



# Évaluation, développement professionnel et organisation scolaire

*Romuald Normand et Jean-Louis Derouet*

## INTRODUCTION

La mondialisation confronte tous les pays à un certain nombre de défis semblables : une obligation de résultats, qui concerne aussi bien les élèves que les enseignants et les établissements scolaires ; une maîtrise des compétences de base nécessaires à l'investissement en capital humain et à la réussite de parcours professionnels tout au long de la vie. Plusieurs ont entrepris des changements importants dans leur politique d'éducation, notamment en définissant des stratégies nationales pour la numératie et la littératie. Ils ont aussi développé des procédures d'évaluation capables d'articuler des données individuelles à des systèmes d'information plus larges utilisés par le management scolaire et les décideurs politiques. Ces réformes reposent souvent sur des démarches d'amélioration du fonctionnement des établissements. Le principe généralement admis est que le progrès des systèmes éducatifs repose sur l'amélioration du fonctionnement de chaque établissement. Les différents partenaires (personnels d'encadrement, enseignants, familles, collectivités territoriales et, pourquoi pas, élèves) doivent être capables de développer des démarches d'auto-évaluation et d'en utiliser les résultats pour réajuster leur action en fonction des défauts qu'elles ont révélés. L'implication des enseignants dans cette démarche suppose un accompagnement

de leur développement professionnel à l'échelle locale. Ces politiques ont valorisé une diversification pédagogique centrée sur les besoins de chaque élève à l'échelle de la classe, sur les demandes d'un environnement situé à l'échelle de l'établissement. Dans le même esprit, les politiques d'éducation ont aussi intégré les TICE dans les classes en ouvrant les établissements sur des structurations en réseaux et des démarches de projet qui mobilisent divers partenaires, à commencer par les parents d'élèves.

La France est engagée dans la méthode ouverte de coordination (ou MOC) définie à la conférence de Lisbonne en 2000. Les objectifs de cette méthode impliquent un effort particulier dans le domaine de la littératie, de la numératie et plus largement dans la maîtrise des compétences de base par les élèves. Les données PISA constituent une partie intégrante du système d'indicateurs qui permet notamment aux pays européens de se rendre mutuellement compte de l'efficacité de leur politique. La loi d'orientation et de programmation sur l'école de 2005 a d'ailleurs posé le principe d'un « socle commun de connaissances et de compétences ». La LOLF (Loi organique sur les lois de finances), votée en 2001, a introduit les principes du nouveau management public dans le domaine de l'éducation. L'annonce de la suppression progressive de la carte scolaire laisse présager de

nouvelles formes de concurrence, en transformant les modes de circulation des élèves d'un établissement à un autre... Toutes ces décisions appellent le développement d'une culture de l'évaluation chez tous les acteurs. Ce développement demeure cependant limité, en dépit des nombreux rapports et avis rendus par le Haut conseil de l'évaluation de l'école. De même, les rapports successifs de l'inspection générale montrent que les établissements sont très loin d'utiliser les possibilités d'autonomie qui leur sont offertes et que les changements dans les pratiques pédagogiques sont relativement lents. Sur ce plan, les difficultés de mise en œuvre des mesures proposées par la loi d'orientation et de programmation sur l'école de 2005 sont très éclairantes. Une des limites de la politique de projet d'établissement est que son impact reste trop souvent à la porte de la classe. Cette politique a eu, entre 1985 et 1995, des effets sur l'aménagement de la vie scolaire, mais elle n'est pas rentrée dans le domaine des apprentissages. Un nouveau souffle impliquerait de poser la question du projet curriculaire d'établissement. Même si les rapports de l'inspection générale révèlent la variété des interprétations locales des programmes nationaux, la démarche est difficile. Tous les indicateurs mettent en évidence un accroissement des inégalités entre les établissements. Officialiser la possibilité de différenciation présente donc un risque. Il serait évidemment possible d'utiliser les nouvelles marges de liberté pour adapter l'enseignement aux besoins des élèves et donc lutter contre les inégalités d'éducation, mais la tâche est difficile, les acteurs ont peu de ressources... Le risque est grand que la logique du marché l'emporte sur les autres. Les enseignants sont donc très réservés. Ils ont refusé que cette question soit traitée par les conseils d'administration : les représentants des parents, des élèves, des collectivités territoriales ne leur paraissent pas qualifiés. En créant un conseil pédagogique composé uniquement de professionnels, la loi de 2005 a levé cet obstacle, mais sa mise en place est extrêmement lente : les résistances ne se limitent pas aux aspects institutionnels. Dans le même esprit, l'article 34 de cette loi a ouvert la possibilité d'expérimentations qui peuvent aller jusqu'à déroger aux programmes et aux règlements nationaux. L'objectif est d'inciter les établissements à s'inscrire dans une démarche de diversification de l'offre d'éducation. Le ministère de l'Éducation nationale a certes reçu à peu près un millier de dossiers présentant cette démarche et validés par les recteurs, mais ceux-ci sortent rarement des sentiers battus et leurs références s'inscrivent dans des courants généraux comme la mise en

place d'un enseignement intégré des sciences. Très rares sont ceux qui partent réellement d'une analyse des besoins d'une population. Au niveau local, la mise en réseau des établissements est rendue difficile par un isolement relatif, la faiblesse des relations avec les partenaires de l'école, notamment les collectivités locales et les entreprises, le manque d'incitations à l'innovation et au changement. Enfin le partage des compétences entre les différents lieux de décision aux niveaux local, régional et national pose des problèmes de coordination et de cohérence dans le processus de décision politique. De même, la formation des enseignants et son éloignement de la recherche sont de plus en plus critiqués.

Il est donc apparu utile de confronter la situation française à l'expérience acquise au plan international. En 2000, Stanley Hilton, chargé des relations internationales à l'INRP, a engagé une coopération avec le Consortium for Policy Research in Education (CPRE), qui rassemble plusieurs universités américaines : l'Université de Pennsylvanie, le Teachers College de l'Université de Columbia, la Northwestern University, l'Université du Michigan, l'Université de Stanford et celle de Wisconsin-Madison. L'UMR « Éducation et politiques » de l'INRP a ainsi été impliquée dans une recherche sur le rôle de l'évaluation formative, soutenue par la fondation Spencer dans le cadre du programme *ROLE (Research on Learning and Education)*. Les partenaires ont tenu un séminaire à Auckland en 2007, puis à Chicago en 2008. Le séminaire de l'année 2009 a été organisé à Lyon par l'INRP, sous le titre : « Quels enjeux stratégiques pour la mise en œuvre d'une obligation de résultats dans les politiques d'éducation ? Les apports de la recherche internationale ». L'objet était de confronter les grandes orientations du monde anglo-saxon à celles de l'Europe et de la France, de constater les convergences et de réfléchir aux différences dans la conduite des réformes. L'évaluation a constitué le fil directeur de ces journées parce que les expériences étrangères montrent qu'elle constitue un vecteur important des changements opérés dans les autres domaines. C'est à partir de l'évaluation que peuvent se penser de nouveaux rapports entre les enseignants et les élèves dans les classes, de nouvelles procédures de concertation à l'échelle des établissements, de nouvelles formes d'articulation entre enseignants, personnels d'éducation, personnels d'encadrement, familles, élus, formateurs et chercheurs, de nouveaux systèmes d'information, de nouvelles conceptions du management et de la gouvernance.

Les discussions se sont centrées sur quelques questions essentielles permettant de réfléchir à la mise en œuvre et à l'accompagnement d'une politique éducative centrée sur l'obligation de résultats : quelles implications sur les formes d'évaluation des élèves et des enseignants, mais aussi sur l'auto-évaluation de l'établissement scolaire ? En quoi l'évaluation peut-elle contribuer à l'amélioration des pratiques pédagogiques ? Quels effets sur l'organisation de l'établissement scolaire, sur le partage des responsabilités entre acteurs, sur la mobilisation des compétences (*capacity building*) au niveau local ? Quels enseignements peuvent être tirés sur le management et l'impulsion du chef d'établissement, sur les relations avec les partenaires ? Quelles attentes peuvent être formulées en termes de développement professionnel des enseignants (*professional development*) ? Quelle place pour la recherche si l'on veut rendre l'enseignement plus efficace (*evidence-based teaching*) ?

C'est à partir de ces questions qu'a été élaboré ce dossier. Mettre à la disposition des lecteurs francophones les acquis des travaux anglo-saxons sur les démarches de transformation de l'organisation scolaire fondées sur ce nouvel usage de l'évaluation nécessite quelques clarifications préliminaires. Les échanges ont mis en évidence des décalages touchant à la fois la chronologie et la culture. Le décalage chronologique est le plus simple à traiter, d'autant qu'il se resserre ; il était très important au début des années quatre-vingt : la France importait, avec les zones d'éducation prioritaire, les pratiques de pédagogie de compensation qui avaient été expérimentées en Grande-Bretagne et aux États-Unis à partir des années soixante. Ces politiques venaient d'être très violemment critiqués au Royaume-Uni par les *white papers* qui ont accompagné le retour au pouvoir des Tories en 1979 et par le rapport *A nation at risk* sorti en 1983 aux États-Unis. Ce renouveau du libéralisme a progressivement imposé un changement de référentiel : l'objectif premier de l'éducation n'est pas l'égalité ou même la cohésion sociale, c'est le rang du pays dans la compétition mondiale. L'écart se resserre depuis le début du <sup>xxi</sup><sup>e</sup> siècle : la MOC mise en place depuis la conférence de Lisbonne en 2000 propose une réorganisation des politiques européennes à partir d'une obligation de résultats. La notion de socle commun définie par la loi de 2005 s'inspire directement des « sept compétences clés pour une formation tout au long de la vie » définie en 2004 par le programme *Éducation et formation 2010* de la Communauté européenne. Les décalages entre les traditions politiques et culturelles sont plus difficiles à combler. Ils ne se limitent

pas à l'opposition tocquevillienne entre les pays centralisés et les pays de *self-government*, il s'agit plus profondément de la conception de l'intérêt général. Sa construction suppose-t-elle une rupture radicale avec les intérêts privés ou peut-elle aussi procéder d'un effet de composition ? On retrouve là une difficulté bien connue. Les organisations internationales ont affirmé que les politiques d'éducation relevaient du domaine du *public interest*, mais la traduction de ce terme en français s'avère difficile. Une seule chose est sûre : cette formule présente un sens assez différent de ce que l'on appelle en français l'intérêt public. Sur un autre plan, la société française commence à peine à se penser comme pluriethnique et pluriculturelle, alors que les pays anglo-saxons ont depuis longtemps intégré cette dimension à leur définition de la citoyenneté et de leur *democracy* : une association de tous aux décisions qui les concernent.

Pour introduire ce dossier, la meilleure démarche est donc de suivre assez précisément la genèse des démarches anglo-saxonnes, d'expliquer la manière dont les politiques d'obligation de résultats sont passés de la *school effectiveness* au *school improvement* et à l'émergence progressive d'une « obligation de résultats intelligente ». Depuis le début des années quatre-vingt, les systèmes éducatifs des pays anglo-saxons ont connu des changements profonds consécutifs à un abandon du modèle de l'école compréhensive promu par les grandes organisations internationales. Le thème de l'équité et de l'égalité des résultats s'est imposé comme nouvelle référence de justice, alors que les établissements devaient tout mettre en œuvre pour améliorer les compétences de base des élèves dans les disciplines fondamentales (Derouet, 2006 ; Derouet & Derouet-Besson, 2009). L'évaluation des résultats scolaires est devenue une priorité pour les chefs d'établissement et pour les décideurs politiques. Ceux-ci utilisent les données chiffrées pour comparer les établissements et réguler les systèmes éducatifs à l'échelon local ou national, en fixant des standards ou des objectifs à atteindre par les enseignants (Normand, 2008, 2010b). Ces évolutions ont été facilitées par l'usage et le développement des tests accompagnant les recompositions du curriculum, même si leur usage intensif a fait l'objet de critiques (Radnor & Ball, 1996). Dans le premier temps des réformes, l'évaluation sommative s'est imposée comme l'instrument d'une obligation de résultats (*accountability*), orientant à la fois les pratiques pédagogiques dans la classe et l'action des responsables de l'éducation dans les départements ministériels (Derouet, Normand & Bessy, 2009 ; Gewirtz, Ball & Bowe, 1995).

Si l'évaluation est devenue centrale, les réformes ont aussi investi d'autres composantes du fonctionnement des établissements scolaires. Elles ont d'abord transformé le rôle des enseignants dans leurs classes et leur environnement professionnel (Hargreaves, 1993, 2003). Alors qu'ils étaient contraints dans le choix de leur progression pédagogique, qu'ils devaient se conformer à des standards, les responsables de l'éducation attendaient qu'ils améliorent les résultats des élèves, travaillent davantage avec leurs collègues et l'équipe de direction, développent le projet d'établissement par des actions transversales et des innovations (Mahony & Hextall, 2000). Les conditions de leur évaluation ont été bouleversées pour inclure des critères de performance et de coopération, de créativité, de prise de responsabilité en dehors de la classe. Les chefs d'établissement n'étaient plus considérés comme des administrateurs mais comme des managers, responsables de la performance des équipes pédagogiques et chargés d'impulser une dynamique pour maintenir le rang de l'établissement dans un environnement concurrentiel (Gewirtz, 2002). Les parents d'élèves furent mobilisés dans le travail avec les enseignants alors que se renforçait l'intégration de l'établissement dans sa communauté locale. Tous ces changements ont été accompagnés par les développements d'une ingénierie du *school improvement* : les établissements ont souvent bénéficié de ressources extérieures et du soutien de professionnels ou d'experts pour développer ces changements dans leur organisation et leurs pratiques pédagogiques (Harris, 2002).

Malgré tout, la mise en œuvre d'une première forme d'obligation de résultats s'est heurtée à un certain nombre de difficultés. Le développement des tests de compétences des élèves à différents niveaux de scolarité a eu des effets pervers. Les pratiques de bachotage et de sélection se sont multipliées avec des effets de stigmatisation et d'exclusion des élèves les plus fragiles (Firestone, Schorr & Monfils, 2004 ; Gillborn & Youdell, 2000 ; Lipman, 2004). Les enseignants, quand ils parvenaient à traiter l'abondance des informations chiffrées à leur disposition, concentraient tellement leurs efforts sur les parties du programme évaluées par les tests qu'ils en oubliaient d'enseigner les autres contenus. La multiplication des évaluations à différentes échelles, avec des objectifs et des formes différentes, brouillait les informations sur les résultats utilisés par les praticiens, comme par les décideurs politiques (Fuhrman, 2001). Les enseignants les plus motivés finissaient par être découragés et certains de leurs collègues fuyaient, comme les élèves, les établissements en difficulté qui n'arri-

vaient pas à améliorer leur performance (Hursh, 2008). Les cadres et les responsables de l'éducation exerçaient une telle pression sur les indicateurs de performance que les objectifs en étaient détournés par des pratiques de tricherie, d'estimation tronquée de certains résultats, de fixation maniaque sur certaines cibles au détriment d'objectifs plus globaux (Nichols & Berliner, 2007 ; Stobart, 2008). Bientôt, il parut évident aux chercheurs comme aux décideurs anglo-saxons qu'il fallait réviser cette politique pour la rendre plus utile et efficace dans l'amélioration de la réussite des élèves, le changement des pratiques pédagogiques des enseignants, la prise de décision et le pilotage du chef d'établissement.

L'idée d'« obligation de résultats intelligente » (*intelligent accountability*) s'est progressivement imposée. Avant d'en préciser les caractéristiques et de montrer la place occupée par l'évaluation dans les transformations de l'organisation scolaire, il est nécessaire de présenter un panorama des réformes entreprises aux États-Unis, en Nouvelle-Zélande et en Grande-Bretagne. Par bien des aspects, ces pays présentent des histoires et des traditions différentes. Ils ont fait certains choix dans l'orientation de leurs politiques éducatives, même si l'évaluation a toujours été placée au centre des réformes. C'est parce que les contextes et les problèmes éducatifs sont spécifiques à une culture et à une société particulière que les réponses apportées sont différentes d'un pays à l'autre. Malgré tout, il est possible de repérer les grandes tendances dans les changements opérés. Elles s'expliquent par une circulation plus rapide des idées et des modèles éducatifs sous l'effet d'un rôle plus influent des grandes organisations internationales et d'une accentuation de la mobilité des chercheurs, des experts, des responsables politiques. De ce point de vue, le modèle anglo-saxon présente une certaine homogénéité, au-delà de l'aspect proprement linguistique, et il exerce une certaine domination, voire une fascination dans les cercles d'expertise proches de la décision politique en France et en Europe. Notre objectif n'est pas d'en faire ici une critique mais d'expliquer comment ce modèle a évolué avec un certain degré de réflexivité malgré ses limites et ses contradictions.

## LES ENJEUX D'UNE POLITIQUE FÉDÉRALE D'OBLIGATION DE RÉSULTATS AUX ÉTATS-UNIS

Aux États-Unis, la politique éducative est d'abord locale. Elle se joue à l'échelle des districts et à travers

la composition des conseils locaux (appelés *boards*) et l'intervention des municipalités. Le district est responsable de l'instruction et des programmes scolaires, du recrutement et de la nomination des enseignants comme des chefs d'établissement, du financement des établissements scolaires, de leur construction et de leur rénovation (Conley, 2003). Un *superintendent*, nommé ou élu par le conseil local, gère l'ensemble des affaires scolaires. Il travaille selon une forte interaction avec la communauté locale quand le district est de petite taille, comme c'est le cas en zone rurale où la recherche du consensus prédomine. Par contre, il s'appuie sur un management et une organisation technocratique pour les grandes agglomérations, en se confrontant à l'expression des intérêts et au regard des parents d'élèves appartenant à différentes communautés (blanches, noires, hispaniques). Il existe plus de 14 000 districts avec des configurations très différentes, ce qui fait du territoire éducatif américain une véritable mosaïque avec des différences et des inégalités très fortes. L'essentiel des ressources publiques des établissements scolaires provient de taxes locales et les quartiers périphériques des villes, à majorité blanche, sont beaucoup plus riches que les quartiers du centre, à majorité noire ou hispanique.

Toutefois, depuis les années soixante-dix, l'intervention des États dans la gestion des établissements scolaires s'est renforcée (McGuinn, 2005). En effet les représentants de la communauté noire et des minorités ethniques ont déposé des recours devant les tribunaux et les cours suprêmes étatiques pour obtenir davantage d'équité dans le financement des établissements publics. Par la suite, les États sont intervenus dans la régulation du système éducatif en définissant des standards à atteindre et en liant l'allocation des ressources publiques à l'amélioration de la performance. Une politique de socle commun et d'obligation de résultats s'est mise en place, alors que chaque État développait son propre système d'évaluation basé sur des tests. Aujourd'hui, la plupart des États spécifient les grandes lignes des programmes scolaires (par des standards), le temps d'enseignement dans les établissements et l'âge de la scolarité obligatoire. Ils sont aussi responsables de la formation et de la certification des enseignants, même s'ils délèguent souvent ce travail à des agences extérieures. Ils jouent un rôle d'incitation pour le développement d'innovations pédagogiques et de projets éducatifs et conçoivent des outils pour accompagner les enseignants dans la mise en œuvre des standards et de l'évaluation. Les établissements et les districts doivent quant à eux se centrer sur leur

performance, collecter et analyser des données sur leur efficacité, développer des stratégies continues d'amélioration de leurs pratiques pédagogiques (*school improvement*). Ces établissements sont comparés entre eux puis financièrement récompensés ou sanctionnés selon leurs résultats.

En matière d'évaluation, l'État fédéral joue un rôle important, notamment pour les établissements scolaires dépourvus de ressources et situés dans les quartiers pauvres. Ce sont d'ailleurs les questions de pauvreté qui l'avaient amené à intervenir dans les années soixante pour mettre en œuvre des politiques de discrimination positive (Silver & Silver, 1991). Un premier programme fédéral, *Head Start*, était mis en place pour aider les élèves défavorisés (majoritairement noirs) et leurs familles durant les premières années de scolarité. L'autre programme, appelé *Titre I*, parce qu'il était le premier chapitre de la loi sur l'enseignement primaire et secondaire votée par le congrès américain, consacrait une politique de discrimination positive pour lutter contre la ségrégation sociale et raciale en donnant plus de ressources aux établissements les plus pauvres. L'arrivée au pouvoir des néoconservateurs et la publication du rapport *A Nation at risk* (1983) ont mis un coup d'arrêt à cette politique d'égalisation des chances en réduisant l'intervention fédérale et les grands programmes publics (Berliner & Biddle, 1995). Mais en 2001, le gouvernement de Georges Bush organisait le retour de l'interventionnisme fédéral à travers la réactivation du *Titre I* et la promulgation d'une loi appelée *No Child Left Behind* (ou « aucun enfant abandonné »). Reste que le programme était bien différent de celui des années soixante : il faisait dépendre les dépenses fédérales en faveur des établissements défavorisés d'une obligation de résultats. Les établissements scolaires les plus pauvres, pour obtenir les fonds publics, devaient montrer des progrès annuels dans les performances des élèves au regard d'objectifs préalablement fixés. Dans le cas contraire, ils s'exposaient à des sanctions allant jusqu'à l'envoi des élèves dans d'autres établissements et une restructuration complète du personnel éducatif.

Les effets de cette loi ont été très débattus aux États-Unis, mais la politique du gouvernement de Barack Obama n'est pas très différente (Hursh, 2007). Pour le comprendre, il suffit de se rappeler que, depuis le moment où Clinton s'est rallié à Bush pour défendre des objectifs communs pour le système éducatif américain (1989), il existe un accord bipartite au congrès américain, c'est-à-dire un consensus entre démocrates et républicains sur la mise en œuvre

d'une politique des standards et de l'obligation de résultats à l'échelle du pays. Depuis, les procédures d'évaluation se sont étendues à tous les États et sont utilisées à différents niveaux de scolarité comme indicateur principal de l'efficacité des systèmes éducatifs. Cette tendance a été renforcée par le vote de la loi *No Child Left Behind*, alors que le système d'évaluation fédérale (*National Assessment of Educational Progress*) devenait une référence commune aux responsables de l'éducation.

Mais les politiques mises en œuvre sont très variables d'un État à l'autre. Les États déterminent le contenu des standards et la législation empêche toute ingérence du gouvernement fédéral dans la définition de l'instruction et du curriculum (Goertz, 2001). De fait, tous les États n'ont pas adopté les mêmes standards et ils ne couvrent pas toujours les mêmes compétences de base. Il demeure des controverses sur leur définition, surtout quand ceux-ci concernent le lien entre des questions religieuses et l'histoire ou la biologie. Les évaluations ont donné parfois des résultats contradictoires et incohérents entre ceux affichés par les États et ceux pris en compte par l'évaluation fédérale NAEP. Bien qu'ils doivent se conformer à la loi *No Child Left Behind*, les États ont établi leurs propres standards en lecture, mathématiques, sciences et leurs propres évaluations, si bien que les critères de performance varient d'un État à l'autre, ce qui crée des conflits avec le gouvernement fédéral. Beaucoup de districts manquent de ressources pour mettre en œuvre des guides méthodologiques et des stages de formation pour les enseignants et ne disposent pas toujours de compétences locales dans le domaine de l'évaluation. Certains districts urbains sont davantage préoccupés de lutter contre les problèmes de pauvreté, de délinquance et de violence, d'usage de drogues, de mobiliser des enseignants volontaires que d'investir les questions d'évaluation des compétences des élèves.

Pour sortir de ces difficultés, beaucoup de chercheurs et de responsables de l'éducation réfléchissent aux moyens d'améliorer les outils et les dispositifs d'évaluation existants, sans toutefois renoncer aux tests (Fuhrman, Cohen & Mosher, 2007). Ils sont déterminants pour l'accès à l'enseignement supérieur et aux universités les plus prestigieuses et ils représentent un marché important. Certains experts et décideurs politiques rêveraient d'aller plus loin, au-delà de la réforme des standards, dans la voie d'une harmonisation des contenus d'enseignement et de la formation des enseignants, mais ils se heurtent à la résistance des États et des districts. L'obligation de

résultats constitue donc le plus petit dénominateur commun d'une politique d'éducation à l'échelle fédérale. C'est à l'échelle des districts que l'amélioration de la gouvernance est recherchée, par une mise en réseau des établissements scolaires et par la définition d'un leadership local suffisamment efficace, capable de coordonner les actions et d'améliorer les résultats des élèves, de mobiliser ressources et partenariats publics comme privés (Supovitz, 2006). Dans ce contexte, le repérage des « bonnes pratiques » et leur diffusion constituent des moyens d'action essentiels pour enclencher une dynamique d'innovation et de changement dans le système éducatif américain. On comprend mieux pourquoi le ministère de l'Éducation américain et son Institut fédéral des sciences de l'éducation s'évertuent, depuis plusieurs années, à développer des protocoles et des méthodes d'évaluation facilitant cette capitalisation des expérimentations et leur généralisation sur le territoire fédéral. L'autre facteur de diffusion des innovations pédagogiques ou des modèles de changement est du ressort des grandes fondations, qui financent des programmes éducatifs à large échelle dans différents domaines de la politique éducative, ce qui permet là aussi de développer des comparaisons et de nourrir la réflexion des États comme des districts.

#### **LA STRATÉGIE NÉO-ZÉLANDAISE POUR LA LITTÉRATIE ET LA NUMÉRATIE : L'ÉVALUATION AU CENTRE D'UNE GOUVERNANCE DÉCENTRALISÉE**

Durant les deux dernières décennies, la Nouvelle-Zélande a connu des réformes importantes en termes de financement et de régulation du système éducatif. Les fonctions du ministère de l'Éducation ont été complètement décentralisées et des agences ont été créées comme l'Education Review Office (ERO) pour l'inspection des établissements et la New Zealand Qualifications Authority pour la définition des programmes scolaires et des standards (Philips, 2000). Avec les développements de l'évaluation, ces changements avaient pour but de réduire les disparités de réussite entre élèves, d'améliorer leurs qualifications afin d'augmenter la compétitivité économique néo-zélandaise. Les contenus scolaires ont été davantage spécifiés en termes d'objectifs à atteindre et les programmes d'évaluation se sont étendus pour encadrer les apprentissages et contrôler la performance des enseignants. De nouvelles prescriptions ont été

adressées aux établissements scolaires pendant que l'accent était mis sur le management local et le rôle des conseils de gouvernance (*boards of trustees*) dans la gestion des affaires scolaires. Des critiques ont alors pointé les dérives entraînées par un renforcement du marché scolaire et la mise en place d'une obligation de résultats. Cette politique ne permettait pas de réduire les inégalités et accentuait la défiance vis-à-vis des professionnels de l'éducation (Thrupp, 1999, 2001). À la différence des États-Unis et de l'Angleterre, les enseignants et les chefs d'établissement se sont opposés à des évaluations nationales obligatoires composées de tests de compétences des élèves.

Alors qu'un curriculum national était mis en œuvre, l'ERO exigeait que les enseignants diversifient leurs pratiques d'évaluation dans un but formatif d'aide aux élèves dans leurs apprentissages. Des objectifs furent affichés pour améliorer les compétences de base des élèves dans les domaines de la littérature et de la numératie, dans la continuité d'une stratégie nationale mise en œuvre par le ministère de l'Éducation (Dale, 2001). Aujourd'hui encore, des standards décrivent précisément ce que les élèves doivent acquérir tout au long de leur scolarité obligatoire (de la 1<sup>re</sup> année à la 13<sup>e</sup> année). Pour chaque niveau sont indiqués des exemples d'enseignement, d'apprentissage et d'évaluation qui peuvent être repris par les enseignants dans leur classe. L'une des fonctions principales de l'ERO est de conduire des cycles d'inspection des établissements scolaires tous les trois ans, pour évaluer si les objectifs nationaux sont ou non atteints. En termes d'évaluation, un programme national a été mis en œuvre pour s'assurer que les élèves atteignent les objectifs fixés à différents niveaux du curriculum. Un système de pilotage national, composé d'outils variés, a été créé afin d'assister les enseignants dans leurs pratiques d'évaluation sommative et formative.

Les établissements et les enseignants disposent d'une grande autonomie dans leur méthode d'évaluation. La seule contrainte est d'atteindre les objectifs fixés. Le ministère de l'Éducation leur fournit une assistance pour les aider à gérer les progrès des élèves, notamment à travers des exemples d'activités d'évaluation articulées au curriculum national et des dispositifs de développement professionnel permettant aux enseignants de s'approprier les démarches et les outils nécessaires. Des évaluations en début de scolarité facilitent le diagnostic des élèves en difficulté et permettent de planifier les dispositifs et les ressources pour le management et le pilotage du sys-

tème éducatif. Ces données nationales sont utilisées aussi pour déterminer les actions en faveur des élèves de la minorité ethnique (la communauté maorie) ou ceux pour lesquels l'anglais n'est pas la première langue. Des banques de données en mathématiques, sciences et anglais sont aussi utilisées pour situer les évaluations faites en classe au regard des objectifs et des standards nationaux. Les enseignants peuvent utiliser ces ressources multiples pour leurs pratiques pédagogiques (groupements d'items par thème, questionnaires à choix multiples, tâches pratiques d'évaluation dans la classe). Pour obtenir des informations sur la réussite des élèves à l'échelon national, le ministère de l'éducation a créé un instrument : le *National Education Monitoring Project* (ou NEMP, projet national de pilotage de l'éducation), qui évalue la réussite des élèves pour différentes compétences dans différents domaines du curriculum des années 4 à 8, mais cette évaluation nationale ne mobilise qu'un petit échantillon d'élèves. Son objectif est avant tout de dresser un tableau général des tendances de la réussite scolaire pour éclairer les responsables de l'éducation sur les changements à introduire en termes de politique éducative.

Au niveau du collège, les enseignants néo-zélandais utilisent un logiciel spécifique dénommé « asTTle » afin de suivre les progrès des élèves en lecture, écriture, mathématiques, conformément aux exigences du programme scolaire. Mais ils peuvent faire passer les tests aux élèves quand ils le souhaitent et peuvent adapter les items à leur progression pédagogique. Le logiciel néo-zélandais permet d'analyser les résultats par niveau selon une configuration ergonomique très adaptée à l'environnement cognitif de l'enseignant. Il ne se contente pas d'élaborer un diagnostic : il permet de suivre les progressions individuelles et collectives des élèves dans le temps. De plus, à partir des profils d'élèves, les enseignants peuvent accéder à des « modèles d'évaluation » qui ont pour fonction d'illustrer les caractéristiques clés de l'apprentissage et de la réussite des élèves à une étape particulière de leur développement cognitif. Ces modèles servent aussi à identifier les différents niveaux de réussite et les prochaines étapes de l'apprentissage selon la planification pédagogique retenue. Ils guident les enseignants dans leur interprétation des programmes scolaires quand ils ont à produire des jugements circonstanciés sur le travail de tel ou tel élève.

Considérons maintenant la façon dont se décline la stratégie néo-zélandaise pour la littérature et la numératie à l'échelle de l'organisation scolaire (Timperley

& Parr, 2009). Le système d'autonomie locale très forte des établissements scolaires et l'absence de hiérarchie intermédiaire entre le ministère de l'Éducation et les autorités locales ne permettent pas la mise en œuvre d'une politique prescriptive. Le ministère se contente de formuler des recommandations et de fournir des lignes directrices, mais les établissements scolaires peuvent choisir leurs démarches en termes d'offre curriculaire et d'outils d'évaluation. Par conséquent, la mise en œuvre de la stratégie nationale s'opère essentiellement selon une ingénierie méthodique et réfléchie offerte comme ressource et soutien aux équipes pédagogiques locales. C'est donc un processus dynamique, multipartenarial combinant de nombreux outils et documents d'accompagnement et qui mise sur un développement professionnel des enseignants centré sur l'évaluation. Le ministère développe des programmes, nourris et accompagnés par la recherche en éducation, pour aider les enseignants à améliorer la réussite et les compétences des élèves. Les enseignants sont considérés comme des enquêteurs sur leurs propres pratiques (Timperley, Parr & Bertanees, 2009). Ils doivent construire et s'approprier des connaissances nouvelles en collectant les informations nécessaires à leurs actions et en bénéficiant de formations adaptées à l'échelle de l'établissement scolaire. Ils sont accompagnés par des équipes d'experts en éducation qui les aident à affiner leurs démarches et à en évaluer les résultats. Cette enquête collective, à l'échelle de l'établissement, et ce partage des connaissances correspondent à un cycle d'apprentissage professionnel orienté vers l'amélioration des résultats des élèves.

## **EN ANGLETERRE : UNE POLITIQUE D'OBLIGATION DE RÉSULTATS INTELLIGENTE**

À la fin des années quatre-vingt-dix, la politique d'obligation de résultats mise en œuvre en Angleterre a fait l'objet de nombreuses critiques (Tomlinson, 2005, 2009). Quelques-unes dénonçaient la stigmatisation de certains établissements scolaires considérés comme « défailants » parce qu'ils ne parvenaient pas à améliorer les résultats de leurs élèves aux évaluations nationales. Cette politique de « *naming and shaming* » était renforcée par l'attitude du corps d'inspection qui exerçait une forte pression sur les chefs d'établissement et les équipes pédagogiques pour améliorer la performance des établissements. Des études ont montré que ces inspections s'appuyaient trop souvent sur des données faussées,

qu'elles étaient faiblement argumentées et que les recommandations des inspecteurs étaient mal exécutées (Wolf & Janssens, 2007). Cette pression à la performance générait aussi un certain nombre d'effets pervers : en plus des comportements de tricherie ou de bachotage, certains établissements passaient tellement de temps à préparer les visites des inspecteurs qu'ils ne s'occupaient plus des élèves et de leurs résultats qui baissaient. D'autres cherchaient systématiquement à se débarrasser des plus mauvais élèves ou à renforcer les procédures de sélection, ce qui augmentait les abandons de scolarité et les sorties sans qualification. L'enjeu était si important qu'en 1999, le parlement anglais reconnaissait l'importance du stress des enseignants et demandait que l'inspection rétablisse la confiance avec les établissements scolaires et cherche à promouvoir des démarches d'auto-évaluation. Le NUT (National Union of Teachers), le principal syndicat d'enseignants, subventionnait alors une étude pour promouvoir l'auto-évaluation des établissements (MacBeath, 1999). Le rapport est devenu rapidement un slogan (« Les établissements parlent par eux-mêmes »), tandis que le syndicat réclamait le boycott des évaluations nationales, c'est-à-dire des tests passés par les élèves à différentes étapes clés de leur scolarité. L'équivalent anglais du Haut conseil de l'éducation exigeait plus de responsabilités pour les établissements en appelant à une obligation de résultats professionnelle donnant plus d'importance aux établissements dans la conduite de leur propre évaluation.

À la suite de ces événements, le ministère anglais de l'Éducation formulait dès 2004 les principes d'une nouvelle relation avec les établissements scolaires pour que le système éducatif anglais adopte une « obligation de résultats intelligente » (MacBeath, 2006). Selon le ministre de l'époque, David Miliband, elle devait correspondre à une démarche d'auto-évaluation de l'établissement scolaire, une inspection resserrée et centrée sur la dynamique de changement et sur l'amélioration des pratiques pédagogiques, la mise en œuvre d'un système d'information performant pour renforcer la communication entre tous les partenaires du système éducatif. En donnant plus d'autonomie locale et en limitant les contraintes bureaucratiques, en s'appuyant sur les technologies de l'information et de la communication les plus récentes et en allégeant le dispositif d'inspection, le ministère souhaitait faire adopter un cadre d'évaluation qui prenne en compte les forces et des faiblesses des établissements scolaires.

La nouvelle relation avec les établissements scolaires supposait une nouvelle stratégie de communication de la part du ministère : chaque établissement pouvait désormais demander des informations en fonction de ses besoins et le chef d'établissement n'était plus envahi par une pléthore de données inutiles. Il pouvait bénéficier de la mise à jour régulière des informations et télécharger de nombreux documents en fonction de ses besoins. Le système d'information était là pour aider les établissements à améliorer leur performance et à éclairer les décisions en termes de pilotage. Aujourd'hui, les données servent au développement professionnel des enseignants, au management des équipes pédagogiques, au suivi des progrès des élèves, à l'attribution de ressources ciblées par les autorités locales, à la régulation des interventions du ministère. Ce système d'information propose une synthèse nationale des résultats aux tests et aux examens, des tableaux de *benchmarks* permettant la comparaison entre établissements, des données sur la valeur ajoutée et les progrès des élèves, des informations et des rapports produits par l'inspection. L'objectif est que l'établissement soit en mesure de se situer dans son environnement local et national, bien que la plupart de ces informations demeurent confidentielles (elles sont seulement connues du ministère et des autorités locales). Toutefois les résultats des élèves et les rapports d'inspection sont rendus publics. Les équipes pédagogiques ont à leur disposition d'autres logiciels pour comparer les progrès des élèves, à titre individuel ou par groupe, à différents cycles de la scolarité. Ces données sont utilisées collectivement par les équipes d'enseignants pour établir des progressions pédagogiques et fixer des objectifs en termes d'apprentissage et de résultats à atteindre.

L'inspection a elle aussi considérablement modifié ses pratiques. Les visites d'établissement sont plus courtes (limitées à deux jours maximum) et les interactions avec l'équipe de direction et les enseignants ont été renforcées. Les délais de notification de l'inspection ont été réduits pour éviter une préparation intempestive des équipes pédagogiques et réduire leur niveau de stress. La délégation d'inspecteurs en visite est moins importante, et parfois composée seulement d'un inspecteur attitré. L'auto-évaluation constitue le point de référence de la démarche d'inspection et les établissements sont invités à mettre à jour leurs informations sur une base annuelle. Le jugement des inspecteurs se veut moins critique et plus attentif à la dynamique de changement de l'établissement scolaire. L'établissement est évalué sur son efficacité globale, c'est-à-dire l'ensemble des

services rendus à la communauté locale mais aussi, en plus de la réussite des élèves et de l'atteinte des standards, sur la qualité de son offre dans la prise en charge des élèves et de leurs besoins particuliers, la conduite des programmes et des activités pédagogiques, l'orientation, le soutien, l'accompagnement des élèves, le management de l'établissement et l'exercice du leadership. Dans leur travail, les inspecteurs doivent respecter un certain nombre de prescriptions en s'assurant que leurs jugements bénéficient à l'établissement et que les personnes sont traitées de manière égalitaire. Ils doivent faire preuve d'objectivité, être fiables et honnêtes. Leurs interlocuteurs doivent être traités avec courtoisie, sensibilité et leur visite générer un minimum de stress. Un dialogue constructif doit être engagé avec les équipes en maintenant la confidentialité des informations recueillies dans les entretiens.

Pour conduire la démarche d'auto-évaluation de l'établissement scolaire, l'inspection a conçu un cadre (appelé SEF comme *Self Evaluation Framework*) qui peut être repris, mais qui n'a pas de caractère obligatoire afin de respecter l'autonomie locale. Les établissements doivent d'abord « s'exprimer par eux-mêmes ». Cependant quelques principes ont été retenus pour inscrire la démarche d'auto-évaluation dans le contexte d'une obligation de résultats intelligente. Elle doit présenter le propre regard de l'établissement sur la manière dont il rend service aux élèves et définit des priorités pour changer et améliorer l'organisation pédagogique. L'auto-évaluation s'ancre dans les pratiques quotidiennes de l'établissement et s'appuie sur des preuves rigoureuses des actions conduites, en évaluant leurs forces et leurs faiblesses. Elle implique l'ensemble de la communauté éducative (enseignants, parents, élèves, membres du conseil d'administration). Elle doit situer l'établissement dans son environnement et au regard de *benchmarks* établis nationalement. Après avoir rempli la grille d'auto-évaluation, les établissements se positionnent pour chaque domaine évalué selon l'échelle suivante : extraordinaire, bon, satisfaisant, inadapté. Pour chacun de ces domaines, le jugement exprimé doit être très précis, s'appuyer sur des données (quand elles existent) et envisager des solutions pour remédier aux défauts qui apparaissent.

Pour compléter cette gouvernance locale de l'évaluation, un partenaire peut être nommé, après accord du ministère, afin d'accompagner la dynamique de changement de l'établissement scolaire (il s'agit du *school improvement partner*). Il assure une médiation entre l'établissement, l'autorité locale et l'inspection.

Ce partenaire est considéré comme un « ami critique » qui aide le chef d'établissement à définir et à conduire sa politique en vue d'améliorer ses résultats. Il est là aussi pour construire les compétences nécessaires aux enseignants pour réaliser la démarche d'auto-évaluation. L'ami critique est le plus souvent un chef d'établissement expérimenté qui appartient à un autre district et qui est formé par le ministère au sein du National College of School Leadership. Il allège le travail de l'inspection dans le respect de la confidentialité des informations et de l'autonomie de l'établissement scolaire. Il doit s'assurer que l'action du chef d'établissement et la démarche d'auto-évaluation conduisent à une amélioration des pratiques pédagogiques et à la mise en œuvre de stratégies efficaces. L'ami critique ne s'intéresse pas seulement au management de la performance, mais aussi au développement professionnel des enseignants et à l'activation des réseaux de l'établissement aux échelons local et national.

#### **DÉVELOPPEMENT PROFESSIONNEL, LEADERSHIP ET ÉVALUATION DES APPRENTISSAGES**

Sans renoncer aux objectifs d'égalité de résultats entre élèves et d'atteinte des standards, les politiques d'obligation de résultats conduites dans les pays anglo-saxons ont singulièrement infléchi leur trajectoire. Alors que le développement des tests et des évaluations standardisées avait été considéré comme l'instrument essentiel du pilotage des systèmes éducatifs, ces réformes ont généré bien des désillusions, de la part des chercheurs comme des décideurs politiques. L'évaluation, d'abord utilisée pour renforcer la logique du marché scolaire et créer une pression sur les enseignants, a creusé les inégalités en démotivant toute une profession (Gewirtz, Mahony, Hextall *et al.*, 2008). La mise en œuvre de stratégies nationales pour la littératie et la numératie, les changements des normes professionnelles pour les enseignants, le soutien apporté aux équipes éducatives, le développement d'une ingénierie pédagogique, la diversification de l'offre de services, les transformations du management scolaire ont ensuite modifié considérablement les configurations de l'organisation scolaire. L'obligation de résultats a continué d'occuper une position centrale, mais elle s'est diversifiée en se faisant plus souple et elle a été accompagnée par d'autres leviers jugés importants dans la conduite des réformes : un développement professionnel continu des enseignants, un nouveau

partage des rôles et des responsabilités dans l'établissement, une évaluation au service des apprentissages des élèves (Elmore, 2004 ; Hopkins, 2007).

Idéalement le développement professionnel des enseignants est l'ensemble des apprentissages et des expériences faites par l'enseignant dans son environnement de travail (Darling-Hammond & Sykes, 1999). Il est mis en œuvre selon un processus continu durant toutes les étapes de la carrière professionnelle, en formation initiale comme en formation continue, mais aussi avec l'accompagnement par un tuteur et la formation dans l'établissement scolaire. Cette notion a pris de l'importance pour les réformateurs et les décideurs politiques alors que les enseignants étaient amenés à prendre en charge des fonctions nouvelles. Durant la dernière décennie, dans les pays anglo-saxons, des standards professionnels ont été créés afin d'améliorer les compétences des enseignants et les résultats des élèves, pour satisfaire aux nouvelles conditions d'obligation de résultats. La recherche en éducation anglo-saxonne s'est alors beaucoup intéressée au développement professionnel des enseignants pour évaluer quelles pratiques d'enseignement amélioreraient la réussite des élèves et quels types de formation étaient les plus efficaces (Alton-Lee, 2003 ; Timperley, Wilson, Barrar *et al.*, 2007).

Ces travaux ont conclu que l'organisation de conférences ou la participation ponctuelle à des ateliers ne parvenaient pas à modifier les pratiques pédagogiques et qu'elles avaient peu d'impact sur la réussite des élèves. Ils ont aussi défendu l'idée que les enseignants, même avec suffisamment de temps et de ressources, ne pouvaient pas s'autoréguler en mettant en œuvre des pratiques pédagogiques efficaces. Ils étaient aussi dans l'incapacité d'appliquer les « recettes » proposées par des experts externes, même si la démarche s'appuyait sur les résultats de la recherche. Aujourd'hui, de nombreux chercheurs considèrent que les enseignants, s'ils veulent améliorer leurs compétences professionnelles dans le but de faire réussir les élèves, doivent être des « enquêteurs sur leurs propres pratiques ». Les connaissances théoriques de la recherche en éducation sont essentielles et elles doivent être intégrées à une démarche réflexive des enseignants permettant d'interroger l'acte pédagogique lui-même, en mobilisant leurs connaissances non seulement disciplinaires, mais aussi pédagogiques (*pedagogical content knowledge*). Ensuite, la formation doit se centrer sur un contenu et un domaine spécifique en prenant en compte les problèmes rencontrés par les ensei-

gnants dans les situations vécues dans la classe. Pour être efficace, elle doit combiner des formations intensives, non pas tant sous la forme de journées de formation, mais à travers des groupes d'études et des réseaux de coopération et de stages sur le lieu de travail. La formation des enseignants doit s'opérer selon un processus continu et cyclique, parce les enseignants ont besoin de deux à trois ans pour intégrer de nouvelles approches dans leurs pratiques pédagogiques. Il est aussi important de développer les pratiques collectives de participation et de coopération dans l'établissement entre des enseignants exerçant dans un même niveau ou une même discipline. Le formateur doit les aider à analyser les situations d'enseignement et d'apprentissage au travers de pratiques d'évaluation et d'observation de classes. L'accent doit être mis sur l'amélioration des connaissances des travaux de recherche, notamment ceux concernant l'efficacité des pratiques pédagogiques.

Parallèlement aux changements dans la formation, les fonctions et les missions des enseignants dans l'établissement scolaire ont été complètement révisées (Butt & Gunter, 2007). Les enseignants ont dû prendre en charge la diversité des élèves, les faire progresser selon un apprentissage individualisé, respectueux de leur culture et de leur niveau d'acquisition, tout en obtenant des résultats mesurables. Ils ont étendu leur rôle en dehors de la classe en devenant des formateurs de leurs collègues dans l'établissement, des chefs de projet ou des coordonnateurs transdisciplinaires, des relais de la direction de l'établissement, des « facilitateurs » pour expérimenter ou innover, des chercheurs enquêteurs sur leurs propres pratiques, des formateurs de terrain, des responsables d'équipes, des animateurs de réseau. Ces enseignants sont évalués non seulement au regard de leur expertise disciplinaire et de l'atteinte de leurs objectifs, dans leur capacité à participer à la politique de l'établissement et au travail en équipe, à être innovant, à entretenir leurs compétences professionnelles et même parfois à mettre en œuvre efficacement les réformes.

Le travail du chef d'établissement s'est singulièrement transformé lui aussi. L'organisation scolaire s'est affranchie d'une séparation stricte entre l'administration et la pédagogie et le pilotage (du seul chef d'établissement) est devenu *self-management* (porté par les personnels de l'établissement), cherchant à concilier une plus grande autonomie pour améliorer les résultats des élèves. Mais ce système a été critiqué pour ses défaillances (Barroso, 2007). Le chef d'établissement était tellement accaparé par des

tâches administratives et financières qu'il en oubliait l'organisation pédagogique, les enseignants et les élèves. Les membres du conseil local, à qui avaient été déléguées de nouvelles compétences, n'avaient pas toujours l'expertise nécessaire pour assurer une bonne gestion. Après cette première phase de décentralisation, le contrôle s'est donc renforcé sur le management et les établissements scolaires, à travers l'adoption de standards, de cibles à atteindre, de *benchmarks*, dans le but d'améliorer leur performance (Gleeson & Husbands, 2001). En plus d'une critique de la pression à la performance, beaucoup de chercheurs ont montré qu'au-delà des aspects fonctionnels et techniques du management, la dimension humaine et les relations interpersonnelles étaient nécessaires à l'amélioration des résultats. La notion de leadership est devenue alors une référence incontournable pour organiser un nouveau partage des rôles et des responsabilités entre équipe de direction et enseignants de l'établissement (Normand, 2010b). Le leadership est généralement considéré par les chercheurs et les responsables de l'éducation comme la propriété émergente d'un groupe, une dynamique collective de travail où s'articulent des initiatives individuelles et différentes formes d'expertise au service d'un projet commun. Les membres de l'équipe éducative sont invités à partager des connaissances et des compétences en les adaptant aux situations vécues, lesquelles changent au rythme du projet d'établissement et en fonction de l'évaluation des résultats atteints.

La notion de « leadership distribué » s'est particulièrement imposée au cours de ces dernières années pour rendre compte du changement des relations de travail dans l'établissement scolaire. Celui-ci est dû au développement des espaces numériques de travail permettant de saisir et d'échanger en temps réel beaucoup d'informations sur les élèves, et à la diversification des formes d'évaluation obligeant les enseignants à rendre compte et à se coordonner selon des progressions pédagogiques et des projets communs (Spillane, Halverson & Diamond, 2008). Ces derniers ne sont pas enfermés dans des rôles ou des statuts définis *a priori* mais ont défini de nouveaux rapports de coopération visant à changer leurs pratiques, à aider les élèves à apprendre et à améliorer leurs résultats. Le leadership distribué n'est jugé efficace que s'il s'appuie sur une communauté d'apprentissage professionnel, c'est-à-dire quand les membres de l'équipe éducative sont invités, dans l'établissement ou à l'extérieur, à enquêter sur leurs propres pratiques, à partager leurs connaissances et leur expertise, à construire des compétences collectives

au service de la réussite des élèves. Évidemment l'action du chef d'établissement est décisive dans la création de ces conditions favorables à la dynamique de changement et à l'amélioration des résultats. Lui-même est considéré comme un leader qui développe et entretient des partenariats avec d'autres établissements scolaires, dans un réseau qui partage les innovations pédagogiques, initie des dispositifs et des programmes transversaux, identifie les bonnes pratiques et organise leur transfert, définit une politique commune à l'échelon local (leadership systémique).

Enseignant enquêteur sur ses propres pratiques, travail des enseignants en équipe à partir des résultats obtenus par les élèves, recueil, partage et analyse des données à l'échelle de l'établissement : il est facile de comprendre que la mise en œuvre de l'obligation de résultats s'est accompagnée d'un développement des systèmes d'information et d'une extension des fonctions de l'évaluation à différentes échelles (la classe, l'établissement, la communauté locale, le territoire national). L'évaluation externe, manifeste à travers des tests standardisés et des inspections prescriptives, s'est assouplie pour mieux s'articuler avec l'évaluation interne à l'établissement (par exemple avec la démarche d'auto-évaluation) et formaliser l'information chiffrée en tenant compte de la diversité des usages et des intérêts en présence. C'est localement que s'organise aujourd'hui ce système d'information, parce que les experts et les décideurs politiques pensent assez largement que cette échelle de gouvernance en réseaux d'établissements permettra d'améliorer les résultats des élèves dans un partage équilibré des compétences entre autonomie locale et régulation par les autorités publiques. Ils attendent beaucoup d'un changement des pratiques pédagogiques des enseignants et d'un renforcement de l'usage de l'évaluation aux côtés des procédures et des dispositifs existants. Deux tendances majeures dans la diversification des approches de l'évaluation formative se dessinent à l'heure actuelle et les contributions présentées dans ce numéro y font plus ou moins explicitement référence.

La première forme d'évaluation, renforcée par les développements des TICE, qualifiée parfois d'évaluation numérique (*e-assessment*), tend à centrer le travail d'enquête de l'enseignant sur la production de preuves fournies par les données collectées sur la réussite des élèves. Elle se réclame d'une épistémologie de la recherche en éducation qui donne priorité à « ce qui marche » (*what works*) dans les pratiques pédagogiques, en termes d'efficacité et d'augmentation des résultats aux tests (Hammersley, 2007). C'est

donc une approche expérimentale de l'évaluation qui est proposée aux enseignants, confortée par le postulat selon lequel la recherche est capable de leur fournir des référentiels ou des bonnes pratiques à partir d'expérimentations et d'évaluations réalisées dans d'autres contextes d'enseignement et d'apprentissage. Le travail de l'enseignant consiste donc à tenir compte de ces données en temps réel pour ajuster ses interventions auprès des élèves en difficulté et se coordonner avec les collègues de sa discipline ou de sa classe afin d'envisager des pratiques plus efficaces. Les enseignants sont aussi entourés d'experts qui les aident à analyser les données, les comparer et décider des moyens d'action selon une démarche considérée comme rigoureuse, fiable, fortement articulée aux résultats produits par la recherche quantitative.

L'autre forme d'évaluation est davantage construite en réaction aux dérives du *testing*, ce qui explique qu'elle est peu développée aux États-Unis, contrairement à la première. Elle s'inspire du travail réalisé par deux chercheurs de King's College, Paul Black et Dylan William, qui ont mis en œuvre une expérimentation concernant le développement professionnel continu des enseignants sur l'évaluation formative dans deux districts anglais, à la suite d'une synthèse internationale de la littérature de recherche sur ce thème (Black & William, 1998). Leur travail répondait aussi aux préoccupations des responsables de l'éducation anglais qui souhaitaient assouplir les conditions d'évaluation des élèves sans renoncer à l'amélioration des résultats aux tests. Alors qu'ils travaillaient avec ces chercheurs, les enseignants étudiés transformaient leurs pratiques d'évaluation selon différentes orientations. La première correspondait à un changement du mode d'interrogation orale des élèves et à une augmentation du temps qui leur était laissé pour organiser leur réponse. Les enseignants ont ensuite révisé la façon dont ils notaient le travail à la maison en concevant des commentaires utiles à la progression des élèves et à leur motivation. Puis la démarche a été étendue aux élèves eux-mêmes au travers de pratiques d'auto-évaluation et d'évaluation entre pairs. Pour finir, les élèves furent invités à concevoir la grille et les critères d'évaluation de leurs travaux avec les enseignants. Cette approche de l'évaluation, qualifiée d'évaluation pour les apprentissages (*assessment for learning*) a produit des effets importants en termes d'amélioration de la motivation et de l'estime de soi des élèves. Les décideurs politiques ont surtout été fortement marqués par les résultats obtenus par les élèves aux évaluations nationales, par comparaison avec d'autres

groupes, au point que cette forme d'évaluation est aujourd'hui inscrite dans les instructions officielles et qu'elle a été prise comme référence par de nombreux pays. Cette approche de l'évaluation est différente de la première même si elle poursuit un but équivalent : améliorer les résultats des élèves aux tests. Elle part du principe que les enseignants doivent être fortement associés à l'expérimentation pédagogique et que celle-ci, en partant de leurs pratiques habituelles, opère des déplacements successifs dont les bienfaits peuvent être appréciés assez rapidement dans la classe. Sans cette implication, l'expérimentation risque d'échouer parce que les contraintes sont trop fortes et que l'expertise des experts semble trop éloignée de l'expérience des enseignants, alors même que la responsabilité et le professionnalisme de ces derniers sont ignorés.

La diversification des méthodes d'évaluation fait débat dans la communauté de recherche anglo-saxonne. Des positions de compromis peuvent être trouvées, comme le montrent les développements du logiciel néo-zélandais qui, au prix certes d'un vrai travail d'ingénierie, est suffisamment flexible pour conjuguer évaluation *des* apprentissages (*assessment of learning*) et évaluation *pour* les apprentissages (*assessment for learning*). L'heure est encore aux expérimentations mais, d'après les contributions de ce dossier, les pays anglo-saxons sont bien à la recherche d'une seconde forme d'obligation de résultats, plus « intelligente » que la première. Cette politique considère les élèves comme des acteurs à part entière, les implique dans des stratégies d'apprentissage et les mobilise dans la dynamique de changement de l'établissement scolaire. Elle demande un effort continu de formation des enseignants et des liens plus étroits avec la recherche et ses résultats. Elle implique un nouveau partage des informations, des rôles et des responsabilités entre équipe de direction et équipes pédagogiques, en rupture avec une ligne de commande hiérarchique. Elle s'organise selon un réseau de partenaires et des initiatives locales qui peuvent être ensuite reconnues et diffusées à plus large échelle.

## PRÉSENTATION DES CONTRIBUTIONS DU DOSSIER

Ce parcours s'ouvre par un article de Brian Annan, qui a animé le programme de développement des établissements en Nouvelle-Zélande : comment un

projet fondé sur l'autonomie des établissements peut-il être piloté par l'administration centrale ? Pour faire face à cette difficulté, la Nouvelle-Zélande a élaboré un dispositif d'accompagnement qui relève de l'ingénierie politique. À partir de l'expérience du Royaume-Uni, Gordon Stobart pose une question inverse : comment des initiatives locales, considérées comme des réussites, peuvent-elles être récupérées par l'administration centrale et par les politiques ? Comment peuvent-ils en diffuser les résultats ? Quels sont les déplacements, les traductions, qui permettent de passer non seulement d'une échelle à une autre (classe, établissement, etc.) mais aussi d'un monde à un autre (univers cognitifs des enseignants, des élèves et de leurs familles, etc.) ?

Un processus fondé sur les initiatives locales suppose évidemment de nouvelles compétences chez les enseignants : savoir analyser collectivement une situation, utiliser les résultats des évaluations des élèves pour mettre en évidence les réussites et les échecs pédagogiques, réajuster les choix curriculaires et les méthodes d'enseignement à partir des résultats de ces analyses, etc. Helen Timperley, qui a suivi le processus néo-zélandais en y apportant ses compétences dans le domaine de la littératie, présente la démarche qui a été choisie pour développer ces compétences : un accompagnement du développement professionnel des enseignants. Celui-ci se différencie sensiblement des opérations de formation continue qui nous sont familières, dans la mesure où il est totalement intégré à la politique d'autonomie des établissements. La dynamique de l'opération repose sur les enquêtes que les acteurs locaux (enseignants, personnels d'éducation, d'encadrement, parents, élèves, etc.) mènent sur leurs pratiques. Les connaissances, les méthodes nécessaires sont apportées par un contact avec la recherche, développé *in situ*. Cette démarche favorise évidemment l'engagement des acteurs et spécialement des enseignants.

Les deux autres articles rendent compte d'expériences américaines, à partir d'études menées par deux jeunes chercheuses du CPRE. Marian Robinson a étudié les effets du programme *Children First* développé dans l'État de New York. Ce programme essaie de traiter de front le défi d'une obligation de résultats intelligente. Si la pression de l'évaluation est forte, elle se double de la mise à la disposition des acteurs, et spécialement des chefs d'établissement, de ressources nouvelles : un système de données et d'évaluations formatives qui les aide à analyser la qualité des apprentissages. Cette démarche est sans doute

promise à un certain avenir et un nouveau marché s'ouvre : des prestataires de service privés sont prêts à intervenir si les ressources fournies par le service public apparaissent insuffisantes. Cette étude est complétée par celle de Leslie Nabors Oláh, qui compare l'usage des évaluations intermédiaires dans les établissements des deux districts scolaires de Cumberland et de Philadelphie. Dans tous les cas, les enseignants croulent sous les propositions publiques et privées, qui s'ajoutent aux épreuves qu'ils ont fabriquées eux-mêmes. L'auteur étudie la manière dont ils s'orientent dans ce maquis et les compétences critiques dont ils font preuve dans l'interprétation et l'usage des informations qui leur sont fournies.

En quoi ces expériences anglo-saxonnes peuvent-elles nourrir la réflexion de l'Europe continentale et son projet de compétences clés pour la formation

tout au long de la vie ? Et de façon plus spécifique, l'étude de la mise en place, en France, du socle commun défini par la loi de 2005 ? Matthis Behrens, qui a travaillé avec son institut (l'Institut de recherche et de documentation pédagogique de Neuchâtel) à la définition des standards pour l'enseignement fondamental en Suisse, présente son analyse dans un article qui conclut ce dossier. Nous remercions Thierry Bessy pour la traduction des articles anglophones de ce dossier.

Romuald Normand  
Romuald.normand@ens-lyon.fr  
Institut français de l'Éducation,  
École normale supérieure de Lyon

Jean-Louis Derouet  
Institut français de l'Éducation,  
École normale supérieure de Lyon

#### BIBLIOGRAPHIE

- ALTON-LEE A. (2003). *Quality teaching for diverse students in schooling. Best evidence synthesis*. Wellington : Ministry of Education.
- BARROSO J. (2007). « L'autonomie et la gestion de l'école publique : entre le marché, le managérialisme et la démocratie ». In J.-L. Derouet & R. Normand (dir.), *L'Europe de l'éducation : entre management et politique*. Lyon : INRP ; Poitiers : ESEN.
- BERLINER D. & BIDDLE B. (1995). *The manufactured crisis: Myths, fraud, and the attack on America's public schools*. Cambridge : Perseus Books.
- BLACK P. & WILLIAM D. (1998). *Inside the black box: Raising standards through classroom assessment*. Londres : King's College Press.
- BUTT G. & GUNTER H. (2007). *Modernizing schools. People, learning and organizations*. Londres : Continuum.
- CONLEY D. (2003). *Who governs our schools? Changing roles and responsibilities*. New York : Teachers College Press.
- DALE R. (2001). « Constructing a long spoon for comparative education: Charting the career of the "New Zealand model" ». *Comparative Education*, vol. 37, n° 4, p. 493-500.
- DARLING-HAMMOND L. & SYKES G. (1999). *Teaching as the learning profession: Handbook of policy and practice*. San Francisco : Jossey-Bass.
- DEROUET J.-L. (2006). « Entre la récupération des savoirs critiques et la construction des standards du management libéral : bougés, glissements et déplacements dans la circulation des savoirs entre recherche, administration et politique en France de 1975 à 2005 ». *Revue française de pédagogie*, n° 154, p. 5-18.
- DEROUET J.-L. & DEROUET-BESSON M.-C. (2009). *Repenser la justice dans l'éducation et la formation*. Berne : Peter Lang ; Lyon : INRP.
- DEROUET J.-L., NORMAND R. & BESSY T. (2009). *Quelles politiques pour l'égalité ? Savoirs, gouvernances et obligation de résultats. Réflexions croisées franco-américaines*. Lyon : INRP.
- ELMORE R. (2004). *School reform from the inside out*. Cambridge : Harvard Education Press.
- FIRESTONE A., SCHORR R. & MONFILS L. (2004). *The ambiguity of teaching to the test: Standards, assessment, and educational reform*. Mahwah : Lawrence Erlbaum.
- FUHRMAN S. (2001). *From the Capitol to the classroom. Standards-based reform in the states*. Chicago : National Society for the Study of Education, University of Chicago Press.
- FUHRMAN S., COHEN D. & MOSHER F. (2007). *The state of education policy research*. Mahwah : Lawrence Erlbaum.
- GEWIRTZ S. (2002). *The managerial school: Post-welfarism and social justice in education*. Buckingham : Open University Press.
- GEWIRTZ S., BALL S. & BOWE R. (1995). *Market, choice and equity*. Buckingham : Open University Press.
- GEWIRTZ S., MAHONY P., HEXTALL I. & CRIBB A. (2008). *Changing teacher professionalism: International trends, challenges and ways forward*. Londres : Routledge.
- GILLBORN D. & YOUDELL D. (2000). *Rationing education. Policy, practice, reform and equity*. Buckingham : Open University Press.
- GLEESON D. & HUSBANDS C. (2001). *The performing school. Managing learning and teaching in a performance culture*. Londres : Routledge.

- GOERTZ M. (2001). « Standards-based accountability: Horse trade or horse whip? » In S. Fuhrman (dir.), *From the Capitol to the classroom. Standards-based reform in the states*. Chicago : National Society for the Study of Education, University of Chicago Press, p. 33-39.
- HAMMERSLEY M. (2007). *Educational research and evidence-based practice*. Londres : Sage et Open University Press.
- HARGREAVES A. (1993). « Teacher development in the postmodern age: Dead certainties, safe simulation and the boundless self ». *Journal of Education for Teaching*, vol. 19, n° 4, p. 95-112.
- HARGREAVES A. (2003). *Teaching in the knowledge society*. New York : Teachers College Press.
- HARRIS A. (2002). *School improvement: What's in it for schools?* Londres : Falmer Press.
- HOPKINS D. (2007). *Every school is a great school. Realizing the potential of system leadership*. Londres : Open University Press.
- HURSH D. (2007). « Assessing No Child Left Behind and the rise of neoliberal education policies ». *American Educational Research Journal*, vol. 44, n° 3, p. 493-518.
- HURSH D. (2008). *High-stake testing and the decline of teaching and learning. The real crisis in education*. Lanham : Rowman & Littlefield Publishers.
- LIPMAN P. (2004). *High stakes education. Inequality, globalization and urban school reform*. Londres : Routledge Falmer.
- MacBEATH J (1999). *Schools must speak for themselves. The case for school self-evaluation*. Londres : Routledge.
- MacBEATH J. (2006). *School inspection and self-evaluation. Working with the new relationship*. Londres : Routledge.
- McGUINN P. (2005). *No Child Left Behind and the transformation of federal education policy, 1965-2005*. Lawrence : University Press of Kansas.
- MAHONY P. & HEXTALL I. (2000). *Reconstructing teaching. Standards, performance, and accountability*. Londres : Routledge Falmer.
- NICHOLS S. & BERLINER D. (2007). *Collateral damage: The effects of high-stakes testing on America's schools*. Cambridge : Harvard Education Press.
- NORMAND R. (2008). « Standards scolaires ». In A. van Zanten (dir.), *Dictionnaire de l'éducation*. Paris : PUF, p. 634-636.
- NORMAND R. (2010a). *Gouverner la réussite scolaire. Une arithmétique politique des inégalités*. Berne : Peter Lang.
- NORMAND R. (2010b). « Le leadership dans l'établissement scolaire. Un nouveau partage des responsabilités entre chef d'établissement et enseignants ». *Administration & éducation*, vol. 127, n° 3, p. 27-35.
- PHILIPS D. (2000). « Curriculum and assessment policy in New Zealand: Ten years of reforms ». *Educational Review*, vol. 52, n° 2, p. 143-153.
- RADNOR H. & BALL S. (1996). *Local education authorities: Accountability and control*. Stoke-on-Trent : Trentham Books.
- SILVER H. & SILVER P. (1991). *An educational war on poverty*. Cambridge : Cambridge University Press.
- SPILLANE J., HALVERSON R. & DIAMOND J. (2008). « Théorisation du leadership en éducation : une analyse en termes de cognition située ». *Éducation et sociétés*, n° 21, p. 121-149.
- STOBART G. (2008). *Testing times. The uses and abuses of assessment*. Londres : Routledge.
- SUPOVITZ J. (2006). *The case for district-based reform: Leading, building, and sustaining school improvement*. Cambridge : Harvard Education Press.
- THRUPP M. (1999). *Schools make difference: Let's be realistic!* Buckingham : Open University Press.
- THRUPP M. (2001). « Sociological and political concerns about school effectiveness research: Time for a new research agenda ». *School Effectiveness and School Improvement*, vol. 12, n° 1, p. 7-40.
- TIMPERLEY H. & PARR J. (2009). « The chain of influence from policy to practice in the New Zealand literacy strategy ». *Research Papers in Education*, vol. 24, n° 2, p. 135-154.
- TIMPERLEY H., PARR J. & BERTANEES C. (2009). « Promoting professional inquiry for improved outcomes for students in New Zealand ». *Professional Development in Education*, vol. 35, n° 2, p. 227-245.
- TIMPERLEY H., WILSON A., BARRAR H. & FUNG I. (2007). *Teacher professional learning and development: Best evidence synthesis iteration (BES)*. Wellington : Ministry of Education. Disponible sur Internet à l'adresse : <<http://www.educationcounts.govt.nz/publications/series/2515/15341>> (consulté le 26 janvier 2011).
- TOMLINSON S. (2005). *Education in a post-welfare society* [2<sup>e</sup> éd.]. New York : Open University Press.
- TOMLINSON S. (2009). « Éducation et justice sociale dans la société de l'État-providence ». In J.-L. Derouet & M.-C. Derouet-Besson, *Repenser la justice dans l'éducation et la formation*. Berne : Peter Lang ; Lyon : INRP, p. 171-186.
- WOLF I. de & JANSSENS F. (2007). « Effects and side effects of inspections and accountability in education: An overview of empirical studies ». *Oxford Review of Education*, vol. 33, n° 3, p. 379-396.

