliorer réellement les choses.

Mots-clés : recherche-action, collaboration, politique éducative, réforme éducative, dispositifs expérimentaux, science de l'amélioration, théorie et changement des organisations, résolution de problème, méthodologie de la recherche, utilisation de la recherche.

Accelerating How We Learn to Improve

A chasm is growing between our rapidly rising aspirations for our educational systems and what schools can routinely accomplish. Education needs an improvement paradigm – one that recognizes the complexity of the work of education and the wide variability in outcomes that our systems currently produce. This article sketches out such a paradigm. It joins together the discipline of improvement science with the power of structured networked communities to accelerate learning to improve. These networked improvement communities (NICs) combine analytic thinking and systematic methods to develop and test changes that can achieve better outcomes more reliably. NICs are inclusive in drawing together the expertise of practitioners, researchers, designers, technologists, and many others. And they organize their activities in ways akin to a scientific community. They develop practice-based evidence as an essential complement to findings from other forms of educational research. The point is not just to know what can make things better or worse; it is to develop the knowhow necessary to actually make things better.

Keywords : *action research, collaboration, educational policy, educational reform, experimental design, improvement science, mixed-methods, organization theory/change, problem solving, research methodology, research utilization.*

Remerciements : Nous remercions vivement la petite équipe de traduction de ce texte, composée de Brigitte Gruson, Carole LeHénaff, Gérard Sensevy, et Tracy Bloor (traduction produite à partir d'une première base fournie par Patrick Picard). Aux notes proposées par A. S. Bryk, nous avons ajouté quelques courtes notes proposées par l'équipe de traduction, différenciées des notes du texte par des chiffres romains, et toutes précédées de la mention NDT. Le texte de Bryk s'adresse à un public américain, dans un contexte américain, même s'il est évidemment de portée bien plus large. Sa version originale est consultable ici : [<http://journals.sagepub.com/doi/full/10.3102/0013189X15621543>]

INTRODUCTION

Cet essai présente une argumentation relative à la façon dont les membres de la communauté éducative pourraient accélérer les progrès dans leurs efforts pour améliorer les possibilités d'éducation offertes aux enfants de notre pays². Je commence par définir un peu le contexte.

À la fin de l'automne 2007, j'ai accepté l'invitation du Conseil d'administration de la *Carnegie Foundation for the Advancement of Teaching* à devenir son neuvième président. Au moment de l'élection du président, j'ai partagé avec le conseil d'administration mes points de vue sur les plus grands défis auxquels le domaine de l'éducation est confronté. Depuis plus de deux décennies, les membres de la communauté éducative, les praticiens, les décideurs, les chercheurs et les responsables des communautés locales ont entrepris activement des réformes. Ces efforts se sont manifestés de multiples façons. Ils ont visé, entre autres aspects, à restructurer la gouvernance des écoles publiques, à transformer la formation des enseignants, à élaborer des modèles de réforme des écoles (*comprehensive school*) et à définir des objectifs et une responsabilité (*accountability*ⁱ) plus ambitieux. Bien que ces réformes aient aidé certaines écoles à s'améliorer, le problème était que nos aspirations pour l'école en Amérique augmentaient à un rythme beaucoup plus rapide. Un gouffre (*chasm*) croissait entre ce que nous recherchions et ce que nous pouvions habituellement atteindre, et ce gouffre était plus grand pour nos élèves et étudiants (*students*ⁱⁱ) et nos communautés les plus défavorisés. Il me semblait invraisemblable que nous puissions combler ce gouffre sans une infrastructure de R & D

i. NDT : la notion d'*accountability* est difficile à traduire, en particulier parce qu'elle est polysémique. Nous l'avons en général traduite par « responsabilité ». L'entrée correspondante dans Wikipedia donne une bonne idée de la complexité du terme : [<https://en.wikipedia.org/wiki/Accountability>]. On pourra également lire l'entrée « *Accountability* », rédigée par Christian Maroy et Annelise Voisin, dans le *Dictionnaire de l'Éducation* (Paris, PUF, 2017), sous la direction d'Agnès Van Zanten et Patrick Rayou.

ii. NDT : dans son texte, A. S. Bryk utilise toujours le terme « *Student* », qui peut référer aussi bien à un élève de première primaire (CP en France) qu'à un doctorant. Nous avons traduit, selon le contexte, par « élève », ou « étudiant », quelquefois en utilisant les deux termes.

(Recherche et Développement) plus forte, qui s'attache explicitement à améliorer l'enseignement et l'apprentissage dans les écoles de notre pays et à renforcer les dispositifs institutionnels dans lesquels les enseignants (*educators*) et les élèves (*students*) ont travaillé jusqu'alors.

Nous voulons que nos écoles soient plus efficaces pour beaucoup plus d'élèves et d'étudiants qu'auparavant. Les écoles doivent devenir plus efficaces dans la façon dont elles utilisent leurs ressources, car les contraintes sur le financement public risquent de persister. Et beaucoup d'élèves ou d'étudiants continuent à quitter le système scolaire sans une préparation adéquate pour leur vie à venir. Le fait de progresser de manière significative sur l'un de ces trois objectifs serait une grande réussite³. La demande est maintenant de faire progresser simultanément les trois. La recherche d'une meilleure façon de relever ces défis extraordinaires est la nouvelle mission que s'est donnée la Fondation Carnegie.

MODUS OPERANDI DE L'ÉDUCATION

Au cours des dernières décennies, de nombreuses tentatives de changement ont traversé l'éducation, mais cela s'est fait sans savoir véritablement comment effectuer les améliorations envisagées par les défenseurs des réformes (voire même sans savoir si ces améliorations étaient possibles). Par exemple, lorsque les réformateurs ont visé la réduction des taux élevés de décrochage scolaire et le faible engagement des élèves dans les lycées, un énorme effort a été lancé pour créer de nouveaux lycées de petite taille. Cependant, il existait peu d'éléments relatifs à la façon exacte de transformer de vastes lycées, accueillant un large public et qui présentaient des dysfonctionnements importants, en petits lycées efficaces.

Lorsque les réformateurs ont concentré leur attention sur l'offre limitée de formation continue en matière de développement professionnel, un nouveau rôle organisationnel – le coach éducatif (*instructional coach*) – a été introduit dans les écoles (voir, par exemple, Elmore & Burney, 1997, 1998; Fink & Resnick, 2001). Mais ce que les « coaches » avaient réellement besoin de connaître et de pouvoir faire, de même que les conditions d'organisation requises pour qu'ils puissent mener à bien ce travail, sont restés largement non spécifiés.

Lorsque les réformateurs ont reconnu l'importance de la gouvernance des établissements scolaires (*principal leadership*), des investissements importants ont été faits pour mettre en place des programmes de développement professionnel pour les chefs d'établissements (par exemple Fink & Resnick, 2001). Les chefs d'établissements ont été invités à devenir des leaders pédagogiques, même s'ils étaient déjà soumis à de fortes contraintes temporelles, et peu (ou pas) de modifications ont été proposées pour soulager ces tensions.

Lorsque l'amélioration du système scolaire n'a pas été à la hauteur des attentes des décideurs politiques, une forte responsabilisation (*high-stakes accountability schemes*) a été introduite. Les conséquences en ont été multiples et imprévues. La triche aux tests s'est accélérée, certains élèves ont été ignorés, les programmes de responsabilisation (*accountability*) ayant centré l'attention sur certains étudiants ou élèves, mais pas sur d'autres (par exemple Jacob & Levitt, 2003 ; Office of the Governor, 2011). L'introduction rapide de méthodes à valeur ajoutée pour évaluer les enseignants a commencé bien avant que les propriétés statistiques et les limites de ces méthodes n'aient été correctement comprises⁴. Il n'est pas surprenant qu'une foule de problèmes ait émergé et que des réactions politiques se soient manifestées.

Dans chaque exemple, il y avait eu un véritable problème à résoudre, un début de bonne idée et, dans la plupart des cas, une assise sur la recherche universitaire qui donnait de la crédibilité à la réforme. Toutefois, comme on pouvait s'y attendre, les professionnels de l'éducation (*educators*) ne savaient pas comment mettre en œuvre ces idées. Les districtsⁱⁱⁱ et les États ne possédaient ni l'expertise individuelle ni la capacité organisationnelle pour soutenir ces changements d'échelle. Et il est probable que de nombreux décideurs aient ignoré sans doute le facteur le plus important pour que tout cela fonctionne : le développement de la volonté et du pouvoir d'agir (*agency*) des enseignants et des chefs d'établissements

de notre pays, pour qu'ils participent de manière productive aux efforts de réforme.

En général, la pression exercée pour mettre en œuvre de bonnes idées à grande échelle fournit rarement les résultats promis (pour des compte-rendus développés sur ce point, voir par exemple, Fullan, 2001 ; Tyack & Cuban, 1995). À la base, c'est toujours la même histoire : on met en œuvre rapidement, mais on en tire des lentement leçons. Dans le domaine de l'éducation, nous sous-évaluons l'importance d'avoir recours à des méthodes systématiques et organisées pour apprendre à améliorer (*learning to improve*). Quand un problème urgent se présente, nous nous lançons dans la mise en œuvre de changements politiques ou programmatiques avant de comprendre parfaitement le problème exact à résoudre. Nous nommons ce phénomène le *solutionisme*. Il s'agit d'une forme de réflexion collective dans laquelle un ensemble de croyances se cristallise sur la base d'une analyse incomplète du problème à traiter et sans tenir compte d'autres alternatives potentielles susceptibles de résoudre le problème⁵. En examinant des questions complexes par le petit bout de la lorgnette, le solutionisme a souvent conduit les décideurs à opter pour des stratégies improductives. Au sein d'établissements scolaires sous un contrôle constant, ce à quoi s'ajoutait la pression pour le changement, les enseignants (*educators*) sont devenus des cibles faciles pour diffuser rapidement de telles « solutions » dans notre domaine⁶. Nous sommes déçus lorsque les résultats positifs n'adviennent pas, puis nous passons à la prochaine nouvelle idée. Cela devrait tous nous perturber. Si nous continuons à chercher des améliorations à la manière dont nous l'avons toujours fait, alors il est probable que nous continuions à obtenir ce que nous avons toujours obtenu.

CHERCHER UNE MEILLEURE FAÇON DE FAIRE

Un certain nombre de pistes visant à trouver une approche plus systématique pour améliorer des résultats sont apparues au cours de la dernière décennie en réponse à cette préoccupation. Je reviens brièvement sur trois d'entre elles ci-dessous.

iii. NDT : aux U. S., un *district* réunit des établissements scolaires (collège, lycée) et des écoles géographiquement proches. Il bénéficie d'un statut de relative indépendance politique, avec des pouvoirs similaires à ceux d'une ville ou d'un comté. Voir la page : [https://en.wikipedia.org/wiki/School_district]

La gestion de la performance (Performance Management Thinking)

Sous cette expression générique, on réfère au fait que des objectifs mesurables spécifiques sont articulés pour chaque district, école, classe et différentes sous-populations d'élèves. Les données, généralement sous la forme de résultats obtenus à des tests normalisés, servent à suivre les progrès vers ces objectifs⁷. Les enseignants (*educators*) sont alors tenus responsables (*accountable*) de la réalisation de ces résultats.

Un aspect positif de cette stratégie est que davantage de données sont maintenant disponibles, plus que jamais auparavant, et que l'apparition de nouveaux outils a permis aux enseignants d'examiner ces données avec efficacité. Les enseignants passent maintenant plus de temps à se pencher sur ces informations et les disparités dans les performances⁸ sont abordées avec beaucoup plus de transparence.

Mais la faiblesse de cette approche, c'est qu'elle manque d'une théorie explicite sur la manière dont les pratiques éducatives pourraient réellement s'améliorer. Les enseignants sont informés des objectifs qu'ils doivent atteindre, mais ils sont en grande partie livrés à eux-mêmes pour définir comment réellement les atteindre.

Si vous pensez que le manque d'attention (*focus*) des enseignants constitue le principal obstacle à l'amélioration, cette approche peut être tout à fait pertinente.

Si, *a contrario*, vous considérez que c'est le manque de connaissances sur la façon d'obtenir de meilleurs résultats pour les élèves qui est le problème essentiel, cette stratégie ne dispose pas d'un mécanisme clair pour générer de tels savoir-comment (*know how*) et n'est certainement pas fiable à grande échelle.

La pratique fondée sur des preuves (Evidence-Based Practice)

La figure 1 ci-dessous illustre une deuxième approche, plutôt différente, de l'amélioration systématique. Cette stratégie s'inspire des essais cliniques en médecine et cherche à transférer de telles enquêtes dans le domaine de l'éducation. Sa force tient à l'accent qu'elle met sur une théorie pertinente et à la rigueur empirique mobilisée pour établir

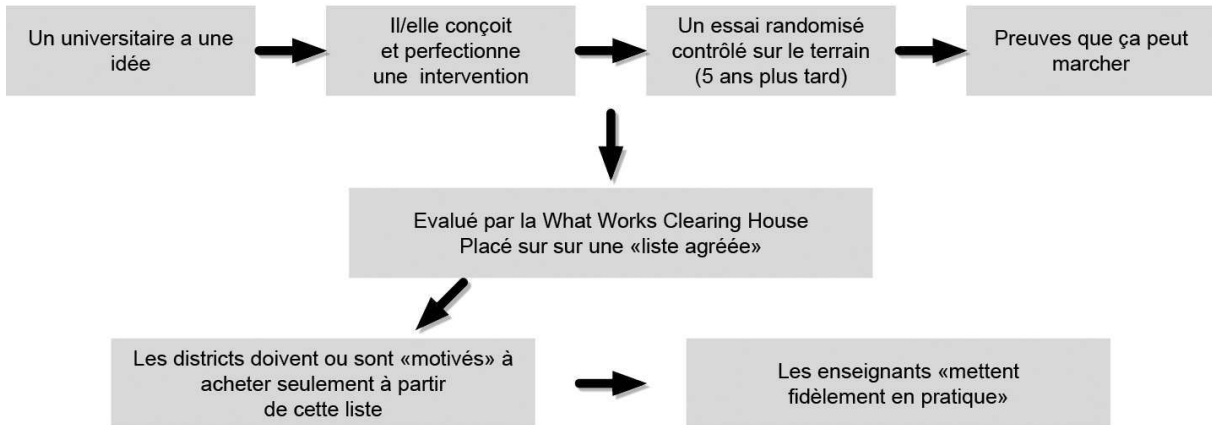
des conclusions quant aux effets de certains projets. Cela représente un antidote vraiment nécessaire dans un domaine où les croyances sont résistantes et les preuves souvent faibles ou inexistantes⁹.

Principalement du fait du leadership de l'Institut des sciences de l'éducation (*Institute of Education Sciences*), nous avons assisté à une expansion considérable de ces études. Environ 1000 entrées concernant les approches fondées sur la preuve (*evidence-based*) existent maintenant dans la base « *What Works Clearing House* » (<https://ies.ed.gov/ncee/wwc/>). Pour mettre cela en perspective, il faut signaler que quelques 27 000 essais cliniques ont été signalés en médecine en seulement un an, en 2010 ; cette base d'enquête a été par ailleurs développée sur plus de cinq décennies supplémentaires. Pourtant, même aujourd'hui, les experts de la santé estiment que 80 % à 90 % de la pratique médicale quotidienne n'est pas ancrée dans de telles preuves, car les informations spécifiques et détaillées dont les praticiens ont besoin n'existent pas encore (Institute of Medicine, 2012). Bien que des essais cliniques vailent certainement la peine d'être menés dans l'éducation, leur processus est très lent et très coûteux. D'un point de vue purement pratique, de telles études ne sont probablement pas non plus une ressource essentielle pour améliorer nos écoles dans un futur proche.

Plus fondamentalement, il est important de reconnaître que les essais cliniques visent une question étroite. Ils nous disent si une expérimentation peut fonctionner. Si un essai sur le terrain produit un effet significatif, cela signifie que l'expérimentation a dû fonctionner quelque part pour quelqu'un. De telles études, cependant, ne sont pas d'abord conçues pour nous dire ce qu'il faudrait faire pour que l'intervention fonctionne pour différents sous-groupes d'élèves et d'enseignants ou dans différents contextes. Ici, au fond, nous sommes confrontés à la différence entre la connaissance que quelque chose peut fonctionner et la connaissance sur la façon de le faire fonctionner de manière fiable sur des contextes et des populations variés. Pourtant, c'est ce que les praticiens veulent généralement savoir : que faudrait-il faire pour que cela fonctionne pour moi, pour mes élèves, dans la situation qui est la mienne ? Malheureusement, les acteurs politiques, qui voient la « pratique fondée sur des preuves » (*evidence-based practice*) comme la réponse à apporter aujourd'hui, manquent généralement cette distinction cruciale.

Figure 1. La stratégie de « Pratique fondée sur la preuve »

Dernière décennie : le mouvement de la pratique fondée sur la preuve



Communautés professionnelles d'apprentissage

L'émergence de communautés d'apprentissage professionnelles, également appelées communautés de pratique, représente une troisième alternative (pour une discussion de ces idées, voir Lave & Wenger, 1991). Celles-ci sont devenues de plus en plus répandues, dans le but de promouvoir la résolution collaborative de problèmes au sein des établissements scolaires, et, plus largement, de relier les enseignants (*educators*) qui partagent des préoccupations communes. Contrairement aux essais sur le terrain, qui ont tendance à se concentrer sur le test de certains nouveaux programmes ou nouvelles politiques, ces enquêtes se focalisent sur des problèmes quotidiens très concrets dans le travail d'enseignement au sein des établissements scolaires. Les enseignants sont alors considérés comme des agents actifs de l'amélioration (*active agents of improvement*) plutôt que comme des récepteurs passifs de connaissances développés par d'autres. Ces communautés cherchent à briser les murs de la pratique solitaire et à créer des espaces sûrs où les personnels (*faculty*) partagent et apprennent les uns des autres.

Par rapport aux essais cliniques, la garantie (*warrant*) réelle, quant à la nature des échanges qui se produisent dans ces conversations, reste moins claire. Le changement visé peut ne pas être bien précisé, et

le succès dépend en grande partie des connaissances tacites personnelles des enseignants (*educators*). Cet état de fait est particulièrement problématique si l'on veut réduire les écarts de qualité (*quality chasm*). En effet, dans de telles communautés de pratique, il n'existe pas de mécanisme formel pour accumuler, détailler davantage, et tester ces connaissances cliniques personnelles afin de pouvoir les transformer au fil du temps en une base de connaissances professionnelles collectivement disponible (sur ce thème, voir Hiebert, Gallimore, & Stigler, 2002 ; Hiebert & Morris, 2009).

Une meilleure voie : la science de l'amélioration dans les communautés en réseau

Chacune de ces stratégies brièvement esquissées apporte une valeur spécifique, mais chacune a également ses défauts. C'est cette observation qui a poussé notre travail à la *Carnegie Foundation for the Advancement of Teaching*. Quelle pourrait être une meilleure voie ? Un paradigme qui fasse, des améliorations mesurables dans les résultats des élèves, son « étoile polaire ». Un paradigme qui intègre les connaissances existantes fondées sur la recherche et s'appuie sur l'orientation analytique et empirique de l'enquête appliquées propre aux sciences sociales. Un paradigme qui accorde également la primauté

aux questions spécifiques auxquelles sont confrontés les enseignants (*practicing educators*) et qui intègre leur manière d'« apprendre-en-faisant » (*learning-by-doing orientation*¹⁰). Tout aussi important, un paradigme qui considère les enseignants comme des enquêteurs actifs liés par des normes et des structures parentes de celles propres à communauté scientifique. Nous avons nommé cette organisation sociale *communauté d'amélioration en réseau* (NIC, *Network Improvement Community*). Ses finalités résident dans la construction de « preuves basées sur la pratique » (*practice-based evidence*¹¹).

Chaque paradigme d'enquête est ancré dans un ensemble d'hypothèses (*assumptions*) relatives aux fondements des phénomènes à propos desquels nous cherchons à apprendre. Par conséquent, je fais un bref détour pour considérer deux caractéristiques centrales des établissements scolaires en tant qu'organisations : la complexité du travail que nous demandons aujourd'hui aux enseignants et, en lien étroit avec cela, la grande variabilité de performance régulièrement constatée¹². Reconnaître et répondre à bon escient à ces deux problèmes – *la complexité des systèmes et la variabilité des performances* – est essentiel pour tout effort visant à améliorer systématiquement la productivité de nos institutions d'enseignement. Les méthodes et l'organisation des NIC prennent leur source dans le travail de ces deux problèmes.

LA COMPLEXITÉ INVISIBLE DANS NOS SYSTÈMES ÉDUCATIFS

Pendant des décennies, la médecine a développé de nouvelles technologies extraordinaires, des médicaments qui sauvent des vies, et des cliniciens bien formés. Atul Gawande (2012b) nous parle de sa spécialité : « en chirurgie, vous ne pourriez avoir de personnes plus spécialisées et vous ne pourriez avoir des personnes mieux formées. Pourtant, nous constatons toujours des niveaux de décès et d'incapacité déraisonnables, qui pourraient être évités. » À l'origine, il y a une énigme moderne pour les professionnels de santé¹³. Leur domaine a plus de connaissances, plus d'outils et plus de ressources que jamais auparavant, mais il reste apparemment incapable de faire fonctionner tout cela de manière fiable pour améliorer la vie (*for advancing human betterment*). Gawande souligne qu'« ajouter plus

de pièces » (c'est-à-dire toute combinaison de plus de personnes, de nouveaux outils et de ressources matérielles supplémentaires), même de manière importante, n'assure pas un meilleur résultat. Au contraire, il soutient que nous devons veiller à ce que ces ressources soient utilisées de manière productive en synergie par les personnes chargées de la réalisation de ce travail, et pour celles qu'ils cherchent à servir. En bref, nous devons améliorer les systèmes de service.

Réfléchir sur ce gouffre de la non-qualité (*quality chasm*) en médecine est instructif (Institute of Medicine, 2001). Comme mentionné précédemment, la médecine bénéficie d'une science puissante, elle investit massivement dans le développement des personnes, qui peuvent s'appuyer sur des technologies extraordinaires. Il n'est pas surprenant que les réformateurs de l'éducation considèrent ces trois facteurs comme essentiels pour améliorer le système éducatif américain. Et en effet, nous avons besoin d'une base de connaissances plus forte, de meilleurs programmes de formation d'enseignants et d'une utilisation plus efficace de la technologie pour favoriser l'apprentissage des élèves. Pourtant, comme en médecine, les développements opérés selon ces seules lignes ne sont pas susceptibles de corriger les résultats scolaires insatisfaisants que nous constatons aujourd'hui.

L'école à classe unique était une entreprise relativement simple : un enseignant instruisait des enfants de différentes classes d'âge, peut-être avec un tableau noir et des textes simples. Les aspirations étaient modestes : acquérir des compétences de base en lecture, certaines compétences en arithmétique et la capacité à écrire son nom (pour une description historique plus détaillée, voir par exemple Graham, 2005). Aujourd'hui, nous attendons des élèves qu'ils maîtrisent bien davantage de contenu académique. Les textes littéraires précédemment enseignés à l'université, par exemple, le sont aujourd'hui au lycée. Les élèves et les étudiants étudient maintenant plus de domaines spécialisés dans les cours de mathématiques et de sciences qu'ils ne l'ont jamais fait. De plus, nous aspirons à ce que tous les élèves n'acquiescent pas seulement des compétences de base, mais se familiarisent plus en profondeur avec les grandes idées qui organisent notre univers de connaissance en explosion (par exemple, William & Flora Hewlett Foundation, 2010). Ces attentes académiques, toujours croissantes, augmentent

considérablement les exigences pour les enseignants, à la fois dans leur connaissance du contenu et dans leurs capacités à enseigner.

Le travail des enseignants devient également plus complexe alors que les classes deviennent plus hétérogènes. Nous nous attendons à ce que tous les enseignants, aujourd'hui, quel que soit le sujet enseigné, prennent en compte les cultures des élèves et leurs langues, différentes des leurs. Dans le même ordre d'idées, les réformateurs ont produit des dispositifs pédagogiques (*instructional initiatives*), tels que la non-répartition des élèves par niveaux (*detracking*), qui rendent les classes plus hétérogènes si l'on considère la formation des élèves, leurs aptitudes, de leurs connaissances antérieures. Ceci revêt une grande importance : la tâche qui consiste à gérer des classes hétérogènes est un défi clé pour l'enseignement (Cohen, 2011).

De même, les enseignants, les chefs d'établissement et les autres responsables scolaires locaux interagissent aujourd'hui avec de nombreuses organisations de services externes. Afin de combler les manques dans la réussite des élèves (*student achievement*), de nombreux programmes supplémentaires ont vu le jour, dont le tutorat individuel, diverses formes d'intervention en petits groupes, et des programmes après l'école et pendant l'été. Les coaches d'enseignement (*instructional coaches*), dans des domaines tels que les mathématiques, les sciences, la littérature, et l'analyse de données relatives aux élèves, sont de plus en plus nombreux à investir les établissements scolaires. Peut-être les plus exigeants (*demanding*) de tous ces dispositifs renvoient-ils aux activités concernant les plans personnels d'éducation requis par l'éducation spéciale, aux niveaux de l'état et au niveau fédéral. Ces dispositifs ont amené un nombre croissant de spécialistes dans les écoles : des spécialistes des troubles de la parole, de l'audition, des difficultés d'apprentissage, des conseillers pour le comportement (*behavioral counsellors*), etc. En principe, tous ces efforts visent à soutenir l'apprentissage des élèves, mais bien les intégrer soulève un grand nombre de problèmes, souvent ignorés.

Un fait particulièrement significatif : la densité de cette activité est plus importante dans les écoles les plus défavorisées de nos pays. Ces écoles sont doublement mises au défi de faire face aux vastes besoins humains de leurs populations scolaires et familiales tout en ayant à gérer un grand nombre de spécialistes, de programmes et d'organisations

externes ciblant leurs élèves. En résulte, pour ces écoles, une augmentation de la complexité organisationnelle qui tend, cependant, à ne pas être reconnue¹⁴.

Enfin, les enseignants font face à une explosion de la connaissance professionnelle. On nous dit aujourd'hui, par exemple, de prendre en compte les multiples formes d'intelligence, les différences dans les styles d'apprentissage des élèves, l'importance des communautés locales, les résultats de la recherche sur le cerveau et des sciences cognitives quant à la façon dont on apprend, et les études socio-psychologiques relatives à l'importance des concepts tels que la résilience, la détermination (*grit*), la persistance dans l'engagement, et les conceptions des élèves. De plus, la liste des nouvelles idées continue de croître à un rythme vertigineux.

En somme, un certain nombre de forces se sont unies pour rendre plus complexe le travail des enseignants et l'organisation des écoles. Comme dans les systèmes de santé, peu de professionnels peuvent répondre efficacement à toute cette complexité dans leur travail quotidien. Aussi il n'est pas surprenant qu'un gouffre existe entre ce que nous cherchons à accomplir et ce que nous atteignons réellement.

Ces observations suggèrent que nous accordions plus d'attention aux tâches que les enseignants accomplissent et aux environnements organisationnels qui façonnent la manière dont ce travail est mené. Plutôt que de laisser croire que la voie vers l'amélioration des résultats consiste à ajouter continuellement de nouveaux programmes (de fait « plus de pièces »), cette perspective nous encourage à nous concentrer d'abord sur l'amélioration de notre compréhension des systèmes de travail qui créent des résultats insatisfaisants. Car c'est dans cette capacité à *voir le système* que les progrès significatifs peuvent s'établir.

Un exemple des Écoles publiques de Baltimore (BPS) illustrera comment *voir* cette complexité opérationnelle peut conduire à un processus d'action très différent. BPS étaient sur le point d'introduire un nouveau processus d'évaluation des enseignants, comprenant notamment un protocole pour que les chefs d'établissement observent et évaluent l'efficacité de chaque enseignant. Ces données devaient informer les décisions prises par le personnel, mais le district voulait les utiliser de surcroît pour améliorer l'enseignement. Classiquement, pour cela, une équipe de l'administration centrale aurait développé

quelque nouvelle procédure pour que les chefs d'établissement fournissent des commentaires à propos des enseignants, à partir des données recueillies. Le district aurait ensuite mené une séance de formation pour les chefs d'établissements, centrée sur l'utilisation du protocole, et leur aurait demandé de l'appliquer, plusieurs fois par an, à chaque enseignant.

Jarrold Bolte, ancien directeur du développement professionnel des enseignants pour les écoles publiques de Baltimore, a eu une idée différente. S'appuyant sur un principe de base de la science de l'amélioration – être centré sur l'utilisateur – Bolte et son équipe ont demandé à un petit groupe de nouveaux enseignants de garder une trace des retours de ceux qui les visitaient pour fournir des conseils et des feedbacks, sur une période de deux semaines. On s'est alors aperçu que certains nouveaux enseignants recevaient déjà beaucoup de feedbacks – dans

certains cas, à partir de 10 sources différentes ou plus (voir la Fig. 2).

Dans le même temps, d'autres nouveaux enseignants n'en recevaient quasiment pas. Donc, Baltimore avait deux problèmes tout à fait différents à résoudre. Pour certains enseignants, le problème tenait au trop grand nombre de feedbacks fournis par trop de personnes différentes – des feedbacks non coordonnés, souvent incohérents, et parfois en contradiction directe les uns avec les autres. Pour d'autres enseignants, le problème résidait dans le manque de feedbacks. Le défi était donc de savoir comment s'assurer que chaque nouvel enseignant recevrait des conseils réguliers qui pourraient réellement l'aider à s'améliorer.

En bref, l'éclair de compréhension (*aha moment*) consista à reconnaître que l'organisation de BPS se caractérisait par un système de processus, déjà en

Figure 2.

De nombreux nouveaux professeurs font l'expérience d'une tempête de conseils (*a blizzard of guidance*)

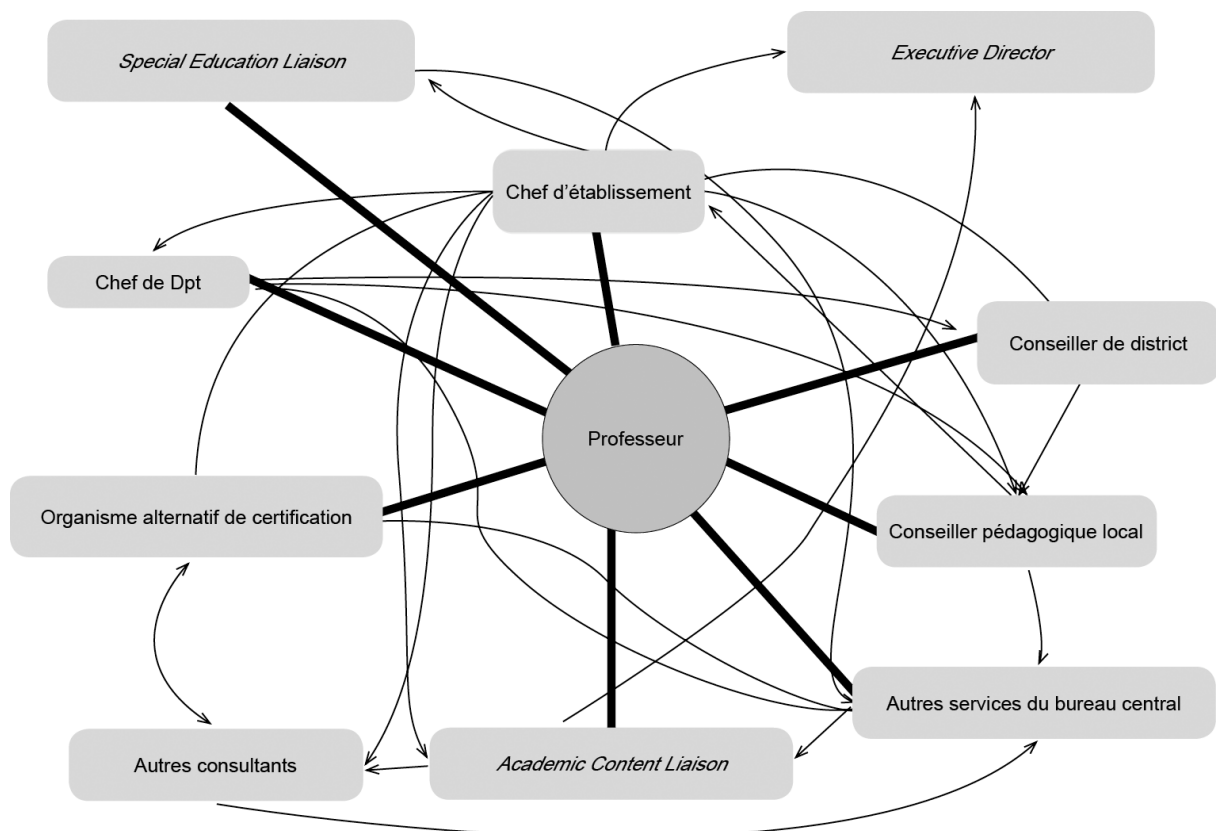


FIG. 2. – Repris de Learning to Improve : How America's Schools Can Get Better at Getting Better (p. 29), by Bryk, Gomez, Grunow, and LeMahieu, 2015, Cambridge, MA : Harvard Education Press. Copyright 2015 by the President and Fellows of Harvard College. Repris avec permission.

fonction, complexe et hautement varié, qui devait être amélioré. On relevait des problèmes importants, et ignorés, de coordination et de communication, ainsi que des opportunités manquées pour améliorer l'efficacité et l'efficience. Assurément, la dernière chose à faire dans un cas comme celui-ci est d'ajouter au système encore un autre processus déconnecté.

Comment peut-on en arriver là? Évidemment, personne n'a consciemment conçu le système pour fonctionner de cette façon. Il s'agit plutôt du résultat naturel de réformes séparées, introduites au fil du temps par différents acteurs institutionnels¹⁵. Par exemple, les programmes incitatifs de l'État et de district possèdent des mentors qui visitent régulièrement les classes des nouveaux enseignants. Si le nouveau professeur est un membre de Teach For America (TFA), il voit également un coach de TFA toutes les deux semaines (les autres programmes alternatifs de formation des enseignants et certains instituts universitaires de formation (*college education*) fournissent un soutien similaire). Ce mentorat est superposé aux réformes antérieures par le Bureau d'enseignement du district, qui a par ailleurs introduit des coachs d'enseignement en littérature, en mathématiques et dans d'autres domaines. Le Bureau de l'apprentissage de la langue anglaise et le Département de l'éducation spécialisée fournissent également d'autres aides. Le Bureau d'instruction a également mis en place des cycles de formation (*school-based learning cycles*) au sein desquels les enseignants sont incités à utiliser les résultats des tests en cours des élèves pour cibler des changements immédiats dans l'enseignement¹⁶. Un nouvel enseignant peut également avoir eu un collègue ou un mentor informel. Et tout cela s'ajoute aux guides des enseignants qui accompagnent les collections de manuels scolaires du district (qui peuvent être en très grand nombre). Au-delà de cet ensemble, il faut prendre en compte une vaste gamme de ressources concernant la formation professionnelle, disponibles sur le portail en ligne du district et plus généralement sur Internet.

Toutes ces aides ont été introduites avec de bonnes intentions – pour aider les enseignants à s'améliorer. Personne, cependant, n'a vraiment porté son attention sur le système de travail global ainsi créé. Plus précisément, personne n'a enquêté sur l'expérience effective que les professeurs faisaient de tout ce qui était censé les aider. En réalité, des problèmes de ce genre ne sont pas uniques à l'éducation. Depuis

plus de 60 ans, on travaille pour améliorer la qualité dans d'autres contextes de travail complexes. Les réalisations de la science de l'amélioration (*improvement science*) dans ces autres contextes fournissent des raisons de croire que ces mêmes principes et méthodes peuvent aider les enseignants (*educators*) à faire de même.

LA VARIATION DE LA PERFORMANCE EN TANT QUE PROBLÈME À RÉSOUDRE

Comme le paléontologue américain Steven Jay Gould (1985) l'écrivait un jour : « La variation elle-même est la seule essence irréductible de la nature. La variation est la réalité même, et non un ensemble de mesures imparfaites pour une tendance centrale. » L'intuition de Gould désigne une deuxième préoccupation qui guide les efforts des communautés d'amélioration en réseau : la tâche et la complexité organisationnelle engendrent une grande variabilité dans la performance. Les résultats obtenus ressemblent souvent à la courbe en cloche familière, illustrée dans la partie gauche de la figure ci-contre (Gawande, 2012a). La recherche pour l'amélioration vise à remodeler cette distribution, à faire avancer la performance globale vers la droite tout en réduisant considérablement la fréquence des résultats négatifs. Ceci est illustré dans la partie droite de la figure 3.

Par conséquent, la *compréhension de la variation* des résultats scolaires (*educational outcomes*) et la réponse efficace à cette variation sont les principaux objectifs de la recherche pour l'amélioration. La science de l'amélioration offre un ensemble de principes et de méthodes très systématiques, mais aussi réellement pratiques, pour faire progresser cet apprentissage. Pour souligner la nécessité de telles enquêtes rigoureuses, considérons un autre exemple éducatif : un essai, sur le terrain, à grande échelle, du programme « *Reading Recovery* » (May et al., 2013).

« *Reading Recovery* » cible les élèves de cours préparatoire (première primaire) qui risquent fort de ne pas apprendre à lire. Il s'agit d'une intervention intensive, individualisée, fruit de plus de 30 ans de recherche et de pratique. Le programme comprend environ 50 séances, d'environ 30 minutes, avec un enseignant spécialement formé.

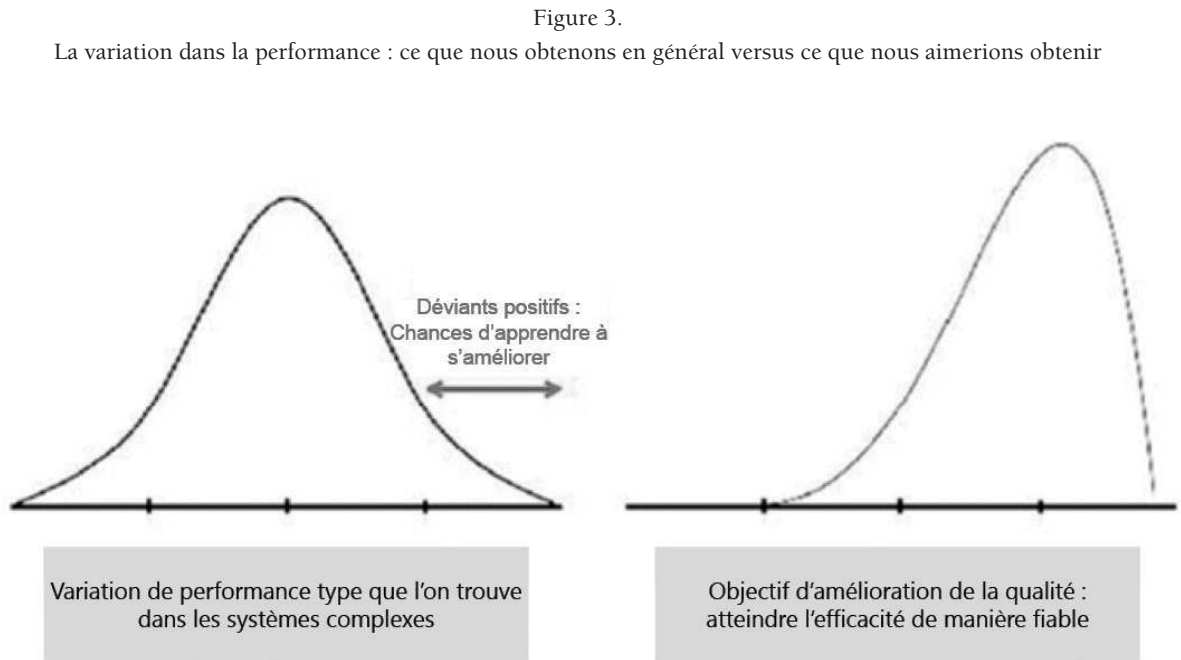


FIG. 3. – Repris de *Learning to Improve : How America's Schools Can Get Better at Getting Better* (p. 55), by Bryk, Gomez, and LeMahieu, 2015, Cambridge, MA : Harvard Education Press. Copyright 2015 by the President and Fellows of Harvard College. Repris avec permission.

Figure 4.
Résultats de l'année 1 de l'essai sur le terrain (May et al., 2013)

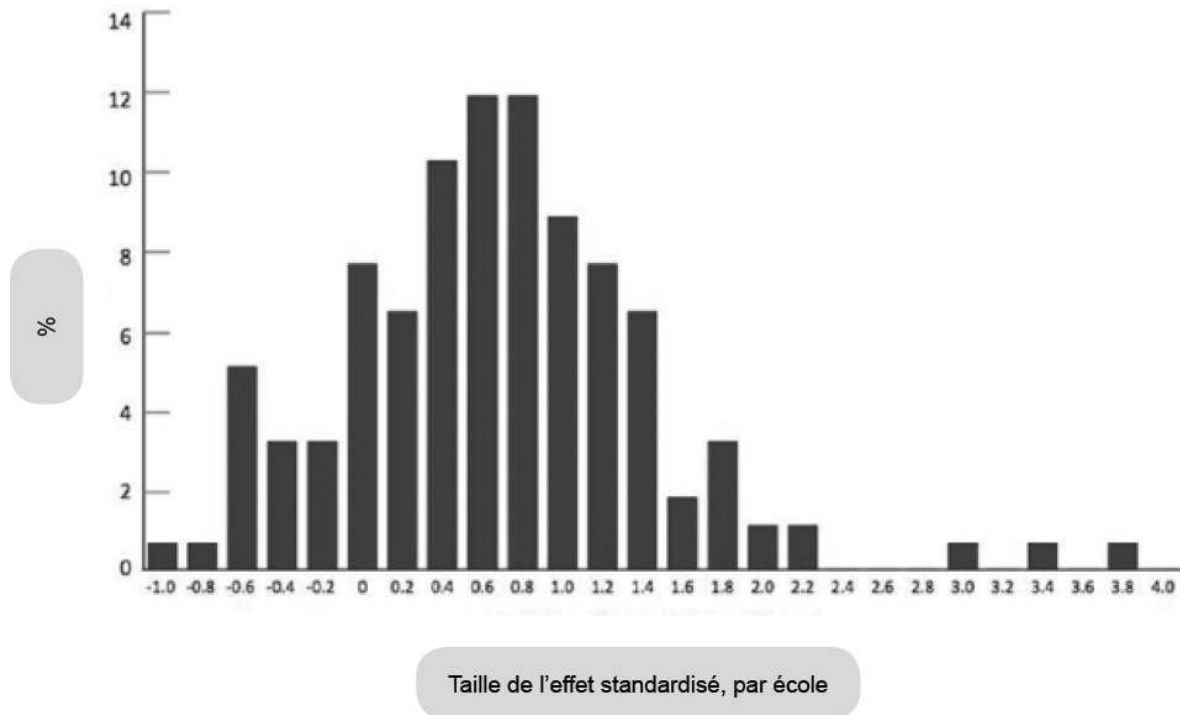


FIG. 4. – Évaluation des impacts, par école, de « Reading recovery ».

En moyenne, la taille de l'effet est assez importante, 0,70. Mais tout aussi importante est la variabilité de ces effets d'une école à l'autre (comme l'étude a été conçue en tant que multisite, cette variation est clairement visible. Dans les essais randomisés en grappes, en revanche, elle serait moins visible). Près de 20 % des écoles avaient des effets négatifs, nuls ou faibles (moins de 0,2). À l'autre extrémité de la distribution, 20 % des écoles avaient des tailles d'effet égales ou supérieures à 1,2. Pour le dire autrement, la variation des effets entre les écoles (c'est-à-dire l'écart-type des effets au niveau des écoles) est presque aussi large que l'effet moyen.

Comprendre les contours de cette variation et les facteurs probables qui y contribuent, c'est essentiel pour la réalisation de meilleurs résultats, plus fiables à grande échelle. L'expérimentation a clairement échoué dans un certain nombre d'écoles. Nous devons examiner ces cas problématiques pour comprendre les facteurs en cause. Le programme fonctionnerait-il mieux, par exemple, pour certains types d'élèves, mais pas pour d'autres ? La variabilité dans la préparation et/ou la compétence des enseignants du programme « *Reading Recovery* » serait-elle déterminante ? Peut-être le programme ne s'est-il pas bien intégré à l'enseignement dans ces écoles ? Ou bien est-ce qu'il y avait un problème logistique dans certaines écoles, de sorte que les élèves-cibles n'ont tout simplement pas bénéficié des interventions ? Et si ce dernier cas était avéré, alors, pourquoi cela s'est-il produit ? L'étude de ces sources possibles de variabilité – pourquoi des effets nuls peuvent-ils se produire ? – peut aider à identifier les points sur lesquels les efforts d'amélioration ultérieurs pourraient se concentrer. De telles enquêtes peuvent également identifier les conditions imites concernant les lieux de l'expérimentation (ou les sous-groupes d'élèves) où celle-ci ne devrait pas être implémentée.

À l'inverse, des effets très intéressants, très importants, ont été accumulés dans d'autres sites de « *Reading Recovery* ». Peut-il y avoir quelque chose que nous pourrions apprendre de ces « déviants positifs » (Pascale, Sternin & Sternin, 2010), qui puisse renforcer l'intervention pour d'autres élèves dans le futur ? En bref, comprendre les résultats négatifs obtenus permet de clarifier les problèmes nécessitant encore des solutions ; l'examen de « déviants positifs » peut aider à établir des hypothèses relatives à des changements possibles qui méritent des tests futurs¹⁷.

Il est important de garder à l'esprit que chacune des méthodes de recherche que nous utilisons aujourd'hui couramment est, originellement, spécifique à un certain problème dans un certain contexte. Herbert Weisberg (2014), dans son livre *Willful Ignorance*, dans lequel il expose en détail l'histoire de la théorie des probabilités, prend l'exemple de l'émergence et l'utilisation des essais randomisés dans l'agriculture. Le dilemme pratique auquel se heurtaient les experts agricoles était que le rendement des cultures était soumis à de nombreux facteurs. Il était difficile, par exemple, de déterminer si l'introduction d'engrais commercial augmentait la production globale (c.-à-d. « en moyenne ») des cultures concernées. La force de la randomisation, c'est qu'elle permet aux chercheurs de discerner l'effet produit par l'ajout d'un nouveau facteur, comme un engrais, en ignorant sciemment l'influence de tous les autres facteurs qui contribuent à la variabilité observée. Cependant, lorsque nous déplaçons notre attention vers la façon d'obtenir des résultats de qualité pour chaque élève et dans tous les contextes, les détails relatifs à « tous les autres facteurs » et à la façon dont ces facteurs peuvent à leur tour *interagir* avec une nouvelle expérimentation (*intervention*) sont vraiment importants. La clé ici est la distinction entre les mots *interaction* et *ajout*. *L'interaction* est une propriété constitutive d'un système. *L'addition* ne l'est pas.

Mon point de vue est simple : les essais randomisés sur le terrain ne sont pas principalement conçus dans le but de nous donner à voir comment il faudrait faire fonctionner les expérimentations de manière fiable pour différents sous-groupes d'élèves ou d'étudiants et de professeurs, travaillant dans des conditions contextuelles variées. La science de l'amélioration, en revanche, accorde la primauté à la variabilité des résultats, en tant que problème central à résoudre. Elle met l'accent sur la façon dont les facteurs relatifs aux tâches et à l'organisation se combinent pour créer cette variabilité.

En termes plus généraux, pour aboutir à des résultats de qualité, de manière fiable, dans des systèmes complexes, tels que les établissements scolaires, il est nécessaire d'adopter des méthodes d'enquête spécifiques¹⁸. L'utilisation d'évaluations itératives, pour guider le développement, est privilégiée, ainsi que la révision et le perfectionnement continu de nouveaux outils et processus, de nouveaux rôles dans le travail, et de nouvelles

relations. Il faut également prendre en compte le fait que souvent, nous ne savons pas si un changement que nous introduisons entraînera réellement une amélioration ou s'il peut aussi engendrer des conséquences négatives non prévues. Une telle ambiguïté est intrinsèque aux systèmes adaptatifs complexes¹⁹. Aussi, cette caractéristique constitutive implique nécessairement de s'orienter vers une approche basée sur l'apprendre-en-faisant (*a learning-by-doing orientation*). Cela nous incite à des débuts d'expérimentation à petite échelle, au cours desquels nous apprenons rapidement, et qui minimisent donc les dommages éventuels en cas d'erreur. Les idées de changement sont continuellement raffinées, ce qui se produit réellement faisant preuve, et progressivement mise en œuvre dans des situations de plus en plus diversifiées.

Fait intéressant, nous apprenons le plus quand une modification ne produit pas les résultats escomptés ou entraîne des conséquences imprévues. Cela nous oblige à une réflexion plus profonde sur le système que nous tentons d'améliorer (« Quelque chose nous échappe ici, qu'est-ce que ça pourrait bien être ? ») Avec les cycles ultérieurs de redéfinition et d'évaluation (*test*), la compréhension du (ou des) problème(s) effectif(s) évolue, et des changements productifs sont davantage susceptibles d'apparaître. Par conséquent, il est judicieux de penser ces enquêtes comme des cycles d'*apprentissage de l'amélioration*.

UN ACCENT SUR L'ÉLABORATION DE PREUVES BASÉES SUR LA PRATIQUE

Au fur et à mesure que le travail devient plus complexe, les travailleurs doivent davantage compter sur de bons processus (*good processes*), des matériaux bien conçus et des normes explicites en matière de communication (Hutchins, 1995). Dans la mesure où au moins certains aspects du travail peuvent être conçus en tant que routines, les organisations créent des conditions plus propices à l'obtention fiable de résultats de qualité. Ces routines permettent à l'attention humaine de se concentrer plus finement sur des conditions spécifiques à la situation immédiate²⁰. En d'autres termes : en l'absence de bonnes routines, la très grande variabilité des résultats est inévitable. Ce constat plaide aujourd'hui pour la nécessité de porter plus d'attention à la recherche sur la formation des

enseignants, par exemple en fournissant des descriptions détaillées des « pratiques d'enseignement efficaces » (« *high-leverage teaching practices* ») pour des domaines de contenus et à des niveaux spécifiques de la scolarité.

De même, à mesure que les organisations deviennent plus spécialisées et plus complexes, les problèmes de communication et de coordination abondent. Les efforts de BPS (*Baltimore Public School*) pour améliorer les feedbacks fournis aux enseignants, que nous avons discutés ci-dessus, l'illustrent avec force. Lorsque les activités éducatives sont peu coordonnées, certains élèves peuvent « passer à travers le système » lorsqu'ils changent d'école, ou lorsqu'ils vont de l'enseignement général à l'enseignement spécialisé. Des problèmes de communication peuvent également survenir lorsque les élèves et les enseignants interagissent avec des conseillers, des travailleurs sociaux, des membres de la famille des élèves, et d'autres personnes à l'intérieur et à l'extérieur de l'école. Face à une telle complexité organisationnelle, la recherche pour l'amélioration vise à synchroniser les efforts de divers individus et organisations, dans le but de prévenir les difficultés et d'augmenter la probabilité d'obtenir de meilleurs résultats.

Pour être clair, afin de traiter avec efficacité la question de la complexité des tâches et des organisations, il ne s'agit pas d'imposer des normes apparemment arbitraires, venues de l'extérieur ou d'en haut. Bien plutôt, cela implique la constitution d'une communauté d'amélioration de l'éducation, qui crée activement ces normes, en appui sur ses enquêtes rigoureuses. Les praticiens doivent s'engager pleinement avec les chercheurs, mais aussi d'autres acteurs de l'éducation, dans le développement, l'évaluation et l'amélioration du travail clinique de l'enseignement scolaire (*the clinical work of schooling*).

Revenons à l'observation de Gawande (2012b) sur la pratique médicale : l'obtention de meilleurs résultats à grande échelle, signifie que « toutes les pièces » fonctionnent conjointement de manière plus constructive. La façon dont les différents processus de travail, les outils, les relations de rôles et les normes interagissent de manière plus productive dans des conditions variées est l'objectif d'apprentissage principal de la communauté. Les connaissances qui en résultent et les garanties empiriques associées sont des *preuves basées sur la pratique* (*practice-based evidence*). La recherche pour l'amélioration, ainsi pensée, serait utile à *Reading Recovery*, par exemple, pour éliminer

les éventuels effets nuls ou négatifs et concentrer la distribution globale des résultats encore plus loin sur la partie droite du graphique (cf. Fig. 3).

C'est la raison pour laquelle l'organisation sociale d'une communauté d'amélioration en réseau est primordiale. Les nombreuses facettes qui peuvent contribuer à la variabilité des résultats scolaires (sur le plan de la tâche et de la complexité organisationnelle dont il a été question précédemment) transforment la question de l'amélioration en un problème d'apprentissage social à grande échelle. En outre, bon nombre des expérimentations que nous cherchons à introduire dans les écoles sont elles-mêmes des systèmes complexes, constitués d'un ensemble interdépendant de principes, de processus, d'outils, de rôles de travail et de normes. Par conséquent, l'introduction de changements productifs dans divers contextes implique des efforts coordonnés et disciplinés auprès d'un grand nombre de personnes et de lieux. Les développements récents dans la science en réseau éclairent la façon dont les réseaux structurés de manière consciente, impliquant de nombreux contributeurs différents, peuvent accélérer les progrès accomplis pour des problèmes complexes (Nielsen, 2012). Les participants à ces réseaux partagent une même théorie en travail. Ils s'appuient sur des mesures communes de l'activité des réseaux, et ils déploient des méthodes d'enquête et des mécanismes de communication pour ancrer la résolution collective de problèmes. Ils établissent des processus relatifs à la manière dont les individus travaillent, et ils établissent des normes de preuve (*evidentiary standards*) pour garantir des affirmations d'efficacité. En travaillant ensemble de cette manière, ils peuvent apprendre plus rapidement comment s'améliorer.

Malheureusement, aucune infrastructure professionnelle n'existe actuellement afin que les enseignants puissent collaborer au développement et à l'évaluation systématique des changements, et à produire et synthétiser des preuves fondées sur la pratique. Ce serait pourtant possible. On pourrait envisager des réseaux nationaux d'enseignants et d'écoles, engagés avec des chercheurs et des développeurs de programmes (*program developers*^{iv}) autour de certains problèmes éducatifs cruciaux

(*high-leverage problems*). Ces réseaux viseraient à faire connaître aux enseignants (*educators*) ce qui est le mieux susceptible de fonctionner, dans quels lieux, où, pour qui, et dans quelles conditions. En outre, comme les enseignants utiliseraient cette connaissance, la connaissance elle-même évoluerait et serait affinée par les applications qui en seraient faites.

UN PARADIGME DE L'AMÉLIORATION

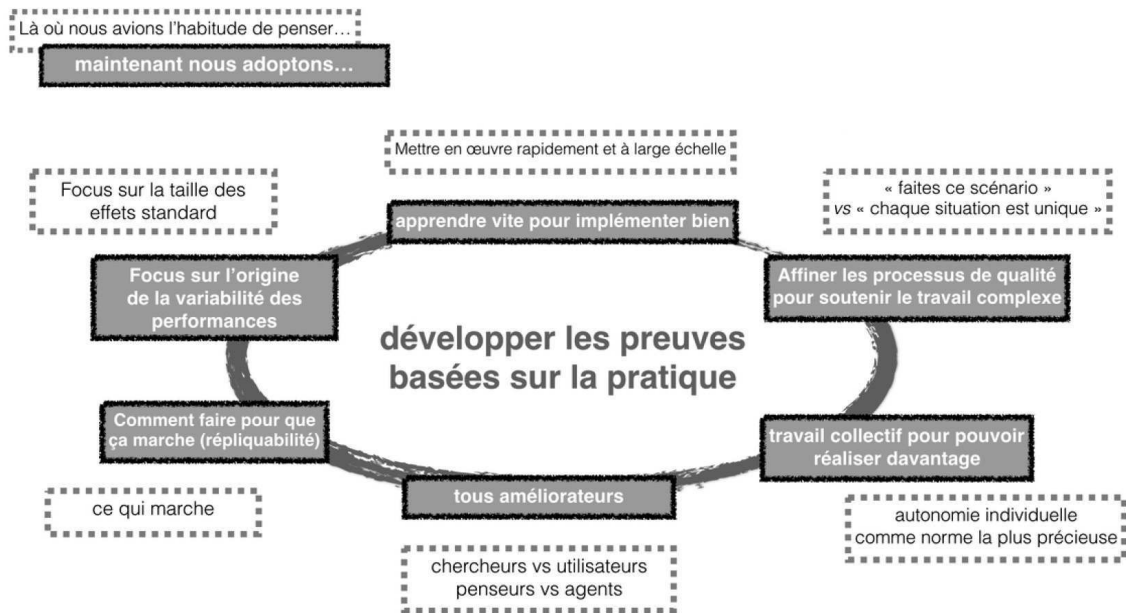
La notion de recherche pour l'amélioration menée par des communautés en réseau vise à accélérer notre connaissance des phénomènes complexes qui engendrent des résultats insatisfaisants. Cette activité de recherche prend forme autour d'un ensemble intégré de principes, de méthodes, de normes organisationnelles et de structures. Elle constitue un ensemble cohérent d'idées sur la façon dont les enquêtes pratiques doivent être envisagées et réalisées. D'une manière formelle, c'est la définition d'un paradigme. Lorsqu'on y adhère, ce paradigme nous encourage à réfléchir et à agir de manière très différente. La figure 5 donne à voir un guide heuristique pour ce paradigme de l'amélioration.

Dans le passé, l'urgence du changement a obligé les réformateurs de l'éducation à mettre en œuvre rapidement des expérimentations à grande échelle. Ce faisant, ils ont sous-estimé considérablement le savoir-comment nécessaire qui devait être développé et le temps nécessaire pour réussir. L'enjeu d'un paradigme de l'amélioration est de pouvoir apprendre rapidement, pour mettre en œuvre des changements, adéquatement, dans une variété de conditions et de contextes. Le fait de continuer à ignorer ce principe, issu des expériences menées dans de nombreuses industries et d'autres secteurs, enverra encore des idées, apparemment bonnes, au pilon sans cesse croissant des promesses non tenues (pour un exemple évocateur de cela dans le contexte de l'armée, voir La discussion de Sutton & Rao [2014] sur « *clusterfug* », p. 24-27).

Là où les chercheurs et les analystes se sont traditionnellement concentrés sur l'estimation précise de la valeur d'un effet – l'écart moyen de résultats entre un groupe ayant bénéficié d'une expérimentation et un groupe de contrôle pour lequel cela n'a pas été le cas – la science de l'amélioration nous permet de contrôler la multitude de facteurs qui contribuent à la variabilité des résultats, et d'utiliser ces connais-

iv. NDT : le sens du terme anglais « *program* » est ici proche du sens français de dispositif, ou d'ingénierie. *Reading Recovery*, auquel A. S. Bryk se réfère dans l'article, est ainsi un tel « programme ».

Figure 5.
Le paradigme de l'amélioration



Educational Researcher, Vol. 44 No. 9, pp. 467-477
DOI: 10.3102/0013189X15621543
© 2015 AERA. <http://er.aera.net>

Fig. 5. – Repris de Learning to Improve : How America’s Schools Can Get Better at Getting Better (p. 186), par Bryk Gomez, Grunow et LeMahieu, 2015, Cambridge, MA : Harvard Education Press. Copyright 2015 par le président et collègues du Harvard College. Reproduit avec permission.

sances pour nourrir l'évaluation continue de ces changements. La capacité d'engendrer des résultats positifs sur une base régulière, dans des conditions et des contextes variés, devient alors la référence absolue.

Le paradigme de l'amélioration représente également une opportunité nouvelle de résoudre un dilemme de longue date pour la gouvernance, les politiques et les pratiques éducatives. Les complexités inhérentes à l'enseignement et à l'apprentissage ont conduit de nombreux enseignants à croire que chaque situation est unique et que, par conséquent, chaque enseignant, chaque école et chaque district doit inventer ses propres pratiques. Bien que ce point de vue repose sur une appréciation pertinente des subtilités de l'enseignement, ceux qui souscrivent à cette « solution » acceptent fondamentalement que la grande variabilité des résultats, si l'on considère une échelle suffisamment grande, est inévitable. D'aucuns ont alors plaidé en faveur d'un enseignement organisé sur des scripts (*scripting teaching*) afin que chaque élève ait la même possibilité d'apprendre. Mais chaque élève est différent, et il en de

même pour les contextes. La complexité est réelle, et ne peut pas être réduite par la standardisation de toutes les activités, qui ferait tendre vers des environnements du type « teacher-proof^v ».

Contrairement à ces deux points de vue, un paradigme de l'amélioration considère la tâche et la complexité organisationnelle comme une préoccupation centrale. Il reconnaît les exigences élevées imposées aux enseignants, qui cherchent à prendre en compte chaque élève, la dynamique de chaque classe, et les caractéristiques spécifiques de chaque communauté scolaire. Un paradigme de l'amélioration prend également en considération l'importance, pour les enseignants, de développer des processus, des outils, des rôles de travail et des relations, qui soient spécifiques, et qui étayent un travail de qualité dans ces différents contextes. Cette approche respecte bien les complexités inhérentes à

v. NDT : dispositif dans lesquels les professeurs suivent un protocole d'enseignement très structuré, avec une marge de manœuvre très faible vis-à-vis des modalités de travail et d'évaluation des contenus étudiés.

l'éducation, tout en affirmant que les aspects importants de l'enseignement (*work of schooling*) devraient suivre des protocoles standard. L'élaboration de ces preuves fondées sur la pratique peut aider à combler les polarités décrites ci-dessus.

Rendre opérationnelle la recherche pour l'amélioration suppose également une attention à ce qui est mesuré, et à la manière dont ces mesures sont rassemblées. Le domaine de l'éducation déborde maintenant de données, principalement des données évaluatives (*accountability data*), sur les élèves, les enseignants, les écoles et les districts. En revanche, la recherche pour l'amélioration construit des données dont le but n'est pas de classer les individus ou les organisations, mais d'apprendre comment les pratiques d'enseignement et les processus organisationnels fonctionnent réellement. Nous avons besoin de preuves (*evidences*) pour nous aider à discerner si les essais de changements spécifiques sont effectivement des améliorations. L'introduction de ce type de mesure pratique pose un défi normatif, car, traditionnellement, ces données sont perçues par les professionnels de l'éducation (*educators*) comme étant destinées à quelqu'un d'autre, à une autorité éloignée qui chercherait à les responsabiliser, ou à un chercheur qui les étudierait. Jamais, les professionnels de l'éducation n'ont pu utiliser des données afin de produire des améliorations. Les choses se compliquent encore du fait que le temps est la ressource la plus restreinte en éducation. Un nouveau défi reste à relever, qui consiste à recueillir, afin de soutenir les efforts d'amélioration, suffisamment d'informations, et d'une manière qui soit la moins envahissante possible (pour une discussion de cette idée, voir Yeager, Bryk, Muhich, Hausman & Morales, 2014).

Enfin, un paradigme de l'amélioration nous invite à reconsidérer notre norme éducative la plus appréciée : l'autonomie de la pratique dans les classes et les écoles. Dans le passé, lorsque nous en savions moins, lorsque les enseignants travaillaient dans des environnements plus simples et (peut-être plus significativement) lorsque nos aspirations étaient plus modestes, il était raisonnable de voir l'enseignant comme un artisan individuel. Aujourd'hui, nous avons clairement atteint les limites de ce qui est réalisable en pensant l'enseignement de cette manière. Nos problèmes sont maintenant trop complexes, et les connaissances et l'expertise nécessaires bien trop variées. Bien que, isolément, certains individus

extraordinaires parviennent à construire individuellement une action positive, la qualité de l'enseignement, laissée à ces seules pratiques, sera toujours peu répandue. Heureusement, l'éducation engage des centaines de milliers de personnes faisant un travail très similaire chaque jour. Si les enseignants se regroupaient dans des réseaux d'amélioration structurés, notre champ démontrerait des capacités extraordinaires pour innover, tester et diffuser rapidement des pratiques efficaces (pour une discussion plus détaillée, voir Bryk, Gomez & Grunow, 2010; Bryk, Gomez, Grunow & LeMahieu, 2015).

Le paradigme de l'amélioration nous appelle à activer ces forces. Il nous appelle à écrire un nouveau récit. Travailler ensemble, et apprendre ensemble, permettra d'accomplir beaucoup plus, pour beaucoup plus de personnes, que nous n'avons jamais pu le faire.

De même, la science de l'amélioration implique que la communauté de recherche en éducation endosse un rôle différent. La division entre recherche et pratique est légendaire. Nous avons tacitement accepté qu'il y ait une petite classe d'« experts » (*knowers*) et une classe beaucoup plus grande de « faiseurs » (*doers*), qui devraient simplement utiliser les connaissances produites par d'autres. La recherche pour l'amélioration, en revanche, appelle les personnes des deux classes à travailler ensemble comme « améliorateurs » (*improvers*). En respectant et en valorisant l'expertise variée nécessaire pour résoudre des problèmes d'éducation, les communautés d'amélioration en réseau embrassent toutes les personnes impliquées en tant que membres à part entière.

Pour résumer : le paradigme de l'amélioration associe la discipline de la science de l'amélioration au pouvoir des communautés structurées en réseau pour accélérer la manière dont nous apprenons à améliorer. Le paradigme utilise des méthodes rigoureuses, analytiques et systématiques pour développer et tester les changements qui apportent des améliorations fiables. Il est inclusif dans la mesure où il met en synergie l'expertise des praticiens, des chercheurs, des concepteurs, des technologues et bien d'autres. Ce paradigme organise délibérément ses activités d'amélioration de manière parente avec ce qui se fait au sein d'une communauté scientifique.

Le paradigme esquissé ici vise à mettre à notre portée un objectif fondamental mais insaisissable : la qualité des résultats des élèves (*students*)

quelle que soit l'échelle considérée. Il met l'accent sur l'élaboration de preuves fondées sur la pratique comme un complément essentiel aux résultats d'autres formes de recherche en éducation. Il ne s'agit pas seulement de savoir ce qui peut améliorer ou faire empirer les choses, mais de développer le savoir-comment nécessaire pour rendre les choses réellement meilleures.

Imaginez un avenir dans lequel ce type d'apprentissage de l'amélioration devient quotidien, dans des milliers de contextes, et engage des milliers d'enseignants, de chercheurs, de concepteurs et d'innombrables autres acteurs. Le champ de l'éducation pourrait devenir une immense communauté d'amélioration en réseau. Nous pourrions accélérer considérablement la façon dont nous apprenons à améliorer. Nous pourrions atteindre les résultats essentiels auxquels nous aspirons maintenant, mais pour l'atteinte desquels nous n'avons aucune stratégie que nous puissions concrètement accomplir : que tous les enfants aient appris à lire au CE2 (3^e primaire) ; que tous soient prêts pour l'université à la fin du lycée ; que tous obtiennent un diplôme (deux ans ou quatre ans après l'université) ; ou que tous les nouveaux enseignants réussissent à bien éduquer leurs élèves (*in educating their students well*).

Ceci suppose un changement profond (*significant shift*) dans notre pensée et dans les façons dont nous travaillons et apprenons ensemble pour faire advenir tout cela et le faire advenir mieux et plus vite.

Je serais heureux d'explorer ce défi avec vous et que nous embarquions ensemble, comme compagnons de voyage, pour apprendre à améliorer.

NOTES

1. Cet article a été adapté de l'AERA *distinguished lecture* donnée le 6 avril 2014 lors du congrès annuel de l'AERA (*American Educational Research Association*) à Philadelphie, Pennsylvanie. Certaines parties de l'article ont ensuite été réélaborées pour un ouvrage (Anthony S. Bryk et al., 2015). Pour plus d'informations, [<http://hepg.org/hep-home/books/learning-to-improve>]. Le présent article a été adapté avec la permission de Harvard Education Press.

2. En essayant de m'adresser à un large public de professionnels de l'éducation, je m'appuie fortement sur des exemples et j'utilise le langage commun. Bien que de nombreuses citations et références soient fournies dans l'article, celles-ci sont également proposées principalement à titre d'exemples de travaux étroitement liés à mon propos, et portés par beaucoup d'autres. Compte tenu de l'ampleur des problèmes présentés ici et des limites inhérentes à un essai dans *Educational Researcher*, je n'ai pas essayé d'atteindre à l'exhaustivité des références.

3. Nous pourrions également ajouter à cette liste que le moral professionnel est à un niveau exceptionnellement bas. Voir le rapport Gallup (2014), *State of America's Schools*. Au sein d'une entreprise aussi dense en ressources humaines et sociales que l'éducation, cela devrait également figurer sur notre liste de grands problèmes à résoudre.

4. Voir, par exemple, le consensus de critiques sur les méthodes à valeur ajoutée [www.carnegieknowledge-network.org].

5. Ce phénomène n'est pas propre à l'éducation, il est répandu dans d'autres secteurs. C'est une tendance naturelle qui se reflète à la fois dans les actions individuelles et organisationnelles. Sur ce point, voir la discussion dans Pfeffer et Sutton (2000). Voir aussi Pentland (2014).

6. Cette idée est soutenue dans la théorie néo-institutionnelle (voir DiMaggio & Powell, 1983). En l'absence d'un statut professionnel largement reconnu, les institutions affirment leur légitimité de manière symbolique – en se tenant à la fine pointe de la réforme. Une nouvelle idée vaut précisément parce qu'elle est nouvelle, quel que soit le problème, et quelles que soient la présence ou l'absence de toute preuve justifiant les affirmations qu'elle permet de produire.

7. Les objectifs annuels de progrès énoncés par le programme « *No Child Left Behind* » offrent un bon exemple de cette façon de fonctionner.

8. Cela peut sembler courant maintenant, mais cela n'a pas été toujours le cas. Lorsque nous avons commencé le Consortium Chicago School Research en 1990, aucune donnée publique n'était régulièrement disponible quant

aux disparités dans les résultats. Tout le monde savait que ces écarts de réussite existaient, mais il y avait une grande ambivalence quant à leur publication.

9. Les exemples abondent ici également. Voir, par exemple, la réinstitution de la rétention des élèves au sein d'un niveau de classe donné, même si plusieurs études ont signalé des effets nuls ou négatifs. Sur ce dernier point, voir la série de rapports de recherche de Melissa Roderick et ses collègues du Consortium sur Chicago School Research, [<http://ccsr.uchicago.edu/publications/ending-social-promotion>]. Voir aussi Roderick (1994).

10. Pour une bonne introduction à cette orientation basée sur la conception (*design-based*), voir Dolle, Gomez, Russell et Bryk (2014). Pour une discussion plus générale sur la recherche fondée sur l'implémentation de dispositifs (*design-based implementation research*), voir le volume édité par Fishman, Penuel, Allen, Cheng et Sabelli (2013).

11. Le choix de l'expression *preuves fondées sur la pratique* (*practice-based evidence*) est délibéré. Notre but est de signaler une différence clé dans la relation entre l'enquête et l'amélioration, par rapport à ce qui est typiquement supposé dans l'expression plus couramment utilisée de *pratique fondée sur les preuves* (*evidence-based practice*). Il y a une idée implicite dans cette dernière : la preuve de l'efficacité existe quelque part hors de la pratique locale et les praticiens devraient simplement mettre en œuvre ces pratiques fondées sur la preuve. La recherche pour l'amélioration, en revanche, constitue une activité d'apprentissage, continue (*ongoing*) et locale. Nous sommes redevables à Lawrence Green de nous avoir présenté ce concept de *preuves fondée sur la pratique*, et d'avoir explicité comment on peut le rapporter au discours relatif à la *pratique fondée sur les preuves*, tout en le contrastant avec celui-ci. Voir Green (2008) et Barkham et Mellor-Clark (2003).

12. Ceci est reconnu dans la littérature comme un caractère fondamental des systèmes adaptatifs complexes. Voir Axelrod et Cohen (2000) et Miller et Page (2007).

13. En réfléchissant à cette énigme, Gawande (2012b) demande : « Que feriez-vous si vous vous vouliez construire une voiture à partir des meilleures pièces existantes ? Eh bien, cela vous conduirait à mettre des freins Porsche, un moteur Ferrari, un corps Volvo, un châssis BMW. Vous mettriez tout cela ensemble et qu'est-ce que vous obtiendriez ? Un tas de déchets très coûteux qui ne va nulle part. Et c'est ainsi que la pratique médicale peut se concevoir aujourd'hui. »

14. Voir la discussion sur les écoles de « quartiers sensibles » (*truly disadvantaged urban schools*) dans Bryk, Sebring, Allensowrth, Luppescu et Easton (2010, p. 209).

15. Cette description se veut le récit générique d'un phénomène propre aux grands districts scolaires. Ses caractéristiques spécifiques, ramenées à l'expérience d'un enseignant particulier, varieront évidemment à la fois entre les écoles et les districts.

16. De nombreux districts déploient maintenant ces tests intermédiaires, qui sont étroitement alignés sur les tests de responsabilisation à fort enjeu (*high-stake accountability tests*) de fin d'année. Bien qu'ils soient souvent décrits comme des évaluations formatives pour améliorer l'enseignement, certains critiques se réfèrent à ces systèmes comme à « enseigner pour les tests » (*teaching for the tests*).

17. Ironiquement, la « *Reading Recovery* » est l'une des rares interventions éducatives qui ont déjà été examinées et approuvées par le centre de documentation *What Works Clearinghouse* comme solidement garanties (*having a strong evidence base*). Par conséquent, on n'a probablement pas besoin de dépenser 40 millions de dollars (coût approximatif de l'essai sur le terrain) pour répéter cela. Apprendre comment améliorer encore *Reading Recovery* pour obtenir des résultats de qualité plus fiables à grande échelle aurait mérité amplement un investissement en recherche. Ce dernier aurait coûté moins cher et nous aurait probablement appris davantage.

18. Voir Bryk *et al.* (2015) ; pour un traitement plus général, voir aussi Langley *et al.* (2009).

19. Pour plus d'informations sur les systèmes adaptatifs complexes, voir Axelrod et Cohen (2000), et Miller et Page (2007).

20. Les psychologues parlent ici d'automatisme. Une telle automatisme joue un rôle clé dans le développement de l'expertise. C'est une caractéristique clé qui distingue les novices et les experts, et son développement marque une étape clé dans l'apprentissage de l'expertise. Voir Ericsson, Krampe et Tesch-Romer (1993). Voir aussi Flyvbjerg (2001).

RÉFÉRENCES

- Axelrod, R. M., & Cohen, M. D. (2000). *Harnessing complexity : Organizational implications of a scientific frontier*. New York, NY : Free Press.
- Ball, D. L., & Forzani, F. (2009). The work of teaching and the challenge for teacher education. *Journal of Teacher Education*, 60(5), 497-511.
- Barkham, M., & Mellor-Clark, J. (2003). Bridging evidence-based practice and practice-based evidence : Developing a rigorous and relevant knowledge for the psychological therapies. *Clinical Psychology & Psychotherapy*, 10(6), 319-327.
- Bryk, A. S., Gomez, L. M., & Grunow, A. (2010). *Getting ideas into action : Building networked improvement communities in education*. Stanford, CA : Carnegie Foundation for the Advancement of Teaching.
- Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G. (2015). *Learning to improve : How America's schools can get better at getting better*. Cambridge, MA : Harvard Education Press.
- Bryk, A. S., Sebring, P. B., Allensworth, E., Luppescu, S., & Easton, J. Q. (2010). *Organizing schools for improvement : Lessons from Chicago*. Chicago, IL : University of Chicago Press.
- Cohen, D. K. (2011). *Teaching and its predicaments*. Cambridge, MA : Harvard University Press.
- DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited : Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48, 147-160.
- Dolle, J. R., Gomez, L. M., Russell, J. L., & Bryk, A. S. (2014). More than a network : Building professional communities for educational improvement. In B. J. Fishman & W. R. Penuel (Eds.), *Designbased implementation research : Theories, methods, and exemplars* (p. 443-463). New York, NY : Teachers College Record.
- Elmore, R., & Burney, D. (1997). *Investing in teacher learning : Staff development and instructional improvement in community school district #2*. New York, NY : National Commission on Teaching & America's Future & the Consortium for Policy Research in Education.
- Elmore, R., & Burney, D. (1998). *Continuous improvement in community school district #2*. New York, NY : University of Pittsburgh, Learning Research and Development Center, High Performance Learning Communities Project.
- Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100, 393-394.
- Fink, E., & Resnick, L. B. (2001). Developing principals as instructional leaders. *Phi Delta Kappan*, 82(8), 598-606.
- Fishman, B. J., Penuel, W. R., Allen, A.-R., Cheng, B. H., & Sabelli, N. (2013). Design-based implementation research : An emerging model for transforming the relationship of research and practice. In B. J. Fishman & W. R. Penuel (Eds.), *National Society for the Study of Education* (p. 136-156). New York, NY : Teachers College Record.
- Flyvbjerg, B. (2001). *Making social science matter : Why social inquiry fails and how it can succeed again*. New York, NY : Cambridge University Press.
- Fountas, I. C., & Pinell, G. S. (2001). *Guiding readers and writers (Grades 3-6) : Teaching, comprehension, genre, and content literacy*. Portsmouth, NH : Heinemann.
- Fullan, M. (2001). *Leading in a culture of change*. San Francisco, CA : Jossey-Bass.
- Gallup. (2014). *State of America's schools : The path to winning again in education*. Washington, DC : Gallup.
- Gawande, A. (2012a). The difference between coaching and teaching [Video file]. Retrieved from [https://www.youtube.com/watch?v=VabtGPVVihA].
- Gawande, A. (2012b). How do we heal medicine [Video file]. Retrieved from [http://www.ted.com/talks/atul_gawande_how_do_we_heal_medicine?language=en#t-3268].
- Gould, S. J. (1985). The median isn't the message. *Discover*, 6, 40-42.
- Graham, P. (2005). *Schooling America : How the public schools meet the nation's changing needs*. Oxford : Oxford University Press.
- Green, L. W. (2008). Making research relevant : If it is an evidencebased practice, where's the practice-based evidence? *Family Practice*, 25(Supplement 1), I20-I24.
- Grossman, P., Hammerness, K., & McDonald, M. (2009). Redefining teaching, re-imagining teacher education. *Teachers and Teaching : Theory and Practice*, 15(2), 273-290.
- Hiebert, J., Gallimore, R., & Stigler, J. W. (2002). A knowledge base for the teaching profession : What would it look like, and how can we get one? *Educational Researcher*, 31, 3-15.
- Hiebert, J., & Morris, A. K. (2009). Building a knowledge base for teacher education : An experience in K-8 mathematics teacher preparation. *Elementary School Journal*, 109(5), 475-490.
- Hutchins, E. (1995). How a cockpit remembers its speeds. *Cognitive Science*, 19(3), 265-405.
- Institute of Medicine, Committee on Quality of Health Care in America. (2001). *Crossing the quality chasm : A new health system for the 21st century*. Washington, DC : National Academies Press.
- Institute of Medicine, Committee on Quality of Health Care in America (2012). *Best care at lower costs : The path to continuously learning health care in America*. Washington, DC : National Academies Press.
- Jacob, B. A., & Levitt, S. (2003). *Catching cheating teachers : The results of an unusual experiment in implementing theory* (Working Paper No. 9413). Retrieved from [http://www.nber.org/papers/w9414].
- Langley, G. L., Moen, R. D., Nolan, K. M., Nolan, T. W., Norman, C. L., & Provost, L. P. (2009). *The improvement guide : A practical approach to enhancing organi-*

- zational performance (2nd ed.). San Francisco, CA : Jossey-Bass.
- Lave, J., & Wenger, E. (1991). *Situated learning : Legitimate peripheral participation*. New York, NY : Cambridge University Press.
- May, H., Gray, A., Gillespie, J., Sirinides, P., Sam, C., Goldsworthy, H. & Tognotta, N. (2013). *Evaluation of the i3 scale-up of Reading Recovery : Year one report, 2011-2012* (CPRE Research Report No. RR-76). Philadelphia, PA : Consortium for Policy Research in Education.
- McDonald, M., Kazemi, E., & Kavanagh, S. S. (2013). Core practices and pedagogies of teacher education : A call for a common language and collective activity. *Journal of Teacher Education*, 64(4), 378-386.
- Miller, J. H., & Page, S. E. (2007). *Complex adaptive systems : An introduction to computational models of social life*. Princeton, NJ : Princeton University Press.
- Nielsen, M. A. (2012). *Reinventing discovery : The new era of networked science*. Princeton, NJ : Princeton University Press.
- Office of the Governor (2011, July 5). *Deal releases findings of Atlanta school probe* [Press release]. Retrieved from [http://gov.georgia.gov/press-releases/2011-07-05/deal-releases-findings-atlanta-schoolprob].
- Pascale, R., Sternin, J., & Sternin, M. (2010). *The power of positive deviance : How unlikely innovators solve the world's toughest problems*. Boston, MA : Harvard Business Review Press.
- Pentland, A. (2014). *Social physics : How good ideas spread – the lessons learned from a new science*. New York : Penguin Press.
- Pfeffer, J., & Sutton, R. I. (2000). *The knowing-doing gap : How smart companies turn knowledge into action*. Boston, MA : Harvard Business School Press.
- Roderick, M. (1994). Grade retention and school dropout : Investigating the association. *American Educational Research Journal*, 31, 729-759.
- Sutton, R. I., & Rao, H. (2014). *Scaling up excellence : Getting to more without settling for less*. New York, NY : Crown Business.
- Tyack, D., & Cuban, L. (1995). *Tinkering toward utopia : A century of public school reform*. Cambridge, MA : Harvard University Press.
- Weisberg, H. I. (2014). *Willful ignorance : The mismeasure of uncertainty*. Hoboken, NJ : John Wiley & Sons.
- The William and Flora Hewlett Foundation. (2010). *Education program strategic plan*. Retrieved from [https://www.hewlett.org/].
- Yeager, D. H., Bryk, A. S., Muhich, J., Hausman, H., & Morales, L. (2014). *Practical measurement* (Working paper). Retrieved from [http://www.carnegiefoundation.org/wp-content/uploads/2014/09/Practical_Measurement_Yeager-Bryk1.pdf].