

LECTURES

Léopold PAQUAY* (1)

QUELLES ÉVALUATIONS DES ENSEIGNANTS ? TENDANCES NORD-AMÉRICAINES

En préparant un ouvrage sur le thème « Évaluation et formation des enseignants », j'ai exploré récemment la littérature anglo-saxonne, particulièrement nord-américaine, partant sur l'évaluation des enseignants. Mon projet est ici de pointer quelques évolutions récentes dans le monde scientifique quant aux conceptions dominantes relatives à ce domaine.

D'emblée, je tiens à préciser le caractère fortement contextualisé des pratiques et des recherches citées. Elles ne prennent sens que dans le contexte socioculturel, politique, administratif et historique des États-Unis d'Amérique. Elles sont d'ailleurs diverses, chaque État ayant une très large autonomie en matière d'éducation (Zuber, 1998). Mais l'on sait par ailleurs que, dans divers domaines relatifs à l'enseignement, bon nombre de concepts et modèles scientifiques développés aux États-Unis sont devenus des références pour la communauté scientifique du monde entier. Dès lors, dans le domaine de l'évaluation des enseignants, il peut être particulièrement utile pour les chercheurs et experts francophones et pour les acteurs clés de l'éducation, de prendre connaissance des conceptualisations, des modélisations et des expérimentations américaines, car elles risquent de devenir demain les références pour les concepteurs de dispositifs nouveaux d'évaluation des enseignants.

Ces notes de lecture n'ont nullement la prétention de proposer un cadre général. J'ai limité mes investigations à quelques numéros de revues spécialisées et à quelques chapitres de synthèse sélectionnés dans les deux *Handbook* les plus récents dans le domaine des recherches sur l'enseignement et de la formation des enseignants (Berliner, Calfee, 1996 ; Sikula, 1996). Dans les limites assignées à ces notes de lecture, je me limiterai aux questions suivantes :

* - Léopold Paquay, Université catholique de Louvain, Faculté de psychologie et des sciences de l'éducation (Belgique).

1 - Je remercie de tout cœur Jean Cardinet, Stefania Casalfiore et Vincent Vandenberghe pour leur lecture attentive du manuscrit et leurs remarques judicieuses et constructives.

1. Quels sont les types d'évaluation des enseignants ? Quelles évaluations des enseignants depuis la sélection des futurs enseignants jusqu'à l'évaluation en cours de carrière ? Cette question permettra tout à la fois de préciser les grands types de pratiques dans le cadre américain et de percevoir la diversité et la complexité du domaine.
2. Ensuite, à titre de prospective, j'explorerai plus à fond une question délicate : quelles évaluations des enseignants au long de leur carrière ? Ce seront surtout les rationalités diversifiées de ces évaluations qui seront mises en évidence ainsi qu'une difficulté à cerner l'objet d'évaluation.
3. Enfin, je ferai état de quelques principes retenus par les spécialistes américains pour développer des dispositifs d'évaluation du personnel enseignant en cours de carrière.

QUELLES ÉVALUATIONS DES ENSEIGNANTS ? UN INVENTAIRE

Dwyer et Stufflebeam (1996) structurent les diverses pratiques d'évaluation autour de quatre grandes phases de la carrière des enseignants : la préparation, l'attribution de la licence pour enseigner, la pratique de l'enseignement et le processus de « professionnalisation ». À chaque fois, ils dégagent les types d'évaluation et les types de décision qui en résultent. En suivant ces auteurs, inventorions ces types d'évaluation.

La préparation au métier

Ceux qui connaissent les ouvrages de référence de D. Stufflebeam et de ses collaborateurs sur l'évaluation des programmes (1971, 1980) ne seront pas étonnés que ces auteurs citent en premier les évaluations du contexte et des intrants (2) qui permettent de fonder et de réajuster les programmes de formation initiale. Au long de la mise en œuvre du programme et au terme de celui-ci, il y a d'ailleurs lieu de recueillir des informations en vue d'ajuster et d'améliorer le programme.

Centrons-nous ici sur les évaluations orientées vers des décisions qui concernent le futur enseignant :

1. d'abord l'évaluation des candidats à l'entrée qui permettra de sélectionner ceux qui pourront entrer dans la formation ;

2 - En francophonie, ces phases de l'évaluation des programmes sont souvent nommées «Analyse des besoins» et «Analyse des ressources». À première lecture, on pourrait croire à un glissement d'objet d'évaluation. Et pourtant, l'analyse des besoins de formation des enseignants constitue une forme d'évaluation des enseignants (voir les deux derniers points ci-après). On peut faire le même raisonnement pour une analyse des besoins de formation de futurs enseignants.

2. les évaluations tout au long de la formation initiale qui permettent les régulations nécessaires ;
3. l'évaluation finale qui débouche sur la qualification et l'attribution de grades.

À rigoureusement parler, les évaluations ci-dessus ne constituent pas des évaluations des enseignants, mais des évaluations des étudiants et des programmes. Toutefois, il est indispensable de les prendre en considération dans une problématique globale car elles constituent les prémices et les fondations de toute évaluation ultérieure. Venons-en maintenant aux évaluations des enseignants en cours de carrière.

Attribution de la licence pour enseigner

Pour enseigner dans l'enseignement public aux États-Unis, un individu doit posséder la licence pour enseigner. Cette licence est accordée par les autorités d'un des États et ne vaut que pour cet État. Traditionnellement, cette licence a une durée indéterminée ; mais de plus en plus elle est accordée pour une durée déterminée (nous reviendrons ultérieurement plus en détail sur cette problématique). Cette licence peut consister simplement dans la reconnaissance qu'un individu a la qualification statutaire déterminée pour pouvoir enseigner (en Belgique, on dirait qu'il a les titres requis pour enseigner telle ou telle discipline) et donc qu'il peut être engagé comme enseignant (pour tel niveau (3) d'enseignement, dans telle discipline...) (Roth, 1996). Parfois ce processus consiste simplement en une opération administrative d'homologation des diplômes obtenus en formation initiale (4). Mais le processus d'attribution de la licence peut être plus complexe.

Dryer et Stufflebeam (1996) décrivent le processus d'attribution de cette licence en trois étapes, chaque décision repasant sur des évaluations :

1. d'abord l'acceptation d'entrer dans le processus d'attribution de la licence ; cette acceptation se fait sur la base des qualifications antérieures (*credentials*) ;
2. attribution d'une licence provisoire et définition de modalités d'évaluation durant la période probatoire ;
3. au terme de cette période probatoire (souvent un an) et sur la base des évaluations de la qualité de l'enseignement, attribution d'une licence à plus long terme ou d'une licence permanente.

3 - Il s'agit de préciser l'ensemble d'années auxquelles on peut enseigner (de la 1^{re} à la 12^e année) ; en France, on distinguerait les enseignants du premier et du second degré ; au Canada, on distingue les ordres d'enseignement...

4 - Roth (1996) distingue d'une part les procédures d'approbation de programmes (*program approval*) en vue d'attribuer une licence pour enseigner et, d'autre part, les procédures d'accréditation de programmes (*accreditation*) qui permettent une certification nationale (voir dernier point ci-après).

On notera avec Rath (1996) que la licence permanente était traditionnellement la plus commune. Cette pratique a été virtuellement balayée lors du mouvement de réforme dans les années 80-90. Certains États, dont la Pennsylvanie, le New-Jersey et New York sont engagés dans un processus de remplacement des licences à vie par des licences renouvelables. Mais ce n'est pas sans mal. À présent, plusieurs États proposent des licences de 2^e degré alors que d'autres n'ont aucune exigence de développement professionnel (5).

Évaluations « ordinaires » en cours de carrière

Ces évaluations sont de plusieurs types, entre autres :

1. une évaluation des besoins du personnel, ce qui permettra de déterminer les priorités pour les formations continuées ;
2. une comparaison entre les compétences des enseignants et les exigences des postes de façon à déterminer les affectations de postes ;
3. une analyse de performances en vue, selon le cas, de prendre des décisions de nomination, de titularisation, de promotion, d'augmentation de traitement, d'attribution d'honneurs... ou encore de sanctions, voire d'éviction du système scolaire.

Évaluations en cours de carrière dans une perspective de plan de carrière (« professionnalisation »)

Ces évaluations sont de plusieurs types, entre autres :

1. une évaluation des besoins du personnel et de l'institution ainsi que des compétences des enseignants, ce qui permettra d'offrir à certains membres des formations continuées, ou d'attribuer des bourses pour des formations complémentaires ;
2. une évaluation des qualifications de base d'un membre du personnel de façon qu'il puisse accéder à un programme national de « certification » ;
3. un examen rigoureux des compétences de ce candidat réalisé par une instance nationale reconnue, ce qui permet d'attribuer au candidat une « certification nationale ».

Ces « certifications nationales » nécessitent quelques commentaires. Selon Roth (1996), on peut les caractériser comme suit. La « certification » est une reconnaissance de compétences professionnelles d'enseignant certifiant que la personne

5 - Le développement professionnel se définit minimalement comme le développement des compétences professionnelles au long de la carrière. Dans une acception plus large, ce concept couvre « toutes les transformations individuelles et collectives de compétences et de composantes identitaires mobilisées ou susceptibles d'être mobilisées dans des situations professionnelles » (Barbier, Chaix, Demailly, 1994, p. 7).

maîtrise certains standards, signe d'un enseignement de qualité. La « certification » est un processus de validation de compétences d'une personne par une organisation professionnelle indépendante telle que le « *National Board for Professional Teaching Standards* ». Le plus souvent, il s'agit de reconnaître et de valoriser un haut niveau d'expertise. Alors que la licence est requise pour tout enseignant et n'est attribuée que pour un État, le processus de « certification » est volontaire et donne droit à un certificat de haute compétence professionnelle pour l'ensemble des états des États-Unis.

Analysons maintenant plus en détail les modalités nord-américaines d'évaluation des enseignants en cours de carrière.

QUELLES ÉVALUATIONS DES ENSEIGNANTS AU LONG DE LA CARRIÈRE PROFESSIONNELLE ?

Nous considérons ici les évaluations des enseignants qui ont pour fonction d'attribuer le droit d'enseigner (*licensing*) ainsi que l'ensemble des évaluations réalisées au long de la carrière qui ont pour fonctions des promotions et/ou le développement professionnel des enseignants. Passons d'abord en revue l'ensemble des modalités d'évaluation des enseignants et dégageons les logiques sous-jacentes.

Logiques sous-jacentes aux pratiques d'évaluation

Le tableau suivant explicite quelques rationalités diverses et des modèles-types d'évaluation des performances des enseignants en cours de carrière. Il s'agit d'une légère adaptation d'un tableau proposé par Dwyer et Stufflebeam (1996, p. 775).

Logiques diverses d'une évaluation des enseignants

Logiques ou rationalités	Description de la logique	Qui propose cette évaluation ?	Modalités et approches concrètes
Amélioration de l'enseignement en classe	L'enseignement est considéré comme un métier ou une profession reposant sur une théorie de base. On peut ranger les pratiques sur un continuum du novice à l'expert. Des observateurs entraînés fournissent aux enseignants des feedback pour améliorer leurs habiletés à enseigner et leurs performances. On met l'accent sur le développement professionnel de l'équipe d'enseignants. Les évaluations sont le plus souvent à visée formative, parfois sommative.	Les responsables de l'enseignement et des spécialistes du curriculum et de l'enseignement.	Modèles divers de supervision clinique, de développement professionnel, de feedback par enregistrement vidéo, micro enseignement, etc.
Responsabilisation (accountability) (6) et développement professionnel	L'enseignement est considéré comme une profession qui exige que les enseignants fassent la preuve de l'étendue et de la qualité de leurs services et qu'ils s'engagent dans une autoévaluation permanente. Les enseignants ou leurs associations professionnelles rassemblent des preuves et les présentent aux autorités. L'accent est mis à la fois sur le fait de rendre compte (accountability) et sur l'amélioration. Selon cette logique, les enseignants contrôlent leurs propres troupes et tentent continuellement d'améliorer leur service aux étudiants.	Les enseignants et leurs organisations professionnelles.	Approches d'autoévaluation (Good). Portfolios. Évaluation par les pairs. Comités professionnels (NBPTS). ...
Contrôle administratif	L'enseignement est considéré comme un job dans un contexte bureaucratique, il doit être régulièrement supervisé et suivi. Les directeurs évaluent les performances des enseignants pour garantir qu'ils rendent annuellement des comptes (accountabilities). Les évaluations sont souvent fondées sur des visites de classe et sur des observations informelles. L'accent est mis sur l'assurance que l'enseignant n'est pas négligent. Ce type d'évaluation fournit une base légale adéquate pour éliminer des enseignants incompetents.	Les administrateurs scolaires.	Évaluation administrative à visée sommative. Visites informelles de classe par le directeur ou par l'inspecteur.

6 - Le terme « *accountability* » a ici été traduit tantôt par « *responsabilisation* », tantôt par « *le fait de rendre compte* ». Nos collègues canadiens francophones parlent d'« *imputabilité* ».

Logiques ou rationalités	Description de la logique	Qui propose cette évaluation ?	Modalités et approches concrètes
Indicateurs d'efficacité basés sur la recherche	L'enseignement est considéré comme un ensemble de variables indépendantes qui a un impact sur la réussite des élèves. Les évaluations sont focalisées sur les comportements des enseignants dont les recherches montrent qu'ils sont positivement corrélés avec la réussite des élèves. Ces évaluations tendent à identifier et à accentuer les styles d'enseignement plutôt que les prises de responsabilité. Les comportements d'enseignement (corrélés avec la réussite) sont souvent évalués au moyen d'observations qui n'exigent que peu d'inférence.	Les chercheurs universitaires en éducation.	« Measurement based teacher performance evaluation ». Le système d'évaluation des enseignants de Géorgie. La plupart des systèmes officiels d'évaluation des différents pays des États-Unis s'inspirent de ces deux approches.
Protection du consommateur et réactions du public	L'enseignement est considéré comme un service public vital qui permet tout à la fois un enseignement efficace et une protection du bien-être des étudiants. Les évaluations doivent être indépendantes ; elles sont focalisées sur la réussite des étudiants et sur la responsabilité de l'enseignant. L'accent est mis sur le caractère pratique et sur le degré de préparation de l'école à de telles démarches d'évaluation. Cette évaluation consiste en feed-back de la part des étudiants et des parents (...). L'accent est mis sur l'importance du maintien de standards de haut niveau pour les performances des professeurs.	Le public, les entreprises, l'enseignant supérieur (qui reçoit les élèves). Condition de base pour l'attribution de la licence pour enseigner.	Évaluation basée sur le devoir (duties-based evaluation selon Scriven). Évaluation de l'enseignement par les élèves. Niveau des étudiants à des tests ou lors de la réalisation de travaux. Évaluation par des instances extérieures (autorités, inspecteurs). EST (Professional assessments for Beginning Teachers).
Paiement au mérite	L'enseignement est considéré comme ayant besoin de reconnaissances et de récompenses. Pour les tenants de cette logique, il est évident que payer les performances engendre un meilleur enseignement. Les évaluations sont fondées sur la réussite des étudiants et sur des jugements par des superviseurs. Les enseignants peuvent être obligés de produire un portfolio relatif à leurs performances et à leurs réussites. Certains auteurs préconisent de considérer le mérite des écoles plutôt que le mérite des enseignants individuels comme un moyen pour éviter des compétitions contre-productives.	Le public, l'État, des fonctionnaires locaux qui doivent être crédibles aux yeux du public.	Plans de carrière. Seules les augmentations dues au mérite sont ajoutées au revenu de base... ou les augmentations dues au mérite consistent des primes qui s'ajoutent aux augmentations régulières annuelles (en plus du revenu de base)... ou bonus annuel. « Merit schools ».

Le contenu de ce tableau nécessiterait bien des commentaires. Les logiques sont multiples et variées. Et l'on peut même s'étonner de la réunion, sous une même « étiquette » d'évaluation des enseignants, de démarches aussi différentes que des analyses de besoins pour déterminer les priorités d'une formation continuée et des démarches de contrôle administratif.

Traditionnellement, aux États-Unis, comme dans la plupart des pays, ces évaluations sont réalisées par des inspecteurs ou par des supérieurs hiérarchiques (directeurs, superintendants...) sur la base de visites de classes, d'une analyse du matériel et d'interviews. Comme le précise Scriven (1995), l'évaluation peut également être réalisée par des pairs (des collègues qui jugent par exemple l'usage approprié des moyens d'enseignement) et/ou se baser sur des jugements des utilisateurs, soit les élèves, soit les parents.

Dans le contexte de responsabilisation des acteurs des années 80, nombreuses ont été les mises en cause du caractère subjectif et partial des décisions prises quant à certaines titularisations ; tout particulièrement, la sensibilité américaine aux droits des minorités a débouché sur la mise en cause devant les tribunaux de la faiblesse des instruments utilisés pour les décisions relatives aux nominations et aux promotions (Dwyer, Stufflebeam, 1996, p. 768). En outre, dans le contexte de grandes réformes relatives à l'éducation, les pratiques évaluatives ont été l'objet de nombreuses expérimentations, mais aussi de bien des débats. Des standards fédéraux pour l'évaluation du personnel ont été définis (*Joint Committee on Standards for Educational Evaluation*, 1988) ; leur but est manifestement d'améliorer la qualité de l'évaluation des enseignants. Des moyens sont mis à la disposition des chercheurs et des évaluateurs pour développer des démarches adéquates.

172

Dans ce contexte, de nombreux dispositifs instrumentés ont été développés et améliorés. Avant de les passer en revue, cernons quelques-uns des plus récents et des plus étonnants tels que présentés par Dwyer, Stufflebeam, 1996, pp. 771-772). « Le *National Board for Professional Teaching Standards (NBPTS)* a été mis en place en 1987 en vue de certifier les enseignants expérimentés. La « certification » renvoie à une manière d'accorder une distinction aux enseignants qui atteignent des standards rigoureux de haut niveau. Ce processus volontaire est contrôlé par la profession. La première évaluation a été opérationnellement mise en œuvre en 1995. Le NBPTS a fourni de grands efforts pour définir des standards pour les enseignants de différents niveaux scolaires et des différentes disciplines et pour garantir un large consensus quant à ces standards. Le NBPTS a aussi fait œuvre de pionnier dans l'usage de simulations et d'enregistrements vidéo pour l'évaluation des enseignants, dans l'usage de portfolios et dans l'utilisation d'exercices originaux de divers types. » Manifestement, cette modalité d'évaluation ouvre des voies nouvelles dans la valorisation des compétences acquises au long de la carrière ; elle s'inscrit clairement dans

des perspectives de mobilité professionnelle. Elle crée de fait une nouvelle catégorie d'enseignants – « *the best teachers* » – qui pourraient monnayer leurs compétences certifiées sur le marché de l'emploi (7).

De telles évolutions dans les pratiques d'évaluation des enseignants en cours de carrière soulèvent de nombreuses questions... Nous esquisserons plus loin certaines questions relatives aux modalités souhaitables d'évaluation des enseignants pour qu'elles participent au développement professionnel des enseignants. Nous aborderons maintenant la question relative aux critères d'évaluation.

La question centrale : qu'évaluer ?

De nombreux travaux déjà anciens portent sur l'évaluation du personnel enseignant. La question centrale est celle des critères de qualité. Le premier critère de qualité auquel pense tout expert en management est celui de la production. Un bon enseignant serait celui qui produit des apprentissages de qualité chez les élèves qui lui sont confiés, c'est celui qui atteint les objectifs pédagogiques assignés. En fait, une analyse plus approfondie rend compte des limites de ce critère. La qualité des apprentissages des élèves dépend de multiples facteurs en interaction, pas seulement des interventions de l'enseignant ; c'est l'élève qui apprend ; le rôle de l'enseignant est d'organiser au mieux les conditions de l'apprentissage, mais il ne peut apprendre à la place de l'élève. Un enseignant ne peut donc jamais avoir d'obligation de résultats ; tout au plus peut-on lui assigner des obligations de moyens. Comme le rappellent Peterson, Kromrey, Smith (1990), le métier d'enseignant est très différent du métier de vendeur qui, lui, peut-être évalué sur sa production (le nombre de ventes réalisées) ; le métier d'enseignant est plus proche du médecin : légalement et professionnellement, celui-ci ne peut être évalué sur le fait que ses patients vivent ou meurent, mais sur la bose des procédures qu'il a mises en œuvre dans son diagnostic ou son traitement. Similairement, on peut attendre d'un enseignant qu'il mette tout en œuvre pour que les élèves apprennent. Les critères premiers de qualité deviendraient ainsi l'investissement de l'enseignant et la qualité des activités enseignantes.

Mais immédiatement, se pose une question fondamentale. Sur quelles boses peut-on évaluer la qualité d'un enseignement ? Les nombreux travaux de recherche menés depuis plus de trente ans dans une perspective de mise en relation du processus d'enseignement et des produits d'apprentissage ont débouché sur un nombre assez

7 - En Angleterre, une démarche similaire voit le jour avec la certification d'« *advanced skills teachers* ». Lorsqu'on apprend par ailleurs que ce modèle de mobilité fondée sur l'excellence résulte de l'échec de tentatives d'organisation d'une mobilité plus statutaire en spécifiant diverses catégories d'enseignants (voir Lessard, 1999), on peut s'interroger en Europe francophone sur les évolutions en cours.

impressionnant de connaissances « générales » quant aux caractéristiques d'un enseignement qui se révèlent efficaces. À titre d'exemple, Rosenshine et Edmonds (1990) ont mis en évidence les caractéristiques d'un enseignement efficace de capacités assez larges telles que résumer un paragraphe, résoudre un problème de mathématique, écrire un texte de recherche, etc. Ce sont des caractéristiques similaires à celles d'un enseignement efficace de matières structurées (Rosenshine, 1986). Précisons-en les principales telles que relatées dans Rosenshine et Edmonds (1990) à partir de quarante études :

- évaluer la maîtrise des prérequis ; les faire acquérir si besoin est ;
- présenter aux apprenants des procédures facilitatrices ; fournir des modèles des bonnes réponses ; anticiper les difficultés en discutant de bons exemples et des exemples limites ; faire pratiquer les apprenants et accroître progressivement la complexité des tâches ;
- accompagner les apprenants tout au long de la résolution de problèmes pratiques ; faire collaborer les apprenants en duo ou par petits groupes ;
- fournir un feed-back en référence aux procédures facilitatrices et aux modèles fournis préalablement ; fournir des outils d'autoévaluation qui permettent à chaque apprenant d'analyser ses propres procédures ;
- faire pratiquer la capacité sur des exemples variés en mettant en place une pédagogie de maîtrise ;
- appliquer à de nouveaux exemples.

Bien évidemment, chacune de ces caractéristiques nécessiterait une description plus fine et plus détaillée. Cependant, une telle liste suffit pour se rendre compte de ce que signifie la mise en évidence de caractéristiques d'un enseignement qui, selon les recherches, ont des effets sur la qualité des apprentissages des élèves. On trouvera dans Gauthier et coll. (1997), une synthèse récente en français de la littérature anglo-saxonne relative à l'ensemble des caractéristiques d'un enseignement dont les recherches ont montré l'efficacité. À partir de plus de 4 000 études scientifiques réalisées dans des classes américaines pour la plupart, ces auteurs dégagent une « base de connaissances » scientifiques relatives à l'enseignement.

C'est en référence à cette base de connaissances relatives à l'enseignement que de nombreux auteurs préconisent une évaluation des enseignants fondée sur la recherche. Le courant des « *research-based teacher evaluation* » est particulièrement développé dans les milieux de chercheurs et a donné lieu à de nombreuses publications (entre autres, Schwab, 1991).

Cependant, autant on perçoit l'intérêt de telles évaluations des enseignements de futurs enseignants, autant on peut questionner le principe même d'une évaluation des enseignants en cours de service réalisée uniquement en référence au modèle normatif qui se dégage des recherches. Dès 1987, Michael Scriven a radicalement

mis en cause la pertinence et les fondements des « *research-based teacher evaluation* ». En fait, précise Scriven, chaque enseignant a son style propre ; il n'est pas question d'imposer à chaque enseignant des types de comportement standards. Plus fondamentalement, Scriven met en cause la logique sous-jacente aux évaluations basées sur les recherches. Même si les recherches montrent en général qu'une caractéristique d'un enseignant est positivement corrélée avec une efficacité des apprentissages, on ne peut tirer la conclusion, pour un enseignant particulier, que son enseignement sera plus efficace s'il développe cette caractéristique. Dans le même sens, Strike (1990) montre que la seule prise en compte de caractéristiques des enseignants pour l'évaluation de la qualité des enseignants serait similaire à une évaluation de joueurs de basket qui serait fondée sur les seuls indicateurs corrélés avec l'efficacité (par exemple, la taille des joueurs, leur degré d'entraînement) mais qui laisserait dans l'ombre les variables les plus déterminantes liées au jeu lui-même : le shoot, les dribbles, etc. Un bon joueur de basket est un joueur qui joue bien et qui marque des paniers, et non pas d'abord le joueur qui a telle ou telle caractéristique générale dont les recherches corrélationnelles ont montré que, en général, elles étaient liées avec l'efficacité. Ainsi donc, Scriven en arrive à proposer une alternative à l'évaluation des enseignants basée sur la recherche : il s'agit d'une évaluation basée sur le devoir (a « *duty-based teacher evaluation* »). « Le devoir des enseignants est de bien enseigner, pas d'enseigner en accord avec les idéaux définis par les chercheurs » (Scriven, 1988, p. 321).

Mais le problème reste entier. Qu'est-ce que « bien enseigner » ? Scriven affirme qu'il y a « quatre dimensions générales de compétence et d'excellence dans l'enseignement : il s'agit de la qualité du contenu enseigné et appris ; de la quantité des apprentissages transmis, de l'éthique et du professionnalisme du processus » (Scriven, 1988, p. 320). Plus spécifiquement, il propose la liste suivante de « devoirs » auxquels devrait satisfaire tout enseignant : la connaissance des exigences du métier, la connaissance de l'école et de la communauté, la connaissance de la matière à enseigner, la capacité à construire un cours (*instructional design*), l'information recueillie à propos de la réussite des élèves, l'information fournie à propos de la réussite des apprenants, les habiletés en classe (*classroom skills*), les caractéristiques personnelles et le service à la profession. À vrai dire, cette liste n'est guère convaincante. Ainsi, Cruickshank et Haefele (1990) montrent que plusieurs composants de cette liste sont très proches des indicateurs basés sur les recherches (*research-based indicators*). Par exemple, Scriven entend par « *instructional design* » le développement des plans de cours, la définition des objectifs, la clarté des options de l'enseignement ; les tenants des indicateurs basés sur la recherche ne disent rien d'autre, tout au plus opérationnalisent-ils davantage les comportements attendus. De même à propos des « *classrooms skills* », Scriven mentionne comme indicateurs : un enseignement en groupes-classe et par petits groupes, le temps centré sur la tâche, la gestion de classe... ce qui correspond à certaines caractéristiques attendues d'un

enseignement selon les tenants des *research-based teacher evaluation*. En définitive, malgré ses affirmations, Scriven n'est pas parvenu à développer un ensemble de « devoirs » qui excluent les listes d'indicateurs d'un enseignement efficace. En réalité, un sous-ensemble des indicateurs basés sur les recherches est niché dans les « devoirs » des enseignants (8). Nous ne prendrons pas ici systématiquement chaque critique. Sachons simplement que ces critiques ont fait l'objet d'un large débat. Peterson, Kromfrey, Smith (1990) démontent point par point chacune des critiques de Scriven.

On ne sort pas du cercle vicieux. Qu'est-ce que bien enseigner ? Comment définir la professionnalité enseignante ? Manifestement, une approche comportementaliste permet de mettre en évidence certaines caractéristiques d'un enseignement qui est généralement efficace, mais cette approche fondée sur le paradigme « processus-produit » a manifestement atteint ses limites. Il est vain de déterminer des comportements standards constituant des critères universels de référence. Tous les experts sont d'accord à ce sujet (voir entre autres, Crahay, Lafontaine, 1986 ; Perrenoud, 1996 ; Simon, Demers, Dahan, Forgette-Giroux, 1994 ; Good, 1996, Paquay, Altet, Charlier, Perrenoud, 1996 ; etc.) Les approches plus récentes de l'activité enseignante inspirée du courant de l'action située (Durand, 1996 ; Derry, Lesgold, 1996 ; Casaliore, Paquay, 1998) sont prometteuses ; mais elles ne sont pas encore assez développées pour déboucher sur des pratiques instrumentées d'évaluation qui soient pertinentes et valides. Ce débat des années 1990 autour de l'intérêt des indicateurs basés sur la recherche continue à rebondir. Il est d'ailleurs à inscrire dans le cadre d'un débat plus large, celui de la définition de l'enseignant-professionnel ; comme le montre Lessard (1998), les conceptions d'une professionnalisation du métier d'enseignant sont nombreuses et diverses.

176

Si même il était possible de déterminer, dans des familles de contextes, des comportements plus particulièrement efficaces, un problème clé de l'évaluation resterait ouvert : celui des critères à privilégier. Ainsi, Scriven (1995) a introduit la distinction entre « évaluation du mérite » (évaluation des compétences manifestées) et « évaluation de la valeur » (correspondance aux besoins). Qu'est-ce à dire ? Un professeur de philosophie peut être d'excellente qualité (et donc avoir du mérite), mais se révéler de peu de « valeur » si l'école n'a nul besoin de professeurs de philosophie ; inversement, dans le cas où l'école manque cruellement de professeurs de mathématiques, des professeurs de mathématiques ont beaucoup de « valeur » même s'ils

8 - L'approche de Scriven a l'avantage de mettre en évidence d'autres facettes du métier d'enseignant. Non seulement l'enseignant enseigne (il organise les apprentissages en classe), mais en plus, il fait partie de l'équipe éducative, il rencontre les parents, il participe aux jury de délibération, il se forme... On est amené ainsi à prendre en compte l'évaluation des diverses fonctions assignées à l'enseignant.

ne sont pas très compétents (s'ils ont peu de « mérite »). Je n'irai pas plus loin ici dans ce débat instauré par Scriven. Qu'on l'adopte ou non, cette distinction a simplement l'avantage de montrer toute la relativité des critères d'évaluation des enseignants surtout lorsqu'on les inscrit dans le contexte du quasi marché scolaire (9). Elle nous invite au minimum à situer toute évaluation des enseignants en cours de carrière dans un contexte plus large... sans doute d'abord le contexte des établissements scolaires.

QUELS PRINCIPES POUR UNE ÉVALUATION DES ENSEIGNANTS ?

Dans la plupart des pays développés, on constate un double processus : le système scolaire accorde une autonomie accrue aux établissements et, dans le même mouvement, exige leur responsabilisation en leur demandant de rendre des comptes et en contrôlant les résultats (Perrenoud, 1998). Les modalités varient évidemment selon le contexte sociopolitique et l'histoire des systèmes éducatifs. Bonami (1996) décrit la double dynamique de professionnalisation (prise en charge de l'école par des enseignants professionnels) et de standardisation (injection de normes par les autorités centrales). Mais quelles que soient les situations antérieures, on constate dans tous les pays qu'une autonomie accrue des établissements va souvent de pair avec une évaluation interactive (10) des établissements et une évaluation externe des acquis des élèves. C'est dans un tel contexte qu'il importe de situer une évaluation des enseignants.

À titre de synthèse d'un numéro spécial de « *Studies in Education and Evaluation* » consacré à l'élaboration d'un modèle général pour l'évaluation du personnel de l'éducation (Nevo, Stufflebeam, Langueheine, 1995), Mc Conney (1995) énonce sept principes à mettre en œuvre pour instaurer une évaluation des enseignants.

177

1. Une articulation nécessaire de cette évaluation avec l'évaluation des besoins, ressources et actions de l'institution (ce qui implique entre autres que l'évaluation des enseignants devrait être liée à l'évaluation des établissements).

9 - Les effets de ce quasi marché scolaire sont particulièrement visibles dans des pays libéraux comme les Pays-Bas ou la Belgique (Vandenberghe, 1998).

10 - Référence est faite ici à la typologie proposée par Cardinet (1990) d'une évaluation interne, externe ou interactive (dite aussi négociée). À noter que Jean Cardinet fait dans cet article un parallélisme particulièrement éclairant entre ces trois types d'évaluation d'établissements et trois types similaires d'évaluation des apprentissages des élèves (autoévaluation, hétéro-évaluation et co-évaluation). Il y aurait lieu d'ailleurs de lire les pratiques d'évaluation des enseignants à la lumière de cette typologie et de développer en conséquence des modalités non piégeantes (voir Saussez, Allal, à paraître et Paquay, Darras et Saussez, à paraître).

2. Une explicitation claire et compréhensible des devoirs et responsabilités des enseignants (et de chaque catégorie de personnel de l'éducation).
3. Une détermination précise des critères et surtout des seuils minima exigés pour chaque critère.
4. Des recueils multiples de données fiables (données objectives, documents, interviews, portfolio, etc.).
5. Une visée qui est d'abord formative (et vise le développement professionnel) avant d'être sommative (et vise le fait de rendre des comptes ou une promotion ou l'attribution de la licence ou une augmentation salariale). Des évaluations à visée principalement sommative, tous les cinq ans par exemple, seraient précédées d'évaluations formatives annuelles de l'enseignement et de l'enseignant).
6. L'utilisation de standards quant aux modalités et quant au contenu (standards à définir au niveau national mais à adapter au contexte local).
7. La réciprocité [ou tout au moins la liaison] des évaluations des acteurs de divers niveaux hiérarchiques.

Dans la ligne de la problématique esquissée ci-avant, un certain nombre de ces principes insistent sur les précautions métriques : définir des normes, des critères, des seuils, recueillir des données fiables, etc.

D'autres insistent sur la fonction d'abord formative de telles évaluations. Ces évaluations sont à inscrire dans le contexte d'une relation ; elles visent d'abord le développement professionnel plutôt que des décisions relatives aux nominations ou aux promotions. « Tout à la fois, ce système fournit aux professionnels une opportunité de s'améliorer de façon telle qu'il (ou elle) soit dans une meilleure position pour contribuer à atteindre les buts de l'école, et en même temps, il assure le droit des parents et des étudiants à des services éducatifs appropriés » (Mc Conney, 1995, p. 108).

En terminant, je voudrais insister sur le premier des principes. Les propositions de Stufflebeam (1995) et de Webster (1995) sont très nettes à cet égard. Une évaluation du personnel d'un établissement scolaire devrait toujours s'inscrire dans un processus global d'évaluation et être liée à l'évaluation globale du système : analyse des besoins, analyse des ressources, analyse des processus et analyse des effets. L'évaluation des enseignants ne prend sens que dans l'évaluation du système. De plus, on voit immédiatement les dérives d'une évaluation qui serait réalisée par des directions (ou par des superintendants) qui eux-mêmes ne feraient pas l'objet d'une évaluation...

Compte tenu de ces principes, il y aurait lieu de préciser les modalités concrètes d'évaluation, d'en discuter les avantages et les limites, d'en cerner les conditions d'efficacité et les pièges et de pointer les antidotes aux dérives potentielles. Une telle analyse devrait faire l'objet d'études approfondies. Elle dépasse en tout cas les visées de cet article.

Perspectives engagées...

Je suis conscient du caractère particulièrement sensible et délicat des questions relatives à l'évaluation du personnel enseignant en cours de carrière. Dans les pays francophones européens, où les enseignants sont avant tout des fonctionnaires publics nommés à durée indéterminée, certains refusent le débat. Cependant, on entend des discours de plus en plus fréquents en faveur d'une mobilité accrue des enseignants. Fonder les promotions sur le seul critère d'ancienneté dans la fonction ne serait guère mobilisateur ; et la prise de conscience du grand nombre d'enseignants plus âgés qui sont blasés amène à penser à des modalités de réorientation en fin de carrière vers d'autres fonctions. Des voix de plus en plus nombreuses se font entendre pour préconiser une mobilité accrue au long de la carrière et pour fonder cette mobilité sur une évaluation sérieuse du personnel enseignant.

Le débat est parfois dur entre les partisans et les opposants d'un système d'évaluation du personnel qui aurait des incidences pour la carrière, que ces incidences soient de l'ordre de valorisations statutaires ou barémiques ou même dans certains cas de l'ordre de la conservation de l'emploi. Les évolutions récentes aux États-Unis et dans quelques autres pays anglo-saxons sont éclairantes à cet égard et permettent de saisir certains enjeux.

La modeste synthèse présentée dans cet article avait simplement pour objectif de pointer quelques tendances lourdes des évolutions des pratiques d'évaluation des enseignants aux États-Unis. J'espère simplement que l'évolution de fait vers la tendance à la création d'une catégorie d'excellence (« *the bests teachers* ») réveillera tous ceux qui en Europe francophone considèrent cette problématique comme un sujet tabou. Le risque est évidemment de voir bientôt surgir des labels de qualité définis par des instances nationales ou européennes, voire par les organismes privés. Comme je suis convaincu qu'il importe que les enseignants soient associés à la définition des critères et des modalités d'évaluation, j'espère que pourront s'élaborer et s'expérimenter, avec tous les acteurs concernés, des modalités pertinentes et constructives d'une évaluation du personnel de l'éducation qui s'inscrivent principalement dans des perspectives de développement professionnel.

BIBLIOGRAPHIE

- Barbier J.-M., Chaix M.-L., Demailly L. (1994). - Éditorial du n° spécial sur « Recherche et développement professionnel », *Recherche et formation*, 17, pp. 5-8.
- Berliner D.C., Calfee R.C. (Ed.) (1996). - *Handbook of educational psychology*, New York, Macmillan.
- Bonami M. (1996). - « Logiques organisationnelles de l'école, changement et innovation », in Bonami M., Garant M., *Systèmes scolaires et pilotage de l'innovation. Émergence et implantation du changement*, Bruxelles et Paris, De Boeck, Larcier, pp. 185-216.
- Cardinet J. (1990). - « Évaluation externe, interne ou négociée ? », in *Hommage à Jean Cardinet*, Neuchâtel, IRDP, Delval, pp. 139-157.
- Casalfiore S., Paquay L. (1998). - « Le courant de l'action située : regard sur l'activité enseignante et réflexion autour de la notion de qualification des enseignants », Communication au symposium du REF à Toulouse (à paraître aux éditions De Boeck, dans l'ouvrage dirigé par M. Cl. Dauvisis).
- Crahay M., Lafontaine D. (1986). - *L'art et la science de l'enseignement*, Bruxelles, Labor.
- Cruikshank D.R., Haefele D.L. (1990). - « Research-based indicators : Is de glass half-full or half-empty ? », *Journal of personnel evaluation in education*, 4, pp. 33-39.
- Derry S., Lesgold A. (1996). - « Toward a situated social practice model for instructional design », in D.C. Berliner, R.C. Calfee (éd.), *Handbook of educational psychology*, New York, Macmillan, pp. 787-806.
- Durand M. (1996). - *L'enseignement en milieu scolaire*, Paris, PUF.
- Dwyer C.A., Stufflebeam D. (1996). - « Teacher evaluation », in D.C. Berliner, R.C. & Calfee (éd.), *Handbook of educational psychology*, New York, Macmillan, pp. 765-786.
- Gauthier Cl. (éd.) (1997). - *Pour une théorie de la pédagogie. Recherches contemporaines sur le savoir des enseignants*, Bruxelles, De Boeck-Université (avec la collaboration de J.-Fr. Desbiens, A. Malo, St. Martineau, D. Simard).
- Good T.L. (1996). - « Teaching effects and teacher evaluation », in J. Sikula (éd.), *Handbook of research on teacher education*, New York, Macmillan, pp. 617-665 (2^e éd.).
- Joint Committee on Standards for Educational Evaluation (1988). - *The personnel evaluation standards*, Newbury Park, CA, Sage.
- Lessard Cl. (1998). - *La professionnalisation comme discours sur les savoirs des enseignants et les nouveaux modes de régulation de l'éducation*, Conférence de clôture du colloque du REF, Toulouse, 30 octobre 1998 (à paraître dans les Actes).

- Lessard Cl. (éd.) (1999). – *La formation des enseignantes et des enseignants : aspects comparatifs et prospectifs*, Étude réalisée pour l'ADEREQ, Montréal, CREFPE-Inter-universitaire, 76 p. (avec la collaboration de Y. Lenoir, D. Martin, M. Tardif, B. Voyer).
- McConney A. (1995). – « Introduction. Common ground for a unified approach », in Nevo, Stufflebeam, Langehein (éd.), *Studies in Education and Evaluation*, 21, 2, pp. 105-110.
- Nevo D., Stufflebeam D.L., Langeheine R. (éd.) (1995). – « Toward a unified model of educational personal evaluation », n° spécial de la revue *Studies in educational evaluation*, 21 (2), pp. 105-254.
- Paquay L., Altet M., Charlier E., Perrenoud Ph. (1996). – *Former des enseignants-professionnels. Quelles stratégies ? Quelles compétences ?* Bruxelles, De Boeck-Université (2^e éd., 1998).
- Paquay L. (1996). – « Quelle formation des enseignants pour améliorer l'école ? Un essai de prospective en Communauté française de Belgique », Communication au symposium du REF, à paraître dans un ouvrage coédité par M. Carboneau et M. Tardif *Les réformes en éducation et leur impact sur la formation des maîtres*, aux éditions De Boeck-Université.
- Paquay L., Darras E., Saussez F. (à paraître). – « Pratiques d'autoévaluation dans les formations professionnalisantes. Enjeux, écueils, perspectives », in G. Figari (éd.), *L'activité évaluative réinterrogée*, Bruxelles, De Boeck-Université.
- Perrenoud Ph. (1996). – *Enseigner : agir dans l'urgence, décider dans l'incertitude*, Paris, ESF.
- Perrenoud Ph. (1998). – « L'établissement scolaire entre mandat et projet : vers une autonomie relative », Conférence d'ouverture du colloque international *Autonomie et évaluation des établissements : l'art du pilotage au temps du changement*, Montréal, 14 octobre 1998, 29 p. (à paraître dans les Actes).
- Peterson D., Kromfrey J., Smith D.C. (1990). – « Research-based teacher evaluation : A response to Scriven », *Journal of personnel evaluation in education*, 4, pp. 7-17.
- Rosenshine B. (1986). – « Vers un enseignement efficace des matières structurées. Un modèle d'action inspiré par le bilan des recherches processus-produit », in M. Crahay, D. Lafontaine (éds), *L'art et la science de l'enseignement*, Bruxelles, Labor, pp. 81-96.
- Rosenshine B., Edmonds J. (1990). – « New sources for improving instruction : the implicit skills study », *Journal of personnel evaluation in education*, 4, pp. 59-73.
- Roth R.A. (1996). – « Standards for certification, licensure and accreditation », in J. Sikula (éd.), *Handbook of research on teacher education*, New York, Macmillan, pp. 242-278 (2^e éd.).

Saussez F., Allal L. (à paraître). - « Réfléchir sur sa pratique ; quel rôle pour l'autoévaluation ? », in L. Paquay, *Évaluation et formation des enseignants*, Bruxelles, De Boeck.

Schab R.L. (dir.) (1991). - *Research based teacher evaluation*, Boston/Dordrecht/London, Kluwer Academic Publishers.

Scriven M. (1987). - « Duty-based teacher evaluation », *Journal of personnel evaluation in education*, 1, pp. 319-334.

Scriven M. (1988). - « Validity in personnel evaluation », *Journal of personnel evaluation in education*, 1, pp. 9-23.

Scriven M. (1990). - « Can research-based teacher evaluation be saved? », *Journal of personnel evaluation in education*, 4, pp. 19-32.

Scriven M. (1995). - « An unified theory approach to teacher evaluation », *Studies in educational evaluation*, 21, pp. 111-129.

Sikula J. (Ed.) (1996). - *Handbook of research on teacher education*, New York, Macmillan (2^e éd.).

Simon M., Demers R., Dahan M., Forgette-Giroux R. (1994). - « Opérationnalisation d'un modèle d'évaluation "sur mesure" du personnel enseignant », *Mesure et évaluation en éducation*, 17 (1), pp. 45-62.

Strike K.A. (1990). - « Is teaching a profession : how would we know? », *Journal of personnel evaluation in education*, 4, pp. 91-117.

Stufflebeam D.L. (éd), Foley W. J, Gephart W. J, Guba E.G., Hammond R.L., Merriam M.O., Provus M.M. (1971). - *Educational evaluation and decision making*, Itasca, Peacock (trad. : *L'évaluation en éducation et la prise de décision*, Ottawa, NHP, 1980).

Stufflebeam D.L. (1995). - « Evaluation of superintendent performance : toward a general model », *Studies in educational evaluation*, 21 (2), pp. 153-226.

Tellez K. (1996). - « Authentic assessment », in J. Sikula (éd.), *Handbook of research on teacher education*, New York, Macmillan, pp. 704-721 (2^e éd.).

Telot Cl. (1993). - *L'évaluation du système éducatif*, Paris, Nathan.

Vandenberghe, V. (1998). - « Educational Quasi-Markets : the Belgian Experience », in W. Bartlett, J.A. Roberts et J. Le Grand (éd.), *A Revolution in Social Policy. Lessons from developments of Quasi-Markets in the 1990s*, Bristol, the Policy Press.

Webster J.W. (1995). - « The connection between personnel evaluation and school evaluation », *Studies in educational evaluation*, 21 (2), pp. 227-254.

Zuber M. (1998). - « Les sciences de l'éducation aux États-Unis », in J.-Cl. Ruano-Borbalan, *Éduquer et former*, Paris, Éd. Sciences humaines, pp. 65-71.