

Le niveau d'huile, le moteur et la voiture : les enjeux d'une évaluation de la qualité de l'enseignement par les indicateurs

NORBERTO BOTTANI

ex-directeur du Service de la Recherche en Éducation (SRED) du Canton de Genève
ex-administrateur principal à l'OCDE et responsable du programme INES
sur les indicateurs internationaux de l'enseignement
norberto.bottani@oxydiane.net

Y a-t-il un enjeu à l'évaluation de la qualité de l'enseignement par les indicateurs ? La question peut étonner car elle lie trois éléments dont la portée et la définition ne sont pas claires : l'évaluation de l'enseignement, la qualité de l'enseignement et les indicateurs de l'enseignement. Ces derniers sont un outil d'analyse du fonctionnement d'un système ; l'évaluation, grosso modo, est une procédure d'examen rationnel des résultats qui aboutit à un diagnostic en vue d'une décision ; la qualité, un état de l'enseignement (produit et/ou démarche) par rapport à un seuil d'appréciation préétabli, conventionnel ou arbitraire. Ces trois éléments ne sont ni complémentaires ni intégrables l'un à l'autre. Leurs caractéristiques, leurs configurations ont varié dans le temps et leur usage univoque ne va pas de soi.

Il convient de spécifier la nature du lien entre indicateurs, évaluation et qualité de l'enseignement avant de déterminer à quelles conditions les indicateurs peuvent fonctionner comme instruments d'évaluation de la qualité de l'enseignement. Si l'évaluation est l'examen rationnel et public des résultats de l'enseignement impliquant le recours à des référentiels exogènes, reconnaissables par les groupes sociaux concernés (compte tenu de la place de l'enseignement dans les sociétés démocratiques), on peut supposer qu'il doit se fonder sur des arguments probants, validés ou invalidés selon des procédures scientifiques. Dès lors, les indicateurs peuvent faire partie de la panoplie d'outils utilisés pour recueillir des preuves documentées sur l'état de l'enseignement et être un élément de l'évaluation. Cette question ne peut cependant être abordée selon un préjugé doctrinaire attribuant aux indicateurs un rôle déterminé dans un type d'évaluation, celle des systèmes d'enseignement. Ce thème mérite un traitement scientifique, attentif

aux dimensions théorique et méthodique : d'abord, identifier d'éventuelles corrélations entre des éléments aussi divers que la qualité de l'enseignement, les indicateurs et les programmes d'évaluation ; ensuite, vérifier si ces corrélations sont causales ou non ; enfin, construire une théorie du fonctionnement de la connexion causale et la mettre à l'épreuve en situation.

Si ma démarche suit ce programme –repérer des corrélations entre ces trois paramètres ; établir la nature causale ou non de ces corrélations ; aboutir à une théorie des indicateurs de l'enseignement au service de l'évaluation–, je me bornerai ici à poser quelques jalons pour un travail à venir. Mon hypothèse est qu'il n'y a pas de corrélation causale entre les indicateurs de l'enseignement et l'évaluation d'un système d'enseignement et que le dispositif d'élaboration des indicateurs n'est pas en soi constitutif d'une théorie de l'évaluation des systèmes d'enseignement, mais je ne peux pas ici l'étayer par des preuves, ce travail présupposant, en effet, une longue analyse et les informations nécessaires à cette démonstration étant encore lacunaires.

La donne du problème

Le macrosystème d'enseignement

Pour tenter de clarifier la question, il est indispensable de délimiter le champ d'application : enseignement en classe ? D'une discipline ? Développement d'une compétence ? Enseignement dans un établissement ? Un département ? Une région ? Un pays ? L'origine de cette réflexion –le colloque international sur l'évaluation des politiques d'éducation et de formation, organisé par l'Association internationale de sociologues de langue française et l'unité mixte de recherche Éducation & Politiques de l'Institut National de Recherche Pédagogique, <http://ep.inrp.fr>, Lyon septembre 2005– amène à retenir les évaluations des systèmes d'enseignement, car ceux-ci sont l'objet principal des politiques d'éducation. Celles-ci n'ont pas comme unique objet les systèmes d'enseignement au sens strict. Elles s'occupent d'une multitude de thèmes qui en sont des composantes – le processus, les modalités, les échelles de prise de décisions sur l'éducation, le fonctionnement du parlement en matière d'enseignement dans un système démocratique, la structure de l'administration scolaire, sa composition, le rôle et la qualité des syndicats d'enseignants, mais aussi la santé des enfants, le degré d'instruction des parents, la politique du logement et celle des équipements culturels, le libre choix des écoles, etc. L'inclusion de ces dimensions dans les politiques et dans les évaluations des systèmes d'enseignement varie selon les finalités d'un système, les buts de l'école ou les objectifs qu'une communauté reconnaît à l'éducation. Il faut donc s'entendre sur une définition du concept de "système

d'enseignement" si l'idée d'une analogie entre évaluation des politiques de l'éducation et évaluation des systèmes d'enseignement est retenue. Les systèmes se différencient par la taille, l'échelle, la configuration, la distribution des centres décisionnels, les compétences attribuées à chaque échelon et les modalités des décisions, les ressources et les objectifs. Une classe, un cours d'histoire et de géographie sont déjà des microsystèmes d'enseignement. Nous savons qu'on peut concevoir, mettre au point, développer et construire aussi bien des indicateurs d'établissement (l'établissement étant un microsystème) que des indicateurs de macrosystème, comme par exemple l'ensemble d'indicateurs rassemblés dans *L'État de l'École en France* (ministère de la Jeunesse, de l'Éducation nationale et de la Recherche). L'évaluation des politiques d'éducation traitée ici ne concerne que des macrosystèmes d'enseignement et les ensembles d'indicateurs évoqués que ceux conçus pour ces systèmes, par exemple celui de la France ou d'un Land allemand comme la Bavière ou d'un canton helvétique comme Lucerne.

L'évaluation avec ou sans indicateurs

Historiquement, l'évaluation des systèmes d'enseignement a anticipé la mise sur pied d'ensembles d'indicateurs. Par exemple, l'évaluation des enseignants, des écoles ou d'un département par les inspecteurs est une opération qui pendant des décennies s'est effectuée sans indicateurs et qui probablement, là où elle se fait encore, continue à s'en passer. Les inspecteurs expriment une appréciation (donc ils évaluent) en se servant de critères ou de paramètres propres à leur corps, qui découlent d'une représentation convenue de la bonne école. Ils savent, croient savoir ou prétendent savoir, ce qu'est une bonne école. Leur appréciation se fonde sur une idée préexistante de la qualité d'une école : ils en ont une intime connaissance et possèdent ou croient posséder une connaissance holistique de la qualité d'un microsystème d'enseignement, comme une classe ou une école, mais aussi, parfois, d'un macrosystème, tel un service d'inspection générale ou nationale. Bien entendu, rien n'empêche que les inspecteurs se servent d'indicateurs pour formuler une appréciation ou une évaluation, mais une démarche de ce genre implique un changement de cap radical dans l'exercice de leur fonction.

Un autre type d'évaluation qui en général se passe d'indicateurs est l'autoévaluation, bien que certaines procédures la combinent avec des indicateurs ou aboutissent à la production d'indicateurs (McBeath 1999, 2000). Pendant longtemps, des évaluations de l'enseignement, dans l'enseignement et pour l'enseignement n'ont pas recouru aux indicateurs. Cela signifie qu'ils ne sont pas indispensables pour mener des évaluations, que celles des politiques de l'éducation ayant le système d'enseignement comme objet peuvent se faire sans indicateurs. Pendant longtemps, on n'a pas eu l'idée d'évaluer les systèmes d'enseignement ni éprouvé le

besoin de développer des indicateurs de ces systèmes, pour les piloter, les suivre ou les évaluer. L'élaboration et la production d'indicateurs des systèmes d'enseignement ne sont qu'une opération relativement récente qui interfère dans des formes multiples avec les procédures d'évaluation. Ce n'est en effet qu'à partir des années 1960 (Bottani 2005) qu'on a commencé à s'occuper d'indicateurs des systèmes d'enseignement.

Le niveau d'huile, le moteur, la voiture

Il n'y a pas a priori de rapport direct entre indicateurs et évaluation, même si souvent un amalgame est fait entre les deux. Pourquoi cette confusion qui réduit l'évaluation des politiques d'éducation à la production d'indicateurs des systèmes d'enseignement ou attribue aux indicateurs une fonction d'évaluation qu'en principe ils n'ont pas ? La réponse exige de clarifier le concept d'indicateur. À cet effet, je voudrais me servir de la métaphore du tableau de bord et d'une anecdote.

Il y a peu, j'ai acheté une voiture, qui après quelques milliers de kilomètres (elle était encore sous garantie) m'a surpris par le clignotement insistant d'un voyant du tableau de bord. Après avoir consulté le manuel, j'ai identifié qu'il s'agissait du voyant du niveau d'huile dans le moteur, et j'ai appliqué la procédure indiquée. Cependant, après une centaine de kilomètres, l'alerte se répète. Je répète les opérations proposées. Le moteur étant neuf, le manque d'huile signalé par le voyant pouvait être considéré, selon le manuel, comme normal. J'accomplis donc une mise à niveau de l'huile pour la deuxième fois et tout semble rentrer dans l'ordre. Malheureusement, le clignotement reprend presque aussitôt. En cas de récurrence, le manuel invite le conducteur à ramener tout de suite la voiture à un garage agréé. Après un contrôle sommaire, le garagiste décide de placer des scellés sur le moteur et des repères étalonnés selon une échelle plus fine que celle de la jauge d'huile pour mieux contrôler sa consommation. Je ne parcours qu'une centaine de kilomètres et le clignotant m'alerte à nouveau. Je repasse au garage où on m'oblige à laisser la voiture : pas question de se remettre en route sans avoir déniché la cause du dérangement signalé par le voyant. L'histoire se conclut par le changement du moteur.

En résumé :

- le voyant du tableau de bord a exercé sa fonction d'indicateur en alertant le conducteur sur une situation anormale ;
- le conducteur a tenu compte du message du voyant et suivi les procédures décrites dans le manuel ;
- le garagiste a cherché la cause du défaut dénoncé par le voyant et interprété le signal ;
- enfin, le constructeur, responsable du système moteur, a pris la décision qui s'imposait et remplacé à ses frais le moteur.

Tout au long du processus, personne n'a exprimé de jugement sur la qualité du moteur ou du véhicule. L'indicateur a attiré l'attention sur un dysfonctionnement, qui s'est révélé réel et grave, mais cette insuffisance n'a pas débouché sur une évaluation négative ni du moteur ni du véhicule ni de la marque. En effet, un moteur, un véhicule ou une marque s'évaluent selon d'autres paramètres que les indicateurs du tableau de bord.

Dans cet épisode, le point central est le niveau d'huile au-dessous duquel le voyant s'allume, c'est-à-dire le niveau d'alerte fixé par le fabricant. L'indicateur ne mesure que la quantité d'huile dans le moteur et ne s'allume que si elle descend au-dessous du seuil de risque de grippage du moteur déterminé par le constructeur. Or c'est lui qui constitue le référentiel de l'indicateur. La responsabilité de la construction des indicateurs ne revient donc qu'aux responsables du système, ici, ceux de la marque et plus particulièrement ceux qui ont conçu, dessiné et testé le moteur. La qualité d'un indicateur du tableau de bord ne dépend que des ingénieurs des équipes de développement, leur compétence technique dans la fabrication des indicateurs du tableau de bord ne peut se confondre avec la responsabilité de la qualité du système-voiture dans son ensemble. La qualité du véhicule ou de la marque n'est pas de leur ressort (Nuttal 1994). Un bon indicateur, c'est-à-dire bien construit, robuste, fiable, clair, concourt probablement à déterminer la qualité d'un système, élément de l'évaluation, il n'est pas et n'effectue pas une évaluation du système.

La construction d'ensembles d'indicateurs des systèmes d'enseignement : une démarche inachevée

A fin de mieux appréhender l'articulation entre indicateurs et évaluation, il est utile de parcourir l'histoire des indicateurs de l'enseignement. Le développement des indicateurs de l'enseignement n'a pas été linéaire : quatre phases se distinguent, avec des acteurs différents dans les milieux de la recherche et ceux de la politique d'éducation et d'enseignement. Cette histoire n'interfère avec celle des évaluations sur une large échelle que dans la phase finale. Pendant des décennies, le discours sur les indicateurs et leur production se sont déroulés indépendamment des évaluations des résultats des acquis des élèves.

Les années 1970 : le rôle des sciences sociales

La première tentative de construction d'un ensemble d'indicateurs de l'enseignement date du début des années 1970 à l'OCDE qui a publié en 1973 un document annonçant dans son titre les intentions de l'initiative : *Un système d'indicateurs de l'enseignement visant à orienter les décisions des pouvoirs publics.*

Pour fournir aux décideurs des informations leur permettant de prendre des décisions appropriées, il fallait produire des indicateurs mesurant l'influence de l'enseignement sur le bien-être social. Quels facteurs retenir ? Les effets internes ou subjectifs, c'est-à-dire les effets sur les individus, ou les effets externes à l'échelle macro, mesurés à l'aune des bénéfices sociaux – croissance économique, intégration des groupes sociaux, partage de valeurs communes, bien-être accru... Le groupe de travail a estimé qu'on ne pouvait pas isoler les effets internes ou externes de l'enseignement, tout en convenant qu'il génère autant d'effets individuels que d'effets dépassant les individus, ce qu'on appelle des externalités. Dans les deux cas, les effets dépendent des variables sociales qui façonnent les systèmes d'enseignement, par exemple les aspirations d'une communauté, les objectifs collectifs et les valeurs éducatives qu'une société tente de promouvoir. De ce fait, la construction d'un système opérant une distinction entre dimensions internes et externes est une absurdité. Pourtant, encore aujourd'hui, on rencontre dans les milieux de l'enseignement une argumentation contestant la pertinence des indicateurs qui s'appuie sur cette distinction, alors que les spécialistes du domaine de l'éducation l'ont écartée depuis longtemps. Tout indicateur se définit par "son appartenance, en tant que paramètre ou en tant que variable, à un modèle de système social" (Land 1970). Si la distinction entre entrées et sorties est problématique et, pire encore, dénuée de signification, selon le rapport, il convient de concevoir un cadre d'organisation des indicateurs sur les effets de l'enseignement qui ne soit pas articulé autour de mesures de l'efficacité interne et externe des systèmes d'enseignement.

Par ailleurs, ce rapport soulève une question problématique dans la construction des indicateurs de l'enseignement, celle des effets indirects sur le fonctionnement des systèmes et le comportement de leurs acteurs à la suite de la production d'un ensemble d'indicateurs : "il ne faut pas oublier que le fait même d'effectuer des mesures introduit un élément d'incertitude : en effet, ceux qui se savent soumis à des évaluations modifient leur comportement mais cet infléchissement est difficile à identifier et a fortiori à mesurer" (3). Ces considérations, formulées dès l'orée de l'histoire des indicateurs, révèlent d'emblée qu'on n'avait pas une conception naïve des enjeux et des problèmes qui se présenteraient au cours de leur production.

Le groupe d'experts, auteur de ce premier rapport de l'OCDE sur les indicateurs de l'enseignement, décida, pour conclure, d'adopter une approche méthodologique inhabituelle pour l'époque, en renonçant à développer un modèle théorique du système d'enseignement, démarche qui n'aurait pas permis de mettre sur pied un ensemble d'indicateurs dans de brefs délais. Les décideurs politiques, qui en général prennent des décisions sans attendre, ont besoin, s'ils adoptent une démarche éclairée, d'outils fournissant des informations à chaud sur l'efficacité des politiques de l'enseignement qu'ils mettent en œuvre ou qu'ils

pilotent. La pression politique sur le groupe d'experts pour qu'il livre un ensemble opérationnel d'indicateurs était donc élevée. Le groupe savait qu'il ne pouvait pas entreprendre un long débat préliminaire sur un modèle universel de système d'enseignement qui risquait de ne pas faire accord. Il opta pour une approche empirique, axée sur l'analyse des objectifs des politiques nationales d'éducation et parvint à composer un ensemble de quarante-six indicateurs. Ces derniers toutefois n'ont jamais été calculés, l'ambition scientifique du projet étant disproportionnée par rapport aux moyens disponibles et aux intérêts des milieux gouvernementaux qui auraient dû financer l'opération. Les spécialistes des sciences sociales n'ont donc pas réussi finalement à accomplir leur programme.

Les années 1980 : pédagogues et pédagogistes en action

La deuxième phase se déroula après la publication aux États-Unis, en 1983, du rapport *A Nation at Risk* sur l'état de l'enseignement primaire et secondaire aux États-Unis (The National Commission on Excellence in Education 1983). Ce rapport a suscité à l'échelon mondial pour les pays ayant les systèmes d'enseignement les plus développés, un large débat sur la qualité de l'enseignement et la façon d'en rendre compte avec des mesures aussi objectives que possible. Dans cette conjoncture il fallait disposer de points de repères sûrs pour confirmer ou invalider la présence d'une crise de la qualité. Le monde de l'enseignement était divisé à ce propos ; il était aussi, dans une certaine mesure, le dos au mur. La production d'indicateurs représenta un point de fixation et de ralliement pour une partie de la communauté scientifique s'occupant d'éducation et soucieuse de déterminer l'état de santé des systèmes d'enseignement. Il convient de souligner que le centre d'intérêt n'était plus le même qu'une dizaine d'années auparavant. À la suite de la crise provoquée par le débat sur la qualité de l'enseignement, il ne s'agissait plus de mesurer l'impact de l'enseignement sur le bien-être social mais d'identifier des renseignements indispensables pour appréhender le fonctionnement, le dysfonctionnement des systèmes d'enseignement et leurs résultats.

Un des rapports phares de cette période, publié aux États-Unis en 1986, fut *Education Indicators. A guide for Policymakers*, rédigé par Oakes (1986) pour le Center For Policy Research in Education –consortium de recherches sur la politique de l'éducation au service de l'amélioration de l'enseignement, financé par le Département américain de l'éducation, regroupant la Rutgers University, la Rand Corporation et l'Université du Wisconsin-Madison. Le but de ce rapport était d'aider les décideurs "à comprendre le rôle légitime (sic) que les indicateurs peuvent jouer dans le 'monitorage' –le suivi– de la condition du système d'enseignement, en appréhendant les changements au cours du temps et en anticipant les changements futurs" (V). Oakes fournit là une définition d'indicateurs de l'enseignement largement reprise ensuite ; elle détaille aussi les applications principales

des indicateurs, décrit leurs limites les plus manifestes et expose succinctement l'état d'avancement de la théorie dans le domaine des indicateurs.

Pour Oakes, le lien entre indicateurs et politique est direct : les indicateurs ont un sens s'ils sont politiquement utiles. D'où la nécessité d'expliquer leur choix, leur construction et la manière de s'en servir. Pour autant, Oakes ne fait aucune association entre indicateurs et évaluation. Elle attire l'attention sur les pressions politiques susceptibles d'émaner de la réalisation d'un ensemble d'indicateurs : "Il doit être clair dès maintenant que les systèmes d'indicateurs ne seront pas neutres, qu'ils ne seront pas des systèmes d'information technologique imperméables aux pressions politiques. Les choix des indicateurs à développer, les buts auxquels ils devraient servir, le type de données à collecter, les comparaisons à effectuer, sont des enjeux aussi bien politiques que techniques. Nous ne pouvons pas ignorer les pressions politiques qui résultent de la simple existence d'un ensemble d'indicateurs" (29). On comprend comment une telle déclaration a pu heurter les statisticiens de l'éducation et braquer une partie de ces spécialistes contre les indicateurs.

Presentant l'évolution, Oakes donne des exemples de pressions engendrées par la publication d'indicateurs : "La pression la plus forte découlant des indicateurs sera exercée par les membres de la communauté enseignante que les indicateurs exposent à un jugement public comme cela ne s'était jamais produit auparavant, en fournissant en surplus des informations sur lesquelles cette communauté n'exerce aucun contrôle. Il est donc logique que les enseignants réagissent, en exerçant à leur tour des pressions pour modeler la sélection des indicateurs, pour influencer le niveau d'agrégation et d'analyse des données, et pour façonner les méthodes d'interprétations, de présentation et publication des données. Une fois les indicateurs en place, les enseignants ne cesseront pas d'être en état d'alerte et se mobiliseront pour biaiser les données dans une direction favorable pour eux et pour les écoles. Cette tentation sera d'autant plus forte si les enseignants ont le sentiment qu'ils n'ont rien à dire dans le processus de développement des indicateurs ou si les informations produites par les indicateurs n'ont que peu d'utilité pour eux. Déjà maintenant on peut observer la manifestation de comportements de ce type. Par exemple, dans certains États des États-Unis qui collectent régulièrement des indicateurs sur les processus scolaires, on a recueilli des témoignages selon lesquels des responsables scolaires incitent les élèves à 'exagérer' leurs réponses, en particulier aux questions sur leurs expériences scolaires. Dans d'autres domaines, des réseaux informels d'enseignants ont fait passer des mots d'ordre parmi les collègues les invitant à sous-évaluer les données sur les salaires et à surévaluer les informations sur la charge de travail, dans le but d'obtenir des politiques plus favorables dans ces deux domaines. Ces pressions vraisemblablement seront proportionnelles à ce qu'on estime pouvoir perdre ou gagner avec les indicateurs" (30). Vision prémonitoire : au Royaume-Uni, les syndicats d'enseignants ont invité leurs adhérents à boycotter

les tests et la participation à l'enquête PISA de 2003. Leur résistance a réussi car le taux de réponse de l'échantillon ayant été inférieur à celui fixé par l'OCDE, les résultats du Royaume-Uni n'ont pas pu être pris en considération.

Dans le même ordre d'idées on peut citer le rapport publié en 1988 par le consortium IEES (Improving the Efficiency of Educational Systems, regroupant la Florida State University, l'Howard University, l'Institut for International Research et la State University of New York à Albany) sur les indicateurs d'efficacité et d'efficience de l'enseignement préparé par Windham (1988). L'intérêt de ce document, commandé par l'Agence américaine pour le développement international (AID), réside dans le fait qu'il a été conçu pour les pays en voie de développement. L'outil proposé, un ensemble d'indicateurs, visait à renforcer les capacités de gestion, de planification et de recherche en éducation des pays en voie de développement pour améliorer les performances de leurs systèmes d'enseignement. Deux aspects méritent d'être soulignés : la conception d'un ensemble d'indicateurs centré sur l'efficacité et l'efficience de l'enseignement ; l'application de ce type d'outil à des systèmes d'enseignement peu développés. La première caractéristique se retrouve dans la phase suivante de mise en œuvre centrée sur l'efficacité interne des systèmes d'enseignement et la deuxième dans la quatrième phase, lorsque l'OCDE lance, avec l'UNESCO et la Banque mondiale, le programme World Education Indicators (WEI, d'où le nom de pays-WEI pour désigner les participants du programme) conçu pour les pays en voie de développement.

Les années 1990 : la volonté des décideurs politiques

Dans la troisième phase, les protagonistes changent. Ce ne sont plus ni les spécialistes des sciences sociales ni ceux des sciences de l'éducation, mais les décideurs politiques eux-mêmes, en responsabilité de la gestion et de l'évolution des systèmes d'enseignement. Leur irruption sur la scène des indicateurs s'opère entre 1987 et 1992 à la suite de la crise et des interrogations concernant la qualité de l'enseignement et des systèmes éducatifs. Deux pays jouent un rôle déterminant : les États-Unis et la France. Leurs délégués sont parvenus à entraîner un groupe substantiel de pays membres pour obliger l'OCDE, dont les spécialistes de l'enseignement renâclaient à mettre sur pied et à produire un ensemble d'indicateurs de l'enseignement.

La production d'un ensemble d'indicateurs internationaux de l'enseignement par l'OCDE en 1992 n'a pas été une promenade de santé. Comme prévu par Oakes, un des experts consultés par l'OCDE, des pressions de toute part se sont manifestées pour essayer d'étouffer ou de modifier le projet dans l'œuf. Dans la coalition des opposants, se côtoient des statisticiens de l'éducation dénonçant scientifiquement la grossièreté des indicateurs et la représentation biaisée qu'ils donnent de l'enseignement ; des représentants des organisations d'enseignants

contestant politiquement la réalité de la crise de la qualité de l'enseignement, en particulier public, un large éventail de pédagogues ou de spécialistes des sciences de l'éducation engagés dans les innovations et les mouvements progressistes d'éducation argumentant éthiquement contre la volonté du pouvoir politique d'instaurer des formes de régulation de l'enseignement calquées sur les principes inspirant la nouvelle gestion publique, la gouvernance du secteur public, la gestion des performances dans l'administration. Nonobstant ces oppositions, l'OCDE parvint à publier en 1992 le premier ensemble d'indicateurs internationaux. Son but était de fournir aux décideurs des informations sûres et comparables sur l'état des systèmes d'enseignement au sens large (y compris le secteur privé). Les données sur les acquis des élèves n'avaient pas une place dominante. Dans la version préliminaire de *Regards sur l'éducation*, présentée en 1991 sous forme de rapport non imprimé et discutée lors d'une réunion internationale des producteurs d'indicateurs et des décideurs organisée à Lugano en septembre 1991, il n'y avait rien ni sur les acquis des élèves ni sur leurs résultats en termes de connaissances. C'est après un intense débat opposant décideurs politiques et scientifiques en réunions plénières à Lugano que les pays membres donnèrent mandat à l'OCDE de réviser le prototype et d'y inclure une section sur les acquis des élèves en utilisant les données fragmentées qui étaient alors disponibles sur le plan international –en particulier la deuxième enquête de l'IEA sur les mathématiques et les sciences (SIMSS 1982-1984) et l'enquête IAEP II de l'ETS, de 1991, qui fournissaient des données comparables sur les performances des élèves de systèmes d'enseignement différents.

Cette troisième phase donne deux points de repère pour notre sujet. Tout d'abord, la production d'indicateurs n'a jamais été associée à l'évaluation de la qualité de l'enseignement. La priorité était de disposer de renseignements éprouvés sur des enjeux politiques cruciaux pour gérer, d'un côté, les répercussions du débat sur la crise de la qualité de l'enseignement et la baisse du niveau d'instruction et, de l'autre, les conséquences budgétaires des politiques d'austérité économique et de contrôle des dépenses publiques. L'évaluation des systèmes d'enseignement n'a pas été abordée –pour preuve, *Making Education Counts: Developing and Using International Indicators*, le recueil des contributions des consultants sollicités par l'OCDE (1994b) pour cadrer la conférence de Lugano en septembre 1991 qui décida de réaliser un ensemble d'indicateurs internationaux de l'enseignement après avoir jugé la tâche impossible, personne n'y parle d'évaluation (OCDE 1994a).

La nouveauté la plus significative du programme fut l'attention dédiée à l'élaboration des indicateurs. D'une manière inhabituelle pour une organisation intergouvernementale comme l'OCDE, la méthode de travail adoptée a été participative et démocratique, axée sur la consultation systématique des producteurs de données pour parvenir à un consensus sur les indicateurs à publier. La consultation a eu un effet indirect non négligeable : la constitution, sur une période

relativement courte, d'un réseau mondial de spécialistes de la production d'indicateurs, visible lors des assemblées générales (terminologie hétérodoxe dans le jargon de l'OCDE). De 1989 à 1995, trois assemblées générales du projet INES (International Indicators of Education Systems) ont eu lieu avec la participation de la quasi-totalité des personnes concernées, à l'échelon mondial, par la production d'un ensemble international d'indicateurs de l'enseignement. La dernière a eu lieu en 1995 à Lahti (Finlande). Il n'y a plus eu ensuite d'assemblée générale véritable. L'OCDE en a, certes, organisé une quatrième en 2000 à Tokyo, mais cette rencontre n'avait que le nom d'assemblée générale, car la participation par pays était réduite à des délégations restreintes et les producteurs de données n'y étaient plus présents, retrouvant les canons d'une Conférence intergouvernementale. Seules les quatre premières versions de *Regards sur l'éducation*, titre du recueil d'indicateurs de l'enseignement produit par l'OCDE (1992, 1993, 1995 et 1996) furent le produit d'une intense coopération entre des centaines de spécialistes de disciplines différentes, actifs à des échelons divers des systèmes d'enseignement.

Les années 2000 : obnubilation pour les indicateurs de résultats et triomphe de la psychométrie comparée

La quatrième et dernière phase, pour le moment, commencée en 1997, est caractérisée par trois points marquants :

- l'amélioration considérable de la qualité des données, imputable en grande partie à la reprise en main du dossier des indicateurs par les statisticiens : l'adoption en 1997 par l'UNESCO de la révision de la CITE (classification internationale type de l'enseignement), l'extension de son emploi au traitement des statistiques nationales de l'enseignement, puis l'adoption en 1995 d'un questionnaire statistique commun à l'OCDE, l'UNESCO et à l'EUROSTAT ont été des facteurs techniques importants ;
- la modification du cadre théorique en 2002 orientant la composition et l'organisation de l'ensemble d'indicateurs produits par l'OCDE ;
- le lancement par l'OCDE du programme PISA d'évaluation cyclique des compétences des élèves de quinze ans.

Ces trois points ont profondément modifié le scénario international des indicateurs de l'enseignement en rendant possible de nouvelles analyses mais aussi en ouvrant les digues qui avaient permis de prévenir les craintes de déformation et d'abus d'utilisation des indicateurs.

Le résultat le plus frappant de cette phase a été l'importance prise par les indicateurs de résultats. De neuf dans l'édition 2001, où on faisait encore la distinction entre une section "résultats au niveau personnel, social et du marché du travail" (cinq indicateurs) et les indicateurs d'acquis des élèves (quatre), ils sont quatorze en 2002 ; quinze en 2003 ; douze en 2004. Dans l'édition de 2002, la

section réservée aux “résultats” qui, pendant une dizaine d’années, était placée en dernière position passe en tête. Rien là d’anodin : c’est la manifestation d’une volonté délibérée de mettre en avant les indicateurs de résultats et de valoriser les informations sur les acquis des élèves, recueillies par une enquête de l’OCDE réalisée au printemps 2000 pour le programme PISA. Ce changement est un tournant décisif dans l’histoire récente des indicateurs. La stratégie de l’OCDE en matière d’enseignement depuis 1995 a été inspirée par cette visée. Dans un document exposant le nouveau cadre d’organisation des indicateurs, le groupe de gestion stratégique avait justifié ce changement d’optique en affirmant qu’“on admettait désormais implicitement que plusieurs dimensions importantes du développement, du fonctionnement et de l’impact des systèmes éducatifs pouvaient être évaluées uniquement si on comprenait les résultats de l’apprentissage ainsi que ses relations aux inputs et aux processus au niveau des individus et des institutions” (OCDE 2001). C’est à la suite de cette nouveauté que le cadre conceptuel et d’organisation des indicateurs de l’enseignement a été changé en 2002, sans qu’il y ait eu de débat démocratique. Tout à coup, une décennie après la publication du premier ensemble d’indicateurs internationaux de l’enseignement, dans lequel avaient été insérés, après beaucoup d’hésitations et de multiples précautions, cinq indicateurs de résultats rubriqués comme des indicateurs expérimentaux et provisoires, on place, au cœur de l’ensemble le plus connu d’indicateurs internationaux de l’enseignement, cinq indicateurs de résultats tirés de données recueillies par un programme d’évaluation de masse conçu pour les besoins de la cause et traitées selon des méthodologies psychométriques particulières, imposées par une communauté influente de scientifiques. Ils avaient été construits à partir des données de l’enquête IAEP (International Assessment of Educational Progress) de l’ETS de 1991 et de la seconde enquête de l’IEA sur l’enseignement des mathématiques (SIMS) réalisée entre 1980 et 1982, ce qui montre par ailleurs l’aridité des informations comparables sur les résultats de l’enseignement disponibles. L’introduction de *Regards sur l’éducation 2002*, première version d’ensemble d’indicateurs de l’OCDE à inclure des indicateurs tirés de PISA, affiche ces intentions : “Le Programme international de l’OCDE pour le suivi des acquis des élèves (PISA), lancé par les gouvernements pour étudier la performance des élèves de façon suivie en se servant d’un cadre conceptuel faisant l’objet d’un consensus à l’échelle internationale –aucune indication n’est donnée sur la procédure suivie pour parvenir à ce consensus– fournit à présent des données comparables sur les résultats de l’éducation et de l’apprentissage ainsi que sur les facteurs essentiels qui conditionnent ces résultats. Ce type d’information comble enfin ce qui aura longtemps été une sérieuse lacune dans le jeu d’indicateurs. PISA entend proposer de nouvelles bases pour alimenter le dialogue politique et relancer la collaboration autour de la définition d’objectifs de l’enseignement à la fois novateurs et réalistes, qui traduisent des jugements sur

les compétences utiles dans la vie adulte. Cette initiative traduit une volonté accrue de se concentrer sur les finalités de l'enseignement plutôt que sur les moyens qui lui sont consacrés. Ce changement de priorités permet de mieux guider les décideurs politiques dans les efforts qu'ils déploient pour améliorer l'enseignement et mieux préparer les jeunes à faire leur entrée dans la vie adulte, dans un monde en pleine mutation placé sous le signe de la mondialisation" (7).

En mélangeant, dans une même grille conceptuelle et un même ensemble d'indicateurs conçus pour observer l'évolution des systèmes d'enseignement, des indicateurs sur les institutions de formation et les fournisseurs des services éducatifs, des indicateurs sur les curricula et l'environnement éducatif à l'intérieur des institutions de formation, des indicateurs sur le rendement de l'éducation à l'échelle individuelle, désignés comme indicateurs sur les participants individuels aux activités d'apprentissage (OCDE 2001), l'OCDE tombe dans un piège qu'elle a elle-même tendu, celui de la confusion entre indicateurs et évaluation. Ce faisant, elle se fait conditionner par sa propre production de données sur les acquis des élèves, y compris celles sur la littératie des adultes, autre grande enquête internationale gérée par l'OCDE entre 1994 et 1998. Première enquête jamais réalisée d'évaluation des compétences adultes, IALS, International Adult Literacy Survey, publiée comme une réalisation conjointe de Statistique Canada et de l'OCDE, elle portait sur les degrés de compétences des adultes dans la compréhension de textes écrits. Reprenant des études des années 1990 aux États-Unis, elle innovait sur plusieurs points, notamment en interrogeant les adultes de l'échantillon chez eux, dans leur cadre de vie privé. Après un premier tour auquel avaient participé neuf pays, l'enquête fut dupliquée deux fois, en 1996 et 1998. Le rapport final présente les résultats concernant vingt-trois pays (OCDE/Statistique Canada 1999). Par ailleurs, l'OCDE a continué à se servir pour la réalisation de ses indicateurs de résultats des enquêtes de l'IEA, notamment de TIMSS et de PIRLS (Progress in International Reading Literacy Study).

Deux problèmes se présentent ici : le premier, déjà apparu au début de cette histoire, est celui de la formation des dirigeants, des décideurs et de la classe politique à comprendre, interpréter et utiliser les indicateurs et les statistiques de l'enseignement ; le deuxième est celui de la coexistence dans un même ensemble d'indicateurs de données de nature différentes, c'est-à-dire de données sur les compétences et les connaissances individuelles et de données systémiques.

Concernant la formation des décideurs, l'OCDE, se rendant compte que le volume d'indicateurs était devenu ingérable pour les décideurs –un volume bilingue de 150 pages en 1992, un volume bilingue de 300 pages en 1993 ; un volume de 378 pages pour chaque langue en 1995– décida en 1996 de lancer une série parallèle au recueil d'indicateurs appelée "Analyse des politiques d'éducation". Si le titre de la série a subi des modifications, le concept est resté stable : produire une brochure d'analyse politique de l'éducation en partant des indicateurs,

qui devait fonctionner comme un dispositif léger d'interprétation des indicateurs à l'intention des décideurs politiques. L'opération n'a réussi que partiellement, car rapidement cette série, qui aurait dû être un outil de divulgation, devint un exercice savant et académique permettant aux producteurs d'indicateurs de rédiger des essais scientifiques sur leurs données. Dans la première version de 1996, les chapitres renvoient directement aux indicateurs de *Regards sur l'éducation* en les citant nommément, pratique abandonnée ensuite. Cependant, cette seule initiative n'aurait pas suffi à aider les décideurs à comprendre les indicateurs et à s'en servir à bon escient, comme l'a démontré la façon dont les cercles politiques de plusieurs pays se sont appropriés des données PISA, avec la complicité des responsables du programme. La construction du test PISA et la manière dont l'OCDE a présenté les scores en ont facilité une exploitation simpliste (Bottani & Vrignaud 2005).

Concernant le deuxième problème, le mélange de données d'ordres différents et d'unités d'analyses multiples, il est indéniable que "les décideurs, comme l'affirme l'OCDE, souhaitent être renseignés sur les connaissances et les compétences des élèves de leur pays et savoir comment ces acquis se comparent à ceux des jeunes d'autres pays" (DEELSA/ED/CERI/CD 2001, 97). La terminologie utilisée ici mérite attention : pour l'OCDE, les systèmes d'enseignement ne sont que des institutions nationales. Dans ses analyses elle réduit souvent, implicitement, les systèmes d'enseignement aux dispositifs étatiques d'enseignement. Pour autant, le lien entre les scores obtenus dans les tests des enquêtes internationales sur les connaissances des élèves et le fonctionnement des systèmes d'enseignement n'est pas évident. Par exemple, la correspondance entre les tests et les programmes d'enseignement a été une pierre d'achoppement permanente pour l'IEA (International Association for Evaluation of Educational Achievement). L'OCDE a préféré trancher dans le vif et a renoncé à construire des tests reflétant les contenus des programmes d'enseignement. Un groupe de travail de l'OCDE interne au projet des indicateurs internationaux (le réseau A) s'est penché pendant quatre ans sur ces enjeux mais, plutôt que d'analyser les relations entre évaluation des élèves et programmes d'enseignement, il a étudié une stratégie de collecte régulière, rapide et sûre des données sur les acquis des élèves. Le rapport livré en avril 1997 est le document fondateur du programme de l'OCDE d'évaluation des acquis des élèves (PISA). Il décrit en détail la structure et l'organisation du dispositif d'évaluation des connaissances, des compétences et des qualifications des jeunes à quinze ans, sans discuter le sens de ces informations quant au niveau de littératie des jeunes par rapport au fonctionnement des systèmes d'enseignement. Même en supposant que les tests soient univoques et non ambigus (ce qui n'est pas), même en admettant qu'ils se déroulent selon des procédures strictes, comparables dans les différents systèmes d'enseignement (ce dont on peut douter), leurs résultats ne peuvent être considérés ni comme parfaitement comparables ni

comme un miroir fidèle de la qualité d'un système d'enseignement. Ces résultats sont peut-être les meilleurs qu'on puisse obtenir, à l'heure actuelle, à l'échelon mondial, compte tenu des rouages de financement, des organisations internationales qui régulent ces entreprises d'évaluation, mais ils ne sont pas parfaits et doivent être traités avec beaucoup de précautions. Des questions demeurent : de quoi ces tests sont-ils le reflet ? Que représentent ces résultats ? Qu'en infère-t-on ? Il ne suffit pas de bâtir une base de connaissances, aussi solide soit-elle, pour produire des indicateurs à intégrer dans un ensemble censé surveiller l'état et l'évolution des systèmes d'enseignement. Ce n'est pas la base de données en soi qui est problématique –encore faut-il que soient respectés les standards techniques de qualité reconnus dans le domaine (Postlethwaite 2004)– mais le passage automatique de la base de données sur les acquis et compétences à des indicateurs intégrables dans un dispositif visant à fournir des renseignements sur la gestion, l'organisation et le fonctionnement des systèmes d'enseignement.

Cet assemblage pose trois problèmes :

- celui de la compatibilité de ces données entre elles et de la validité de leur inclusion dans un ensemble unique ;
- celui de la validité des inférences suggérées par les coïncidences de ces données ;
- celui de l'évaluation du système (l'équivalent d'une marque dans le système commercial), c'est-à-dire son efficacité, son efficience, la concordance entre ses performances et ses objectifs, ses modalités de fonctionnement par rapport à sa taille et à ses ressources avec des indicateurs d'acquis individuels mis en relation avec des facteurs multiples sur la base de suppositions ou d'hypothèses précaires.

Ces questions ne mettent en jeu ni la pertinence ni l'utilité des indicateurs en soi, mais leur exploitation et leur emploi.

Les indicateurs dans la pratique

J e me suis concentré sur les indicateurs internationaux produits par l'OCDE car ce cas est emblématique : premier élaboré dans le domaine de l'enseignement, il est un des plus connus et la qualité statistique des données utilisées pour construire les indicateurs est remarquable. Sur le plan international il y a néanmoins des démarches qui aboutissent ou pourraient aboutir à des ensembles alternatifs fondés sur des systèmes différents de valeurs, d'autres définitions de la qualité de l'enseignement, qui ne prétendent pas à la neutralité de l'information sous le manteau de la rigueur statistique et technique. Ainsi les indicateurs de développement humain dans le monde, publiés par le PNUD depuis 1990, qui contiennent une batterie d'indicateurs économiques, sociaux et environnementaux.

La méthode utilisée par le PNUD a été amplement discutée dans les milieux spécialisés (Gadrey & Jany-Catrice 2005) car le cadrage des indicateurs du PNUD est très différent de celui adopté par l'OCDE.

Un autre ensemble alternatif d'indicateurs des systèmes d'enseignement, sur l'équité, est proposé par le GERESE (Groupe européen de recherche sur l'équité des systèmes éducatifs), dans le cadre d'un projet soutenu par la Direction générale de l'éducation et de la culture de la Commission européenne (Socrates SO2-610BGE). Ce groupe a élaboré un cadre conceptuel inspiré par différentes théories de la justice appliquées à l'éducation et l'a validé en calculant un ensemble de vingt-neuf indicateurs et en proposant des démarches originales dans la perspective de la construction d'indicateurs de l'enseignement en fonction d'un ordre de valeurs sociales susceptibles de mettre en évidence les effets des techniques de gouvernement sur la justice sociale (GERESE 2005)?

Dans différents pays, Canada, États-Unis, France, Suisse, Belgique (communauté néerlandophone) par exemple, au cours de ces derniers quinze ans, ont été produits des ensembles d'indicateurs de l'enseignement conçus pour le pilotage et le suivi (monitorage) des systèmes d'enseignement. Dans certains systèmes politiques fédéralistes on a aussi commencé à produire des ensembles d'indicateurs à l'échelon régional –en Suisse pour les Cantons de Genève ou du Tessin ; en Italie, dans une perspective politique particulière, où des régions produisent leur ensemble d'indicateurs de l'enseignement en accompagnement d'un processus de dévolution et décentralisation des compétences en éducation dans un système autrefois hyper-centralisé. Ces ensembles ne sont ni une copie ni une reproduction de ceux de l'OCDE dont ils s'écartent car ils ont été élaborés à partir d'autres cadres théoriques. Conçus en fonction soit de l'organisation du système d'enseignement (par exemple en France), soit de ses priorités de fonctionnement (par exemple à Genève), soit de ses finalités propres (par exemple au Tessin). Ces ensembles n'ont pas été développés pour évaluer les systèmes d'enseignement.

Pour conclure ce travelling, un mot sur *L'État de l'École*, ensemble d'indicateurs produit tous les ans par le ministère français de l'Éducation nationale depuis 1991. Il comprend trente indicateurs calqués sur la structure du système français. Dans un article qu'il lui consacre, Meuret (2001) discute les effets de la production de ces indicateurs et remarque "qu'il est assez facile de s'apercevoir que le débat sur l'éducation en France porte depuis dix ans sur des thèmes qui ne sont pas abordés dans la brochure". Force est de constater que la production d'ensemble d'indicateurs n'engendre pas en soi des changements dans les politiques de l'éducation. Il n'y a pas de retombées automatiques imputables ni à une plus grande transparence sur la nature et le fonctionnement du système d'enseignement ni à une production abondante de données ni à une amélioration de la qualité des informations. Le système peut rester imperméable à ces changements

ou à ces efforts, comme le constate Meuret. Il ne faut pas se faire d'illusion sur la valeur des indicateurs ni l'importance qui leur est attribuée dans ce domaine. C'est la politique qui se sert des indicateurs et des évaluations et non les indicateurs ou les évaluations qui conditionnent ou déterminent les politiques. Parfois les intérêts des uns rencontrent ceux des autres, mais cette coïncidence est plus le fruit du hasard ou de la volonté de pouvoir que d'une stratégie démocratiquement délibérée pour améliorer des systèmes d'enseignement. Une meilleure connaissance du système n'est pas en soi un avatar de changement. Celui-ci s'enclenche seulement si d'autres ingrédients sont présents. Pour Meuret, ce qui a manqué en France c'est "un débat sur l'éducation auquel les usagers, et plus largement les citoyens, seraient admis à participer" (25). Le contenu et la conception de l'ensemble d'indicateurs concourent en partie aussi à expliquer cet échec : l'absence de certaines données – "par exemple sur la qualité de la vie à l'école" précise l'auteur– "considérées impertinentes dans le débat sur l'enseignement tel qu'il se déroule entre professionnels" peut expliquer la faiblesse de l'impact de l'effort de clarification et d'amélioration de l'information.

Au moins pour un système d'enseignement, le paramètre décisif pour que les indicateurs aient une retombée semble la procédure de production et plus particulièrement l'appropriation par les réseaux d'enseignement du cadre conceptuel de choix, des hypothèses sous-jacentes relatives à l'évolution de l'école et, *dulcis in fundo*, des indicateurs eux-mêmes. À cette condition, il est plus difficile de réduire la production des indicateurs à une opération de pouvoir ou d'en faire un acte de terrorisme intellectuel imposant aux usagers un modèle éducatif venu d'ailleurs, sans se soucier de leurs préoccupations ou de leurs aspirations. L'Office fédéral de la statistique (OFS) de Suisse en donne un exemple paradoxal. Il a publié en août 2005 un rapport comparant les prestations de l'enseignement en Suisse avec les objectifs adoptés par l'Union européenne dans le cadre de la stratégie de Lisbonne ("Progress Towards the Lisbon Objectives in Education and Training. 2005 Report" <http://europa.eu.int/comm/education/policies/2010/doc/progressreport05.pdf>).

L'étude visait à démontrer que l'enseignement en Suisse ne se portait pas si mal, malgré les résultats peu gratifiants obtenus par les élèves de quinze ans dans l'enquête PISA. Cet office, en s'appuyant sur des données internationales qu'il ne maîtrise pas et sur ses données nationales, a publié un rapport, antidote pour tranquilliser l'opinion, les médias, les enseignants, les décideurs sur l'état de l'enseignement en Suisse. Face à cette publication, exemple d'exploitation acritique des données et des indicateurs internationaux, les responsables cantonaux de l'enseignement, qui n'avaient pas été consultés, réagirent, en prenant leurs distances par le communiqué ci-dessous adressé dès août 2005 aux responsables de l'enseignement en Suisse :

“Mesdames, Messieurs,

Nous vous faisons parvenir ci-joint, à titre d'information préalable, le communiqué de presse qui va être publié par l'Office fédéral de la statistique pour présenter le rapport intitulé “Le système d'éducation suisse en comparaison européenne”. (Nous ne possédons pas le rapport lui-même).

Si les médias s'adressent à vous pour des questions de rang, classement et autre place de podium, nous vous recommandons de ne pas entrer en matière. Ce sont des questions qui n'ont pas d'intérêt pour cette étude et qui sont difficiles à comprendre sur la base du communiqué de presse.

En revanche, ce qui nous paraît important, c'est que le système éducatif suisse :

a) est mesuré à l'aune des objectifs fixés dans le cadre de la stratégie de Lisbonne adoptée par l'UE, b) est comparé à la moyenne européenne et c) est mis en parallèle avec certains pays européens en particulier.

Il ressort en l'occurrence de ce rapport que les résultats obtenus par la Suisse se situent, dans tous les domaines évoqués dans le communiqué de presse, au-dessus de la moyenne européenne et que la Suisse a déjà atteint ou dépassé les niveaux de référence (*benchmarks*) fixés par l'UE pour 2010. Il va de soi que, dans certains domaines, certains pays se révèlent meilleurs que la Suisse et il est clair également que la Suisse doit poursuivre ses efforts en vue d'améliorer son système éducatif.

Avec mes salutations les meilleures.

XY... Chef de l'Unité de coordination Développement de la qualité CDIP / EDK

(Conférence des chefs des départements cantonaux de l'Instruction publique)

Conclusion

L'interrogation sur le rôle des indicateurs dans les politiques de l'éducation est dominée par les enjeux de l'évaluation et concerne une question cruciale pour la recherche en éducation, celle de l'espace de liberté dont jouit la science. En effet, le développement des indicateurs de l'enseignement, aussi bien à l'échelon international, national voire régional, oblige à faire face aux questions relatives à l'emploi de cet outil et à son utilisation. Jonas (1903-1993), philosophe allemand, élève de Heidegger, a approfondi cette distinction dans son œuvre et tout particulièrement dans *The Imperative of Responsibility: In Search of Ethics for the Technological Age* (1979). Il propose de distinguer la science “dévouée essentiellement à la connaissance de ce qui nous entoure” et “la non-science, c'est-à-dire la technique, dévouée au contraire à modifier la réalité”. Est-ce qu'on peut soutenir que les indicateurs sont une démarche cognitive neutre qui ne prétend pas modifier les systèmes d'enseignement ? Est-ce qu'ils font partie de la science selon la définition de Jonas ou de la non-science ? N'y aurait-il pas un degré de naïveté à prétendre que les indicateurs ne sont qu'un outil de connaissance, sans incidence sur la réalité de l'enseignement, sur les modalités de fonctionnement de l'appareil scolaire ? La production d'ensemble d'indicateurs, soutenue et encouragée par les autorités scolaires, est-elle neutre ?

La dichotomie de Jonas n'est pas défendable. La caractéristique de la réalité physique, celle dans laquelle nous sommes immergés, celle qui nous constitue ou que nous façonnons, est qu'elle peut être connue uniquement si on la modifie. Cet acquis théorique nous a été légué par les avancés scientifiques du XX^e siècle. Ce principe s'applique aux sciences humaines et aussi à la connaissance des systèmes d'enseignement. Posons donc comme valable l'axiome selon lequel aucune démarche cognitive ne laisse indemne l'objet qu'elle observe, mesure, analyse, compare, déchiffre. Il n'y a pas de ligne de démarcation ontologique claire entre le champ de la science et celui de la technique ni entre la connaissance du réel et sa modification. Il est impossible à la fois d'attribuer un statut éthique distinct à ces deux opérations et d'estimer qu'existe une liberté indéfinie pour les processus cognitifs et une responsabilité limitée, lorsque la connaissance s'applique à des processus ou à des actes visant à modifier la réalité. Les indicateurs produisent de la connaissance sur les systèmes d'enseignement et, en même temps, leur conception, installation et production les modifient. La remarque vaut pour l'évaluation. Cette situation spéculative ne signifie pourtant pas une similitude entre développement des indicateurs et programmes d'évaluation.

Malheureusement, nous ne connaissons que de manière imprécise et limitée les effets du développement des indicateurs sur l'évolution des systèmes d'enseignement. Ils sont probablement multiples et différenciés, mais pour le moment nous ne pouvons pas aller plus loin dans les inférences et les suppositions. Il n'y a pas d'imperméabilité entre développement et construction d'un ensemble d'indicateurs d'un côté et changements des systèmes d'enseignement de l'autre. Il serait faux de prétendre que les indicateurs n'ont ni effets ni répercussions sur les politiques d'éducation et la gouvernabilité des systèmes d'enseignement. Si un ensemble d'indicateurs n'est pas uniquement le produit d'une politique déterminée, il ne se situe pas non plus en dehors du champ politique et des processus visant à changer l'organisation et le fonctionnement des systèmes d'enseignement et les techniques de gouvernement. Établir un lien direct entre le développement d'ensembles d'indicateurs et l'évaluation de la qualité de l'enseignement est par ailleurs problématique. Ce sont deux opérations différentes, greffées différemment sur le tronc de la politique d'enseignement et des réformes scolaires, qui n'ont pas les mêmes fonctions. La production d'indicateurs et l'évaluation ont une incidence sur le fonctionnement des systèmes d'enseignement, sans être nécessairement l'expression d'une même politique d'éducation. Nous ne pouvons non plus exclure qu'il y ait des affinités entre ces deux démarches et que l'une exploite ou utilise l'autre. Les stratégies nécessaires pour extrapoler des données adaptées pour la construction d'indicateurs de réussite en partant des évaluations sur grande échelle sont complexes d'autant que ces évaluations posent à leur tour des problèmes de fiabilité, qui ne sont qu'en partie résolus. Ainsi, l'articulation

entre évaluation des acquis des élèves sur une vaste échelle et développement d'ensembles d'indicateurs ne va pas de soi et doit être abordée avec précaution (Porter & Gamoran 2002).

Références bibliographiques

- BERGER E., BOTTANI N., SOUSSI A., OSTINELLI G., GAUVRAU C. & RHYN H. 2005 "De l'émergence du courant *School improvement* et exemples d'applications", *Revue française de pédagogie*-148, 119-133
- BOTTANI N. & VRIGNAUD P. 2005 *La France et les évaluations internationales*, Paris, Haut Conseil de l'évaluation de l'école
- BOTTANI N. 2005 Des indicateurs dans l'enseignement : une contribution à la comparaison internationale. Intervention lors du colloque sur la comparaison en éducation organisé par la FPSE, Université de Genève, 8-9 décembre 2005, sous presse
- COMMISSION EUROPÉENNE 2005 "Progress Towards the Lisbon Objectives in Education and Training" Report, Bruxelles
- GADREY J. & JANY-CATRICE F. 2005 *Les nouveaux indicateurs de richesse*, Paris, La Découverte
- GERESE 2005 *L'équité des systèmes éducatifs européens. Un ensemble d'indicateurs*, Service de pédagogie théorique et expérimentale, Université de Liège
- JONAS H. 1979 *The Imperative of Responsibility: In Search of Ethics for the Technological Age*, Chicago, University of Chicago Press 1984 ; pour la traduction en français : *Le principe de responsabilité*, Flammarion, 1998
- LAND K.C. 1970 "Social Indicators", *Social Science Methods*, New York, Free Press
- MCBEATH J. 1999 *Schools Must Speak for Themselves : The Case for School Self-Evaluation*, London, Routledge
- Self-Evaluation in European Schools: A Story of Change
- MEURET D. 2001 "De la contribution des indicateurs au débat sur l'éducation. Une étude de cas : L'État de l'École", *Politiques d'éducation et de formation. Analyses et comparaisons internationales*, Bruxelles, De Boeck
- MINISTÈRE DE LA JEUNESSE, DE L'ÉDUCATION NATIONALE ET DE LA RECHERCHE, DIRECTION DE LA PROGRAMMATION ET DU DÉVELOPPEMENT Publication annuelle, *L'État de l'école*. Paris
- NUTTAL D. 1994 "Choosing Indicators", *Making Education Count : Developing and using international Indicators*, Paris, OECD
- OAKES J. 1986 *Education Indicators. A guide for Policymakers*, Center For Policy Research in Education
- OCDE 1992, 1993, 1995, 1996 *Regards sur l'éducation*, Paris, OCDE
- OCDE 1994a *Évaluer l'enseignement. De l'utilité des indicateurs internationaux*, Paris, OCDE
- OCDE 1994b *Making Education Counts: Developing and Using International Indicators*, Paris, OCDE
- OCDE 2001 *Indicators of Education Systems: Scope of INES Activities*, EELSA/INES/SMG (2001)12, Paris (document interne)

- OCDE/STATISTIQUE CANADA 1999 *La littératie à l'ère de l'information*, Paris, OCDE
- PORTER A. & GAMORAN A. 2002 *Methodological Advances in Cross-National Surveys of Educational Achievement*, Washington D.C., National Academy Press
- POSTLETHWAITE N. 2004 *Monitoring Educational Achievement*, UNESCO, IPE, rapport 81 de la série *Fundamentals of educational planning*, Paris
- THE NATIONAL COMMISSION ON EXCELLENCE IN EDUCATION 1983 *A Nation at Risk: The Imperative for Educational Reform. A Report to the Nation and the Secretary of Education* United States Department of Education, Washington, D.C.
- WINDHAM D. 1988 *Indicators of Educational Effectiveness and Efficiency, Improving the Efficiency of Educational Systems*, United States Agency for International Development, Washington, D.C.