Les politiques d'*accountability*

Responsabilisation et formation continue des enseignants

Accountability policies

Die Politiken von Accountability

Las políticas de accountability

Nathalie Mons et Vincent Durpiez

Édition électronique

URL : <http://journals.openedition.org/rechercheformation/131>

DOI : 10.4000/rechercheformation.131

ISSN : 1968-3936

Éditeur

ENS Éditions

Édition imprimée

Date de publication : 1 décembre 2010

Pagination : 45-59

ISBN : 978-2-7342-1180-8

ISSN : 0988-1824

Référence électronique

Nathalie Mons et Vincent Durpiez, « Les politiques d'*accountability* », *Recherche et formation* [En ligne], 65 | 2010, mis en ligne le 01 décembre 2012, consulté le 01 mai 2019. URL : <http://journals.openedition.org/rechercheformation/131> ; DOI : 10.4000/rechercheformation.131

Les politiques d'*accountability*

Responsabilisation et formation continue des enseignants

> **Nathalie MONS**

Université Paris-Est Marne-la-Vallée, LATTS (Laboratoire techniques, territoires et sociétés, UMR 8134 CNRS)

> **Vincent DUPRIEZ**

Université catholique de Louvain, Faculté de psychologie et des sciences de l'éducation

RÉSUMÉ • La majorité des systèmes éducatifs des pays de l'OCDE a été marquée, depuis la fin des années quatre-vingt, par l'introduction concomitante de deux politiques majeures : le développement de nouveaux dispositifs de formation continue des enseignants dans une perspective de formation tout au long de la vie et la mise en place d'évaluations standardisées des élèves, enchâssées dans des politiques plus larges d'*accountability*. Cet article interroge les liens qui existent entre ces deux phénomènes, en mobilisant à la fois les cadres théoriques qui ont lié ces deux grandes familles de réformes et en analysant les politiques réellement mises en œuvre dans les pays de l'OCDE.

MOTS-CLÉS • Formation continue des enseignants, évaluations externes, éducation comparée

Les systèmes éducatifs des pays de l'OCDE ont été marqués depuis la fin des années quatre-vingt par des mutations d'ampleur mettant en avant l'avènement de nouvelles politiques éducatives en nette rupture avec la forme traditionnelle de l'organisation scolaire qui avait dominé depuis la fin du XIX^e siècle (Maroy, 2008 ; Mons, 2007). Parmi les traits saillants de ces nouvelles politiques, se sont imposées, dans la très grande majorité des pays de l'OCDE, deux réformes majeures : le développement de nouveaux dispositifs de formation continue des enseignants et la mise en place d'évaluations standardisées des élèves (Carnoy & Loeb, 2002 ; Eurydice, 2007 ; Mons, 2009). Ces dernières sont clairement enchâssées dans des politiques plus larges d'*accountability* que l'on peut définir comme le couplage de mesures de la production du système éducatif (le plus souvent les acquis des élèves) et de responsabilisation des établissements scolaires, eu égard aux résultats obtenus par leurs élèves.

Dans la littérature d'origine anglo-saxonne, qui domine le discours politique et la réflexion scientifique sur l'évaluation standardisée, ces deux dimensions du *testing* et de la formation continue des enseignants sont souvent associées. Les évaluations standardisées des élèves qui s'imposent désormais comme les instruments clés d'une meilleure reddition des comptes par les établissements scolaires et les équipes pédagogiques – acteurs privilégiés de la nouvelle responsabilisation – doivent permettre, entre autres, d'orienter le développement de plans de formation continue pour les enseignants afin d'accompagner ces derniers dans l'acquisition du professionnalisme nécessaire à l'atteinte des objectifs cognitifs testés dans le cadre des nouvelles évaluations. Ces dispositifs d'évaluations standardisées ne se veulent pas seulement incitation, contrôles et sanctions mais aussi signaux de déclenchement d'actions de remédiation et d'orientation des ressources pour l'établissement scolaire, entre autres sous forme de formation continue pour les équipes pédagogiques. Ce modèle croise ainsi deux dimensions s'inscrivant pour l'une dans une régulation par les résultats – le *testing* – et se présentant pour l'autre comme une modalité de la régulation par les normes et les ressources – le développement de plans de formation continue.

Au-delà de ce discours politique et scientifique dominant, cet article vise à interroger les cadres « théoriques » et les politiques réellement mises en place dans les pays de l'OCDE. L'analyse nous amènera à identifier dans quelle mesure ce modèle politique et scientifique s'est imposé dans les faits, c'est-à-dire dans quelle mesure cette relation entre évaluation standardisée et formation continue des enseignants est pensée et mise en œuvre dans les pays de l'OCDE.

Dans une première partie de ce texte, l'analyse détaillée des cadres de réflexion « théoriques », à la fois anglo-saxons et européens continentaux, montre que les premiers, orientés sur des dispositifs à forts enjeux de sanctions et récompenses, intègrent le plus souvent le couplage de l'évaluation standardisée avec la formation continue alors que les seconds révèlent une relation plus lâche entre ces deux dimensions. Dans une seconde partie, construite principalement à partir des résultats de deux enquêtes menées par Eurydice (2008 et 2009) auprès d'experts des systèmes éducatifs, l'étude des réformes mises en œuvre dans les pays de l'OCDE montre la même opposition entre un modèle d'*accountability* anglo-saxon qui peut s'avérer intrusif, et un régime européen continental à la fois moins directif, moins centré sur la responsabilisation de l'établissement et des enseignants, et moins préoccupé par la formation continue. En conclusion, l'article présentera quelques réflexions relatives à la signification de ces réformes, en les replaçant dans un cadre d'analyse politique plus large.

1. Évaluation standardisée et formation continue : le modèle théorique anglo-saxon concurrencé par des alternatives européennes continentales

Si d'autres politiques éducatives comme l'autonomie scolaire, la décentralisation ou encore le libre choix de l'école ont suscité des recherches théoriques significatives, bien que toujours très controversées, l'analyse théorique de l'évaluation standardisée, centrée sur les mécanismes en jeu dans l'établissement et la classe, reste à ce jour peu développée. Nous faisons toutefois le choix de considérer qu'il y a bien une ébauche théorique relative à la régulation par l'évaluation standardisée, dans la mesure où certains auteurs ont tenté d'en décoder les modalités d'influence et d'action, au-delà d'une prescription des voies d'intervention. C'est à une mise en évidence de ces principes généraux, souvent plutôt supposés qu'observés, que nous allons consacrer cette première partie. Au niveau des apprentissages, l'objectif de l'introduction du *testing* est une amélioration des acquis des élèves. Pour autant, de nombreux auteurs (Linn, 2001 ; Nichols, 2007...) pointent du doigt le fait que les modèles utilisés expliquent rarement par quels processus intermédiaires l'évaluation standardisée est censée conduire à une progression des acquis des élèves. Cette fragilité du cadre théorique explique que la liaison entre évaluation standardisée et formation continue soit pensée à la marge.

Une série d'arguments, assez vagues, sont avancés par les défenseurs de ces dispositifs. Suivant les types d'*accountability* développés – « *accountability* dure » fondée sur l'organisation de tests à forts enjeux (*high stakes testing*) pour les élèves et les écoles sur le modèle américain ou anglais, et ce que l'on pourrait qualifier d'« *accountability* réflexive » (le modèle de l'école européenne continentale) –, les modes d'explication diffèrent, voire s'opposent quant à la place donnée à la formation continue comme processus médiateur entre l'évaluation externe et les pratiques enseignantes.

1.1 Le modèle de l'« *accountability* dure » dans le monde anglo-saxon

Le modèle de l'« *accountability* dure » (Goodwin, Eglert *et al.*, 2002 ; Carnoy & Loeb, 2002 ; Hanushek & Raymond, 2005 ; Mc Donnell, 2005 ; Phelps, 2005 ; Wöessmann, 2007) attache aux résultats des tests une série de sanctions et de ressources qui peuvent avoir de lourdes conséquences pour les écoles, les enseignants ou les élèves (financement de l'école, diplôme ou redoublement des élèves...). Ce modèle est fondé sur l'existence des mécanismes suivants qui, déclenchés par le *testing* et fonctionnant de concert dans une dynamique positive, permettraient de rendre le système éducatif plus efficace :

- il est attendu des élèves un travail plus important et donc, au final, de meilleurs apprentissages du fait à la fois d'une définition plus claire des objectifs scolaires à atteindre, de la désignation de champs d'enseignement jugés prioritaires et

de la pression psychologique exercée par le test lui-même puisqu'il conditionne leurs carrières scolaires ;

- pour les personnels administratifs et politiques encadrant localement les établissements scolaires, l'évaluation standardisée permettrait une meilleure connaissance des difficultés locales, la mise à disposition de ressources quand l'établissement présente des défaillances en termes de résultats des élèves et la mise en œuvre de mesures efficaces de sanctions quand l'établissement ne parvient pas à améliorer ses résultats dans les délais impartis (changement de leadership, restructuration ou fermeture de l'école) ;
- les équipes pédagogiques constituent la cible privilégiée de ces nouveaux dispositifs. Les enseignants, rendus responsables des résultats des élèves devant leurs supérieurs hiérarchiques et la société civile par la publication des performances de leur école, travailleraient davantage à la réussite des élèves et amélioreraient leur professionnalisme aussi bien par des formations que des échanges entre collègues au sein des écoles, les résultats aux tests donnant lieu à des analyses réflexives dans les établissements.

Pour Carnoy et Loeb (2002) notamment, dans le monde anglo-saxon, et en particulier aux États-Unis dans les années quatre-vingt, ces politiques incorporent bien deux dimensions complémentaires : d'un côté, des standards et des évaluations qui permettent de guider les acteurs locaux en termes de contenus d'enseignement et, de l'autre, des dispositifs d'amélioration des capacités des écoles (*capacity building*). « Les réformes fondées sur les standards ne partent pas du simple supposé que les standards et évaluation permettent à eux seuls d'améliorer l'éducation. Les réformes éducatives doivent également améliorer les capacités des enseignants et administrateurs à fournir une bonne éducation » (Carnoy & Loeb, 2002, p. 307), notamment à travers la formation continue défendent des auteurs qui se fondent sur des recherches empiriques liant *accountability*, formation et performances scolaires (Darling-Hammond, 1997).

Si certains cadres théoriques nord-américains font bien une place à la formation continue comme composante des dispositifs d'*accountability*, leurs auteurs soulignent eux-mêmes que cette vision large de la réforme scolaire ne s'est pas imposée à toutes les époques. Les années soixante-dix ont été marquées par des réflexions théoriques fondées sur une vision plus rustique des effets de l'évaluation standardisée. Associés à de forts enjeux, c'est principalement par les contraintes incorporées que ces dispositifs agiraient sur les acteurs en rétablissant à la fois un même niveau d'information entre le donneur d'ordre et le fournisseur de service, et un stimulant pour les équipes pédagogiques. Ces écrits rejoignent notamment le courant de la *Personnel Economics*, dérivée de la théorie de l'Agence, selon laquelle, très schématiquement, la mesure des performances des prestations des enseignants (à travers les résultats de leurs élèves aux tests) permet une meilleure connaissance de la réalité de leur travail limitant ainsi les effets d'opportunité. Fondé sur une

défiance face au travail des agents publics et dérivé de la théorie de l'Agence, ce mouvement de pensée anglo-saxon n'incorpore pas de relation entre évaluation standardisée et formation continue, le seul contrôle social des acteurs opérationnels du système devant permettre une amélioration des performances des élèves.

Si les écrits théoriques anglo-saxons n'incorporent qu'à certaines époques le lien entre l'évaluation des élèves et la formation continue des enseignants, les réflexions européennes sur le sujet font encore moins de cas de cette relation.

1.2 Le modèle de l'« accountability réflexive » en Europe continentale

Nous présentons ici les deux modèles principaux, qui pour l'un a irrigué la réflexion sur les standards et leur évaluation en Allemagne, en Suisse et en Autriche, dans le cadre du rapport Klieme et, pour l'autre, fut développé en France par Thélot sous le nom de l'« effet miroir ». À nouveau, davantage qu'une réelle construction théorique, il s'agit à vrai dire d'ébauches de mises en ordre du social, à travers des métaphores et des schèmes de relations non soumis à une évaluation empirique.

Dans le cadre du rapport *Le développement de standards nationaux de formation : une expertise*, Klieme, chercheur à l'Institut allemand de recherche pédagogique internationale (DIPF), et son équipe préconisent la mise en place de « standards de résultat » (Klieme, 2004, p. 48). Ces « *outputs standards* » décrivent des objectifs d'enseignement dont l'atteinte doit être validée par des évaluations standardisées par opposition aux « standards de contenus » ou « *input standards* » qui sont axés sur les contenus à enseigner. Selon l'équipe dirigée par Klieme, ces dispositifs qui allient standards et *testing* auraient une mission centrale : « ils font ressortir sous une forme claire et concise ce qui importe dans le cadre d'un système scolaire. Cette fonction d'orientation est utile tant pour les élèves que pour leurs parents ; mais elle sert aussi au professionnalisme des enseignants et au développement de la qualité au niveau institutionnel. Concrétisés en procédés de test, les standards sont utilisés dans le cadre du monitoring de la formation et de l'évaluation des établissements scolaires. [...] Leur objectif est d'analyser les effets (et les effets secondaires) des méthodes pédagogiques et de permettre ainsi un agir professionnel, rationnel. Les enquêtes sous forme de tests ne sont donc utiles que si elles contribuent à développer le professionnalisme du personnel enseignant ainsi que la qualité de l'école et de l'enseignement » (Klieme *et al.*, 2004, p. 46). Par contre, contrairement au modèle de l'« *accountability* dure » qui rend les équipes pédagogiques redevables devant leurs concitoyens grâce à la publication des performances des établissements auprès du grand public, le modèle Klieme préconise que « les feed-back [soient] adressés au corps enseignant et aux organes d'un établissement scolaire et non au grand public. [...] Nous voyons dans la mesure des acquis scolaires (compétences des élèves) une opportunité pour les écoles de vérifier les résultats du travail réalisé et d'y réagir professionnellement » (p. 53).

L'accent n'est mis que très marginalement sur la formation continue des enseignants. Le standard de performance ne sert pas prioritairement à orienter la conception de plans de formation. Quand le rapport Klieme évoque la formation continue, il s'agit principalement d'une formation à la compréhension et aux usages des standards, ou au mieux d'une formation didactique disciplinaire en lien avec les standards et les tests. « Accompagnés d'une formation continue adéquate, ils [les standards] peuvent faire progresser la réflexion en didactique des disciplines et la compétence de diagnostic des enseignants (p. 50) ».

Autre cadre « théorique » développé autour de l'évaluation standardisée en France, le modèle de « l'effet miroir », proposé par Thélot (2002), ancien directeur de la direction chargée de la prospective et des statistiques au ministère de l'Éducation, s'oppose aussi à la vision d'une « *accountability* dure », imposant à la fois des sanctions et des bonnes pratiques véhiculées *via* la formation continue. Pour lui, l'évaluation standardisée doit réussir l'« effet miroir ». Cela signifie que l'évaluation ne doit pas nécessairement donner des schémas explicatifs des méthodes pédagogiques efficaces, parce que le politique n'en détient pas toujours la connaissance mais qu'elle doit permettre de confronter les acteurs du système éducatif avec les résultats de leurs actions. Il faut renvoyer à ces derniers une image de leurs pratiques afin qu'ils puissent les améliorer si elles ne donnent pas lieu aux résultats escomptés. L'« effet miroir » est une modélisation qui s'appuie sur l'incitation à une pratique réflexive face aux résultats des élèves et sur une sanction symbolique. C'est le constat d'une incapacité à définir des normes pédagogiques efficaces, et donc des bonnes pratiques qui pourraient être apportées aux enseignants par le biais de la formation continue, qui conduit l'auteur de ce modèle théorique à préconiser une régulation par les résultats, seule capable, selon lui, de faire évoluer dans le sens d'apprentissages plus efficaces les acteurs opérationnels du système éducatif.

Au total, cette analyse de la littérature théorique développée autour de l'évaluation standardisée montre que le modèle anglo-saxon dominant, centré sur la responsabilisation des équipes pédagogiques, des dispositifs de sanctions et de récompenses, et des actions de soutien (et de contrôle social) des équipes au travers de la formation continue est concurrencé par des modèles européens continentaux qui inscrivent davantage cet outil dans un cadre d'autonomie scolaire locale, sur le plan pédagogique du moins.

2. Réformes dans l'OCDE : des liens entre formation continue et évaluation standardisée très variables, selon les pays

Les réformes mises en œuvre dans les pays de l'OCDE autour de l'évaluation standardisée et de la formation continue des enseignants reflètent bien ces visions théoriques différentes voire contradictoires. Selon les pays, les relations entre le dispositif de *testing* et la formation continue peuvent suivre trois principales

configurations : premier cas – ultra-minoritaire –, la formation continue, dans une perspective généraliste, est une composante essentielle des programmes d'*accountability* ; seconde situation, plus courante, la formation continue est bien intégrée à la réflexion sur l'évaluation standardisée mais dans une acception techniciste (elle est centrée sur l'apprentissage des standards) ; troisième figure, largement dominante, les dispositifs d'évaluation standardisée ne comprennent aucune composante de formation.

Pour conduire cette analyse, nous avons mobilisé le corpus réuni dans deux études empiriques menées avec le réseau européen Eurydice autour des responsabilités et de l'autonomie des enseignants (2008) et des dispositifs d'évaluation standardisée (2009)¹.

2.1 Première configuration : des relations fortes entre formation continue et évaluation standardisée

Nous retrouvons bien, influencée par le cadre théorique anglo-saxon, cette vision d'un lien fortement présent entre évaluation standardisée et formation continue dans certains pays de cette ère culturelle. La politique menée en Angleterre dès la fin des années quatre-vingt, dans le cadre des orientations « thatchériennes » puis « blairistes », va dans ce sens.

En 1988, l'*Education Reform Act* (ERA) marque, en effet, au Royaume-Uni un tournant majeur en introduisant simultanément le premier programme scolaire national – le National Curriculum –, assorti de standards, une politique de reddition des comptes explicite vis-à-vis des écoles, une nouvelle liberté de choix de l'établissement scolaire ainsi qu'une dose accrue d'autonomie scolaire, hors pédagogie. La loi spécifie que chacun des cycles scolaires (les *Key stages*) doit être clos par une évaluation, mais laisse toute latitude au ministre de l'Éducation de l'époque quant à leur organisation concrète. C'est donc le *Secretary of State for Education*, Kenneth Baker – conservateur – qui décida que le *National Curriculum* récemment promu ne suffirait pas à améliorer les résultats des écoles sans un dispositif de *testing* complet, visant à donner des informations précises aux élèves,

1 Le Réseau européen Eurydice réunit des unités opérationnelles dans chacun des pays de l'Union Européenne. Il vise à travers des études thématiques à analyser le fonctionnement des systèmes éducatifs et les politiques mises en œuvre. À partir d'une grille de contenu standardisée élaborée par l'unité européenne, les experts des différents pays renvoient des contributions nationales qui consistent en un corpus conséquent pour chaque étude. Pour cet article, nous analysons la partie des contributions de l'étude *Responsabilités et autonomie des enseignants en Europe* (Eurydice, 2008) qui porte sur la formation continue des enseignants. Les questions posées dans la grille de contenu visaient à appréhender non seulement le cadre légal du développement de la formation continue mais aussi à travers certains critères à appréhender la réalité de la mise en œuvre (heures réellement affectées à cette mission, caractère obligatoire en lien avec les carrières, capacité de l'institution à libérer l'enseignant pour le temps nécessaire à la formation, etc.). Cet article est également fondé sur l'analyse des contributions des pays européens à l'étude *Les évaluations standardisées des élèves en Europe : objectifs, organisation et utilisation des résultats* (Eurydice, 2009) qui visait à la fois à mieux connaître les outils développés dans les différents pays sous ce label générique, la genèse de ces politiques et leurs usages notamment en termes de formation continue.

parents et enseignants. À l'origine de sa mise en œuvre, le mandat du groupe opérationnel – le *Task Group on Assessment and Testing* (TGAT) – chargé de définir ce dispositif mit clairement en évidence le lien entre les tests et la formation continue des enseignants. C'est ainsi que se mirent en place dès le début des années quatre-vingt-dix les évaluations SATS (Standard Assessment Task) dans trois domaines : mathématiques, anglais et sciences.

Ce dispositif rencontra immédiatement une forte opposition des enseignants. Dès 1993, s'organisa un boycott qui conduisit à une forte réduction du nombre de tests.

Aujourd'hui, malgré la poursuite de cette opposition virulente des syndicats d'enseignants, les SATS demeurent un des points d'ancrage du système éducatif en général et de la réflexion sur les politiques d'établissement en particulier, notamment en matière de formation continue. Un des rôles de l'OFSTED², l'inspection anglaise, est de définir des actions dans ce domaine sur la base, entre autres, des résultats des élèves aux tests dans le cadre de l'évaluation des établissements.

Un même lien entre formation continue et résultats aux évaluations standardisées est aussi mis en évidence par la législation américaine bien que moins explicitement et de façon variable selon les époques. En effet, ce lien n'a pas toujours existé. Les programmes d'évaluation standardisée des années quatre-vingt par exemple, fondés entre autres sur le concept de *basic skills* attendaient surtout du *testing* une remotivation des équipes enseignantes et des élèves conduisant à une amélioration générale des performances scolaires. L'absence de résultats probants ainsi que la montée d'une nouvelle rhétorique autour de la culture d'évaluation à partir des analyses scientifiques initialement prometteuses des expériences du Texas (l'État dont le président Georges Bush était originaire) et du district de Chicago a donné une nouvelle impulsion à cette politique (Carnoy & Loeb, 2002 ; Mons, 2009). En 2001, la loi *No Child Left Behind* (NCLB), adoptée par les républicains et les démocrates, impose l'obligation aux États fédérés de mettre en place des dispositifs d'évaluation standardisés en lien avec les financements fédéraux. Alors que se développent des standards de contenus nationaux notamment en mathématiques, des normes de qualification pour les enseignants sont également imposées afin d'améliorer le faible niveau de certification de certains enseignants, notamment ceux exerçant dans les écoles publiques des districts les plus démunis. Ce niveau de qualification peut être atteint par la formation continue et notamment certains programmes de master spécialisés en éducation. Cependant, si la formation continue est bien intégrée comme composante de ces nouveaux dispositifs d'*accountability*, la politique conçue par la loi NCLB est tout autant l'imposition d'une certification initiale qu'une promotion de la formation continue des enseignants. Le président Obama a poursuivi cette politique d'*accountability* depuis son élection (la politique

2 Office for Standards in Education, Children's Services and Skills.

du district de Chicago influençant fortement sa réflexion et celle de son ministre de l'Éducation issu de ce secteur géographique). Le nouveau programme *Race to the Top*, fondé sur des appels d'offres, qui s'organise autour d'un ensemble de fonds dédiés à des actions spécifiques (restructuration des écoles dites déficientes, poursuite du fonds évaluation et récompenses des enseignants, etc.), ne met cependant plus l'accent aussi fortement sur la formation continue comme input nécessaire à l'amélioration du système. Il prévoit davantage, à travers le fonds *Teacher Incentives funds*, une liaison entre les performances des élèves, l'évaluation des enseignants et des rémunérations au mérite.

2.2 Seconde configuration : une formation continue centrée sur l'apprentissage des standards

Au-delà du monde anglo-saxon, dans une seconde configuration plus commune, des actions de formation continue sont prévues comme une composante des systèmes d'*accountability* mais elles s'inscrivent dans une vision centrée sur l'apprentissage par les enseignants de la politique de standards au sens large. Il s'agit d'initier les enseignants à ces nouveaux dispositifs : compréhension des outils (standards, tests, modèle de compétences) et des usages pertinents de ces instruments supposés étrangers à la culture professionnelle traditionnelle des enseignants, sommés ainsi d'entrer dans une nouvelle ère de professionnalisme. Au mieux, la formation continue peut être étendue à la didactique des matières testées mais l'accent n'est aucunement mis sur des actions de formation généralistes en psychologie, pédagogie, sociologie ou politique de l'éducation. L'Autriche et l'Allemagne sont emblématiques de ce modèle.

L'Autriche n'a rejoint qu'au début des années deux mille le mouvement des évaluations standardisées qui font l'objet d'opérations pilotes approfondies étendues sur plus d'une décennie. Les enseignants sont considérés comme les agents principaux de cette réforme et sont donc fortement accompagnés dans le cadre de cette nouvelle politique. Il est ainsi prévu que les institutions de formation continue consacrent 40 % de leur investissement à des formations exclusivement dédiées à l'apprentissage des standards. Des formations généralistes ne sont pas explicitement prévues de façon à ne pas s'inscrire en rupture avec le développement parallèle d'un mouvement d'autonomie scolaire qui s'oppose par nature à la transmission de bonnes pratiques éducatives généralistes.

L'Allemagne a adopté une vision du lien entre la formation continue et l'évaluation standardisée relativement similaire depuis le développement de standards dans les années quatre-vingt-dix par la Conférence des ministres de l'Éducation et de la Culture des Länder, instance en charge d'une harmonisation inter-Länder de la politique éducative.

Bien que classée dans l'ère culturelle anglo-saxonne, l'Écosse est un cas intéressant car elle se démarque du grand frère anglais et peut être aussi rangée

dans cette seconde famille de pays. Dès 1991, cette région du Royaume-Uni a, en effet, adopté une politique de *testing* moins axée sur la reddition des comptes à la société civile que sur l'évaluation formative, outil d'apprentissage au service des enseignants. Alors qu'en Angleterre, le dispositif des SATS donna lieu à une régulation standardisée qui laissait peu de latitude aux enseignants (les tests sont élaborés nationalement, administrés selon un protocole standardisé et corrigés en externe), les enseignants écossais ont eu le choix du timing du test et le corrigent eux-mêmes. Depuis 2007, toutes les écoles sont intégrées dans le dispositif *Assessment for Learning Programme* qui prévoit entre autres des actions de formation continue des enseignants à l'usage de l'évaluation standardisée, comme outil formatif.

2.3 Troisième configuration : les évaluations standardisées sans conséquence sur la formation continue

Troisième et dernière configuration de politiques dominante dans l'OCDE, certains pays ne font aucun lien entre l'évaluation standardisée et la formation continue. C'est le cas de la Finlande dont le programme de tests national repose sur un échantillon d'élèves et vise principalement à évaluer les performances scolaires du système éducatif national et celui des collectivités locales, qui peuvent aussi développer leur propre dispositif d'*accountability*. La législation nationale préconise au mieux que les écoles participant au test utilisent les résultats dans un processus d'autoévaluation qui peut inclure, selon leurs souhaits, des actions de formation continue. Par ailleurs, bien que pionnière en matière d'évaluation avec le lancement du premier test standardisé en 1975 et disposant d'un dispositif national très complet de tests sommatifs et d'épreuves de diagnostic jusqu'à très récemment, la France n'a pas non plus prévu de liens explicites entre ces évaluations et la formation des enseignants, du fait certainement de l'influence dans ce pays du modèle de l'« effet miroir » développé par C. Thélot (voir partie précédente). De façon générale, la quasi-totalité des pays européens comme le Danemark, la Suède ou la Belgique flamande ne font pas non plus ce lien.

Une analyse des dispositifs de formation continue conduite dans le cadre de l'étude *Responsabilités et autonomie des enseignants en Europe* (Eurydice, 2008) confirme cette conclusion. Il n'y a pas de lien fort entre l'évaluation des performances des élèves et la formation continue de leurs enseignants non seulement parce que ces actions de soutien aux enseignants sont rarement décidées au niveau de l'école mais aussi plus fondamentalement parce que ces dispositifs demeurent très en retrait par rapport aux déclarations de principe incluses dans les législations nationales. Il apparaît que, d'un côté, dans la quasi-totalité des pays de l'OCDE, des systèmes d'évaluations standardisées ont été développés (Eurydice 2009, Mons, 2009) et que, de l'autre, derrière des rhétoriques politiques mettant en avant le nécessaire développement de la formation continue – par exemple, vingt pays des trente et un pays ou régions inclus dans l'enquête Eurydice présentent la formation

continue des enseignants comme obligatoire statutairement –, les législations n'imposent pas de conditions réellement favorables à son développement, et ce, pour plusieurs raisons. Ainsi, tout d'abord, seule une minorité de pays européens (treize pays ou régions) dédient un nombre d'heures annuel légal à ce soutien aux enseignants. C'est le cas notamment de l'Angleterre, (l'équivalent de 35 heures annuelles maximum³), la Norvège (35 heures), la Belgique francophone (21 heures), l'Autriche (15 heures) ou la Finlande (18 heures). De façon générale, les plans de formation ne sont que rarement établis au niveau de l'école ou des collectivités locales, rompant ainsi le possible lien entre l'analyse des résultats des élèves et la formation de leurs enseignants. On retrouve cette situation minoritaire dans douze pays sur les trente et un étudiés, notamment dans les pays nordiques comme au Danemark, en Suède ou en Norvège, ainsi que dans les Communautés francophone et flamande belges. Dans le reste des pays, les directives concernant le contenu des formations sont exclusivement élaborées de façon centralisée (à un niveau national ou régional), comme en France, en Espagne ou en Allemagne. Cette situation rend donc le lien entre les résultats des évaluations et la formation des enseignants difficile à construire. Autre élément est révélateur du peu d'intérêt national pour une réflexion autour de la formation continue qui serait conçue au niveau de l'établissement scolaire : ces actions ne sont que rarement incluses dans le plan de développement de l'école, puisque cette situation ne concerne que six pays ou régions (les communautés linguistiques belges, la Lituanie, l'Angleterre et Malte). Même dans ces cas-là, « il est cependant difficile, souligne le rapport Eurydice (2008), d'évaluer si les besoins des enseignants en termes de formation continue sont bien intégrés dans ces plans » (p. 49). Enfin les financements de formation continue sont aussi rarement entre les mains des établissements scolaires.

Autant de signes qui montrent que le lien entre les résultats des évaluations standardisées des élèves et la formation continue de leurs enseignants est difficile à établir à la fois parce que, malgré la rhétorique à l'œuvre sur la nécessité de la formation continue des enseignants, les ressources dont les systèmes éducatifs disposent sur le sujet sont rares et parce que les écoles n'en ont que très rarement la maîtrise.

Conclusion

En somme, cet article a montré que tant du côté des politiques effectivement déployées que des modèles « théoriques » auxquelles elles s'adossent, il existe une grande diversité de pratiques et de références. Nous avons dans ce texte largement fait référence au « modèle anglo-saxon » de régulation et attiré l'attention sur le fait qu'il incorpore conjointement des dispositifs de régulation s'appuyant sur des mesures externes des apprentissages des élèves, sur une exigence de « redevabilité »

3 Cinq jours de présence à l'école, en dehors des cours et dans le cadre de la fermeture de l'école aux élèves sont prévus entre autres pour assurer les formations.

adressée aux établissements et sur une exigence de formation (initiale et continue) pour les équipes éducatives, en particulier lorsqu'elles semblent être défaillantes. Nous avons également mis en évidence que ce « modèle anglo-saxon » est loin d'être un modèle pur – au contraire, il se frotte en Amérique du Nord, notamment, à des visions plus mécanistes de la régulation, faisant abstraction des plans de formation – et que sa mise en œuvre est à vrai dire très variable en fonction des périodes et des lieux considérés.

Ce modèle, et son application dans le contexte anglais en particulier a par ailleurs fait l'objet de critiques importantes. Au-delà de la rhétorique de professionnalisation à laquelle il se réfère, certains auteurs (Ball, 2003 ; Cattonar & Maroy, 2000) questionnent la logique managériale qui l'anime et s'interrogent sur une possible déprofessionnalisation des enseignants qu'il entraîne, si du moins on considère qu'un groupe professionnel est un corps collectif qui a du pouvoir sur son métier et dont la légitimité collective fonde la relative autonomie d'action dont va bénéficier chacun des membres de la profession. Dans une critique semblable, le sociologue portugais Lima (2010) parle, lui, d'une hyper-bureaucratization des systèmes éducatifs, évoquant par là la vision instrumentale et rationaliste qui anime ces nouveaux modes de régulation et qui enferme les systèmes éducatifs et les établissements scolaires dans des cadres d'action de plus en plus restrictifs, présentés comme des instruments neutres au service d'objectifs non problématiques.

Une dimension-clé de ces analyses qui, sur le plan théorique en tout cas, devrait permettre de trancher entre divers modèles d'*accountability* est la place qui est réservée à une intervention sur les processus d'enseignement (que ce soit à travers la formation à de bonnes pratiques, l'imposition de manuels de référence ou par d'autres voies encore). En effet, dans la foulée du raisonnement de Thélot exposé ci-dessus, un point de vue peut être de plaider pour un basculement dans les modes de régulation des systèmes éducatifs, à travers un remplacement de l'intervention sur les processus (qui vise à mettre en place des « manières d'enseigner » considérées comme souhaitables) par une intervention sur les résultats, sollicitant en retour les compétences des professionnels invités, si nécessaire, à améliorer leurs pratiques. Comme nous l'avons illustré, l'Autriche est probablement un des systèmes éducatifs en Europe le plus proche d'un tel modèle : le suivi rapproché des performances scolaires des élèves s'accompagne d'une importante autonomie locale dans le choix des procédures et dispositifs d'enseignement, et la formation continue veille davantage à rendre explicite les dispositifs d'évaluation externe qu'à diffuser des pratiques de référence. Un tel modèle est très différent du système anglais où, dans une large mesure, les prescriptions relatives aux méthodes d'enseignement s'additionnent à une régulation par les résultats, conduisant (potentiellement, en tout cas) à ce que Lima qualifie d'hyper-bureaucratization.

La question clé, à nos yeux encore partiellement en suspens à l'issue de cet article, est la place de la formation des enseignants au regard de l'opposition que

nous venons de formuler. Nous l'avons vu, dans le modèle anglo-saxon, d'un point de vue théorique en tout cas, les formations initiale et continue des enseignants font partie du dispositif de régulation et des enseignants hautement qualifiés sont censés être davantage capables de « lire » les évaluations externes et d'ajuster leur travail dans la perspective d'une amélioration de leurs performances. Toutefois, plusieurs auteurs tendent à considérer que la formation continue, dans un tel contexte, est essentiellement de nature instrumentale et devient une voie de diffusion des pratiques éducatives souhaitées par les autorités plutôt qu'une voie d'*empowerment* des professionnels. Paradoxalement, nous avons constaté que quand les dispositifs d'*accountability* sont plus modestes (à plus faibles enjeux diraient les Anglo-Saxons) et parient davantage sur la réflexivité des professionnels que sur des menaces de sanctions, le lien avec la formation continue semble absent. Bien entendu, comme nous l'avons évoqué ci-dessus, de telles politiques s'adosent à un principe d'autonomie des unités éducatives locales et invoquent précisément cette liberté pédagogique locale comme facteur d'amélioration des performances des élèves. Mais les données à notre disposition ne laissent pas entrevoir de conception alternative de la formation continue, comme si celle-ci ne pouvait pas être conçue dans une voie alternative, professionnalisante dans le sens fort du terme, c'est-à-dire, participant au renforcement d'une profession, à la définition en son sein de normes professionnelles et de savoirs de haut niveau, que chaque professionnel est ensuite apte à mobiliser en fonction de son environnement de travail spécifique.

Dans ce domaine comme dans bien d'autres, cette analyse fait *in fine* ressortir que derrière l'adossement à un modèle normatif de référence (les politiques déclarées de reddition de compte et de formation continue), on retrouve un phénomène d'adaptation nationale des politiques éducatives. Mais l'analyse déployée dans l'article fait également ressortir qu'au-delà des spécificités nationales, il y a probablement des divergences de fond relatives au sens même des politiques d'*accountability*, et de leur cohabitation avec le développement de la formation continue des enseignants.

Nathalie MONS

nmons@wanadoo.fr

Vincent DUPRIEZ

Vincent.Dupriez@uclouvain.be

BIBLIOGRAPHIE

BALL S.J. (2003). « The teacher's soul and the terrors of performativity », *Journal of Educational Policy*, vol. XVIII, n° 2, p. 215-228.

- CARNOY M. & LOEB S. (2002). « Does External Accountability Affect Student Outcomes? A Cross-State Analysis », *Educational Evaluation and Policy Analysis*, vol. XXIV, n° 4, p. 305-331.
- CATTONAR B. & MAROY C. (2000). « Rhétorique du changement du métier d'enseignant et stratégie de transformation de l'institution scolaire », *Éducation et Sociétés*, n° 6, p. 21-42.
- EURYDICE (2008). *Responsabilité et autonomie des enseignants en Europe*, Bruxelles : Commission européenne.
- EURYDICE (2009). *Les évaluations standardisées en Europe : objectifs, organisation et utilisation des résultats*, Bruxelles : Commission européenne.
- GOODWIN B., ENGLERT K. & CICCHINELLI L. F. (2003). *Comprehensive Accountability Systems. A Framework for Evaluation*, Aurora (CO) : Mid-continent Research for Education and Learning.
- HANUSHEK E.A. & RAYMOND M.E. (2005). « Does school accountability lead to improved student performance? », *Journal of policy analysis and management*, vol. XXIV, n° 2, p. 297-327.
- HOOD C. (1991). « A Public Management for All Seasons? », *Public administration*, vol. LXIX, n° 1, p. 3-19.
- KLIEME E. *et al.* (2004). *Le développement de standards nationaux de formation : une expertise*. Bonn : Ministère fédéral de l'éducation et de la recherche (BMBF).
- LIMA L. (2010). *Évaluation, compétitivité et hyper bureaucratie*. Conférence d'ouverture du colloque de l'ADMEE-Europe, Braga, janvier 2010.
- LINN R. L. (2001). *The design and evaluation of educational assessment and accountability systems*. Los Angeles : University of California, Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- MCDONNELL L. M. (2005). « Assessment and accountability from the policy-maker's perspective », in E. H. Haertel & J. L. Herman (dir.), *Uses and misuses of data for educational accountability and improvement*, Chicago : NSSE : Malden (MA) : Blackwell Publ., Part II, p. 1-34.
- MAROY C. (2006). *École, régulation et marché : une comparaison de six espaces scolaires locaux en Europe*. Paris : PUF.
- MAROY C. (2008). « Vers une régulation post-bureaucratique des systèmes d'enseignement en Europe ? », *Sociologie et Société*, vol. XL, n° 1, p. 31-55.
- MONS N. (2007). *Les nouvelles politiques éducatives : la France fait-elle les bons choix ?*, Paris : PUF.
- MONS N. (2009). *Les effets théoriques et réels de l'évaluation standardisée*, Bruxelles : Commission européenne.
- NICHOLS S. L. (2007). « High-Stakes Testing: Does It Increase Achievement? », *Journal of Applied School Psychology*, vol. XXIII, n° 2, p. 47-64.

- PHELPS R. P. (2005). *Defending standardized testing*, Mahwah (NJ) : Erlbaum Associates.
- THÉLOT C. (2002). « Évaluer l'École », *Études*, n° 10 (tome 397), p. 323-334.
- WÖßMANN L. (2007). « International Evidence on School, Competition, Autonomy and Accountability: a Review », *Peabody Journal of Education*, vol. LXXXII, n° 2-3, p. 473-497.

Abstracts • Zusammenfassungen • Resúmenes

Accountability policies

Since the end of the 1980's two major trends in educational policies have gained momentum across most OECD countries: the development of new in-service teacher training systems based on life-long learning and the implementation of pupils' standardized assessment within the framework of wider accountability policies. This paper sheds light on the links between these two trends by analyzing the theoretical research conducted in these two fields of reform and the educational policies implemented in OECD countries with regard to both policies.

Die Politiken von Accountability

Seit den 80er Jahren haben zwei Hauptpolitiken die Mehrheit der Erziehungssysteme der OECD-Länder beeinflusst: die Entwicklung von neuen Einrichtungen in der Lehrerfortbildung, die als lebenslange Ausbildung gedacht wird, und die Einführung standardisierter Evaluationen der Schüler im Rahmen weiterer Politiken von "accountability". Dieser Artikel beschreibt die Beziehungen zwischen den beiden Phänomenen und analysiert dabei gleichzeitig die theoretische, in diesem Feld gemachte Forschung und die in den OECD-Ländern wirklich geführten Erziehungspolitiken.

Las políticas de accountability

La mayoría de los sistemas educativos de los países de la OCDE han sido marcados desde el fin de los 80 por la introducción concomitante de dos políticas mayores : el desarrollo de nuevos dispositivos de formación continúa de los docentes dentro de una perspectiva de formación a lo largo de la vida y la instauración de evaluaciones estandarizadas de los alumnos, engastadas en unas políticas mas amplias de accountability. Este artículo interroga las relaciones que existen entre estos dos fenómenos, movilizando a la par los marcos teóricos que han relacionado estos dos grandes grupos de reformas y analizando las políticas realmente puestas en práctica en los países de la OCDE.