

Portée et limites d'une évaluation des politiques et des pratiques éducatives par les résultats

MARIE DURU-BELLAT
IREDU, Université de Bourgogne, France

JEAN-PIERRE JAROUSSE
IREDU, Université de Bourgogne, France

Alors que les institutions éducatives consacrent une part importante de leur temps à l'évaluation des connaissances des élèves, l'évaluation des dispositifs et des réformes pédagogiques reste, particulièrement en France, une pratique peu répandue et souvent considérée avec quelque suspicion. Pourtant, l'évaluation des élèves elle-même pourrait ne pas être considérée seulement comme une mesure, en soi intéressante, de ce qu'ils savent à un moment donné mais aussi comme le résultat d'un processus pédagogique particulier. On dépasserait ainsi "l'évaluation constat" pour une évaluation visant à éclairer la genèse du phénomène constaté, pour mieux en comprendre les ressorts, et orienter de manière adéquate les actions ultérieures à des fins de régulation. Cette dualité des fonctions de l'évaluation, entre objectif de connaissance et objectif d'optimisation de l'action, vaut particulièrement concernant l'évaluation des politiques éducatives, qui peut, au-delà de ses finalités théoriques de régulation de l'action, constituer un analyseur du fonctionnement des institutions d'enseignement et des jeux d'acteurs qui y prennent place (Demaillly 2000b). Si l'évaluation suscite de fait tant de réticences et de débats, c'est notamment parce qu'elle conduit à expliciter les finalités éducatives et leur hiérarchisation à un moment et dans un contexte particuliers. Nous présenterons dans un premier temps les justifications et les méthodes d'évaluation des pratiques et des politiques éducatives, en adoptant la posture du chercheur ou de l'expert, sans analyser en eux-mêmes les dispositifs d'évaluation étatiques, et l'utilisation qu'en font les acteurs, comme, par

exemple, Dupriez et Maroy (2000) ou Demailly (2000a). Puis, nous analyserons leurs effets en terme de connaissance et de régulation des systèmes d'enseignement, ainsi que les problèmes et les réticences rencontrés.

Évaluer, puisqu'il faut choisir

L'évaluation des pratiques et des politiques éducatives s'impose si l'enfant est effectivement "au centre du système éducatif", l'essentiel étant alors ce qu'il "gagne" à fréquenter l'école. Si on prend au sérieux cette perspective, alors, dans des pays qui consacrent une part importante des fonds publics au financement de l'éducation, parce qu'on en attend des effets positifs sur les individus mais plus largement sur la collectivité, on va chercher à évaluer la manière optimale d'organiser le contexte scolaire pour maximiser les "résultats" de l'action pédagogique sur l'élève. Réciproquement, il serait irresponsable de se contenter d'appréhender les moyens mobilisés, sans se soucier de vérifier que ces moyens ont bien quelque effet sur les élèves. L'objet de l'évaluation sera donc d'instruire les choix entre différents modes d'organisation alternatifs en se référant à leurs effets sur le plan pédagogique et sans ignorer les incidences économiques de ces choix, dès lors qu'on entend servir l'intérêt des élèves avec un maximum d'efficacité.

Car dans le domaine éducatif comme dans d'autres, on n'échappe pas à la double nécessité de définir dans un premier temps ce qu'on souhaite "produire", puis, dans un second temps, de choisir la manière la plus efficace d'organiser sa "production". Un premier stade correspond aux orientations générales assignées aux systèmes éducatifs (transmission de savoirs, développement d'attitudes, insertion sociale et économique des citoyens, traitement équitable des différents groupes...), ou à tel ou tel dispositif pédagogique spécifique à visée plus limitée. Un second stade, plus technique, est celui des objectifs intermédiaires; il renvoie à la très grande variété des facettes du fonctionnement des systèmes éducatifs (niveau et type de qualification des formateurs, —mode— de groupement des élèves, méthodes pédagogiques...). Le choix de ces modes d'organisation est lourd de conséquences, car ils ne sont a priori ni également efficaces eu égard aux objectifs poursuivis ni également coûteux. Or une perspective comparatiste convainc aisément à la fois de la diversité des modes d'organisation et du caractère très lâche de la relation entre —les performances des systèmes de formation et— les moyens qui leur sont alloués (Jarousse & Mingat 1992, Jarousse 1999). Est ainsi mise en cause l'opinion commune selon laquelle la mobilisation de ressources nouvelles est la clef du développement des systèmes éducatifs. Cette faiblesse de la relation "moyens/résultats" témoigne de l'absence d'un mécanisme de justification des politiques éducatives dans la plupart des pays, pour deux types de raisons.

La première raison est de nature structurelle et renvoie au mode d'administration, dans la plupart des cas publics, des systèmes éducatifs. Sans discuter ici du bien-fondé de ce choix, force est de reconnaître qu'il prive les gestionnaires d'informations comparables à celles qu'obtient le producteur privé grâce au fonctionnement courant du marché. Soumis à la concurrence, ce dernier, sous peine de disparaître, est fortement incité à s'orienter vers la production de biens ou services susceptibles de répondre à une réelle demande et à sélectionner les modes de production les plus efficaces. À l'inverse, dans un système public tel que l'école, rien n'empêche des dispositifs peu efficaces voire nuisibles pour les élèves de perdurer. Cela ne veut pas dire que l'instauration d'un "marché" en éducation, par l'instauration du choix de l'école notamment, débouche sur plus d'efficacité, dès lors que les familles ne se règlent pas forcément sur ce critère pour choisir (Meuret & al. 2001).

La seconde raison tient au processus de décision lui-même qui souvent accorde une très large place aux opinions. Ce poids des opinions, sans doute plus important dans le domaine de l'éducation que dans celui de la politique énergétique, par exemple, s'explique par la familiarité que confère à tout un chacun la fréquentation de l'école et aussi par le fait que nombre de décideurs du système sont également d'anciens enseignants qui en ont une connaissance spontanée. En outre, les choix pédagogiques ne peuvent guère se fonder sur des théories bien stabilisées des apprentissages : la psychologie cognitive ou les didactiques sont loin de proposer des modèles validés, qui seraient valables en toutes circonstances. Enfin, l'enseignement prend place dans des "unités de production" décentralisées, dotées de fait d'une grande autonomie, et on sait combien les "effets maître" vont s'avérer importants (par exemple, Mingat 1991). La "technologie de production" reste donc à la fois opaque et artisanale.

Cette situation ne serait pas en elle-même problématique si ces pratiques s'accompagnaient systématiquement d'une évaluation des résultats des actions engagées, suivie de procédures de réajustement. En fait, la logique de l'expertise a priori (doublée ou non de la foi pédagogique dans le caractère intrinsèquement positif de telle ou telle action) n'intègre pas ce type d'évaluation. Le résultat des actions étant supposé "connu" ou ne pouvant qu'aller dans le "bon sens", on s'en tient à un contrôle administratif et réglementaire portant sur les moyens mobilisés. La mesure systématique des conséquences des actions mises en œuvre constitue alors un changement radical d'optique, et peut constituer une réponse, certes imparfaite, aux difficultés posées par l'absence d'un mécanisme de justification des choix de politiques éducatives. C'est en outre particulièrement nécessaire dès lors qu'on convient qu'il n'est plus possible d'édicter "d'en haut" les bonnes formes pédagogiques et que c'est au contraire aux établissements à imaginer et à mettre en place les modalités

de fonctionnement les mieux adaptées à leur public. La question de la justification des choix diversifiés qui vont être faits devient particulièrement cruciale, l'évaluation instaurant de fait de nouvelles modalités de régulation du système.

Démarches et écueils méthodologiques de l'évaluation par les résultats

Une première difficulté lorsqu'on cherche à évaluer le fonctionnement des systèmes éducatifs est le caractère relativement vague et général des objectifs qui leur sont assignés. Il est clair que la mesure du "résultat" des activités éducatives suppose une opérationnalisation et une quantification, à l'évidence simplificatrices, mais on peut arguer qu'elles offrent de très nombreux avantages, au plan instrumental d'abord, mais également au plan conceptuel. La seconde difficulté, d'ordre technique, renvoie à la complexité des processus d'apprentissage qui mettent en cause simultanément de nombreux éléments et revêtent une dimension temporelle.

Une évaluation objective des politiques éducatives impose tout d'abord de traduire les grandes orientations en objectifs concrets dont le degré de réalisation confirme ou infirme le sens et l'amplitude des effets qui étaient attendus par la collectivité et définissent donc de fait "l'intérêt général". Pour évaluer le degré d'atteinte de ces finalités éducatives larges, qui concernent en général l'insertion économique et sociale des citoyens, l'accès aux savoirs nés du progrès des sciences et des techniques, la réduction des inégalités sociales, il faut nécessairement en donner une définition opérationnelle en se référant à des indicateurs mesurables. Le recours à des objectifs quantifiés est volontiers considéré comme exagérément réducteur. Pourtant, il est rare qu'on ne puisse concevoir, par rapport aux visées générales des systèmes de formation, des objectifs susceptibles d'être traduits en indicateurs acceptables de "résultat". Qui contesterait, même si ceci n'en traduit pas toute la complexité, qu'on puisse effectivement classer les options de politiques éducatives en fonction de leur impact sur les acquisitions et/ou les comportements des élèves, mais aussi en fonction de leurs conséquences en termes d'insertion économique des diplômés? Certes, on peut toujours contester le caractère partiel de la mesure et se demander si les compétences délivrées par le système scolaire sont bien adaptées aux exigences de l'emploi. De même, de nombreuses dimensions des activités éducatives restent difficilement quantifiables. Mais pour autant, cette référence à des "produits" clairement identifiables est une véritable nécessité dans un débat où dominent le plus souvent des arguments rhétoriques.

En outre, la quantification des missions des systèmes de formation est en elle-même heuristique, en ce qu'elle renouvelle la problématique traditionnelle de l'allocation des moyens. Cette dernière se réfère en général à un mode de fonctionnement idéal de l'école. Il y aurait ainsi une taille de classe "idéale", tout comme il existerait un "bon" niveau de recrutement et un "bon" mode de formation des enseignants. La juxtaposition de ces éléments serait alors garante de la qualité et de l'efficacité du système. Cette façon d'appréhender les problèmes ne tient compte ni des conditions qui prévalent actuellement dans le système, ni du caractère substitutif des différents éléments; elle omet enfin toute référence aux contraintes budgétaires.

Sur tous ces aspects, c'est l'adoption d'une perspective économique (rappelons que l'économie est une science des choix), plus que la quantification des objectifs et des produits, qui s'avère heuristique. Pour traiter de l'influence de chacun des facteurs d'organisation sur la qualité de l'école, elle substitue le "raisonnement à la marge", cher aux économistes, à la logique binaire (bon/mauvais) qui prévaut traditionnellement dans les choix éducatifs. Plutôt que de dire, par exemple, qu'il importe que le maître soit bien formé ou que les effectifs d'élèves ne soient pas trop élevés, le recours à des objectifs quantifiés (l'importance des apprentissages des élèves par exemple) permet de mesurer précisément ce qu'on gagne et ce qu'on perd à modifier graduellement la formation des enseignants, à partir de la situation actuelle (à la marge, dira-t-on). De même, plutôt que de définir a priori à partir de quel seuil un groupe d'élèves sera considéré comme "pléthorique", on évaluera dans quelle mesure les acquisitions des élèves sont affectées par la variation de la taille de la classe. En d'autres termes, on considère qu'aucun "facteur de production" n'est bon en soi, mais que, vu son amplitude actuelle, il peut être bon ou au contraire inutile voire nuisible d'en augmenter la dotation.

Le recours à des objectifs quantifiés permet par ailleurs d'introduire la dimension économique des choix en mettant en regard dépenses et résultats. Car à moins de poser qu'il n'existe aucune contrainte budgétaire, on admettra que choisir c'est renoncer, et qu'il est alors important de comparer des solutions concurrentes. Si on choisit de réduire la taille de classe, par exemple, il faut mettre en regard ses conséquences pédagogiques éventuelles et le supplément de dépense par élève que cela entraîne, et voir comment évolue le rapport coût-efficacité associé à chaque unité du facteur considéré (ici, chaque fois qu'on abaisse la taille des classes d'une unité). On estimera ainsi jusqu'à quelle taille il apparaît coût-efficace de descendre, et également, ce qui est tout aussi important, s'il n'aurait pas été plus coût-efficace de retenir d'autres solutions. Répliquée pour chacun des composantes de l'organisation scolaire, cette opération qui consiste à associer effet sur la qualité et conséquences financières, introduit de fait un véritable "taux de change"

entre des activités autrement incommensurables. L'allocation des moyens n'est plus pensée facteur par facteur, mais dans un cadre d'ensemble où des substitutions sont possibles en fonction des rapports coût-efficacité des différents facteurs considérés.

On peut ainsi évaluer l'opportunité des décisions de politiques éducatives en considérant non seulement ce qu'on gagne à prendre une décision quelconque, mais ce qu'on pourrait gagner en investissant un montant comparable dans une activité concurrente. Ces simulations sont d'autant plus instructives que dans ce domaine les possibilités d'expérimentation sont des plus limitées. On ouvre par là un espace de négociation, structuré par les notions d'arbitrage et de priorité, notions qui n'ont rien de spécifiquement économique, puisqu'il s'agit d'affecter des ressources, qui ne sont pas extensibles, ici plutôt que là, pour atteindre le mieux possible les objectifs qu'on s'est fixés.

Il reste que ces arbitrages reposent sur une évaluation des différents facteurs d'organisation scolaire en référence aux résultats produits, et une mesure du coût de chacun des facteurs. Si ce dernier point pose peu de problèmes, l'analyse de l'impact des dotations de facteurs est par contre beaucoup plus délicate. Elle suppose la collecte et le traitement de données nombreuses, concernant à la fois le contexte scolaire et les élèves eux-mêmes. Et surtout, il n'est pas immédiat d'isoler dans les productions, les comportements, les parcours scolaires des élèves, ce qui tient à leurs caractéristiques personnelles, familiales et sociales, et ce qui tient spécifiquement aux modes d'organisation des classes et des établissements eux-mêmes. De plus, la mesure de l'impact des facteurs d'organisation impose de raisonner sur des élèves strictement comparables. Cette comparabilité ne pouvant être assurée par un dispositif expérimental, elle est obtenue à travers un dispositif empirique dans lequel on s'efforce de contrôler tous les éléments qui, à côté des facteurs d'organisation scolaire, affectent les acquisitions des élèves. Cela suppose d'une part de définir a priori les éléments qui affectent les apprentissages et de collecter les informations qui s'y rapportent, d'autre part de recourir à des méthodes statistiques qui permettent de séparer l'influence des différents éléments considérés.

Ainsi, pour analyser le processus d'apprentissage en classe de CM2, la dimension temporelle du processus d'apprentissage impose, si on entend évaluer l'impact des modes d'organisation scolaire en CM2, de contrôler le niveau des élèves à l'entrée dans cette classe, et de mesurer l'efficacité de ces modes d'organisation en termes de progression au cours de l'année, et non pas en termes de niveau final atteint. Cependant, cette progression n'est pas uniquement imputable à l'organisation scolaire au cours de l'année étudiée. Si la prise en compte du niveau des élèves à l'entrée en CM2 permet de tenir compte de l'influence qu'ont pu avoir les caractéristiques personnelles et

familiales, avant l'entrée dans la classe étudiée, ces dernières, du fait de leur permanence, marqueront également la progression des élèves en cours de CM2. La dimension temporelle peut être évidemment étendue au-delà, pour prendre en compte d'éventuels effets différés des pratiques pédagogiques.

L'évaluation des facteurs d'organisation scolaire consiste donc à mesurer leur contribution spécifique à l'explication du niveau final des élèves, en contrôlant à la fois leur niveau initial et l'ensemble des variables personnelles et familiales susceptibles d'affecter les performances scolaires en cours d'année. L'analyse se base donc le plus souvent sur la progression des élèves: le niveau initial des élèves est introduit dans le modèle explicatif du niveau final comme une caractéristique individuelle particulière qui permet d'examiner l'influence des autres caractéristiques à "niveau initial identique".

Au plan empirique, ces analyses sont relativement longues du fait de la période incompressible consacrée aux apprentissages; elles sont également lourdes et délicates à mettre en œuvre du fait de la masse de données à collecter. Au plan technique, l'influence spécifique des différents modes d'organisation scolaire est évaluée grâce à des modélisations multivariées, ce qui permet empiriquement de définir a posteriori une comparabilité des individus étudiés relativement proche de celle qui résulterait d'un plan expérimental si celui-ci était applicable dans le domaine de l'éducation. C'est ce qu'évoque la formule "toutes choses égales par ailleurs" qu'on doit interpréter évidemment en référence aux variables de contrôle introduites dans les modèles. La modélisation permet, à un premier stade, d'évaluer l'influence respective globale des caractéristiques des élèves et du contexte scolaire sur la variabilité des résultats individuels (de même que la capacité globale du modèle à rendre compte de la variance du phénomène étudié). À un second stade, elle évalue l'impact spécifique de chacun des facteurs relevant de ces deux groupes: pour chacune des variables, un coefficient estime son effet sur le niveau final des élèves, toutes les autres situations considérées dans le modèle étant à un niveau comparable. On dispose en outre aujourd'hui de modèles multiniveau qui autorisent une meilleure séparation des niveaux individuels et contextuels (Bressoux et al. 1997).

Au travers de cette démarche, des connaissances sur le fonctionnement du système peuvent s'accumuler. Certes avec prudence, puisqu'on sait bien que le propre des sciences humaines est de produire des résultats contextualisés, et la recherche d'une interprétation, pour ne pas s'en tenir à un coefficient "boîte noire", n'a pas un caractère accessoire.

Quelle pertinence des évaluations pour la régulation du système ?

S'il est intellectuellement aisé de convaincre que l'évaluation des systèmes éducatifs a un rôle important à jouer dans la définition des politiques éducatives, il reste que pour des raisons techniques et sociales ce rôle est encore aujourd'hui relativement modeste. Un premier obstacle à la prise en considération des résultats d'évaluations de ce type est le degré important et croissant de technicité qui les caractérise. Cette technicité risque de détourner des évaluations ceux-là mêmes qu'elles concernent. Par ailleurs, cette technicité ne reçoit pas (en France du moins) le support d'un milieu scientifique assez large pour garantir aux non-initiés le respect des conditions scientifiques de production des connaissances, qui rendraient les résultats acceptables à leurs yeux.

Au-delà de cette difficulté technique, l'évaluation des politiques éducatives par les résultats fait l'objet de réticences parfois très vives, notamment dans les pays francophones. Certaines mettent en exergue l'obsession "industrielle" que révéleraient ces mesures du fonctionnement de l'école uniquement basées sur les apprentissages scolaires des élèves; d'autres encore craignent que ceci n'exprime un "déli paradoxal des finalités (éthiques, valeurs et visées) considérées comme trop vagues" (Demailly 2000b). Une chose est sûre, la limitation fréquente des "résultats" de l'école aux apprentissages des élèves est plus une commodité qu'une nécessité. Techniquement, il est possible de prendre en compte de nombreuses dimensions des apprentissages, des attitudes et des comportements, le véritable frein étant davantage à rechercher du côté de la lourdeur et du coût déjà importants des analyses. Cela dit, il semble y avoir une certaine substituabilité entre les diverses mesures des acquis des élèves, qui tendent à converger rapidement (cf., au baccalauréat, Jarousse et al. 1998). De manière générale, toutes les transformations de l'élève, entre un état initial et un état final sont a priori susceptibles d'être évaluées, il n'y a que les transformations ineffables qui ne le sont pas...

D'autres critiques, tout aussi vives, se centrent sur les évaluations des politiques éducatives et en questionnent le sens, eu égard à ce que la sociologie des organisations ou la sociologie de l'éducation révèlent du fonctionnement des systèmes éducatifs. On souligne ainsi, pour la critiquer (notamment Perrenoud, 1996) que la démarche d'évaluation par les résultats fait comme si tous les enseignants partageaient des finalités identiques, celles mises en avant par les textes. Il est sans doute vrai que face à des textes, des directives, des programmes souvent ambigus, les enseignants s'adaptent, voire pratiquent une résistance passive, quand ils n'en partagent pas la philosophie. Il serait alors simpliste de prétendre évaluer les politiques affichées dans les textes dans des contextes concrets où les objectifs initiaux n'ont plus forcément cours, tant ils ont été réinterprétés, détournés... ou ignorés.

Il est vrai aussi que certaines réformes ont des objectifs souvent implicites, perçus parfois comme peu avouables, des objectifs de remobilisation ou encore de paix syndicale par exemple, plus importants que ceux toujours mis en avant d'amélioration des résultats des élèves (on pense à l'élévation du niveau de recrutement et de rémunération des instituteurs par exemple). D'autres réformes ne recouvrent pas forcément un consensus général (l'objectif de démocratisation du système par exemple), au-delà du "politiquement correct" de rigueur. Mais cela n'empêche en rien d'examiner si les dites politiques ont été accompagnées d'effets visibles sur les résultats des élèves, car la paix syndicale ou l'amélioration de la situation des enseignants constituent a priori des objectifs intermédiaires d'autant plus respectables qu'ils auraient des retombées sur l'objectif final du système éducatif, à savoir les transformations des élèves.

Ce que ces débats mettent utilement en exergue, c'est qu'on n'évalue jamais que les réalisations effectives d'une politique, telle que les acteurs ont bien voulu la mettre en œuvre sur le terrain. On n'évalue donc pas des principes, mais tel ou tel degré de leur traduction concrète. Il faut aussi souligner qu'on n'évalue jamais que comparativement —telle politique semble meilleure que telle situation de référence où elle n'est pas ou moins mise en œuvre—, par rapport à la gamme des situations existantes. Un dispositif peut apparaître non seulement "bon" mais meilleur que toute autre alternative, sans que cela interdise d'imaginer d'autres solutions, qui seraient encore meilleures. On n'évalue que ce qui existe, et cela ne doit pas brider l'imagination pédagogique.

Toujours est-il que la transparence est exigeante, et on comprendrait que les enseignants, comme dans tout milieu de travail, aient intérêt à un certain flou autour des objectifs qui ont réellement cours et des résultats censés être atteints. Si ce flou les protège dans des contextes parfois très difficiles, faut-il le tolérer sous prétexte de réalisme sociologique (comme semble le suggérer Perrenoud, 1996), ou le considérer comme une forme d'hypocrisie préjudiciable aux élèves (Dubet et Duru-Bellat, 2000)? Si on se rallie à l'idée d'objectifs explicites, assortis de bénéfices sociaux, on ne peut admettre que dans une institution financée par les fonds publics, où œuvrent des fonctionnaires, des dérives importantes se manifestent dans certains contextes par rapport aux objectifs fixés à tous. Il est certain qu'au plan politique aussi, le flou a quelque avantage en termes de paix civile : on fait semblant d'avoir un système uniforme. C'est donc le coût social de l'opacité et le coût social d'une plus grande transparence qu'il faut mettre en balance, au-delà de la "compréhension sociologique" qu'on peut en avoir.

L'évaluation introduit un regard externe de la collectivité qui finance les actions rappelant que le système n'appartient pas aux seuls enseignants. On comprend alors que ces derniers aient tendance à vivre ces évaluations comme un contrôle, voire une remise en cause, de leur activité et de leur

autonomie. En attestent des témoignages comme celui de Meirieu (2000), qui affirme que “l’obligation de résultats n’a pas de sens en éducation”. En arrière-plan, il y a la crainte que l’évaluation n’affecte les rapports de force au sein du système éducatif. En France, il est clair que le contrôle de conformité pratiqué actuellement à doses homéopathiques par les inspecteurs est une forme bien plus “douce” d’évaluation que ce qui serait une évaluation externe et publique des résultats des élèves, impulsé par un “État évaluateur” ou, mieux encore, une agence autonome spécifiquement chargée de l’évaluation des politiques éducatives.

Enfin, les résistances multiformes à l’évaluation tiennent aussi, sinon plus, aux difficultés de la diffusion d’un modèle “technocratique” du changement dans une institution comme l’école. Dans nombre de domaines, la connaissance procède de l’accumulation de résultats au sein d’un milieu scientifique, où le choix des méthodes et les résultats eux-mêmes font l’objet d’une confrontation qui en garantit la fiabilité. En matière d’évaluation des systèmes éducatifs, en France, un tel milieu fait encore largement défaut (de même que le volume des connaissances engrangées reste très limité). Si ce milieu existe à l’échelon international, il contribue davantage à la confrontation des méthodes qu’à l’accumulation de résultats, compte tenu notamment de la spécificité des systèmes de formation. Il n’en inspire pas moins, notamment au travers des grandes organisations internationales, les pistes à suivre par des politiques éducatives qui seraient “evidence informed” ou “knowledge based” (pour une discussion de cette tendance au Royaume-Uni : Atkinson 2000).

Mais ce modèle du changement technique où la science définit la voie à suivre, à laquelle les “exécutants” n’auront qu’à se rallier, peine à s’appliquer à l’éducation. Certes le Centre peut s’appuyer sur les recherches pour décider d’éradiquer telle ou telle pratique dont on aurait démontré les effets nocifs (le redoublement ou les classes de niveau par exemple). Mais pour le reste... On connaît la faible portée des directives insufflées d’en haut, vue l’autonomie importante dont jouissent les enseignants dans leur classe (notamment Suchaut, 1996). On sait aussi que l’efficacité d’une pratique pédagogique est souvent relative à un contexte (efficace avec tel type de public, ou à tel niveau scolaire seulement), ce qui rend les évaluations inévitablement indexées aux conditions dans lesquelles elles ont été réalisées. L’idéal serait de se poser d’emblée —dès la conception des dispositifs— et systématiquement la question de la manière dont les évaluations peuvent/pourront s’intégrer aux pratiques courantes des acteurs. C’est d’ailleurs ce qu’on observe à l’échelle européenne, avec la montée de dispositifs d’évaluation de la “qualité” de la vie scolaire et des résultats de l’enseignement par les établissements eux-mêmes. Ces évaluations utilisables par les acteurs sont en elles-mêmes susceptibles de participer d’une “recomposition des professionnalités” enseignantes (Demailly 2000b).

Reste la question sans doute la plus délicate de l'utilisation de ces évaluations. En Belgique par exemple (Dupriez & Maroy 2000), une évaluation externe des compétences des élèves est proposée à partir de 1997 aux établissements, mais à titre indicatif seulement (il s'agit de rendre disponible un étalon de ce qui peut être exigé d'un élève à tel stade de sa scolarité). Mais si l'État entend bien faire prendre conscience de fortes disparités entre les élèves, il ne s'engage pas pour autant dans la gestion de mécanismes de contrôle allant contre le mode de régulation dominant (faisant une grande place aux "réseaux" d'enseignement). La situation n'est pas très différente en France, où les audits d'établissement, pratiqués sous une forme expérimentale dans l'Académie de Lille (Demailly et al. 1998) se sont souvent contentés de lever des lièvres, sans parvenir à exercer un impact régulateur, ce qui était pourtant leur visée. À l'inverse, dans les pays anglo-saxons, les évaluations de l'efficacité des écoles sont parfois utilisées de manière très offensive dans le cadre de véritables campagnes de dénonciation qui condamnent les écoles plus qu'elles ne les aident (Thrupp 1998). De même, certaines évaluations peuvent être "sur-utilisées", quand par exemple des acteurs tels que les collectivités locales s'en emparent pour justifier leur politique (Dutercq 2000).

Ces écarts entre l'évaluation et ses usages sociaux ne sont pas étonnants dans une perspective sociologique dénonçant le postulat (simpliste) sous-tendant les pratiques d'évaluation, comme quoi "l'action organisée peut être pensée comme lien simple et instrumental entre des objectifs clairs et des résultats mesurables, ce qui est une vision très administrative de la réalité sociale" (Demailly 2000a). Mais faut-il attendre que les sociologues aient mis en pleine lumière la complexité de l'action sociale et toutes les dérives possibles de l'utilisation des évaluations, l'écart entre leur fonction manifeste et leur fonction latente, pour engager ce type d'analyse ? Car des décisions sont prises tous les jours, dans les systèmes éducatifs, dont il n'est pas exclu qu'elles aient quelque impact sur les élèves. Toujours est-il que le chercheur pratiquant des évaluations, sauf à accepter que ses travaux se cantonnent (ce qui n'est pas rien) dans le domaine de la production de connaissances, doit, ce qui n'est pas toujours le cas, s'interroger sur leur efficacité en termes d'action, jamais acquise d'avance quelle que soit la qualité scientifique du travail.

Conclusion

Quel peut être l'avenir de l'évaluation des systèmes éducatifs en tant que champ de recherche (et de pratique) particulier ? Au-delà des résistances et des obstacles à l'évaluation du système éducatif, certaines circonstances poussent au contraire à son développement.

— La première est sans aucun doute la généralisation progressive de l'évaluation des politiques publiques qu'alimente la prise de conscience d'une nécessaire maîtrise des dépenses de l'État. Les dépenses d'éducation qui mobilisent une part importante des ressources publiques échapperont de moins en moins à une demande générale de justification concernant l'opportunité des politiques suivies tant au plan financier qu'à celui de la qualité du service. Cette demande de justification est aujourd'hui largement présente dans une opinion de plus en plus consciente de l'hétérogénéité du système éducatif sur le territoire, et prête à sanctionner les inégalités de qualités de l'offre éducative qu'elle perçoit par le contournement de la sectorisation.

— La seconde concerne la décentralisation effective d'une part importante des activités de gestion et de financement de l'éducation. Ceci devrait entraîner une double demande de justification des politiques mises en œuvre, des collectivités territoriales vers l'État d'un côté, de l'État vers les collectivités territoriales de l'autre. La prise en charge d'une part croissante des dépenses éducatives, leur plus forte sensibilité aux mouvements de l'opinion conduiront naturellement les collectivités territoriales à discuter les politiques nationales tout en ayant aussi à justifier les politiques locales auprès de leurs électeurs. De façon réciproque, le mouvement de décentralisation devrait conduire l'État à s'intéresser aux conséquences de l'implication de plus en plus large des collectivités territoriales dans des activités ayant une incidence sur le fonctionnement du système éducatif. Les responsables nationaux du système éducatif se doivent de capitaliser les innovations qui sont nées de la gestion décentralisée de ces activités et de contrôler l'homogénéité des services éducatifs dont ils demeurent les garants.

Plus largement, l'évaluation devrait constituer l'élément pivot de la gestion et de la régulation d'un système éducatif national relativement hybride. Il y a en effet une contradiction de plus en plus manifeste entre les missions du service public qui supposent une offre d'éducation homogène (ou compensatrice des inégalités) et le fonctionnement effectivement très décentralisé de ce système. Outre la décentralisation administrative et financière des activités d'enseignement qui touche à présent la dimension pédagogique, l'acte d'enseignement lui-même reste fondamentalement décentralisé comme en atteste l'importance des "effets-maître" et des effets établissements sur les apprentissages. Face à cette situation, on peut renoncer à l'évaluation pour maintenir la fiction d'un système éducatif homogène, en acceptant les effets pervers d'une telle attitude, ou bien en faire un outil d'observation privilégié permettant une gestion du système public d'éducation conforme à ses orientations. Mais il faut alors une volonté politique claire puisque c'est bien la régulation du système qui est en jeu.

Bibliographie

- ATKINSON E. 2000 "In Defence of Ideas, or Why 'What Works' is not enough", *British Journal of Sociology of Education*, 21, n3, p317-330
- BRESSOUX P. 1994 "Les recherches sur les effets-école et les effets-maître", *Revue Française de Pédagogie*, n108, p91-137
- BRESSOUX P., COUSTÈRE P., LEROY-AUDOUIN C. 1997 "Les modèles multiniveaux dans l'analyse écologique: le cas de la recherche en éducation", *Revue française de Sociologie*, vol. 38 n1, p67-96
- DEMAILLY L. 2000a Ce que peut apporter une approche sociologique des dispositifs d'évaluation, in Solaux G. (dir) 2000 *L'évaluation des politiques d'éducation*, Dijon, CNDP, p67-78
- DEMAILLY L. (dir.) 2000b *Évaluer les politiques éducatives*, Bruxelles, De Boeck
- DEMAILLY L., DEUBEL Ph., GADREY N., VERDIÈRE J. 1998 *Évaluer les établissements scolaires. Enjeux, expériences, débats*, Paris, L'Harmattan
- DUBET F., DURU-BELLAT M., 2000, *L'hypocrisie scolaire*, Paris, Seuil
- DUPRIEZ V., MAROY C. 2000 Les dispositifs d'évaluation comme construction sociale, in Solaux G. (dir) 2000 *L'évaluation des politiques d'éducation*, Dijon, CNDP, p95-108
- DURU-BELLAT M., MINGAT A. 1993 *Pour une approche analytique du fonctionnement du système éducatif*, Paris, PUF
- DUTERCQ Y. 2000 Évaluation des politiques éducatives et usages stratégiques de l'innovation, in Demailly (dir) 2000, *Évaluer les politiques éducatives*, Bruxelles, De Boeck, p47-64
- JAROUSSE J.-P. 1999 Évaluer les systèmes éducatifs: de quoi parle-t-on?, in Paul J.-J. (dir) *Administrer, gérer, évaluer les systèmes éducatifs*, Paris, ESF, p159-183
- JAROUSSE J.-P., MINGAT A. 1992 "La formation du capital humain: gestion par le marché ou gestion par l'État?", *Revue Économique*, n4, p739-753
- JAROUSSE J.-P., MINGAT A., OGET D. 1998 "Réflexions pour un changement de l'organisation des épreuves du baccalauréat", *Savoir*, vol. 10, n2-3-4, p81-105
- MEIRIEU P. 2000 *L'école et les parents*, Paris, Plon
- MEURET D., BROCCOLICHI S., DURU-BELLAT M. 2001 "Autonomie et choix des établissements scolaires: finalités, modalités, effets", *Cahier de l'IREDU* n62
- MINGAT A. 1991 "Expliquer la variété des acquisitions au Cours Préparatoire: les rôles de l'enfant, la famille et l'école", *Revue Française de Pédagogie*, n95, p47-63
- PERRENOUD P. 1996 "Évaluer les réformes scolaires, est-ce bien raisonnable?", *Mesure et Évaluation en Éducation*, vol.19, n2, p53-97
- SOLAUX G. (dir.) 2000 *L'évaluation des politiques d'éducation*, Dijon, CNDP
- SUCHAUT B. 1996 "La gestion du temps à l'école maternelle et primaire", *L'Année de la Recherche en Sciences de l'Éducation*, PUF, p123-153
- THRUPP M. 1998 "Exploring the Politics of Blame: school inspection and its contestation in New Zealand and England", *Comparative Education*, vol.34, n2, p195-210