

A large, bold, black letter 'V' is centered within a light gray rectangular box. This box is part of a vertical gray bar that runs down the left side of the page.

# ARIA

## En Europe : l'évaluation contre la crise des systèmes scolaires, l'évaluation en crise

---

LISE DEMAILLY

CLERSE/IFRESI, USTL

(Centre Lillois d'études et de recherches en sociologie et économie,

Université des sciences et techniques de Lille)

2, rue des Canonniers

59800 – Lille

Dans le paysage international du management des systèmes éducatifs, une “technique sociale” (Demailly 2000) semble s'inscrire durablement et y constituer un outil central des nouveaux modes de régulation : l'évaluation. Nous entendons par là l'évaluation des politiques, des organisations et des pratiques scolaires.

La comparaison internationale permet de dégager quelques grandes lignes de convergence en Europe quant aux usages de l'évaluation. Ces convergences confirment des analyses antérieures sur la nature sociologique de l'évaluation, comme pratique composite, à l'articulation de jeux d'intérêts et de production de valeurs et de connaissances (Demailly 1995). Elles appellent des hypothèses sur son enjeu politique dans l'espace européen. La réflexion proposée ici interroge un paradoxe. Jamais l'évaluation n'a été aussi légitime, autant perçue comme un remède à la crise des systèmes scolaires et jamais pourtant les dispositifs d'évaluation n'ont été aussi unanimement critiqués pour leurs outils (manque de rigueur scientifique des indicateurs), leurs usages (lourdeur bureaucratique, illisibilité, coût), leurs effets (déstabilisations stériles, stigmatisations liées aux palmarès, inapplicabilité des préconisations...). Comment comprendre les raisons de ce paradoxe ?

Il s'agit ensuite de repérer les différences nationales dans les usages de l'évaluation, car celles-ci se maintiennent malgré les pressions à la convergence européenne. Ce problème politique pose la question théorique de la nature, de la force et des limites de l'“effet sociétal” dans les fonctionnements des systèmes scolaires. Réciproquement, la manière de modéliser ceux-ci est un problème scientifique, qui n'est pas sans enjeu proprement politique.

Enfin, l'article s'attache aux évaluations de proximité, d'une grande diversité, mais qui présentent quelques aspects spécifiques transnationaux. Il prend en compte les différents types de dispositifs d'évaluation : aussi bien les évaluations internationales fondées sur des indicateurs quantitatifs comme PISA (Programme for International Student Assessment) que, à une autre extrémité du spectre, les petites évaluations locales qualitatives conduites par un groupe d'enseignants ou de chefs d'établissement dans un bassin scolaire ou une ZEP (Zone d'éducation prioritaire). Ceci pour la raison épistémologique suivante : la segmentation des perspectives subjectives des acteurs impliqués dans les dispositifs d'évaluation en est une caractéristique constante. Il est donc important pour le chercheur de s'intéresser aux différents échelons du système, sous peine d'être victime de la même étroitesse de perspective que les acteurs sociaux.

De fait il est possible et intéressant de dégager une cohérence d'analyse sur l'ensemble des segments institutionnels du phénomène de l'évaluation. Et l'intérêt pour chaque segment permet à la fois de percevoir ses enjeux propres (un dispositif local d'évaluation ne peut pas être construit comme un critère standardisé international, et c'est beaucoup plus difficile) et d'identifier les enjeux théoriques ou politiques liés à l'"universelle" émergence de l'évaluation.

L'article s'appuie, pour les données empiriques originales, sur les résultats de "Reguleducnetwork"<sup>1</sup>, recherche pionnière sur l'évolution des modes de régulation des systèmes scolaires, qui a procédé en étudiant les politiques nationales, les niveaux intermédiaires, les espaces locaux d'interdépendance entre établissements et les établissements eux-mêmes. Il mobilise aussi, de manière synthétique, les travaux personnels antérieurs de l'auteur qui ont porté depuis dix ans sur certains dispositifs d'évaluation en France (audits d'établissements scolaires, contrat de réussite ZEP...).

## **Généralisation de l'évaluation comme stratégie de pilotage des enseignements scolaires européens**

L'évaluation de l'action publique se développe effectivement dans tous les pays en Europe. C'est même un élément important de convergence des systèmes européens. Elle se généralise sur la base d'une rhétorique commune de modernisation des systèmes et elle présente, sur le plan sociologique, trois aspects transnationaux :

---

1 Reguleducnetwork est l'acronyme d'une recherche du 5<sup>e</sup> PRCD, coordonnée à l'échelon international par C. Maroy : "Changement des modes de régulation et production sociale des inégalités dans les systèmes d'éducation : une comparaison européenne". Pour la France, j'ai coordonné l'équipe lilloise du CLERSE-IFRESI-CNRS. Je remercie la commission européenne qui a permis ce travail de trois ans, mené en Angleterre, Belgique francophone, France, Hongrie et Portugal.

- elle est préconisée avec insistance par le haut ou par l'extérieur des systèmes nationaux ou locaux ;
- les dispositifs sont en crise récurrente et ont une légitimité scientifique et politique souvent faible ;
- le développement de l'évaluation est cependant peu résistant car il s'inscrit dans des transformations du management du travail humain et des modes actuels de 'gouvernementalité'" (Foucault 2001), toutes choses qui excèdent largement la question scolaire.

Ces trois points sont développés successivement.

## **Une pression générale au développement de l'évaluation**

La généralisation de l'évaluation se joue à plusieurs degrés et mobilise de nombreux acteurs, mais l'impulsion vient en général du haut et du centre des systèmes, même si les usagers sont généralement aussi demandeurs.

Sous un premier aspect, la généralisation résulte d'une pression d'acteurs internationaux ou nationaux pour que les États participent à la fabrication des évaluations internationales (évaluations psychométriques sur les compétences et connaissances des élèves, avec comparaison internationale et classement entre pays) et prennent en compte ces évaluations dans l'élaboration de leurs politiques (cf. les textes du Haut comité de l'évaluation et de l'OCDE, Bottani & Vriegnaud 2004, Cyterman & Demeuse 2005). La commission européenne et les gouvernements, comme sur d'autres dossiers, pratiquent un management qui n'est pas de contrainte ouverte mais de persuasion insistante, avec mises en compétition symboliques entre pays, qui visent à ce que les acteurs, objets de cette persuasion, se disent : "il ne faut pas que nous soyons en retard sur les autres". La nécessité de participer aux évaluations internationales s'argumente des résultats mêmes de l'évaluation (en Allemagne et en Communauté Française de Belgique) ou des blasons de la modernité : savoir évaluer, accepter l'évaluation, c'est être moderne. La persuasion s'appuie sur la crise des systèmes scolaires dans ses différentes dimensions.

Sous un second aspect, la pression à l'évaluation prend la forme d'une incitation internationale à la mise en place d'outils nationaux ou régionaux d'évaluation dans les États qui n'en disposent pas.

Enfin, une troisième impulsion se joue du côté des médias. Ils relaient l'incitation à ce que ces évaluations internationales soient prises en compte par les établissements et les enseignants. Ils font de même pour les palmarès d'établissements, dans les pays où ils existent et où ils sont publics, et s'appuient sur la demande d'information émanant des classes moyennes et supérieures ou les inquiétudes des milieux populaires, voire les instrumentalisent.

Qui sont donc les initiateurs des dispositifs d'évaluation ? Les organisations internationales, les États (pas tous), les niveaux intermédiaires (de manière iné-

gale), les services pédagogiques privés marchands (en Angleterre et en Hongrie), les chercheurs en sciences humaines et sociales. Il est très rare que les enseignants ou les chefs d'établissements soient à l'initiative d'un dispositif d'évaluation.

L'évaluation présente plusieurs intérêts pour les dirigeants politiques ou administratifs.

Avantage politique : c'est une tentative de sortie de crise et de relégitimation de l'action publique (Monnier 1992 montre bien le lien entre le développement de l'évaluation de l'action publique et les moments de crise des États).

Avantage idéologique : c'est l'occasion de développer une rhétorique commune, transnationale, qui articule évaluation, pilotage, management, modernisation, nouveau mode de régulation. Sa particularité est de faire de la question des dysfonctionnements des systèmes un problème de management et du management un problème de connaissance, de mesure (et non pas une question de finalité politique, ou un problème économique ou de choix culturel ou de société).

Avantage pratique : c'est un renforcement du contrôle sur les enseignants. Dans le cadre de Reguleducnetwork, nous avons pu observer une tendance commune aux systèmes scolaires européens (Maroy 2005) au renforcement des régulations de contrôle. Sans doute les textes praxéologiques sur l'évaluation distinguent-ils souvent normativement l'"évaluation" du "contrôle". Mais ce qui se développe sous le nom d'évaluation en Europe est un mode de contrôle social des enseignants et des établissements. La montée générale des régulations intermédiaires, une autre des tendances convergentes que nous avons identifiées (Maroy & Demailly 2004), favorise l'émergence d'un contrôle de proximité quelle que soit la situation de départ des différents pays. Dans les pays anciennement centralisés, où les services centraux de l'État étaient distants des terrains, comme la France ou le Portugal ou la Hongrie, dans le cadre de la décentralisation les évaluateurs se rapprochent des professionnels de première ligne, maîtrisent mieux le recueil de données grâce à la sophistication croissante des outils d'enregistrement. Dans les pays anciennement décentralisés, l'évaluation (recentralisée comme en Angleterre) vient diminuer les autonomies et libertés professionnelles.

## **Une deuxième constante : des dispositifs segmentés, à la légitimité fragile**

La comparaison entre pays montre d'autres ressemblances transnationales dans la mise en place des évaluations : leur efficacité est faible en matière de rationalisation des modes de pilotage. Une première raison en est l'émiettement concurrentiel des évaluations. La segmentation de ces dispositifs, leur empilement peu coordonné sont la règle.

Segmentation disciplinaire : à l'international, les évaluateurs font de la psychométrie ; à l'échelon local de la sociologie des organisations et des sciences de l'éducation ; dans la classe de la pédagogie ; dans l'établissement du marketing (on évalue l'image, l'attractivité) ; à l'échelon intermédiaire de la stratégie de changement organisationnel ; nationalement de la gestion de flux.

Segmentation des objets : on évalue les résultats des élèves, les programmes, les dispositifs, les personnels, les établissements, les actions, l'effet de telle mesure (le redoublement, la diminution du nombre d'élèves par classe...) et il est difficile que ces évaluations cloisonnées donnent des pistes cohérentes et incontestables pour la décision publique.

Segmentation des initiateurs : en France, plusieurs niveaux hiérarchiques fabriquent des outils qui se superposent et se concurrencent. En Angleterre, chaque direction de programme tend à développer ses propres outils d'évaluation. Autour de ces évaluations, se constituent dans chaque pays des métiers différents : inspecteurs pédagogiques, inspecteurs administratifs, consultants privés (Angleterre et Hongrie), statisticiens, psychométriciens, animateurs de "bassins" et territoires de types divers, qui développent des méthodologies spécifiques, plus ou moins formalisées, des attentions différenciées aux objets potentiels d'évaluation et des points aveugles différents. Ces concurrences entretiennent la vitalité du champ de production de l'évaluation et en légitiment l'idée, mais elles produisent une forte segmentation des résultats.

De ce fait, les organisations sont amenées à construire des dispositifs transversaux pour essayer d'articuler les différentes pratiques, les outils, les résultats, les points de vue, les cultures professionnelles des évaluateurs. Des lieux sont créés pour opérer des synthèses, en une recherche tâtonnante du "bon territoire" de pilotage, qui reste introuvable. Opérer cette synthèse a été par exemple une des motivations de l'audit des établissements secondaires mis en place dans l'académie de Lille, expérience innovante à grande échelle en France puisqu'elle a pour la première fois concerné la totalité des collèges et lycées d'un territoire régional. Notre observation nous a permis de constater que la synthèse est difficile : malgré une méthodologie soignée de saisir l'établissement dans sa globalité, malgré un travail en concertation mobilisant différents types d'évaluateurs, l'évaluation de ce qui se fait dans les classes est restée déconnectée de l'évaluation du management et de la communication interne dans l'établissement (Demailly, Gadrey & al. 1998).

D'autre part, les évaluateurs perdent toujours en route un segment de l'évaluation globale visée. En ce qui concerne les audits lillois, c'est la pédagogie qui a été sous-étudiée, ainsi que l'usage des crédits. Pour les audits d'académie réalisés par l'inspection générale, de l'aveu même des responsables (Inspection Générale de l'Éducation nationale 2003), ce sont les finalités politiques qui ont été perdues de vue, par la trop grande attention portée à l'usage des outils du management.

La segmentation rationnelle du réel produite par les dispositifs d'évaluation fait toujours courir le risque d'une perte de sens d'une part, d'une perte de capacité analytique d'autre part, d'une redondance paperassière enfin.

Une cause et une conséquence de cette segmentation du travail d'évaluation sont que les résultats produits par chacun de ces niveaux intéressent peu les autres acteurs. Par exemple, dans l'ensemble de l'Europe, les enseignants s'occupent très peu du coût de l'éducation ou de PISA. Chaque ensemble évalué est surtout retenu par les données qui peuvent servir sa propre conduite d'action et pas par les autres. En France, la DEP (Direction de l'évaluation et de la prospective) a fini par reconnaître que les établissements ne faisaient pas grand-chose des Indicateurs Pour le Pilotage de l'Établissement scolaire. Ce sont surtout les parents, les chefs d'établissement et les élus locaux qui consultent les palmarès d'établissements. Mais les élus nationaux les analysent peu, ils préfèrent PISA. Les États, eux, s'intéressent très peu aux évaluations qualitatives produites par les établissements, ils les méprisent comme un bricolage peu scientifique, les soupçonnent de complaisance et d'esquiver la question des résultats scolaires des élèves (Entretiens pour Reguleducnetwork et Verdière 2002).

La communication autour de l'évaluation est un grand dialogue de sourds en somme, avec une circulation parallèle de flux de commandes/réception/ utilisation de l'évaluation, qui se rencontrent très peu.

Du coup, à la base des systèmes –c'est-à-dire dans leurs petites composantes, établissements, bassins, etc.– la légitimité de l'évaluation est faible. Partout dans les systèmes scolaires européens, les experts et les cadres se plaignent de l'absence de culture de l'évaluation des enseignants ou de la faiblesse de l'outillage (quel que soit d'ailleurs le degré –fort inégal– de développement réel de l'évaluation dans les différents pays), alors même que les enseignants disent qu'ils n'attendent pas les évaluateurs professionnels pour estimer l'efficacité de leur propre action. Les évaluations sont le jeu de délégitimations croisées à toutes les échelles et de toutes les sources, des professionnels de première ligne –il leur arrive de ressentir le monde des comptes rendus produits par l'évaluation comme un double fictionnel du monde réel qui se met à vivre de sa vie propre– comme du champ scientifique : les chercheurs peuvent délégitimer les méthodes de l'administration (Felouzis 2004 sur les travaux de la DEP). Il ne manque pas de critiques sur la façon dont les indicateurs de PISA ont été construits. Il y a autant de production de discours intellectuel, expert ou profane pour délégitimer l'évaluation que pour la légitimer, ce chassé-croisé occupe beaucoup de temps social de travail.

### *Évaluation et nouvelle gouvernamentalité*

Comment rendre compte du paradoxe de la légitimité de l'évaluation et de l'illégitimité des dispositifs d'évaluation ?

Dans sa rhétorique, l'évaluation est liée à la modernisation du management, à la mise en place de nouvelles modalités de régulation. Elle a été rendue nécessaire par la crise des systèmes d'enseignement, qui semblait impliquer, partout, de modifier les modes de pilotages. Ce qui en a favorisé le développement, ce qui a permis l'explosion de cette technologie du social déborde largement la question éducative. C'est pratiquement le développement des réseaux informatisés qui en est le support matériel. Et il est encouragé idéologiquement par le développement du post-fordisme (augmentation du contrôle sur le travail, recherche d'efficacité et d'efficacités) et la globalisation (développement de la concurrence entre États ou entre ensembles d'États).

Les nouveaux modes de gouvernementalité passent par la psychologisation, l'individualisation, l'extorsion de signes de bonne volonté (Aballéa & Demailly 2005 sur les nouvelles formes de mobilisation des salariés et de management des subjectivités, Fassin & Memmi 2004). L'autoévaluation spontanée ou la bonne réaction face à l'évaluation externe sont un de ces signes de soumission que réclame le management moderne. Ces signes valent en tant que tels, indépendamment souvent du contenu de l'évaluation qui finalement peut n'avoir pas beaucoup d'importance. Ainsi en ZEP où on sait d'avance que les résultats ne vont pas être très bons, nous avons pu l'observer lors de la première mise en place de l'évaluation du contrat de réussite ZEP (Demailly 2003) : l'important est que les acteurs acceptent l'évaluation et participent à la mise en scène publique du jeu de l'évaluation.

L'évaluation est une pratique stratégique (politique), une pratique de pouvoir et une pratique de production de connaissances. C'est très précisément un dispositif de gouvernementalité (pratique de production de pouvoir et de savoir) au sens de Foucault : multiplication des sources et des objets de pouvoir, remplacement de la loi par la norme explicitement négociable, multirégulation fragmentée. De ce fait, les dispositifs d'évaluation ne peuvent bénéficier de la stabilité d'un cadre normatif de type disciplinaire et substantif. La négociation constante, les jeux d'acteurs ininterrompus érodent leur valeur cognitive et instrumentalisent leur valeur normative.

## Des modèles sociétaux différents

Malgré le travail d'influence et de mise en circulation de maîtres-mots assuré par les instances internationales et malgré les points communs mis en évidence ci-dessus, il serait illusoire de croire que l'évaluation accomplit dans les faits une unification et une homogénéisation des modes de pilotage, à l'échelle de l'Europe, ni même que l'outil est utilisé de manière similaire. Le poids des modèles sociétaux est fort, tant sur les modalités des évaluations que sur leurs usages sociopolitiques.

## **Des implantations différentes dans des systèmes différents**

Tout d'abord, le développement de l'évaluation, qui implique l'existence d'appareils statistiques, est fort inégal. Dans l'échantillon de Reguleducnetwork, le pays qui a les systèmes d'évaluation les plus développés est l'Angleterre, suivi de la France. La Hongrie, qui ne disposait pas de tradition en la matière, rattrape à grands pas son retard, suivie du Portugal. La Communauté française de Belgique présente une configuration extrême, elle est la plus dépourvue d'instruments d'évaluation : pas d'examens nationaux pour comparer le niveau des élèves, appareil statistique très faible, évaluation des établissements très peu légitime.

L'Angleterre et la Hongrie ont vu se développer des services privés pour le conseil et l'évaluation externe, alors que la France et le Portugal font surtout travailler l'administration publique. Mais alors que l'Angleterre dispose d'une expertise forte et innovante en matière d'outillage de l'évaluation, la Hongrie connaît un déficit de compétence de son administration car l'ancienne a été détruite avec le changement de régime politique. Les palmarès publics d'établissement n'existent qu'en France et en Angleterre. Les évaluations d'établissements sont inégalement pratiquées en Europe. Elles sont aussi, quand elles existent, très inégalement suivies de sanctions concrètes contre les équipes ou les établissements ; c'est le cas en Angleterre, mais pas en France.

Complétons ce tableau en dehors de l'échantillon de Reguleducnetwork. Dans les pays scandinaves, l'évaluation est moins développée, notamment il n'y a pas d'évaluation d'établissements mais des municipalités (Wastiau-Schlüter 2004).

On peut trouver des cohérences, que nous n'avons pas la place de détailler ici, entre la manière dont fonctionne l'évaluation et d'autres caractéristiques des systèmes telles que les formes et les degrés d'autonomie de l'établissement scolaire. Il y a par exemple une cohérence entre le principe constitutionnel de liberté pédagogique en Communauté française de Belgique (les établissements choisissent leurs élèves, les familles choisissent leur établissement, il n'y a pas de curriculum national) et le très faible développement de l'évaluation, sous toutes ses formes.

## **Comment modéliser les différences ?**

La question épistémologique de la modélisation des différences sociétales s'inscrit dans celle de la méthodologie des comparaisons internationales (Vigour 2005, Lallement & Spurk 2003). Il y a deux types de modélisation possibles : nomologique (agencement de variables) ou typologique. La modélisation typologique peut construire des idéaux types à utilisation combinatoire, utilisant un raisonnement par traits, par dimensions ; la réalité empirique est alors une combinaison



de deux trois idéaux types à dosages variables (chez Weber, la typologie de l'action sociale). Ou bien des idéaux types à utilisation historique, qui renvoient à des entités empiriques idiosyncrasiques (chez Weber, occident, capitalisme).

Notre souci a été de penser la diversité des usages de l'évaluation en rapport avec les autres caractéristiques institutionnelles et organisationnelles des systèmes, telles l'autonomie de l'établissement et le caractère plus ou moins démocratique de l'enseignement. La première solution méthodologique, autrement dit les combinaisons de variables statistiques pour spécifier les systèmes scolaires nationaux et articuler leurs caractéristiques internes et leurs capacités démocratiques, s'est révélée décevante, vu l'extrême fragilité des indicateurs. Construire un indicateur international d'autonomie de l'établissement scolaire par exemple a déclenché de fortes critiques au sein de l'équipe de recherche. La démarche typologique a semblé finalement plus pertinente.

Elle permet d'identifier deux modèles de régulation en déclin : le modèle bureaucratique-professionnel (France et Portugal dans une moindre mesure, car le fonctionnement bureaucratique est beaucoup plus facile et efficace dans ce petit pays), qui se caractérise par une évaluation nationale des élèves et une évaluation des personnels faisant une large part à l'ancienneté. Le deuxième modèle (Belgique) est libéral au sens strict du terme, caractérisé par la totale liberté pédagogique de l'établissement et la quasi-absence d'évaluation des élèves et des établissements.

En ce qui concerne les modèles émergents, on peut présenter les constats de deux façons, selon une typologie par dimension ou selon une typologie par cas.

Par dimension, il aurait deux idéaux types de dimensions en émergence : le quasi-marché (la concurrence entre établissements permet d'augmenter les performances des établissements), l'État évaluateur (montée de l'évaluation externe centralisée organisée par l'État), qui se combineraient de façon variable et indépendante l'un de l'autre, à faibles ou hautes doses. Les deux dimensions combinées et à fortes doses caractériseraient l'Angleterre. La prédominance du quasi-marché avec une montée de l'État évaluateur, la Belgique. La France relèverait surtout de l'État évaluateur avec une montée du quasi-marché.

Si on utilise des idéaux types de cas, on pourrait identifier un modèle émergent "autoritaire marchand" où le rôle de l'État évaluateur et celui du marché scolaire formeraient ensemble un modèle consistant (l'Angleterre fournirait l'accomplissement le plus approchant de cet idéal type). Un deuxième modèle émergent serait le "modernisme de service public", combinaison de traits bureaucratiques et post-bureaucratiques, avec une méfiance vis-à-vis de la concurrence entre établissements. Plusieurs académies françaises en fourniraient une réalisation (Lille, Clermont-Ferrand, Rennes), même si par ailleurs elles connaissent une certaine extension du processus de marché scolaire et de la concurrence entre établissements dans les grandes villes. Les pays scandinaves relèveraient de ce modèle.

Regulateducnetwork a plutôt soutenu la première modélisation (rapport final 2005), mais notre préférence personnelle va plutôt à la seconde car, nous semble-t-il, la première masque le lien aujourd'hui consubstantiel entre le marché et le contrôle. Le modèle de quasi-marché seul ne fonctionne pas comme modèle émergent dans la réalité, c'est une dimension analytique abstraite, car un marché sans État évaluateur (c'est le cas du modèle traditionnel belge) connaît, malgré la concurrence entre établissements, peu de compétition sur les résultats et donc finalement une performance globale faible, les établissements se créant des niches d'adaptation avec leur clientèle et s'y routinisant. Le fonctionnement performant du marché scolaire implique aujourd'hui, dans le cadre concurrentiel international en management moderne, le recours à l'évaluation (ce qui est d'ailleurs en train d'advenir en Belgique française ou en Allemagne, après le choc des mauvais résultats à l'enquête PISA). Le modèle "autoritaire-marchand" a donc bien une cohérence historique propre.

Nous trouvons que la seconde modélisation rend mieux compte des évolutions historiques différenciées des pays et des conflits de modèles, notamment à la base –à l'échelle des petites composantes–, dont il ne faut pas sous-estimer l'impact historique (comme on vient de le voir avec la force d'une opinion anti-néolibérale en France). La première modélisation en revanche rend mieux compte des convergences et des diffusions par les élites d'experts internationaux, puisqu'elle permet de construire des sortes d'échelles avec des degrés de contamination d'idées. Le débat scientifique sur ce point reste donc ouvert.

## **Comment expliquer les différences ?**

L'explication pose aussi problème dans les comparaisons internationales. On peut brièvement évoquer quelques cadres explicatifs : l'analyse sociétale, anti-fonctionnaliste, (Maurice, Sellier & Sylvestre 1992) qui considère les sociétés nationales comme des systèmes spécifiques singuliers, l'analyse culturaliste (d'Iribarne 1991), différentes variétés de déterminisme qui accordent la priorité explicative à l'organisation économique, au niveau de vie, au développement technologique ou à la morphologie des sociétés (leur démographie). On peut aussi essayer de tout penser en termes de jeux d'acteurs : locaux, nationaux, internationaux, mais pour ce type de démarche, on est tout de même à un moment obligés de passer par des concepts globalisants.

Pour l'Europe scolaire, la diversité des structures administrativo-politiques paraît jouer un rôle important dans l'effet sociétal (Maroy & Demailly 2004). D'un côté, sur notre échantillon, nous trouvons deux pays anciennement centralisés sur le mode rationnel légal, la France et le Portugal, dans lesquels les modes de raisonnement faisant appel à la tradition bureaucratique démocratique restent

puissants ainsi que la référence à l'éthique du service public, même si des formes post-bureaucratiques de management émergent dans certaines régions françaises.

De l'autre, nous avons trois pays plus libéraux où la gestion de l'enseignement est beaucoup plus laissée à l'initiative privée (familles, organisations, intérêts religieux...). Ce libéralisme n'est certainement pas homogène. Un libéralisme autoritaire est identifiable en Angleterre, avec le développement de la culture de l'objectif et de la performance, une centralisation progressive qui diminue fortement le pouvoir du local en matière d'éducation. Un libéralisme communautaire en Belgique, où le principe de liberté de l'éducation s'inscrit au sein de ces réseaux communautaires qui sont les piliers de la société belge, une valorisation très forte du local (pouvoirs publics locaux ou initiatives locales) dans un contexte d'État faible en matière d'enseignement. Enfin, la société hongroise développe le libéralisme peut-être le moins régulé où, au moment du changement de régime, l'autonomie des écoles et des collectivités locales a été portée aux nues tandis que les récents efforts visant à encourager la concertation et la planification ont eu peu d'effets.

Les principes de légitimité historique opposent donc d'un côté celles de l'État, de la nation, du volontarisme politique, de l'autre celles du local, de l'initiative privée (familiale, non marchande ou marchande), du marché dont la main invisible serait productrice de qualité. Ces principes restent puissants et informent les représentations et actions des agents à tous les échelons, même s'ils sont en voie de déstabilisation.

## Des processus locaux diversifiés

Conjointement avec les deux modèles institutionnels d'évaluation en émergence, "autoritaire-marchand" et "moderniste de service public", il y a deux pratiques idéal-typiques de l'évaluation d'origine locale qui portent sur des actions pédagogiques, des équipes d'enseignants ou de petits territoires géographiques :

- une évaluation "inamicale" aux établissements et aux équipes, où il s'agit de trier le bon grain de l'ivraie, avec comme cadre normatif une obligation de résultat impérative, des sanctions à l'égard de ceux qui ne fonctionnent pas bien, des mises en concurrence et compétitions entre établissements ;
- une évaluation "amicale" : formatrice, solidaire, qui vise à soutenir ceux qui ne marchent pas de manière satisfaisante. Ceci, avec, comme cadre normatif, une obligation incitative d'attention aux résultats de l'action, le rappel à une solidarité nécessaire entre établissements pour lutter contre les mécanismes spontanés de concurrence.

Ces deux modèles praxéologiques se rencontrent rarement de manière pure. Les pratiques d'évaluation sont influencées par les choix politiques et donc

forcément ambiguës puisque les politiques le sont. Par exemple, dans l'académie de Lille, même si on peut estimer que c'est le "modernisme de service public" qui domine, il coexiste avec des fonctionnements traditionnellement bureaucratiques et des tentations néolibérales, fortes dans les années 1985 (Demailly, van Zanten & al. 2004), dont on voit encore les traces dans le positionnement acquis de certains établissements.

Mais le lien avec les politiques peut être distendu car la technicité de l'évaluation amène à mettre en jeu des valeurs politiques ou professionnelles qui ne sont pas forcément celles énoncées par les responsables politiques locaux. Un consultant, un universitaire, un enseignant en promotion professionnelle (Demailly & Tondellier 2004) ont des intérêts propres, qu'ils font jouer dans leur conception de l'évaluation. De plus, les acteurs de base ont des marges de manœuvre : dans un dispositif qui relèverait en principe d'une évaluation amicale, tel cadre peut jouer au petit chef et tenter une évaluation dure avec menace de sanction. Le phénomène inverse est vrai : par deux fois, sur les audits et sur le contrat de réussite ZEP, nous avons observé comment une évaluation annoncée au départ comme sanctionnante s'est finalement construite, en deux ans, comme évaluation formative (Demailly, Gadrey & al. 1998, Demailly 2003). L'approche microsociologique des dispositifs d'évaluation est donc nécessaire, au cas par cas, pour saisir leurs enjeux politiques et leurs effets sur les pratiques éducatives.

Mais, paradoxalement, pour les pratiques de proximité, au-delà de l'extrême diversité qui implique d'étudier les situations comme des configurations politiques singulières, certains éléments de convergence transnationale se retrouvent. Tout d'abord sur la professionnalité des agents locaux de l'évaluation de proximité.

## **Les professionnalités des agents de l'évaluation**

Entendons par agents locaux d'évaluation l'ensemble de ceux qui dirigent les dispositifs d'évaluation, ou réalisent des évaluations de proximité, ou mobilisent les professionnels de première ligne pour qu'ils coopèrent à des démarches d'autoévaluation collectives ou encore cherchent à diffuser et expliquer les évaluations quantitatives réalisées à l'échelon régional, national ou international.

Ces agents sont de statuts extrêmement divers, mais ils gagnent leur vie au moins partiellement en contribuant, de façon elle aussi diverse, à l'évaluation de proximité. Leur pratique est, de façon transnationale, soumise à deux contradictions : elle oscille entre contrôle et intéressement, entre enracinement et expertise (Maroy & Demailly 2004).

La tension entre contrôle et intéressement concerne surtout les cadres de proximité. Ils présentent des caractéristiques communes : expériences positives

et bien vécues d'enseignants, intérêt vif pour la pédagogie, démarche de formation et de développement professionnel, voire de mobilité professionnelle. Ils constituent bien une élite professionnelle au sein du corps enseignant, élite qui peut à la fois se réclamer de l'expérience du terrain des professeurs, mais aussi s'en détacher. Ils doivent dans leur travail maintenir l'équilibre entre l'instauration d'un rapport d'intéressement (Callon 1986) et un souci d'évaluation, qui ne mette pas en péril l'enrôlement des acteurs de base. Leur travail est bien apprécié par leurs pairs dès lors qu'ils ne construisent pas une distance techniciste à la pratique pédagogique quotidienne.

Une deuxième tension fait hésiter les agents d'évaluation entre un statut symbolique d'expert et celui d'animateur. La position de l'expert peut être illustrée, en France, par la figure de l'inspecteur pédagogique régional ou la "personne-ressource", en Hongrie et en Angleterre, par celle du consultant qui travaille en libéral ou comme salarié selon un contrat à relativement court terme. Elle lui permet de ne pas devoir assumer avec les terrains des relations de travail dans la longue durée. L'expert peut donc tenir le discours de la qualité, de l'efficacité, de l'objectivité, du changement. À l'inverse, il risque de faire des préconisations sans prise réelle sur la réalité. L'animateur de terrain, lui (chef d'établissement, animateur ZEP), connaît bien un espace singulier dont il maîtrise les jeux, les habitudes, l'histoire, il peut y identifier les possibilités d'évolution ou d'alliance. Mais il risque de s'y engluier.

Les figures de la professionnalité des agents de l'évaluation, y compris dans leurs tensions, manifestent donc une forte convergence d'un pays européen à l'autre. L'évolution de leur travail, également.

## **La rationalisation de l'encadrement**

Le travail des agents d'encadrement est actuellement l'objet d'un processus de rationalisation, qui s'exerce par trois voies : la réflexivité, l'outillage, la mise en concurrence.

La réflexivité est le fait que les cadres de proximité et les cadres moyens sont incités à prendre le temps de discuter entre eux et avec d'autres de leur manière d'agir. Les hauts cadres commanditent à des chercheurs universitaires des études qui leur renvoient une image de leur mode de fonctionnement et de ses effets. On peut penser aussi au développement important de la formation de formateurs, des multiples séminaires pour les agents de régulation où une "culture commune de l'encadrement" et une "culture de l'évaluation" sont présentées comme un idéal pratique, où la remise en cause de ses propres pratiques est une exigence qu'on ne saurait imposer aux autres (les enseignants, les établissements) sans se l'imposer à soi-même.

Le développement de l'outillage est assez général, avec un minimum en Belgique et un maximum en Angleterre.

La sous-traitance est la troisième méthode possible pour rationaliser le travail des agents de régulation, pratiquée en Hongrie et Angleterre. Elle implique une forme d'externalisation, de privatisation de services (pédagogiques ou de management scolaire) par le recours à des appels d'offres.

## **Les résistances de la base et les possibilités d'intéressement**

Une autre donnée transnationale est la résistance des personnels de base à l'évaluation. Cette résistance est d'autant plus facile que les statuts des personnels sont plus solides, d'autant plus difficile que les statuts des personnels sont flexibilisés, mais dans ce cas il s'agit d'un management par la contrainte (Courpasson 1997).

Il est intéressant de se demander dans quelles conditions peut se développer une attention aux résultats de l'action qui soit autonome, non contrainte, un intéressement réel à l'évaluation. L'observation de processus où se manifestent des résistances ou bien à l'inverse des mobilisations pour un travail d'évaluation (Demailly 2003) montre que les résistances ont plusieurs sources, qui concernent à la fois les intérêts des personnels, leurs valeurs et aussi les ressources cognitives perçues comme utiles. Il semble bien que la dimension décisive soit celle des valeurs partagées et que l'intéressement des professionnels à l'évaluation de leur propre pratique ne puisse se faire que sur fond d'une réelle confiance politique. Il est également nécessaire que les évaluateurs acceptent d'écouter et d'apprendre des évalués, y compris sur la fabrication des outils d'évaluation et sur les méthodes de dialogue autour de ces évaluations.

## **L'impact des processus locaux d'évaluation**

En ce qui concerne les évaluations locales, l'évaluation externe peut se construire en dialogue avec l'évaluation interne et vice-versa. À cette condition, elle peut contribuer au développement d'un espace public et réflexif de débat sur les pratiques.

Rappelons que les évaluations internationales ont une très faible influence directe sur les pratiques. Les palmarès nationaux, quand ils existent, jouent un rôle, celui d'influencer l'image que les enseignants ont de leur établissement, donc, de manière perverse, ils favorisent les processus de ghettoïsation.

Les évaluations qui ont un impact sur les pratiques professionnelles, les pratiques de classe, ressortissent essentiellement de l'évaluation de proximité effectuée par les cadres intermédiaires agents publics ou privés ou de l'autoévaluation assistée réalisée par les établissements ou de petits ensembles d'établissements ou des équipes.

Sous les deux conditions énoncées plus haut, confiance politique et processus complexe de co-apprentissage, l'évaluation peut contribuer au développement de la réflexivité de l'action des enseignants et chefs d'établissements.

## Conclusion

L'évaluation est un remède à la crise des anciens modes de pilotage des systèmes éducatifs, modes de management aveugles, régulés de manière routinière, sous-informés sur leur propre fonctionnement. Elle correspond au développement d'une réflexivité des États, des pilotes, des établissements, des équipes enseignantes et à ce titre est promise à une généralisation peu résistible et à un avenir certain, car elle s'inscrit dans un mouvement général de rationalisation des activités humaines et de recherche de qualité.

Mais ce n'est pas une pratique de connaissance pure et désintéressée, c'est une pratique stratégique, politique. Elle est donc prise dans les enjeux d'une part des modernisations des appareils d'État, de la construction européenne dans une période de forte poussée des modes de réflexion néolibéraux sur la gouvernance des sociétés et leur mise en compétition tous azimuts.

Elle est prise aussi dans les enjeux liés à la redéfinition des priorités, des finalités, des valeurs centrales des politiques éducatives. Bref, son sens politique n'est pas fixé et elle peut parfois être mise au service d'une action professionnelle éthiquement orientée contre les inégalités sociales et scolaires et les dominations ordinaires, d'autres fois au service d'une rationalité gestionnaire ou technocratique ou des intérêts sociaux de parents de classes supérieures.

Du coup, traversés par ces contradictions, les dispositifs d'évaluation sont eux-mêmes en crise permanente, que celle-ci s'exprime sur le plan épistémologique ou directement politique, que soient mises en cause en cause leur objectivité et leur rigueur, ou leur utilité et leur efficacité, voire même leur innocuité sociale.

## Références bibliographiques

### Sur le Reguleducnetwork 5e PCRD

- BARROSO J., AFONSO N., BAJOMI I., BERKOVITS B. & IMRE A., EROSS G. 2002 (Deliverable 3), *Synthèse des études de cas nationales*
- DEMAILLY L., MONFROY B., TONDELIER M. & VERDIÈRE J. 2003 (Deliverable 8 France) *Les processus et les agents de régulations intermédiaires dans le système éducatif français. Le cas de l'académie de Lille*
- DEMAILLY L., MAROY C. 2004 (Deliverable 9) *Convergences et divergences dans les régulations intermédiaires des systèmes scolaires européens*

- DEMAILLY L., van ZANTEN A., DA COSTA S., MONTFROY B. & TONDELLIER M.  
2004 *Rapport final France*
- MAROY C. & DEMAILLY L. 2004 "Les régulations intermédiaires des systèmes éducatifs en Europe quelles convergences ?", *Recherches Sociologiques-XXXV-2*, Belgique
- MAROY C. 2005 (Final Report) *Regulation and Inequalities in European Education Systems*

### **Bibliographie générale**

- ABALLÉA F. & DEMAILLY L. 2005 "Les nouveaux régimes de mobilisation des salariés", in Durand J.-P. & Linhart D. *Les ressorts de la mobilisation au travail*, Octares
- BOTTANI N. & VRIGNAUD P. 2004 *La France et les évaluations internationales des acquis des élèves*, Paris, HCEE, OCDE
- CALLON M. 1986 "Éléments pour une sociologie de la traduction", *L'année sociologique-36*, 169-208
- COURPASSON D. 1997 "Régulation et gouvernement des organisations. Pour une sociologie de l'action managériale", *Sociologie du travail-1/97*, 39-61
- CYTERMAN J.R. & DEMEUSE M. 2005 *La lecture des indicateurs internationaux en France*, Paris, HCEE, OCDE
- DEMAILLY L. 1995 "Observer les pratiques d'évaluation", *Clés-25*, 79-100
- DEMAILLY L. 2000 "Le mode d'existence des techniques du social" *Cahiers internationaux de sociologie-CVIII*, juillet, 103-124
- DEMAILLY L. 2003 "L'évaluation comme apprentissage et négociation", *Revue française de Pédagogie-142*, 115-129
- DEMAILLY L., GADREY N., DEUBEL P. & VERDIÈRE J. 1998 *Évaluer les établissements scolaires : enjeux, expériences, débats*, Paris, L'Harmattan, Logiques sociales
- DEMAILLY L. & TONDELLIER M. 2004 "Les transactions entre l'institution scolaire et ses agents : le cas des coordonnateurs des réseaux d'éducation prioritaire", *Les Sciences de l'éducation-37-2*, 7-33
- d'IRIBARNE P. 1991 "Culture et effet sociétal", *Revue Française de Sociologie-XXXII*, 599-614
- FASSIN D. & MEMMI D. (éds) 2004 *Le gouvernement des corps*, Paris, Éditions de l'EHESS
- FELOUZIS G. 2004 "Les indicateurs de performance des lycées, une analyse critique" *Éducation & formations-70-décembre*, 83-95
- FOUCAULT M. 2001 *L'herméneutique du sujet*, Paris, Gallimard
- INSPECTION GÉNÉRALE DE L'ÉDUCATION NATIONALE 2003 <[ftp://trf.education.gouv.fr/pub/edutel/syst/igen/rapports/evaluation\\_academies\\_2003.pdf](ftp://trf.education.gouv.fr/pub/edutel/syst/igen/rapports/evaluation_academies_2003.pdf)>
- LALLEMENT M. & SPURK J. (éds) 2003 *Stratégies de la comparaison internationale*, Paris, CNRS
- MAURICE M., SELLIER F. & SILVESTRE J.-J. 1992 "Analyse sociétale et cultures nationales. Réponse à P. d'Iribarne", *Revue Française de Sociologie-XXXIII-1*, 75-86
- MONNIER E. 1992 *Évaluation de l'action des pouvoirs publics*, Paris, Economica
- McGAW B. 2004 mars "Vers une éducation de qualité", *L'Observateur de l'OCDE-242*, Paris, OCDE
- VERDIÈRE J. 2002 *L'évaluation du travail d'enseignement*, Thèse, Université de Lille
- VIGOUR C 2005 *La Comparaison dans les sciences sociales*, Paris, Repères
- WASTIAU-SCHLÜTER P. 2004 *L'évaluation des établissements scolaires sous la loupe*, Bruxelles, Eurydice