

MINISTÈRE DE L'ÉDUCATION NATIONALE,
DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE

Inspection générale de l'éducation nationale

*Inspection générale de l'administration
de l'éducation nationale et de la recherche*

L'évaluation des collèges et des lycées en France :

bilan critique et perspectives en 2004

*rapport à monsieur le ministre
de l'éducation nationale,
de l'enseignement supérieur
et de la recherche*



Juillet 2004

N° 2004-118

MINISTÈRE DE L'ÉDUCATION NATIONALE,
DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE

Inspection générale de l'éducation nationale

*Inspection générale de l'administration
de l'éducation nationale et de la recherche*

Groupe thématique « Evaluation »

L'évaluation des collèges et des lycées en France :
bilan critique et perspectives en 2004

JUILLET 2004

Jean ÉTIENNE
*Inspecteur général
de l'éducation nationale*

Roger-François GAUTHIER
*Inspecteur général de l'administration
de l'éducation nationale et de la recherche*

SOMMAIRE

Introduction 1

1. L'évaluation des établissements publics locaux d'enseignement : quels enjeux ? 5

1.1. Pourquoi s'intéresser à l'« évaluation » des lycées et collèges en France en 2004 ? 5

1.1.1. *Une interrogation sur une exception française : la faible attention portée à la question de l'évaluation des établissements..... 5*

1.1.2. *L'évaluation des établissements : un cadre institutionnel qui devrait s'imposer..... 8*

1.1.3. *L'évaluation des établissements : une relative faiblesse au sein de l'appareil d'évaluation 10*

1.1.4. *L'évaluation des établissements : une nécessité renforcée par les difficultés que rencontre le système 11*

1.2. Les difficultés de l'exercice 12

1.2.1. *Les difficultés venant du manque de visibilité des établissements en tant que tels 13*

1.2.2. *Les difficultés venant des réticences classiques aux démarches d'évaluation dans le champ de l'éducation 15*

1.3. Précautions pour que cet exercice ne demeure pas lettre morte : définir avec clarté les finalités et le cadre de l'évaluation des établissements 20

1.3.1. *Appréhender l'établissement dans sa globalité..... 20*

1.3.2. *Clarifier la démarche d'évaluation..... 21*

1.3.3. *Privilégier les résultats des élèves 22*

2. L'évaluation de l'EPLÉ aujourd'hui, un paysage marqué par son caractère disparate 23

2.1. Les expériences étrangères d'évaluation des établissements : acquis méthodologiques et problèmes ouverts 23

2.1.1. *La diversité des facteurs à l'origine du développement de démarches d'évaluation 23*

2.1.2. *Positionnement des instances chargées des évaluations externes (inspections générales, notamment) 25*

2.1.3. *Méthodes et critères des évaluations externes 25*

2.1.4. *La question du rapport entre l'évaluation et les usagers 27*

2.1.5. *Développement et ambiguïtés des démarches d'auto-évaluation..... 28*

2.2. Les évaluations d'EPLÉ à initiative nationale en France	28
2.2.1. <i>Un héritage riche mais devenu partiel.....</i>	29
2.2.2. <i>Du contrôle à l'analyse systémique.....</i>	30
2.2.3. <i>Limites de la situation actuelle.....</i>	33
2.3. Les évaluations d'EPLÉ sur initiative académique.....	35
2.3.1. <i>La mise en place des évaluations en académie : une gestation plus ou moins difficile</i>	35
2.3.2. <i>Des processus d'évaluation tiraillés entre des objectifs et des attentes contradictoires</i>	40
2.3.3. <i>L'impact des évaluations reste limité en dehors des personnels d'encadrement.....</i>	43
3. L'évaluation de l'EPLÉ demain : perspectives méthodologiques et stratégiques	47
3.1. Revisiter l'évaluation de l'EPLÉ	47
3.1.1. <i>L'évaluation de l'EPLÉ : une démarche spécifique.....</i>	47
3.1.2. <i>Evaluation interne et évaluation externe : entre tensions et complémentarités.....</i>	50
3.1.3. <i>Construire une problématique de l'évaluation.....</i>	53
3.2. Améliorer l'évaluation de l'EPLÉ dans le cadre de l'existant.....	57
3.2.1. <i>Assurer une meilleure transparence à l'évaluation externe et associer davantage l'établissement à ses différentes phases</i>	57
3.2.2. <i>Recentrer l'objet de l'évaluation sur la construction des performances au sein de l'EPLÉ.....</i>	59
3.2.3. <i>Mieux utiliser les indicateurs d'évaluation disponibles.....</i>	64
3.3. Repenser a terme l'évaluation des EPLÉ	67
3.3.1. <i>Revoir l'articulation entre les différents niveaux d'évaluation : constituer une chaîne de l'évaluation.....</i>	67
3.3.2. <i>Réformer notre système d'informations statistiques</i>	69
3.3.3. <i>Élargir et confronter divers modes d'investigation.....</i>	73
Conclusion : l'évaluation, une ambition pour le service public....	77
Annexes	79

Ce rapport a été élaboré dans le cadre du travail du groupe thématique « évaluation : méthodes, objectifs, résultats » conjoint à l'inspection générale de l'éducation nationale et à l'inspection générale de l'administration de l'éducation nationale et de la recherche, mis en place en septembre 2003, qui regroupait les inspecteurs généraux dont les noms suivent :

- M. Alain -Marie BASSY, IGAENR
- M. François -Yves CANEVET, IGAENR
- M. Yvon CEAS, IGAENR
- Mme Michèle CHEVALIER-COYOT, IGEN
- Mme Gisèle DESSIEUX, IGEN
- Mme Annie DYCKMANS, IGEN
- Mme Françoise DUCHENE, IGEN
- Mme Lucienne DUTRIEZ, IGAENR
- M. Jean ETIENNE, IGEN
- M. Roger-François GAUTHIER, IGAENR
- M. Jean-René GENTY, IGAENR
- Mme Béatrice GILLE, IGAENR
- M. Michel LEBLANC, IGEN
- M. Tristan LECOQ, IGEN
- M. Pierre MALLEUS, IGEN
- Mme Myriem MAZODIER, IGAENR
- M. Jean-Louis MICHARD, IGEN
- M. Jacques MOISAN, IGEN
- M. Jean-Paul PITOORS, IGAENR
- M. Jean-Claude RAVAT, IGAENR
- M. Yvon ROBERT, IGAENR
- Mme Christine SAINT-MARC, IGEN
- M. Bernard TOULEMONDE, IGEN
- Mme Brigitte WICKER, IGAENR

Introduction

En France, aujourd'hui, un proviseur de lycée, un principal de collège peuvent recevoir dans l'établissement qu'ils dirigent des « visites » qui ont toutes en commun d'être conduites par des membres de « corps d'inspection », mais qui diffèrent, tant par leurs objectifs que leurs modalités : c'est ainsi que parcourent le territoire avec une mission d'« évaluation » des établissements scolaires ou au moins une mission comprenant certains aspects d'évaluation des établissements scolaires :

- pour le niveau national, aussi bien des inspecteurs généraux de l'administration de l'éducation nationale et de la recherche (qui parlent volontiers d'un « audit » de l'établissement, mais plus communément de son « suivi »), des inspecteurs généraux de l'éducation nationale de la spécialité « établissements et vie scolaire » (dont la démarche, dans ses modalités sinon dans ses finalités, s'apparente assez souvent à la première), et des équipes formées d'un ou plusieurs inspecteurs généraux de l'éducation nationale et d'un ou plusieurs inspecteurs généraux de l'administration de l'éducation nationale et de la recherche, dans le cadre souvent d'opérations d'évaluation de l'enseignement à l'échelle d'une académie entière... ; des inspecteurs généraux de l'éducation nationale peuvent aussi se trouver faire des observations portant sur la globalité d'un établissement à l'occasion de l'inspection individuelle d'un professeur exerçant en classes préparatoires ou en sections de techniciens supérieurs ;
- pour le niveau académique, des inspecteurs pédagogiques régionaux, qui, de façon relativement informelle, à l'occasion d'une « inspection » individuelle traditionnelle de professeurs d'une discipline, adoptent parfois une attitude d'évaluation partielle de l'établissement, mais aussi les mêmes inspecteurs intervenant de façon groupée, et parfois assistés d'agents extérieurs à des corps d'inspection, à l'initiative des autorités déconcentrées de l'académie ou du département pour procéder à l'évaluation de l'établissement.

Qui interrogerait ces différents « visiteurs » sur les motifs de leurs visites et l'articulation de leurs démarches, constaterait qu'elles s'inscrivent dans des logiques différentes, qu'elles s'ignorent largement, et que les commandes auxquelles elles répondent semblent peu reliées *a priori*.

Est-ce à dire que l'appareil français d'éducation nationale « en fait trop » pour l'évaluation de ses établissements, y investit trop de forces de travail qui pourraient être employées à autre chose ? Rien n'est moins sûr sans examen, car on pourrait aussi repérer sans difficulté tel collège ou lycée qui n'a reçu aucune « visite » à portée globale depuis des années.

Qui se place par ailleurs « du côté » des usagers, et au premier chef des parents et des élèves, découvre alors un autre aspect de l'état de l'évaluation des établissements en France : si les autorités nationales ou académiques, parfois les responsables de l'établissement sont les commanditaires ou les bénéficiaires des évaluations qui sont conduites, ce n'est que très rarement le cas des usagers, à la différence de ce qui se passe dans bon nombre de pays. Les établissements français diffusent pourtant de plus en plus souvent de l'information à propos d'eux-mêmes, en particulier sur des sites internet, mais cette information, le plus souvent simple « carte de visite », est tournée vers le compte-rendu d'activités plus que vers les résultats, vers le récapitulatif de « moyens » plus que vers la mesure de l'atteinte d'objectifs : il ne s'agit pas d'évaluation.

Si l'évaluation des établissements est donc actuellement le fait de plusieurs acteurs dont on voit mal le positionnement réciproque, si les usagers ne sont pas les destinataires ordinaires des évaluations que des membres de l'éducation nationale produisent sur les établissements d'enseignement, on peut aussi constater que la question de l'évaluation des établissements ne correspond pas actuellement à un problème de politique éducative évoqué fortement au plan national¹, et donc s'interroger sur l'opportunité qu'un groupe d'inspecteurs généraux se saisisse *hic et nunc* de ce sujet qui peut effectivement n'apparaître pas aussi brûlant que d'autres en matière de politique éducative.

Ils l'ont fait, pourtant, pour des motifs assez simples : il leur est apparu en effet, dans l'exercice de leur métier, que les responsables des établissements qu'ils sont appelés à visiter ont souvent des difficultés à mesurer eux-mêmes les résultats et les effets de l'action collective de leur établissement au sein du système éducatif dont ils font partie. Si certaines pratiques récentes, comme celle de l'auto-diagnostic préalable à la lettre de mission que l'autorité hiérarchique adresse aux chefs d'établissements, font évoluer la situation de façon positive, ces évolutions sont apparues aux inspecteurs généraux comme lentes au regard de la situation qu'ils constatent souvent : des établissements peu mobilisés par les démarches d'évaluation. Le système éducatif français anciennement dirigé par des normes et un contrôle centralisé a certes progressivement laissé à l'établissement une marge de manœuvre plus importante pour répondre à la diversité des questions qui lui sont posées, mais cela s'est fait bien souvent sans que les responsables, les agents et les usagers aient été placés ou se soient

¹ A l'occasion du « débat national sur l'avenir de l'école », le sujet qui interrogeait sur « le couple autonomie et évaluation des établissements scolaires » a été « assez peu choisi » (*Les Français et leur école, le miroir du débat*, p. 157). Les résultats de ce débat sur ces points sont repris en annexe 2.

placés en position de réfléchir suffisamment à l'évaluation de leur activité collective aussi bien qu'individuelle.

La lente émergence de l'« établissement public local d'enseignement »² dans un paysage éducatif français que cette émergence est appelée à renouveler assez profondément implique sans doute que soient remises en perspective les fonctions traditionnelles du contrôle, qui ne sont pas à abandonner sans examen, et des fonctions nouvelles d'évaluation, voire d'auto-évaluation répondant à la situation nouvelle, qui méritent une définition spécifique.

Les questions à poser étaient assez simples :

- pourquoi s'intéresser à ce thème de l'évaluation d'établissement en France en 2004 ?
- à quels enjeux de politique éducative renvoie le thème de l'évaluation des établissements d'enseignement ? Sont-ils importants ?
- comment l'institution y répond-elle aujourd'hui, à ses différents niveaux ?
- quelles sont les questions politiques et techniques qui se dégagent du sujet et auxquelles une réponse devrait être apportée ?
- est-il plus généralement possible, et à quelles conditions, de concilier culture d'évaluation et culture d'éducation ? ou plus précisément objectifs et pratiques d'évaluation et objectifs et pratiques d'un établissement chargé d'éducation ? d'imaginer que les établissements soient aidés dans leur tâche par la diffusion de pratiques d'évaluation ou d'auto-évaluation qui soient conformes au génie et aux valeurs du système éducatif français et qui agréent à ses acteurs ?

Les auteurs du présent rapport sont en pleine conscience du fait qu'« évaluer » une école est un sujet qui fait depuis toujours émerger quantité de questions sur la légitimité ou le sens de la démarche. Ils considèrent toutefois qu'il peut y avoir de bonnes et de moins bonnes évaluations, des évaluations qui passent à côté de la réalité comme des évaluations qui la découvrent quand elle était ignorée des acteurs, des évaluations qui n'apportent aucun progrès, comme des évaluations qui, suivies après un an, montrent à tous ceux qui y avaient passé du temps que c'était un investissement gagnant : il faut être d'abord vigilant sur la qualité d'une démarche qui prétend précisément s'intéresser à la qualité de l'activité d'une école, mais la seule erreur serait de se priver de progresser pour de mauvaises raisons quand c'est possible pour de bonnes.

² Depuis le décret du 30 août 1985 qui le crée.

Car que peut au fond apporter l'« évaluation³ » ?

- c'est d'abord, pour l'établissement et ceux qui y œuvrent, qu'il s'agisse d'une évaluation externe ou d'une auto-évaluation par laquelle les acteurs de l'établissement se chargent eux-mêmes de tout ou partie de la démarche, une occasion unique de sortir de la solitude de l'activité individuelle ou de l'action collective mais limitée à un seul collège ou un seul lycée par les voies de l'analyse et de la comparaison avec d'autres ;
- c'est ensuite, pour les responsables⁴ du système, une façon aussi irremplaçable de connaître dans la perspective de l'action les réalités sur lesquelles les décisions qu'ils prennent vont porter : le pilotage du système s'effectue moins souvent que par le passé par les normes, qui appelaient une simple vérification de conformité, plus souvent à la fois par les résultats, ce qui est parfois récusé comme réducteur, et par les valeurs, ce qui est souvent décrit comme insaisissable. L'évaluation considère que tout cela est en tension et que c'est précisément sur ces tensions qu'un regard objectif est nécessaire et bénéfique, parce que c'est complexe ;
- c'est enfin, pour les familles et les usagers de l'école, un moyen pour trouver ou retrouver une confiance dans le service public sans laquelle il ne peut rien, même avec les meilleures intentions du monde⁵ : si la métaphore de la « boîte noire » a souvent cours entre pédagogues, avec plus ou moins de zéloteurs, pour parler de la salle de classe, rien ne justifie que l'établissement en soit une, aussi bien sur les acquisitions et les résultats de ses élèves, que sur leur devenir scolaire et professionnel à l'issue de l'école ou que sur la qualité de la vie qu'ils y mènent au quotidien.

³ Evaluation : jugement porté sur l'activité d'un agent ou d'une organisation, en considération de ses résultats, proposant des explications du niveau d'atteinte de ces résultats dans le but de permettre à l'agent ou à l'organisation de les améliorer. On trouvera en annexe 1 un certain nombre de définitions des termes utilisés dans le présent rapport.

⁴ S'il est clair que le présent rapport ne considère pas l'évaluation des établissements comme un élément de stratégie pour diminuer la dépense publique, il est en revanche certain que l'objectif d'amélioration des résultats et du fonctionnement du système par l'évaluation devrait en toute logique conduire à dépenser mieux chaque fois que possible.

⁵ Il n'est pas rare de constater un écart entre la qualité objectivement bonne d'un établissement et l'idée que s'en fait, fondée sur des rumeurs, telle ou telle partie de la population qui a fait le choix de l'éviter.

1. L'évaluation des établissements publics locaux d'enseignement : quels enjeux ?

1.1. Pourquoi s'intéresser à l'« évaluation » des lycées et collèges en France en 2004 ?

1.1.1. *Une interrogation sur une exception française : la faible attention portée à la question de l'évaluation des établissements*

▪ 1.1.1.1. Dans la plupart des pays, autonomie et évaluation vont de pair

Une récente recension par Eurydice de l'« état du dossier »⁶ dans une vingtaine de pays européens fait apparaître qu'en beaucoup de ces pays, la question de l'évaluation des établissements scolaires a été l'objet d'une attention plus constante et le lieu de débats plus intenses qu'en France au cours des dernières années, l'Angleterre étant sans doute le cas le plus connu : il serait inexact d'affirmer que la plupart des pays sont « plus avancés » que la France en ce domaine (certains, en effet, comme les pays scandinaves, sont engagés vers une décentralisation en direction des municipalités qui n'a pas jusqu'ici toujours trouvé son contrepoids en termes d'évaluation), mais il semblerait que la France fasse partie de ces pays où l'autonomie des établissements, quels que soient les textes qui la définissent, n'est pas une réalité encore assez active pour que le thème de l'évaluation des établissements ait été jusqu'ici perçu comme un enjeu majeur des politiques de l'école.

En bien des pays, en effet, des débats qui ne se sont pas toujours limités à la sphère des spécialistes ou des responsables politiques ont conduit à des décisions ayant tenté de répondre à des questions simples comme : évaluer l'établissement, pour quoi faire ? évaluer l'établissement, pour qui ? évaluer l'établissement, qui peut en être chargé, et avec quelle déontologie ?

Plus remarquable est le fait que le débat se soit porté en bien des pays sur des questions très concrètes, qui ont marqué une volonté de clarification du champ, comme celle de la périodicité requise pour les évaluations d'établissements (6 ans aux Pays-Bas ou au Royaume-

⁶ http://www.eurydice.org/Documents/Evaluation/terminology_fr.html

Uni, 7 ans en Ecosse avec la précision que chaque élève au cours de sa scolarité obligatoire, verrait l'établissement où il étudie évalué deux fois), celle des protocoles qui doivent être utilisés pour cette évaluation (obligations par exemple de rendre compte aux « évalués », voire de recueillir l'avis des évalués sur l'évaluation dont ils ont été l'objet), celle de la définition des instances adéquates d'évaluation et de leur ouverture à d'autres qu'aux spécialistes (recours au Royaume-Uni au « layman », ou « profane ») ou celle de la prise en compte du point de vue des usagers (interrogations diverses portant sur le degré de satisfaction des usagers et des agents de telle école), ou encore celle de l'accès des usagers aux rapports d'évaluation (sur internet, par exemple, sur décision de l'instance évaluatrice, et sans visa des autorités administratives).

En France, en revanche, ces questions ont peu intéressé le public, en apparence, au cours des années récentes, et les responsables administratifs et politiques, aux différents niveaux géographiques, n'en ont pas fait de leur côté⁷ un point important de leur action, de même qu'ils ont rarement fondé leurs décisions sur les résultats préalables d'évaluations.

L'impression dominante est sans doute, même si elle est peu exprimée en ces termes, que l'existence d'un système d'éducation centralisé pour l'essentiel de ses fonctions, avec des cursus et des programmes d'enseignement nationaux, des examens eux aussi à définition nationale, une « inspection » individuelle des personnels enseignants (phénomène que connaissent très peu de pays), deux « inspections générales »⁸ assurant régulièrement ces missions « de terrain » justifie la croyance que les établissements français, ou tout au moins leurs fonctions les plus stratégiques, comme celle d'enseignement, sont suffisamment « contrôlés » pour qu'il ne soit pas urgent de se poser des questions relatives à leur « évaluation ».

▪ **1.1.1.2. Sur la difficulté particulière à interroger l'école de France sur ses résultats**

L'explication est toutefois insuffisante : dans la plupart des pays, en effet, qui, dans le discours au moins, comme c'est le cas en France, mettent assez récemment en avant l'idée d'autonomie des établissements, comme dans ceux où cette autonomie est ancienne et fondamentale, le thème de l'évaluation des établissements est important. En bien des cas, l'existence d'une inspection générale elle-même (la France est la seule à en avoir deux), loin d'être un obstacle au développement de l'évaluation des établissements au nom du traditionnel « contrôle de conformité », en est souvent le fer de lance, et cela d'autant plus qu'elle est de fondation récente, comme en République tchèque. L'existence enfin de programmes et d'examens nationaux est par ailleurs plutôt ressentie par bien des pays qui en

⁷ A l'exception de quelques académies, à certaines périodes tout au moins.

⁸ Inspection générale de l'administration de l'éducation nationale et de la recherche et inspection générale de l'éducation nationale (voir plus loin le point sur leurs activités en matière d'évaluation d'établissements).

sont dépourvus, et les Etats-Unis sont bien sûr l'un des plus souvent cités, comme une facilitation à la mise en place de procédures satisfaisantes d'évaluation, auxquelles ces programmes fournissent un référent.

Si l'explication est à chercher ailleurs, peut-être est-il intéressant de remarquer que dans la plupart des pays l'évaluation de l'établissement, même si elle ne se limite pas à cela, prend d'abord en compte les « résultats » des élèves, tels qu'ils sont disponibles. Peut-être faut-il dès lors se demander si le peu d'attention politique portée en France à la question de l'évaluation des établissements n'est pas à relier à la difficulté d'interroger l'école de France sur ses « résultats ».

Simple hypothèse explicatrice à ce stade : étant donné que l'école de France est l'une des rares à avoir barre aussi bien sur la définition des formations, leur mise en œuvre, la collation des grades, et l'organisation des examens qui y conduisent⁹, toute interrogation sur ses résultats, au delà de taux de succès à des examens qu'elle organise d'ailleurs elle-même, apparaît comme relativement « en boucle » (Claude Pair, dans le rapport précité) et peu significative. Ce serait donc peut-être moins en raison de quelque résistance idéologique à ce qu'on a appelé la culture du résultat, et moins en raison de l'importance de son ancienne tradition du « contrôle » que pour des motifs à trouver dans sa structure même que l'école de France éprouve plus encore que d'autres difficultés et réticences à s'ouvrir à des démarches d'évaluation.

En beaucoup de pays, par ailleurs, l'évaluation de l'établissement va davantage de soi qu'en France, soit que parents et citoyens y demandent plus fortement l'information qui en résulte, soit que, même si des « objectifs », ou « curricula » nationaux existent, certaines décisions relatives aux contenus à enseigner restent à prendre au niveau de l'établissement lui-même et sont donc matière à une évaluation essentielle (les protocoles commencent souvent par interroger l'établissement sur « ce qu'il enseigne », question qui n'a guère de sens en France dans le cas général), soit encore que l'établissement se prête spontanément à une évaluation externe par son affiliation à un examen « externe » (concours d'entrée à l'université, examen proposé par une instance indépendante).

Comparaison n'est, pas plus en matière d'éducation qu'en aucune autre, raison ni leçon : elle nous permet seulement de repérer qu'il ne semble possible en aucun pays de parler d'autonomie d'établissement sans que cette autonomie s'accompagne d'une politique active et structurée d'évaluation des établissements et que les obstacles à une telle démarche, qui paraît en la plupart des cas aller de soi, pourraient bien, en France plus qu'ailleurs, relever de motifs

⁹ On pourrait de la même façon relever que l'administration de l'éducation nationale, en France, par ses services déconcentrés, est non seulement financeur, mais aussi gestionnaire, ce qui ne rend pas non plus nécessairement attentif à la question des résultats.

structurels profonds. En avoir conscience peut permettre, le moment venu, de tenter de les dépasser.

1.1.2. L'évaluation des établissements : un cadre institutionnel qui devrait s'imposer

▪ 1.1.2.1. En raison d'obligations légales, présentes et à venir

Le cadre législatif est pourtant clair sur le sujet, puisque la loi d'orientation sur l'éducation de 1989 mentionne en son article 25, au chapitre « évaluation », « le rapport annuel des établissements publics locaux d'enseignement qui rend compte, notamment, de la mise en œuvre et des résultats du projet d'établissement¹⁰ ».

Le décret du 30 août 1985, qui avait défini l'EPL, était déjà allé loin sur une voie qui aurait dû être à l'origine du développement dans les établissements de véritables procédures d'auto-évaluation, puisqu'il dispose en son article 16 que « le Conseil d'administration établit chaque année un rapport sur le fonctionnement pédagogique de l'établissement, qui rend compte de la mise en œuvre du projet d'établissement, des objectifs à atteindre et des résultats obtenus ».

Au-delà, le travail réglementaire de même que les directives ministérielles données sous forme de circulaire relativement à des thèmes qui ont connu depuis le début des années quatre-vingt la faveur des responsables comme le « projet d'établissement » rappellent l'importance de l'évaluation de l'établissement : c'est, plutôt, d'ailleurs, sous la forme d'une évaluation interne (circulaire du 17 mai 1990 sur le projet d'établissement) qu'en se préoccupant des rapports entre cette évaluation interne et les formes de contrôle ou d'évaluation externe déjà à l'œuvre dans le système (notamment sous la responsabilité des inspections générales).

Les diagnostics précédant les lettres de mission que les autorités académiques adressent désormais aux chefs d'établissements (aux termes du décret n° 2001-117 du 11-12-2001, art.21) font aussi partie de ce paysage qui fait obligation aux responsables, nationaux, locaux ou académiques, de se préoccuper à divers titres de l'évaluation des établissements.

En outre, la mise en place généralisée de la LOLF à partir de 2006 fixe une obligation de résultats et non plus seulement de moyens à l'ensemble du système : même si le législateur définira naturellement ses programmes et ses indicateurs de façon agrégée et nationale, les acteurs locaux seront dans l'obligation de les décliner chaque fois que possible au niveau où

¹⁰ A noter qu'il s'agit plutôt d'auto-évaluation.

se construit l'action, notamment à celui de l'établissement : la question de l'évaluation est bien entendu posée aussi dans ce cadre.

▪ **1.1.2.2. En raison des évolutions institutionnelles et notamment du développement de la responsabilité des établissements impliquée par l'accroissement de leur autonomie**

Au-delà de ces textes spécifiques, on peut considérer que les trois évolutions institutionnelles qui marquent l'évolution de l'organisation des pouvoirs au sein de l'éducation nationale plaident pour la mise en place de procédures d'évaluation des établissements :

1. c'est vrai de la « déconcentration », qui en donnant plus d'importance à l'échelon académique, permet une gestion plus proche des besoins locaux et, du même coup, crée pour les autorités déconcentrée (recteurs, IA-DSDEN) un besoin véritable d'évaluer les établissements, pour des motifs de gestion aussi bien que de pilotage, en particulier dans le cadre de la volonté de ces autorités de mieux être « au service » des établissements relevée lors de la conférence de clôture du colloque de Clermont-Ferrand¹¹ ;
2. c'est vrai de la « décentralisation » en direction des collectivités territoriales : partenaires de l'Etat avec un certain nombre de compétences déterminées par la Loi, ces collectivités peuvent elles-mêmes éprouver le besoin de connaître l'efficacité des ressources qu'elles mettent à la disposition des établissements (par exemple pour savoir si les efforts consentis en matière de technologies de l'information se révèlent efficaces). L'Etat aussi, de son côté, préoccupé de la bonne marche des collèges et lycées, mais aussi du caractère équitable de leur traitement par les collectivités territoriales, peut souhaiter connaître les effets des interventions¹² de ces collectivités sur ces établissements ;
3. c'est plus vrai encore de la décentralisation dite « fonctionnelle » en faveur des établissements eux-mêmes et de leur autonomie : on a vu que si dans la plupart des pays la décentralisation n'appelait pas nécessairement le développement de pratiques d'évaluation, il n'en allait pas de même de l'autonomie des établissements qui allait presque toujours de pair avec lui. De fait le développement de l'autonomie a pour contrepartie l'application d'un principe de responsabilité et la mise en place de formes nouvelles de régulation, substituant à des degrés divers selon les champs négociation et relations contractuelles avec les autorités à la simple conformation à des normes uniformes. Au contrôle de conformité, qui ne disparaît pas nécessairement pour autant, mais qui doit être redéfini, se substituent aussi des relations qui appellent la mise en place de modalités d'« évaluation », telle qu'elle a été définie plus haut.

¹¹ Pratiques d'inspection, pratiques d'évaluation, colloque de Clermont-Ferrand, 5 et 6 décembre 2001, coord. J-P. LAUBY, CRDP d'Auvergne 2003.

¹² Parfois au-delà de ce qui est obligatoire, y compris en payant des enseignants d'Etat pour des interventions pédagogiques dans leur établissement au-delà de leurs obligations statutaires.

1.1.3. L'évaluation des établissements : une relative faiblesse au sein de l'appareil d'évaluation

▪ 1.1.3.1. ...que plusieurs rapports ont fustigée

Les rapports d'experts¹³ n'ont pas manqué au cours des années récentes qui fustigeaient le faible développement du niveau intermédiaire de l'évaluation au sein du système français d'évaluation : Claude PAIR (rapport établi à la demande du Haut Conseil de l'Évaluation de l'École en octobre 2001) parle par exemple de ce niveau de « méso-évaluation », beaucoup moins développé à ses yeux que le niveau « macro » (évaluation du système dans son entier) ou le niveau « micro » (évaluations diagnostiques ou « sommatives » des élèves à tel ou tel moment de la scolarité). Même si les difficultés rencontrées par le développement des évaluations d'établissements ne sont peut-être pas sans lien avec certaines faiblesses plus générales de l'appareil français d'évaluation du système à ses différents niveaux, la remarque n'en a que plus de force et ne peut manquer d'interroger.

▪ 1.1.3.2. et à laquelle les académies ont souvent tenté de répondre par différents essais d'évaluation des établissements

Les injonctions légales, les évolutions des compétences, les rapports d'experts seraient toutefois peu de choses si les établissements eux-mêmes, et leurs responsables, ainsi que les autorités éducatives locales, n'éprouvaient pas aussi les effets négatifs de cette carence.

C'est ainsi que dans un rapport au ministre remis en avril 2001 les inspections générales¹⁴ estimaient que les chefs d'établissements, « confrontés dans le domaine pédagogique à la succession des inspecteurs pédagogiques... expriment le regret que ces "regards extérieurs" n'assurent pas davantage un rôle d'évaluation plus générale du système... » sur des questions comme la « constitution des divisions », les « difficultés propres à l'hétérogénéité »...

Le besoin créant la fonction, on a assisté, dans beaucoup d'académies, après ce qui a été mis en place dans l'académie de Lille dès 1991, à diverses tentatives d'évaluations d'établissements à l'initiative des recteurs ou des IA-DSDEN, avec des objectifs différenciés et des modes opératoires disparates. Curieusement, cet apparent besoin du « terrain » n'a pas été relayé par l'attention des responsables nationaux, à la fois pour étudier ce besoin et voir comment le niveau national pouvait tenter d'aider méthodologiquement les académies dans leurs essais, pour gagner du temps, et veiller à ce que ce besoin apparemment répandu ne reste pas sans réponse dans celles qui ne s'en saisissaient pas effectivement. L'échelon central n'a même pas tenté de mettre en harmonie ces évaluations d'établissements à initiative

¹³ Rapport DASTE sur l'évaluation des lycées, réalisé aussi à la demande du HCEE.

¹⁴ Gérard SAURAT et Philippe DUVAL, « La fonction conseil aux EPLE ».

académique avec celles qui sont traditionnellement conduites par les corps nationaux des inspections générales.

Un problème véritable de cohérence nationale s'est donc créé, auquel il est maintenant urgent de répondre, sauf à considérer qu'évaluer un établissement dans l'académie de Bordeaux n'a rien de commun avec l'activité parallèle dans celle de Rennes, et que les inspections générales peuvent développer dans les académies des activités devenues concurrentes de celles développées par les recteurs.

1.1.4. L'évaluation des établissements : une nécessité renforcée par les difficultés que rencontre le système

- **1.1.4.1. Les difficultés contemporaines du système éducatif suscitent de fait des « évaluations » subjectives que l'institution ne saurait laisser se développer sans intervenir**

A la suite de la publication par un hebdomadaire du palmarès « brut » des lycées, établi sur les seuls taux de succès au baccalauréat, le ministère de l'éducation avait publié de son côté son propre classement tentant de faire droit non seulement aux résultats, mais aussi aux efforts des établissements pour faire face à leur mission dans la diversité des contextes, notamment sociaux, de leur action. Si l'on peut considérer qu'il est regrettable en l'espèce que sur la question cruciale des résultats des établissements l'institution n'ait pu en définitive apparaître que comme réactive à des enquêtes journalistiques, il est évident que les écarts qui se sont objectivement creusés entre les établissements, et notamment les collèges, depuis une décennie au moins, nourrissent en permanence les interrogations des usagers et les rumeurs, et que l'institution ne peut défendre le service public et ses usagers qu'en prenant les devants pour analyser objectivement la réalité des établissements et anticiper sur les ravages de l'opinion.

Il n'est pas question ici d'avoir la naïveté de prétendre que la « bonne » information chasse par son seul surgissement la mauvaise, mais elle donne aux acteurs les moyens de construire sur le long terme des stratégies de communication qui peuvent influencer sur les choix. Que le système français reste fidèle à une pratique aujourd'hui majoritaire de « carte scolaire » ne laissant pas aux familles le choix de leur établissement ou qu'il adopte d'autres règles, dans l'un et l'autre cas l'attention des familles sur cette question¹⁵, rend nécessaire une reprise de parole de l'institution sur la qualité de ses établissements, fondée objectivement et opposable sur des critères reconnus en dehors du cercle des seuls spécialistes.

¹⁵ Les allers-retours pour un même élève au cours de sa scolarité entre l'enseignement public et l'enseignement privé ou les observations issues du « grand débat sur l'école » mettent également en évidence cette question.

- **1.1.4.2. L'évaluation peut permettre à un établissement de conquérir une autonomie le mettant en position de mieux répondre aux problèmes locaux de l'éducation**

Faire de l'évaluation une démarche qui doit nécessairement accompagner, voire contrebalancer l'autonomie de l'établissement comme nous le proposons plus haut est à la fois partiel et maladroit : le système éducatif rencontre en effet des turbulences telles qu'il ne peut y faire face par des réponses uniformes ; les établissements sont mis à contribution plus fortement qu'autrefois parce que les problèmes sont là, et que la réponse n'est fournie ni par une norme qu'il suffirait d'appliquer, ni par la foi en quelque réforme miracle à laquelle il suffirait de croire, ni par l'appel à des apports externes en termes de moyens supplémentaires ou de « formations » rédemptrices. Dans ce contexte, la pression qui s'exerce sur l'établissement est telle qu'il ne peut résoudre les questions qui lui sont posées que par lui-même, en fondant son action sur la connaissance la plus objective, la plus précise et la plus partagée de son contexte d'action, de ses atouts, de ses défauts et de ses possibles. L'évaluation, et si possible l'auto-évaluation, est sans doute pour l'établissement, au moment où il lui est beaucoup demandé, une étape indispensable pour prendre ou reprendre du pouvoir sur sa destinée : loin de devoir être opposée au « projet » d'établissement, cette démarche est en fait la seule qui soit susceptible de lui donner une portée véritable.

Ce qui est irremplaçable avec la démarche d'évaluation, c'est qu'à la fois elle est totalement immergée dans le concret de l'établissement et qu'elle l'invite constamment à échapper à l'autisme parce qu'elle cherche en permanence des références et des termes de comparaison¹⁶ : elle est d'abord acte de foi dans l'établissement et devrait lui permettre, après la remise en cause qu'elle implique, d'en sortir plus assuré sur son action. Elle est en cela un enjeu stratégique pour le système dans son entier, qui, sans un outil d'autorégulation dont la définition fine pose des questions politiques que nous tenterons de cerner plus loin, risque de ne plus pouvoir répondre, dans la diversité des situations, de la mise en œuvre des principes qui sont les siens.

1.2. Les difficultés de l'exercice

Que le système éducatif français soit en train ou non de rattraper son retard et de combler ses lacunes en matière d'évaluation d'établissements, il appartiendra à la deuxième partie du présent rapport de l'établir : on ne peut toutefois ignorer à ce stade la prégnance de tous les motifs contraires, de tous les obstacles qu'a rencontrés et que rencontrera sur sa route la démarche d'évaluation des établissements scolaires. Ils sont de plusieurs ordres et niveaux :

¹⁶ L'expérience du terrain apprend du reste que chaque établissement est inséparable d'un espace plus vaste, celui du territoire en fonction duquel se définissent des logiques éducatives, mises en œuvre par plusieurs établissements : les résultats finaux devraient être approchés à ce niveau qui est sans doute le plus pertinent pour une évaluation.

1.2.1. Les difficultés venant du manque de visibilité des établissements en tant que tels

▪ 1.2.1.1. Une relative incertitude sur l'« importance » effective de l'établissement pour le succès des élèves

Des sociologues¹⁷ ont récemment remarqué que les parents ont en général spontanément l'idée que l'établissement où sont scolarisés leurs enfants est moins important que la classe qu'ils fréquentent ou la qualité des professeurs qui y enseignent : l'attention des évaluateurs devrait donc se porter davantage sur cela plutôt que sur l'introuvable « établissement ». Nombre d'études (sous la plume de Denis MEURET, par exemple) font toutefois apparaître que si l'effet-établissement n'est pas toujours le plus apparent, il devient tout à coup premier aux extrêmes, c'est-à-dire à la fois dans le cas d'établissements très bons ou très déficients, ainsi que pour les élèves en difficulté, qui trouveront ou ne trouveront pas à ce niveau de l'établissement une organisation des réponses à leurs besoins.

Ce seul argument, s'il devait nous donner un accès au cœur des problèmes de l'école, c'est-à-dire sa production excessive d'échec scolaire et social, suffirait bien sûr à justifier que l'institution se donne les moyens d'évaluer précisément ce qui « marche » ici et qui pourrait « marcher » là.

▪ 1.2.1.2. Des doutes sur l'intérêt d'évaluer une entité dont l'autonomie apparaît encore souvent comme marginale

D'autres doutes peuvent être émis sur la pertinence d'évaluer une institution peu autonome, qui doivent être cités et soumis à la réflexion critique :

➤ *L'établissement français : une consistance incertaine*

L'établissement français d'aujourd'hui n'échappe pas toujours à la description classique de la coexistence d'une bureaucratie, qui répercute et édicte des « normes », et d'un corps de « professionnels », vivant au quotidien dans un régime de très grande autonomie. Sans qu'aucun jugement de valeur soit ici porté¹⁸, un établissement est-il spontanément autre chose ? L'objection est aisée : s'il est, le cas échéant, autre chose que cette juxtaposition, n'est-ce pas précisément le résultat d'une construction ?

¹⁷ in *Education et Sociétés*, n° 9, 2002/1, « A quoi sert la sociologie de l'éducation ? ».

¹⁸ On oublie parfois que la répétition justifiée par l'institution de l'importance du travail collectif des maîtres ne doit pas cacher que d'excellents professeurs, au moins pour une partie de leurs tâches, peuvent exercer sans collaboration autre que formelle avec leurs collègues.

➤ *Une confiance ambiguë en son autonomie*

En même temps qu'était diffusé un discours sur l'autonomie de l'établissement et qu'effectivement des marges tangibles de cette autonomie étaient mises en place par l'institution, de nombreux dispositifs pédagogiques ayant la plupart du temps comme facteur commun le développement de l'autonomie de l'élève et le travail entre les disciplines (Travaux personnels encadrés au lycée, Projet personnel à caractère professionnel au lycée professionnel, Itinéraires de découverte au collège) sont venus dicter aux établissements des règles pédagogiques qui n'étaient pas en tous points conformes à la vieille tradition de la liberté du maître dans le choix des méthodes. L'administration centrale, mais aussi l'administration rectorale, ont voulu être plus interventionnistes en matière pédagogique au moment même où elles prônaient l'autonomie ; l'évaluation d'établissement risque alors de devenir évaluation de dispositifs, évaluation du succès d'une politique, laissant alors l'« établissement » échapper.

On peut toutefois, sans refuser à l'institution de jouer son rôle en assurant la promotion de méthodes susceptibles de mieux répondre aux objectifs de formation des élèves, estimer que l'évaluation de l'établissement est alors d'autant plus nécessaire qu'il faudra mesurer l'effet « en contexte » des orientations nationales ou académiques, et y compris de leur éventuelle non-observation. En des matières complexes qui méritent cette approche, l'évaluation permettra de dépasser le contrôle.

➤ *Une autonomie de fait limitée*

On peut aussi estimer que même en dehors de ces exemples, l'autonomie de l'établissement français est inconsistante, puisqu'il n'a le choix ni des programmes, ni des maîtres, ni des horaires, ni la plupart du temps de ses élèves, que le choix même des formations qu'il offre est largement décidé par les autorités académiques et que l'établissement n'a même pas de pouvoir de négociation avec les autres établissements de l'éducation nationale qui constituent son « amont », son « aval », ou son éventuelle « concurrence ». On a pu faire observer¹⁹ que les administrations gestionnaires, en changeant leurs règles d'une année sur l'autre, par exemple sur le nombre « seuil » d'élèves nécessaires à ouvrir une classe, réduisent encore facilement une autonomie déjà restreinte.

¹⁹ Cf. Marie-Martine BOISSINOT, *Pilotage pédagogique et autonomie des établissements scolaires*, CRDP d'Auvergne, 2003, 160 p.

On peut toutefois objecter à ce qui prétendrait être un argument pour démontrer la vanité de l'évaluation d'établissement, que précisément la réponse à la question de savoir si l'établissement se saisit ou non de son autonomie²⁰ pour obtenir des résultats favorables à ses élèves est nécessaire à tous, aux acteurs pour se comparer, aux décideurs externes pour piloter sans risque le déploiement de l'autonomie des établissements. On peut surtout surenchérir en rappelant que l'autonomie récemment consentie par les autorités aux établissements n'est en tout état de cause pas grand chose à côté de l'autonomie pédagogique dont jouissent traditionnellement les maîtres dans leur activité, et qu'il existe donc depuis longtemps sinon une autonomie de l'établissement, au moins des autonomies dans l'établissement qui se conjuguent en une autonomie plus large. Cette autonomie nécessite ses contrepoids en termes d'évaluation de résultats et d'effets.

1.2.2. Les difficultés venant des réticences classiques aux démarches d'évaluation dans le champ de l'éducation

- **1.2.2.1. Un faisceau ancien de motifs qui mettent en cause l'évaluation d'une école à partir de ses « résultats »**

L'école est presque partout dans le monde et l'histoire ce lieu étrange où une large partie du temps est utilisée à évaluer des élèves (soit de façon ordinaire pour suivre les apprentissages, soit à l'occasion de tests ou de sondages portant le bilan d'une période ou d'une situation individuelle, soit pour se préparer ou subir les épreuves d'un examen ou d'un concours organisé ou non par les autorités de l'école) et où pourtant l'évaluation est la plupart du temps l'objet d'une réticence spontanée, si elle entend s'exercer à l'égard de l'école elle-même.

Plusieurs motifs à cela, que la littérature pédagogique a depuis longtemps décrits²¹ :

➤ *L'évaluateur évalué*

La confusion même, qui brouille les pistes, entre l'évaluation de l'activité et des résultats d'un service public et l'activité ordinaire d'évaluation des élèves qui fait partie intégrante de l'enseignement : il est parfois difficile d'imaginer que les deux « évaluations » ont bien sûr en

²⁰ En particulier au plan des fonctions administratives, au sens large, qui sont plutôt caractérisées par l'uniformité que par l'innovation. Quant aux moyens d'enseignement dont bénéficie un établissement, ils lui apportent une marge d'autonomie facile à appréhender et qui est loin d'être toujours utilisée : renforcer par exemple l'effectif des classes pour mettre en place avec les heures ainsi récupérées des enseignements en petit groupe (groupes de soutien) ou optionnels.

²¹ On peut se reporter par exemple à la note de synthèse de Jean-Marie DE KETELE, in « Revue française de pédagogie », n°103 ; avril -mai -juin 1993 : *L'évaluation conjugée en paradigmes* ; on peut lire aussi François DUBET « Pourquoi ne croit-on pas les sociologues ? » in « Education et Sociétés », *op. cit.*.

commun certaines attitudes et l'objectif de formuler un « jugement », mais dans des scénarios radicalement différents ; le miroitement autour du mot d'« évaluation » se résoudra s'il apparaît clairement que précisément l'activité d'évaluation des élèves (fréquence, typologie, modalités chiffrées ou littérales, déontologie de correction et de délais, critérium, échelle, dispersion, corrélations appropriées, conséquences internes et externes...) fait partie de ce qui doit être « évalué » dans un ensemble pédagogique. Ce n'est pas toujours le cas.

➤ *L'hétérogénéité de l'évaluation des élèves et de l'évaluation du système*

Les enseignants, connaissant la complexité de l'activité d'évaluation des élèves, la part de non-dit (cf. la note dite « d'encouragement » attribuée à un élève avec la conscience partagée entre évaluateur et évalué qu'elle ne reflète pas une performance²², cf. aussi l'échelle de notation restreinte par tradition de telle discipline, qui n'est pas non plus censée refléter une performance en dehors d'un cadre de connivence initiée, etc.), d'implicite, d'exercice plus ou moins transparent d'un pouvoir, ne peuvent qu'être sceptiques *a priori* sur une évaluation d'établissement qui, ponctuelle et ne s'inscrivant pas dans un cadre pédagogique, ne pourra pas rester aussi souvent dans l'implicite ou la connivence. Des doutes s'expriment dès qu'il est question de donner aux résultats des élèves des significations étrangères aux scénarios pédagogiques divers où chaque maître les a conçus et en particulier d'en tirer des enseignements externes, par exemple dans le cadre d'une évaluation d'établissement. Ces objections sont importantes et méritent toute notre attention, pour chercher les voies et moyens pour que les notes « indigènes » d'un établissement (par exemple les notes attribuées en contrôle continu au brevet des collèges) soient lisibles au niveau de cet établissement.

➤ *La crainte de l'évaluation appauvrissante*

La crainte réciproque existe : si un système éducatif veut précisément se doter d'éléments de mesure des résultats des élèves comparables dans l'espace et le temps, il risque de privilégier des modalités d'évaluation simples et appauvrissantes, laissant d'une part échapper des dimensions essentielles des programmes et objectifs de l'école, rétroagissant d'autre part de façon négative sur les enseignements : l'appauvrissement de l'enseignement des curricula fustigé par beaucoup d'universitaires américains à la suite de l'introduction de tests nationaux, ne manque en effet pas d'inquiéter et d'inviter les évaluateurs français à se prémunir contre ce type de dérive : l'enseignement vise certes l'acquisition de compétences et de connaissances qui peuvent être vérifiées facilement, mais aussi l'acquisition d'attitudes intellectuelles et morales, ainsi que la construction d'une culture qui impliquent des précautions fines.

²² C'est souvent le cas des notes attribuées en contrôle continu au brevet des collèges, à des élèves qu'il ne faudrait pas « désespérer ».

➤ *L'intériorisation des doutes sur les résultats scolaires*

D'une autre façon encore, les mêmes maîtres qui attribuent des notes (chacun plusieurs milliers chaque année) en permanence sont réticents à ce que des jugements externes se fondent sur les résultats des élèves car la circulation de certains savoirs scientifiques issus de la docimologie, de la sociologie ou de l'économie leur a fait savoir à quel point le « résultat », qui a besoin dans l'histoire des scolarités individuelles d'apparaître à un moment comme un absolu, est en fait relatif, non seulement dans les limites d'implicite que les maîtres croient dominer, mais bien au-delà d'une façon dont ils n'ont pas toujours conscience au quotidien²³.

➤ *Le spectre de la marchandisation*

Une crainte complémentaire peut venir de la perspective que l'évaluation d'établissement soit un cheval de Troie d'une entreprise de type « consumériste » en matière éducative : de fait, la recherche de transparence servira d'abord ceux qui sont le plus à même de se saisir de ses résultats pour développer des stratégies personnelles : l'outil d'évaluation, en théorie au service d'un développement de la démocratie, peut devenir facilement un instrument de plus au service des « intérêts » opaques. En dehors même de ces dérives, l'évaluation, qui veut pour un service public faire un peu le travail d'ajustement et de sélection dont se charge ailleurs le marché, n'est-elle pas l'introduction plus ou moins clandestine d'une logique de marché au cœur de l'activité d'éducation ? Il est pourtant facile de voir que l'argument peut être retourné et que le service public ne peut que gagner à se doter de modes d'observation de lui-même et de régulations de qualité.

▪ **1.2.2.2. Un contexte institutionnel qui n'a pas toujours favorisé la diffusion de démarches d'évaluation ou d'auto-évaluation d'établissements**

➤ *Des obligations peu respectées*

On peut tout d'abord s'étonner que l'administration de l'éducation nationale n'ait pas toujours favorisé comme elle l'aurait pu la démarche d'évaluation d'établissements. Si (cf. supra), des textes officiels invitaient en effet à développer ces démarches, on ne peut toutefois que constater que ceux par exemple qui en appelaient au développement de procédures d'auto-évaluation ou de diagnostic partagé interne aux établissements ont été peu appliqués, ou de

²³ Or les maîtres, parce que leur art s'exerce sur autrui, ne peuvent se permettre, sauf à ne plus pouvoir enseigner, d'adopter une posture trop critique à l'égard de quelques articles de foi (cf. François DUBET, op. cit.) parmi lesquels figure la foi dans l'objectivité, l'équité et l'utilité « morale » des notes.

façon superficielle (décret de 1985 sur l'EPLE, par exemple). Il était d'ailleurs peut-être étonnant dès l'origine qu'on invite les établissements à se lancer dans des opérations difficiles d'évaluation interne sans définir par ailleurs les modalités et objectifs d'évaluations externes, qui, au cours de la même période, on l'a vu, sont inscrites plus nettement dans le cahier de charges des inspections générales, voire de l'inspection pédagogique régionale. Dans la plupart des pays en effet qui souhaitent voir se développer des procédures d'auto-évaluation, ces procédures sont définies par référence, le cas échéant pour s'en distinguer, à des procédures publiques fortement cadrées d'évaluation externe.

➤ *Les effets pervers d'une mauvaise gestion de l'idée de projet d'établissement*

La gestion qui semble avoir trop souvent prévalu dans les établissements de la notion de projet d'établissement n'a pas non plus favorisé la démarche d'évaluation : si la notion de projet d'établissement n'était pas en son principe contraire à celle d'évaluation ni d'auto-évaluation, et qu'elle l'appelait même comme une phase indispensable, il faut bien convenir que cette phase a souvent fait les frais de l'improvisation fréquente de l'élaboration des projets, et que comme c'était sans doute la phase la plus difficile, les établissements, dont les acteurs étaient peu formés à la complexité de l'affaire et à qui l'administration demandait parfois de façon incantatoire de « se projeter », l'ont souvent escamotée. Faut-il aujourd'hui continuer à répéter indéfiniment l'article de foi sur le projet d'établissement rédempteur ou convenir que cette démarche a si souvent recouvert une logomachie sans efficacité qu'elle est devenue un obstacle à la qualité du regard que les établissements peuvent porter sur eux-mêmes ? A la complexité de la démarche d'évaluation d'un établissement, on substitue souvent l'évaluation de la mise en œuvre d'un projet : comme ce dernier lui-même ne résultait pas d'un regard construit avec méthode, on ne met pas de son côté toutes les chances d'agir sur ce qui est crucial.

➤ *L'ambiguïté des inspections individuelles*

Les comparaisons internationales nous montrent quelle quasi-exception française est l'inspection des personnels des établissements par des inspecteurs, en particulier s'agissant des personnels enseignants. Il n'appartient pas à ce rapport de se prononcer sur des pratiques qui sont peut-être indispensables au fonctionnement systémique de l'école française, mais il convient sans doute de se demander comment la perspective des « inspections individuelles » peut cohabiter avec celle d'évaluation d'établissements, surtout si ce sont les mêmes acteurs qui s'en chargent. Il y a au moins des clarifications à apporter, et peut-être à réfléchir à la question de savoir si en dehors de cas particuliers, le vrai jeu de l'autonomie des établissements ne consisterait pas à évaluer ses activités individuelles (dont celle de ses maîtres) uniquement en référence aux activités d'ensemble (le travail de tel professeur de mathématiques au sein de l'enseignement des mathématiques dans tel collège, puis la fonction

de l'enseignement des mathématiques parmi les apprentissages des élèves du collège), puis ces activités d'ensemble uniquement en rapport avec leurs résultats et effets ;

➤ *L'absence de logique d'établissements dans le premier degré*

Une autre difficulté provient aussi des contradictions apparentes de l'institution qui considère que dans le second degré, dès le collège, l'établissement est la structure porteuse de l'adaptation de l'école nationale aux réalités locales, alors que le premier degré reste étranger à cette logique, avec des écoles qui demeurent de simples services déconcentrés : comment expliquer que vérité au-delà de la sixième soit inexistence en deçà ? Dans la plupart des pays, s'est imposée une unicité et une cohérence des modes d'évaluation au moins pour l'ensemble de l'enseignement obligatoire, parfois au delà, qui fait défaut dans le cas français : autant on pourrait imaginer assez facilement que deux évaluations bien conçues d'un collège et d'un lycée qu'il fournit en élèves apportent des analyses directement complémentaires, dues à l'identité de procédure, autant le dialogue indispensable avec le premier degré est appelé à rester délicat en raison-même des différences dans les modes de gestion et d'évaluation. Pour importante, cette objection ne semble toutefois pas rédhibitoire, car des travaux coordonnés d'évaluation sont d'ores et déjà souvent rendus possibles par l'action des inspecteurs de l'éducation nationale du premier degré ou la mise au point d'indicateurs d'évaluation dans des projets d'école intégrant en particulier les résultats de l'évaluation de 6^{ème} ;

➤ *Les faiblesses des outils institutionnels d'évaluation*

Il n'est pas certain non plus qu'en leur état actuel les outils d'évaluation fournis aux établissements par l'administration soient très incitatifs et facilitent la tâche de ceux, responsables et acteurs locaux, mais aussi inspecteurs, qui veulent « évaluer » un établissement : on reviendra sur ces questions de façon plus précise dans le présent rapport, mais il est certain que les indicateurs fournis au plan national, même dans les cas de plus en plus fréquents où ils étaient « fiables », ont eu l'effet de faire croire indûment aux chefs d'établissement que là étaient les meilleurs éléments de connaissance à considérer pour agir dans un établissement. Or, et cette réflexion sera reprise plus tard, ces indicateurs avaient la caractéristique d'être tournés davantage vers le contexte social de l'établissement, avec une insistance marquée sur les catégories socioprofessionnelles des familles des élèves que vers les « compétences de départ » et les « résultats d'arrivée » des élèves : l'appareil informatif du ministère, conforme à la culture scolaire dominante en France (cf. 1.1.1.2. « Sur la difficulté particulière à interroger l'école de France sur ses résultats »), n'a pas toujours favorisé le développement de la culture d'évaluation dans les établissements scolaires comme cela eût été souhaitable.

1.3. Précautions pour que cet exercice ne demeure pas lettre morte : définir avec clarté les finalités et le cadre de l'évaluation des établissements

1.3.1. Appréhender l'établissement dans sa globalité

Le premier argument à énoncer de façon claire doit être celui-là. En effet il s'agit de marquer la différence par rapport aux évaluations (de personnes, comme les professeurs, ou de parties de services, comme celui de « restauration scolaire ») qui ont cours traditionnellement, en disant que l'évaluation d'un établissement le considère dans sa globalité, comme un système qui associe des fonctions très hétérogènes associées pour la gestion d'un service complexe.

L'unité de l'établissement est d'abord à considérer par rapport à l'habitude prise dans beaucoup d'établissements français de distinguer des fonctions d'administration, des fonctions d'enseignement et des fonctions d'éducation.

Cette distinction est tentante en ce qu'elle paraît correspondre à la différenciation entre les catégories de personnels : les personnels d'administration, techniques, ouvriers et de services d'une part, les personnels enseignants d'autre part, les personnels dits « de vie scolaire » d'autre part. Elle est toutefois un obstacle à la perception juste de ce qu'est l'établissement :

- qui précisément fait appel à des compétences et à des moyens divers pour les combiner de façon judicieuse au service des objectifs communs ;
- dont l'activité spécifique, surtout en ce qu'elle s'exerce en direction d'« élèves », ne doit pas être tronçonnée en fonctions spécifiques, mais cohérente et intégrée.

Pour ces motifs devra être rejetée toute forme d'évaluation que sous-tendrait une conception sectorisée de l'établissement : les concepts parfois rencontrés par exemple d'évaluation « pédagogique » ou « administrative » de l'établissement sont sans doute à éviter, et les personnes en charge de l'évaluation, interne ou externe, devront se garder d'aborder l'établissement autrement que comme une globalité, poursuivant des objectifs divers subordonnés aux apprentissages des élèves.

Mais la globalité de l'établissement n'est pas seulement à définir face aux divisions fonctionnelles dont il vient d'être question : au sein -même des activités d'enseignement, cohabitent souvent plusieurs entités facilement autonomes ou autonomistes, comme les différentes disciplines ou les différents niveaux de classes. Il est certain que l'évaluation de

l'établissement ne saurait se limiter à constater l'activité de chacune de ces autonomies sans s'interroger d'abord sur la façon dont elles coopèrent au bénéfice des apprentissages des élèves. Les liens entre les apprentissages des diverses disciplines, par exemple, font donc partie de ce que l'évaluation d'un établissement doit approcher.

L'objectif de saisir l'établissement dans sa globalité est aussi le seul qui permette d'insister sur toutes les surfaces d'échange entre cet établissement et les autres établissements scolaires, de l'amont, de l'aval ou de la simultanéité, et l'ensemble des partenaires associatifs, politiques, sociaux ou économiques.

1.3.2. Clarifier la démarche d'évaluation

On observe d'ordinaire qu'aucune évaluation ne réussit quand ceux qui en passent commande, ceux qui la réalisent et ceux qui en sont l'objet (en admettant qu'ils soient distincts, ce qui n'est pas le cas s'il s'agit d'auto-évaluation) ne sont pas totalement au clair quant au sens de la démarche engagée. Une démarche d'évaluation intervient dans un établissement dont les personnels sont évalués et notés, y compris pour certains d'entre eux par des évaluateurs externes dans le cadre d'« inspections » : l'évaluation de l'établissement devra se distinguer de ces procédures, notamment individuelles.

De la même façon, les choix, quels qu'ils soient, entre les différents types possibles d'évaluation²⁴, doivent être parfaitement indiqués et clairs pour tous: qui a été à l'initiative de l'évaluation ? pourquoi l'établissement a-t-il été choisi (parce que l'autorité qui a décidé de l'évaluation avait connaissance de problèmes spécifiques ? à titre d'échantillon ? parce que c'est son tour selon une périodicité affichée ?), quel est le statut des évaluateurs (sont-ils indépendants ? sont-ils une émanation de l'autorité hiérarchique ?), quel doit être le protocole de l'évaluation (est-il public ? général ou adapté à l'établissement ? intégrateur ou limité à certaines fonctions ? qui l'a élaboré ? en fonction de quels objectifs ? en référence à quels critères ?) ? quel doit être le degré d'association des acteurs de l'établissement (où se situe-t-on, entre une évaluation totalement externe, se limitant à interroger les acteurs locaux, et ce que serait une auto-évaluation, dont ces derniers pourraient se charger totalement ?) les évalués pourront-ils s'exprimer à propos de l'évaluation ? sur quoi l'évaluation doit-elle déboucher (un rapport d'analyse ? des recommandations ? un plan stratégique ? des changements initiés par l'autorité hiérarchique dans l'affectation des moyens ou la désignation des responsables ?), les résultats de l'évaluation seront-ils connus des acteurs professionnels ? des usagers ? doit-il y avoir un suivi postérieur à l'évaluation ?

²⁴ On en trouvera un rappel récent dans Sylvie TROSA, *L'évaluation des politiques publiques*, in Les notes de benchmarking international, Institut de l'entreprise, 2003.

Aucune de ces questions ne doit être laissée dans l'ombre et il appartient même à l'évaluateur, qui rencontre des personnes moins familières que lui de ces problèmes, de prendre les devants en les amenant lui-même.

1.3.3. Privilégier les résultats des élèves

S'intéresser à l'évaluation d'un établissement scolaire en n'évaluant les apprentissages des élèves que comme un élément parmi d'autres, pédagogiques (les programmes sont-ils enseignés ? tel dispositif pédagogique est-il en place ?) ou non (l'établissement est-il bien géré au plan financier ? le « climat » de l'établissement est-il bon ? les élèves se sentent-ils bien ?, etc.) a longtemps conduit et conduit encore à une méprise complète sur ce qu'est un établissement scolaire et sur la hiérarchisation de ses priorités.

Plusieurs points doivent être rappelés :

- Un établissement d'enseignement est un établissement au service des apprentissages de ses élèves ;
- Ces apprentissages sont d'ordre divers, incluant des apprentissages sociaux ;
- Les résultats de ces apprentissages sont partiellement lisibles dans les performances des élèves eux-mêmes, à des examens, dans leurs attitudes comportementales au sein de la société scolaire et au dehors, dans des études supérieures à celles de l'établissement considéré, ou encore en termes d'insertion professionnelle pour les établissements dont c'est l'un des débouchés.

« Privilégier les résultats des élèves » n'a donc en aucun cas ici le sens de ne considérer que les résultats aux examens, ni de ramener l'activité de l'établissement à une simple activité de transmission ou encore de ne vouloir évaluer que des connaissances vérifiables sous forme de tests : cette attitude, conçue comme il a été énoncé, consiste au contraire à prendre l'établissement pour ce qu'il est : un lieu de passage où se jouent les apprentissages.

Loin donc d'être inspiré par quelque utilitarisme réducteur ou de transformer les maîtres en exécutants, cette attitude consistant à interroger d'abord un établissement scolaire à partir des résultats de ses élèves (qui encore une fois ne se ramènent pas à ce qu'on appelle d'ordinaire « résultats scolaires ») consiste en fait à prendre au sérieux ce qui s'y passe d'une part, à laisser faire les acteurs d'autre part sans leur demander des comptes de façon systématique sur ce qu'ils font et la façon dont ils procèdent.

2. L'évaluation de l'EPLÉ aujourd'hui, un paysage marqué par son caractère disparate

2.1. Les expériences étrangères d'évaluation des établissements : acquis méthodologiques et problèmes ouverts

L'apparence est que la question de l'évaluation des établissements est perçue dans la grande majorité des pays comme une question stratégique : la réalité est que l'importance accordée à cette question est non seulement très variable d'un pays à l'autre, mais que la signification politique des évaluations d'établissements mises en place et les motifs qui président à cette mise en place divergent beaucoup entre différentes situations, en fonction de l'histoire, de la phase « centralisatrice » (Angleterre) ou « décentralisatrice » (pays scandinaves) que traversent les systèmes, du type de décentralisation dont il est le cas échéant question (vers les collectivités locales, vers les écoles elles-mêmes), de la part que prend la collectivité publique à compétence générale (la nation, le land...) à la définition des politiques scolaires (programmes et examens nationaux ou non), etc.

On s'intéressera successivement aux facteurs de développement des démarches d'évaluation, au positionnement des instances chargées des évaluations externes, aux méthodes et critères auxquels il est recouru, à la question du rapport entre l'évaluation et les usagers, ainsi qu'au développement progressif et ambigu des démarches d'auto-évaluation.

2.1.1. La diversité des facteurs à l'origine du développement de démarches d'évaluation

- **Le poids des évaluations internationales**

Les évaluations internationales ont joué un rôle dans des pays auparavant assez indifférents à ces procédures et ont déclenché un intérêt pour les évaluations des systèmes nationaux et des établissements, par exemple dans le cas de l'Allemagne. En République tchèque, ce lien avec les évaluations internationales fait explicitement partie du cahier des charges de l'inspection générale.

▪ **L'autonomie plus que la décentralisation**

L'évaluation d'établissement ne semble guère progresser dans des pays de décentralisation extrême en direction des autorités locales, comme dans les pays scandinaves : l'Etat évalue le travail des municipalités en matière d'enseignement, mais l'établissement n'est pas évalué spécifiquement. Il n'a pas d'autonomie.

Les pays en revanche qui développent des procédures riches d'évaluation des établissements sont ceux qui leur reconnaissent une grande autonomie, comme les Flandres belges ou la République Tchèque, mais il existe bien sûr aussi des exceptions avec des pays pour lesquels le développement de l'évaluation des établissements s'est fait dans un contexte contraire au renforcement de l'autonomie des établissements, et c'est le cas notamment de l'Angleterre.

Beaucoup de pays sont dans une position intermédiaire, comme l'Allemagne, où la préoccupation de l'amélioration de la qualité amène à chercher les voies d'une plus grande autonomie des établissements, et à s'intéresser progressivement à leur évaluation. La même préoccupation de qualité conduit d'autres pays sur la voie de la centralisation (Angleterre).

Schématiquement, la même préoccupation de restauration de qualité face aux crises des systèmes scolaires²⁵ conduit selon les pays à des réponses centralisatrices, ce qui est propice au développement d'un certain type d'évaluation d'établissements (R-U), à des réponses favorables à l'autonomie des établissements (France, République tchèque), avec le développement éventuel d'un autre type d'évaluation d'établissements²⁶, ou à des réponses en termes de développement des responsabilités des collectivités locales (pays scandinaves), évolution qui ne conduit en général pas au développement de l'évaluation des établissements. La préoccupation de restauration de la qualité dans les pays où les établissements sont depuis longtemps autonomes (Flandres belges) conduit à une forte préoccupation d'évaluation d'établissements.

Si l'autonomie des établissements est donc souvent accompagnée du développement de procédures d'évaluation d'établissements, on ne peut pas dire que la centralisation soit par essence contraire au développement de ces procédures (Angleterre) ni que la décentralisation leur soit favorable (pays scandinaves).

²⁵ La crise des systèmes scolaires se manifestant d'abord par le fait que beaucoup de ces systèmes donnent l'impression de ne plus progresser, par exemple en matière de taux d'accès ou de taux de diplômés, pour un coût qui ne se stabilise pas.

²⁶ L'Italie représente une exception puisque ce pays se retrouve dans cette catégorie sans disposer d'inspection générale ni de modèle national d'évaluation.

Le tableau suivant peut être proposé :

<i>Degré d'autonomie/ de centralisation</i>	<i>Attitude par rapport à l'évaluation des établissements</i>
Autonomie ancienne des établissements	Intérêt certain pour évaluation des établissements
Autonomie des établissements en développement	Intérêt plus aléatoire pour l'évaluation des établissements
Ancienne centralisation	Peu de culture d'évaluation, mais peut se développer
Centralisation de reconquête	Fort intérêt pour l'évaluation d'établissements
Pouvoir aux collectivités territoriales	Peu d'intérêt pour l'évaluation

2.1.2. *Positionnement des instances chargées des évaluations externes (inspections générales, notamment)*

Les inspections générales, dans les pays, assez nombreux, où elles existent (elles ont pu être supprimées, comme en Grèce, remplacées par des « conseillers pédagogiques ») sont parfois très anciennes (1832 en Angleterre, époque napoléonienne aux Pays-Bas), mais il est certain que les pays où elles sont de création récente, comme la République tchèque, (où cette création s'accompagne d'une suppression des échelons déconcentrés), peuvent écrire la page avec plus d'imagination, ce qui confère à la définition de ce corps nouveau une place d'emblée stratégique dans la politique éducative générale.

En plusieurs pays le besoin a été ressenti de donner à l'inspection un statut indépendant des autorités ministérielles, qu'il s'agisse des Pays-Bas où le programme de travail de l'inspection est suivi de près et peut être amendé par le Parlement, ou de l'Angleterre où l'OFSTED, qui reste formellement une inspection générale, est rattaché directement au gouvernement, et dont les rapports sont publics sans passer par le ministère.

A noter que certaines inspections jouent un rôle pédagogique fort dans la diffusion de la culture d'évaluation des établissements : cf. l'inspection écossaise publiant « How good is our school ? », best seller en matière d'auto-évaluation.

2.1.3. *Méthodes et critères des évaluations externes*

Si en certains pays, comme la Pologne ou la Slovaquie, on peut considérer que les clarifications nécessaires sur le rôle de l'instance évaluatrice n'ont pas été suffisantes (selon le cas, mélange

du rôle de l'évaluateur et de l'opérateur, distinction contre-productive étanche entre évaluation pédagogique et administrative, ou entre un « contrôle » qui proclame ne pas s'occuper de qualité et une qualité qui n'est pas contrôlée), on a la plupart du temps une spécification du rôle de l'évaluateur externe.

Cela se marque par un travail important et public sur les procédures et critères d'évaluation, la publicité desdits critères et procédures, la périodicité de ces évaluations, la publicité de leurs résultats (cf. point 2.1.4.) ou encore les procédures de suivi.

▪ **2.1.3.1. Sur la périodicité**

Beaucoup de pays s'attachent à définir une périodicité rigoureuse des évaluations externes, comme le Portugal (6 ans), l'Angleterre (6 ans) ou encore l'Ecosse (tous les 7 ans, avec intégration des rythmes du primaire et du secondaire afin que durant une scolarité obligatoire chaque enfant ait son établissement inspecté 2 fois). S'agissant du temps de l'évaluation externe et de sa préparation, toutes les situations existent, avec la situation singulière de la Catalogne et du Pays Basque où les évaluateurs (universitaires au Pays Basque) évaluent l'établissement dans la durée de son activité, suivant par exemple une même cohorte d'élèves sur 4 ans.

▪ **2.1.3.2. Sur la méthode**

Les méthodes des évaluations externes des établissements sont diverses mais définies publiquement, dans des textes parfois longuement discutés au niveau parlementaire (Pays-Bas), et on voit par exemple souvent s'établir des distinctions entre « inspections de routine », ou « short » et « intégrales », ou « full », entre évaluation exhaustive ou focalisation sur des faiblesses apparentes (Ecosse). La question du rapport avec les évalués est l'objet d'une attention spéciale dans les protocoles tchèques où il est prévu d'interroger chaque fois le responsable de l'école sur ce qu'il a pensé des procédures et résultats de l'évaluation externe ; c'est aussi en République tchèque qu'existe, au sein de l'inspection générale, un groupe de méthodes permanent sur les questions d'évaluation ainsi qu'une obligation faite à l'inspection d'une auto-évaluation annuelle de ses propres rapports.

▪ **2.1.3.3. Sur les questions curriculaires**

S'agissant de la part des questions pédagogiques dans les évaluations d'établissements, elle est partout très forte, tournée d'abord vers la mesure des résultats des élèves : cette mesure n'est pas facilitée dans les pays sans programmes nationaux ni examens nationaux (on assiste alors, comme en Hongrie, au développement inflationniste du marché des tests privés), elle l'est en revanche dans les pays qui s'en dotent, comme le Royaume-Uni, (le cas de la France, dotée de programmes et d'examens nationaux, mais peu focalisée sur les résultats des élèves étant assez spécifique). A noter aussi que dans la plupart des pays, même quand il existe un

« curriculum » national, l'établissement doit expliciter ce qu'il en fait et que la première question des évaluateurs externes est en général : « quels programmes enseignez-vous ? quels curricula ? ».

▪ 2.1.3.4. Sur le suivi des évaluations

Les procédures de suivi sont en général assez systématiques, comme en Ecosse ou en République Tchèque (2 ans), mais elles vont parfois, comme en Angleterre, bien au-delà d'un suivi, selon la catégorie (« mesures spéciales », ou « faiblesses spécifiques ») à laquelle l'établissement a été rattaché à l'issue de l'évaluation externe : toute une procédure est lancée qui peut aller, au moins de façon exceptionnelle en théorie, jusqu'à la fermeture de l'établissement.

2.1.4. La question du rapport entre l'évaluation et les usagers

Les questions de l'interrogation des usagers et des retours qui leur sont faits de l'évaluation font partie des plus débattues. Parfois comme en Hongrie, certaines évaluations sont directement faites pour répondre à la demande des familles, par exemple sur les taux de succès des lycées aux concours d'entrée à l'Université.

Beaucoup de pays adressent des questionnaires aux usagers, soit à tous (Roumanie), soit sur échantillon (Ecosse), allant parfois (Norvège) jusqu'à imaginer la figure de l'élève -inspecteur s'exprimant sur Internet. Des méthodes diffusées en plusieurs pays, comme l'European Foundation for Quality Management, prévoient une interrogation sur la « satisfaction » à la fois des usagers et des agents.

Les rapports sont adressés à tous les parents sous 10 semaines en République Tchèque, mais beaucoup de pays (Angleterre, Ecosse, notamment), les mettent systématiquement sur Internet. S'il est décidé au Portugal que le rapport sera publié, il doit l'être in extenso. Dans beaucoup de pays, en revanche, les recommandations qui accompagnent les rapports, ou encore les notes mises aux établissements (Angleterre) ne sont pas publiées.

Le souci de ne pas enfermer l'évaluation dans les logiques professionnelles de l'éducation a conduit l'Angleterre et l'Ecosse à ouvrir les équipes d'inspection à des profanes (Laymen, Lay members), dont la présence est parfois contestée, parfois jugée insuffisante en nombre pour être significative.

2.1.5. Développement et ambiguïtés des démarches d'auto-évaluation

Il est question dans presque tous les pays de démarches d'auto-évaluation, et la Roumanie paraît avoir pris la mesure de l'objectif en décidant d'y former tous les chefs d'établissement avant 2004, ou l'Ecosse, qui a intégré à ses évaluations externes l'obligation d'évaluer la mise en place et la qualité du regard auto-évaluateur de l'établissement sur lui-même.

En certains pays, comme l'Italie, où les établissements se sentent un peu livrés à eux-mêmes dans un contexte d'autonomie sans évaluation externe, les établissements de plusieurs régions ont initié des démarches d'auto-évaluation (Province de Trente, notamment).

Retenons aussi la situation de l'Angleterre, où le thème de l'auto-évaluation est prôné par les autorités ministérielles au nom des mêmes objectifs que l'évaluation externe réalisée par l'OFSTED, qui a d'ailleurs publié un manuel d'auto-évaluation très proche de son propre manuel ! La question est alors posée de savoir si l'auto-évaluation est la simple intériorisation d'un contrôle externe, ou bien si, à la fois pour être crédible et pour avoir sa valeur ajoutée propre, elle ne pourrait pas être davantage prise en compte par... l'évaluation externe, en étant par exemple à l'origine de l'inscription de chapitres spécifiques au menu de cette dernière.

Le rapport officiel d'Eurydice considère que l'auto-évaluation est un des meilleurs moyens de développer l'autonomie, et que si l'établissement est très autonome, un contrôle-qualité peut s'avérer suffisant (on ne considère ni les inputs, ni les processus, mais les outputs en termes de résultats, de satisfaction, d'effets...).

La conclusion mérite d'être lue : « Du point de vue de l'amélioration de la qualité et de l'innovation, c'est sans doute le degré de développement de l'évaluation interne (« auto-évaluation ») qui va déterminer les contours de l'évaluation externe. Plus la première est systématique, autonome, plus la seconde peut se faire discrète, complémentaire. Il faut libérer du temps pour que l'évaluation se centre sur la supervision des processus d'évaluation interne et le suivi des résultats ».

2.2. Les évaluations d'EPLÉ à initiative nationale en France

Au sein d'un système éducatif qui est traditionnellement décrit comme « centralisé », et dont la centralisation reste une des caractéristiques essentielles, il n'est pas surprenant de constater qu'existent des procédures nationales d'observation des EPLÉ, selon diverses modalités qui vont du « contrôle » à ce qu'on pourrait appeler effectivement « évaluation », et que s'y soient investis, diversement, des « inspecteurs généraux », puisqu' aussi bien dans presque tous les

pays dotés d'une inspection générale cette dernière est appelée à porter son regard sur les établissements scolaires. Ce n'est bien sûr pas ici « du point de vue » de l'une ou l'autre des inspections générales que la question des évaluations à initiative nationale sera abordée, mais du point de vue de l'exercice de la fonction d'évaluation d'établissements à l'échelle du système éducatif en son entier.

On distinguera donc :

- le rappel de l'existence d'un riche héritage en matière de regard porté sur les établissements, paradoxalement peu tourné vers les aspects pédagogiques ;
- la recherche d'évolutions délicates entre procédures de contrôle et analyses systémiques ;
- la réflexion sur les limites contemporaines de la situation face au développement éventuel de démarches d'«évaluation» ou d'auto-évaluation des établissements.

2.2.1. Un héritage riche mais devenu partiel

▪ 2.2.1.1. Du regard global des origines...

Aux termes de la Loi du 11 floréal an 10, l'inspection générale est tellement liée dès son origine à la nécessité d'un regard porté depuis le niveau national sur les établissements du second degré qu'elle paraît même créée pour cela, pour « visiter, une fois au moins l'année, les lycées, en arrêter définitivement la comptabilité, examiner toutes les parties de l'enseignement et de l'administration, et en rendre compte au gouvernement ». On retient, dans le cahier de charges initial d'un métier qui a plus de deux siècles d'expérience, l'exigence de fréquence des « visites », leur lien avec des exigences de gestion (« arrêter la comptabilité »), et surtout le caractère global de ces visites, pour lesquelles le mot d'« examen » est préféré à celui de « contrôle ». L'appellation des « inspecteurs généraux de l'instruction publique », faisant suite à celle d'« inspecteurs des études » qui avait été initialement retenue en 1802, montre encore cette visée globale, centrée sur les objectifs généraux de l'action du ministère.

Cette définition des tâches de l'inspection générale est bien entendu tracée à l'époque dans le cadre d'un système qui traite spécifiquement du second degré, et dont le contrôle ne peut être que national dans la mesure où la gestion elle-même l'est aussi.

▪ 2.2.1.2. ...à une spécialisation à laquelle échappent des aspects essentiels

Ultérieurement, toutefois, on a assisté à une spécialisation de certains inspecteurs généraux dans des tâches spécifiques (création en 1920 des « inspecteurs généraux des services administratifs », auxquels revient notamment aux termes du décret du 12 mars 1920 « le contrôle de l'économat »).

Dans ces conditions, on constate que, tandis que l'IGAEN, créée en 1965, se voyait précisément fixer comme mission celle de « constater dans tous les organismes sur lesquels s'exerce son contrôle, l'observation des lois, règlements, décisions et directives ministérielles qui en régissent le fonctionnement », et qu'une partie seulement des inspecteurs généraux de ce corps se spécialisait dans le contrôle des établissements du second degré en se centrant sur leur fonctionnement administratif, financier et matériel, le regard étendu « à toutes les parties de l'enseignement et de l'administration » de la Loi de floréal an 10 n'existe plus avec toute sa portée initiale : la façon dont se construisent les enseignements et les apprentissages des élèves n'est pas l'objet d'une attention spécifique au niveau d'un établissement scolaire, l'inspection générale à vocation pédagogique (« inspection générale de l'éducation nationale ») se centrant pendant des décennies²⁷, dans les établissements, sur le contrôle individuel de l'activité des maîtres.

2.2.2. *Du contrôle à l'analyse systémique*

▪ 2.2.2.1. Une position de plus en plus claire d'analyse de système...

Le développement même de l'enseignement du second degré à partir des années soixante est à l'origine de nouvelles évolutions dans l'organisation nationale des regards portés sur les établissements : en 1984, la mission d'évaluation institutionnelle dévolue à l'IGAEN inclut l'appréciation de la pertinence et de l'efficacité des activités observées. De fait, dans les établissements, les contrôles réglementaires sont exercés, notamment par des inspecteurs généraux de l'IGAENR, de plus en plus souvent dans le domaine de l'organisation pédagogique où les différentes réformes sont à l'origine de textes abondants et fréquemment renouvelés. Parallèlement à cette entrée dans le domaine pédagogique d'un regard qui s'en tenait à distance, la pénétration plus régulière de ce même regard dans le domaine de l'opportunité se fait de plus en plus fréquente, puisque sont étudiées efficacité et efficacité des dispositifs que met en œuvre, dans tous les domaines dont le domaine pédagogique (évaluation de la répartition des moyens à l'interne, des emplois du temps, de la constitution des classes, des résultats scolaires), l'établissement désormais plus autonome.

²⁷ Il est remarqué dans le rapport d'Eurydice déjà cité que si dans la plupart des pays on est parti d'évaluations externes essentiellement tournées vers la matière pédagogique, on a été progressivement amené à considérer aussi les domaines du management et de la gestion. L'évolution française a été largement différente, et presque inverse.

A la fin des années quatre-vingt-dix, l'IGAEN élabore à l'intention de ses membres un guide du « suivi permanent » des établissements, intégré ensuite dans un « guide méthodologique des missions d'inspection générale », qui est une avancée incontestable de méthode vers une approche systémique de l'établissement, distinguant les phases de préparation de la mission, de visite sur place, de rédaction des notes, listant les interlocuteurs à rencontrer et les sujets d'entretien avec chaque catégorie. Insistant sur le passage par une démarche problématisée, le guide précise que la finalité n'est pas de décrire un établissement scolaire de façon exhaustive, mais d'en mentionner les principales caractéristiques, d'en analyser les points-clés, d'en dégager forces et faiblesses et de formuler des recommandations adaptées.

A noter qu'au cours de la même période, à partir d'un regard initialement tourné vers l'évaluation personnelle des chefs d'établissements, les inspecteurs généraux du groupe « établissements et vie scolaire » de l'inspection générale de l'éducation nationale suivent un chemin assez similaire, s'avancant de plus en plus vers la rédaction de notes analysant elles aussi la globalité de l'établissement.

Un certain nombre d'opérations nationales, par ailleurs, d'évaluation de tout ou partie du système éducatif par un travail conjoint à l'inspection générale de l'éducation nationale et à l'inspection générale de l'administration de l'éducation nationale et de la recherche ont conduit aussi à des démarches plus globales d'évaluation d'établissements :

- à plusieurs reprises a été réalisé un travail d'analyse nationale tant de l'état des établissements que de l'avancée de réformes dans le cadre de l'« évaluation » en un bref laps de temps d'un nombre important de collèges, lycées ou lycées professionnels (opération dite des cent collèges, par exemple) ;
- à l'occasion de l'« évaluation de l'enseignement » dans des académies, démarche initiée à partir de 1998 et ayant à la date du présent rapport concerné 14 académies, des équipes issues des deux inspections sont amenées à visiter des établissements de façon globale, mais avec la problématique spécifique de saisir à leur propos les caractéristiques du fonctionnement et des résultats de l'académie où ils sont situés.

▪ 2.2.2.2. ...qui n'entend pas rompre avec une logique de contrôle

Malgré ces évolutions en direction de logiques d'évaluations d'établissements, totales ou partielles, il semblerait que des choix n'aient pas été faits, avec des arguments en certains cas intéressants, entre posture d'évaluation et posture de contrôle.

Le développement quantitatif de l'enseignement secondaire à partir des années soixante a fait surgir des problèmes inédits relatifs à la gestion des établissements, aux compétences des

agents chargés de les gérer, puis de les manager, au fur et à mesure que davantage de décisions, dans des domaines nouveaux, revenaient à l'établissement au lieu d'être prises aux niveaux national ou académique. La redéfinition des missions de l'IGAENR en 1965 était aussi la conséquence d'une plus grande urgence du contrôle sur un système qui changeait d'échelle et l'actuelle configuration de la « mission d'inspection générale » telle que la définit l'IGAENR, alliant des contrôles, de l'évaluation et des techniques d'audit, est le reflet de l'ambition d'un regard porté sur l'établissement comme système, avec une préoccupation légitime pour la rigueur de la gestion, sans quoi aucune politique managériale ne pourrait réussir, mais de manière plus large pour que dans tous les domaines où la norme est omniprésente, le contrôle de son respect continue de constituer l'étape préalable à toute autre analyse.

A noter que l'IGAENR, malgré l'élargissement de sa démarche, n'a pas échappé aux réticences de nombreux acteurs du système à utiliser précisément le mot d'« évaluation », préférant l'appellation *sui generis* de « suivi permanent »²⁸, reprise presque chaque année dans le programme de travail de l'IGAENR tel que le signe le ministre.

Cela dit de l'ancienne posture du seul contrôle, alors même que l'inspection générale a mis au point au fil des années une approche de l'établissement autrement plus riche, sont conservés un certain nombre de traits dont il est plus difficile de justifier le maintien à l'identique dans la configuration actuelle :

- le protocole même de l'IGAENR est un outil utilisé par le corps de façon de plus en plus régulière, mais il n'est ni public, connu par exemple des chefs d'établissements ou des recteurs, voire des usagers, ni même respecté de façon obligatoire par l'ensemble des inspecteurs généraux du corps ;
- les critères du choix des établissements visités ne sont pas clairs, ni dans l'ensemble du système, ni même pour la totalité des groupes d'organisation géographique qui composent le corps : si des critères relatifs au poids des différentes catégories d'établissements, à leurs caractéristiques sociologiques ou géographiques (rural *versus* urbain, par exemple), à leurs résultats aux examens, à leur appartenance ou non à une zone de politiques territoriales sont souvent pris en compte, l'IGAENR a toujours aussi dans sa logique que cette incertitude sur les critères présente un côté positif puisque tous les établissements restent des cibles

²⁸ L'appellation de suivi permanent n'est par elle-même ni bonne ni mauvaise : elle fait toutefois comme si en matière d'évaluation le mot de suivi n'avait pas un sens précis que de toute évidence l'IGAENR ne retient pas au moins d'ordinaire (le « follow up », le fait de revoir un établissement évalué après un certain laps de temps) et comme si la « permanence » était à observer du point de vue du seul ministre, pour lequel elle observe « en permanence » un système à partir de la diversité d'établissements qui, eux, ne sont visités que de façon relativement aléatoire et en tous cas espacée.

potentielles et que tous les établissements savent, ou doivent savoir qu'ils sont susceptibles d'être ainsi visités ;

- la question du rapport avec les autorités administratives déconcentrées n'est pas claire : sans poser ici la question qui sera abordée plus loin de la différence entre le commanditaire et le destinataire du rapport élaboré à propos de l'établissement visité, notons que la jurisprudence varie d'un inspecteur général à un autre, ou d'un groupe géographique à un autre, à la fois sur l'existence d'une commande locale de visite (venant du recteur, ou du DSDEN) et sur les destinataires du rapport, catégorie qui peut encore se réduire au chef d'établissement (le C.A. ne sera destinataire que si ce dernier le souhaite, sans qu'on sache quelle proportion des cas cela représente).

2.2.3. *Limites de la situation actuelle*

▪ 2.2.3.1. Les ambiguïtés sur les usagers, les effets et le suivi de l'évaluation

Parmi les questions qui marquent les limites des évaluations d'établissements à initiative nationale, et en plus des ambiguïtés sur la place de la fonction de contrôle dans un ensemble qui s'est désormais ouvert à d'autres considérations plus systémiques, et du coup plus proches d'une évaluation, on peut citer :

- les incertitudes sur le suivi des rapports d'inspection générale (dans le cas général appelées « notes », en tant que telles non destinées au ministre) : s'ils sont régulièrement utilisés pour élaborer des synthèses à thème général ou spécifique à destination du ministre, leur suivi par l'institution auteur du rapport à l'échelle de l'établissement est en l'état exceptionnel : là encore la démarche de rapport au ministre sur un établissement ou un échantillon d'établissements, pour vérifier l'application d'une norme ou la mise en œuvre d'une politique, prime encore sur une logique d'évaluation telle que définie plus haut²⁹, tournée vers l'amélioration, même si cette dimension est de plus en plus privilégiée dans la réalité des rapports eux-mêmes ;
- de la même façon, indépendamment du « suivi » proprement dit (« qu'a-t-on fait des conclusions et préconisations du rapport ? ») à l'échelle de l'établissement, la connaissance des effets du rapport et de l'évaluation sur l'établissement ou sur

²⁹ Evaluation : jugement porté sur l'activité d'une organisation, en considération de ses résultats, proposant des explications du niveau d'atteinte de ces résultats dans le but de permettre cette organisation de les améliorer.

l'attitude des autorités déconcentrées à l'égard de l'établissement n'est pas rassemblée³⁰ ;

- la question est posée du fait que, dans un système désormais fortement déconcentré, le commanditaire de la mission dans un établissement reste, selon une coordination assez aléatoire avec ces autorités (recteurs, DSDEN), le ministre, de façon assez fictive d'ailleurs puisque l'IGAENR arrête seule son propre programme de visite d'établissements dont elle s'auto-saisit. Le destinataire n'est, lui, pas le ministre, mais, comme il a été dit, le chef d'établissement et, le plus souvent, les autorités déconcentrées. On ne doit dès lors pas s'étonner du manque de visibilité, dans le système, souvent, de travaux qui mériteraient meilleur sort, ni des incertitudes précédemment citées sur leurs effets et leur suivi : l'IGAENR elle-même ignore peut-être ce qu'elle a préféré ne pas toujours savoir d'une efficacité en pointillé et d'une efficience critiquable si l'on met les résultats en face des talents et efforts déployés.

▪ **2.2.3.2. Les problèmes d'échelle et la prise en compte insuffisante des acquisitions des élèves**

Les deux questions principales posées par le fonctionnement actuel, en regard de ce que pourrait être un bon système d'évaluation des établissements demeurent toutefois les suivantes :

- le manque de visibilité d'un fonctionnement qui ne prétend pas couvrir l'ensemble des établissements dans un délai raisonnable : même l'argument, peut-être discutable, de l'efficacité de la « crainte » du contrôle, perd de son efficace si cette éventualité est trop incertaine : or l'IGAENR n'est guère en position que d'évaluer environ 150 établissements par an, alors qu'existent plus de 7500 établissements du second degré publics et 4000 privés sous contrat, ce qui signifie de façon toute schématique qu'un établissement ne serait visité sous la forme nationale que deux fois par siècle, alors que dans le programme de travail d'institutions similaires à l'inspection générale dans d'autres pays européens, la périodicité des visites fait partie du cahier des charges des instances chargées de l'évaluation. Ni une démarche de « contrôle », ni une démarche qui privilégierait une attitude d'évaluation ne peuvent s'accommoder d'un plan de charge aussi peu dense quant à la périodicité des regards portés sur les établissements ;
- peu de progrès ont été accomplis s'agissant de l'évaluation des acquisitions des élèves : l'IGAENR a certes élargi son champ d'investigation de même que

³⁰ Un rapport peut être suivi dans ses préconisations avec des effets pervers inédits ; un rapport peut n'être pas suivi alors même que l'évaluation a déclenché un mouvement, positif ou négatif, dans l'établissement.

l'IGEN-EVS, et les inspecteurs généraux faisant partie des groupes disciplinaires de l'IGEN ont certes participé à un nombre croissant d'évaluations d'établissements dans des cadres divers mais aucun travail de mise au point d'une démarche nationalement conjuguée, d'un protocole, ou d'un plan de charge n'a véritablement modifié la donne. On peut dire que dans les rapports élaborés par la seule inspection générale qui réalise de façon continue des visites d'établissements (l'IGAENR), les questions centrales des acquisitions des élèves et de leurs résultats, ainsi que de l'utilisation fine des moyens pédagogiques, parmi lesquels les talents et pratiques des maîtres, pour en arriver à ces résultats, ne sont pas traitées de façon centrale ni systématique, faute des éléments préalables d'analyse qui permettraient d'y arriver.

2.3. Les évaluations d'EPLÉ sur initiative académique

2.3.1. La mise en place des évaluations en académie : une gestation plus ou moins difficile

Sous diverses appellations, audit (Lille), audit pédagogique (Toulouse), inspection globale d'établissement (Rouen), inspections coordonnées (Paris), concertées (Clermont-Ferrand), plusieurs académies ont mis en place, au cours de la décennie 1990, des évaluations d'établissements scolaires, sans pour autant utiliser cette dénomination. A l'exception de l'expérience lilloise, ces opérations n'ont généralement concerné qu'un petit nombre d'établissements, d'ailleurs variable d'une année à l'autre, et n'ont fait qu'exceptionnellement l'objet d'un projet global plaçant l'évaluation des EPLÉ au cœur du pilotage académique.

- **2.3.1.1. Les audits d'EPLÉ dans l'académie de Lille : une expérience pionnière mais aujourd'hui en quête d'identité**

- *Les audits de la première génération*

Sur l'initiative du recteur Pair, l'académie de Lille, a mis en place dès 1991 une évaluation systématique de tous les établissements publics du second degré. Assez curieusement, le terme audit – défini comme une démarche à caractère participatif – a été préféré à celui d'évaluation considérée comme un acte « purement hiérarchique ». Présentés avant tout comme une aide à l'établissement, les audits lillois s'inscrivent dans la perspective de mise en place des projets d'établissement prévus par la loi d'orientation de 1989 et poursuivent explicitement trois objectifs : favoriser une auto-analyse du fonctionnement de l'établissement ; constituer une base d'informations sur les établissements pour attirer

l'attention des autorités académiques sur leurs problèmes et les réponses à y apporter ; fournir des éléments pour élaborer les contrats à passer avec les établissements.

Après une expérimentation dans quatre établissements volontaires, 128 collèges et lycées seront audités dès la première année par 32 équipes d'auditeurs. En quatre ans, la quasi-totalité des collèges et des lycées de l'académie sera visitée par 235 auditeurs, parmi lesquels on compte, pour moitié des inspecteurs territoriaux, pour un quart des chefs d'établissements et pour un autre quart des généralistes (cadres du rectorat, responsables de l'IUFM ou de la MAFPEN...). Les opérations ont été méticuleusement préparées en amont par un groupe de pilotage académique qui a programmé un plan solide de formation des auditeurs et réalisé un guide d'audit comprenant plusieurs documents volumineux : une liste des démarches à entreprendre, un référent qui fixe les éléments sur lesquels doivent porter les observations dans les établissements, un guide pour constituer les dossiers d'information, enfin un guide d'entretien ainsi que des fiches annexes concernant l'élaboration des synthèses disciplinaires. L'entreprise a été conduite avec ténacité par une équipe de pilotage convaincue de l'intérêt de la démarche d'audit et avec le soutien indéfectible du recteur en place.

La réalisation d'une opération de cette envergure dans la deuxième académie de France par la taille constitue déjà en soi un véritable succès. Pour parachever l'ensemble, un bilan académique a été établi entre 1995 à 1997 à partir des réponses à un questionnaire envoyé à tous les établissements audités et des comptes rendus d'une seconde visite dans 48 de ces établissements.

➤ *Les nouveaux audits lillois en quête d'identité*

Divers facteurs vont cependant conduire les responsables académiques à infléchir sérieusement le dispositif envisagé par le recteur Pair à partir de 1998 : d'une part, les conclusions du bilan précité font apparaître que les audits n'ont pas produit les effets que l'on escomptait ; par ailleurs, les inspecteurs qui ont porté l'expérience ressentent un certain essoufflement, d'autant que leur investissement massif dans les audits est de plus en plus difficile à concilier avec la poursuite de leurs missions traditionnelles ; enfin et surtout, la mise en place des nouveaux dispositifs pédagogiques transversaux requiert désormais une mobilisation accrue des corps d'inspection territoriaux dans le domaine pédagogique.

Sous l'impulsion de l'inspecteur général en académie Jean-Louis Piednoir, une nouvelle campagne d'audits est donc engagée au cours des deux années scolaires 1999-2000 et 2000-2001. Le nouveau dispositif est à la fois recentré sur l'observation des pratiques pédagogiques disciplinaires et interdisciplinaires, allégé quant au nombre d'établissements audités (40 par année scolaire) et élargi à l'évaluation de deux écoles primaires pour prendre en compte l'articulation CM2/6^{ème}. Dès 2001, une nouvelle impulsion est donnée qui privilégie l'audit de réseaux d'établissements et d'écoles dans le cadre des bassins. L'analyse se centre désormais

sur le positionnement des établissements et leurs relations, plus que sur le fonctionnement interne de chacun d'entre eux. Des observations dans des établissements choisis pour leurs actions innovantes sont également entreprises.

A son arrivée, dans une circulaire envoyée aux chefs d'établissement le 22 décembre 2002, le recteur Desneuf dresse un bilan plutôt contrasté des opérations d'audit dans l'académie de Lille. Parmi les points positifs, il relève le croisement des compétences entre chefs d'établissement et inspecteurs, la mise en place d'une culture de l'évaluation et la transformation des pratiques des corps d'inspection. Mais il pointe également plusieurs limites : l'audit ne sert pas à l'animation pédagogique locale ou académique ; il n'a pas alimenté une base de données utilisable pour le pilotage rectoral et n'a guère servi à éclairer la politique des établissements eux-mêmes. Une clarification des objectifs et de la méthodologie de l'audit lui apparaît nécessaire. L'année 2002-2003 sera donc une année de transition dans l'attente d'une redéfinition de l'architecture des audits : consigne est donnée aux auditeurs de se limiter à l'analyse de la seule dimension pédagogique et de renoncer à l'évaluation globale des établissements.

Le nouveau protocole, élaboré à la demande du recteur, pour la campagne des audits désormais dénommés « pédagogiques » (2003-2004) se caractérise par l'accent mis sur l'utilisation des synthèses disciplinaires rédigées par chaque inspecteur et par la volonté de simplifier la structure du rapport³¹. Un soin tout particulier est attaché à la détermination des établissements audités (réduits à 18, dont 6 écoles) qui sont choisis en fonction de leur niveau de valeur ajoutée, de manière à comprendre à la fois des établissements « très performants, des « moyennement performants » et des « non performants ». Il est encore trop tôt pour porter un jugement tranché sur les nouveaux audits lillois. A la lecture du protocole, on pouvait craindre que la dimension systémique de l'évaluation de l'établissement soit abandonnée au profit d'une entrée plus strictement disciplinaire. La lecture des premiers rapports atteste qu'il n'en est rien : si l'analyse est plus ramassée que dans les premiers audits, tous les compartiments de la vie de l'établissement, à l'exception des dimensions comptables et financières, continuent à faire l'objet d'un examen approfondi. En revanche, l'apport des synthèses disciplinaires à la compréhension du fonctionnement pédagogique de l'établissement ne semble pas encore avoir donné tous ses fruits.

Après la démarche assurée de la première moitié de la décennie 1990, marquée par le caractère systématique et obligatoire de l'audit, celui-ci semble aujourd'hui à la recherche d'un second souffle.

³¹ Les rapports d'audit rédigés entre 1991 et 1995 comprenaient 6 parties : caractéristiques de l'établissement, politique mise en oeuvre, qualité de la prise en charge des élèves, système de pilotage, performances de l'établissement, recommandations. Les actuels audits n'affichent plus que trois parties : présentation de l'établissement, pilotage pédagogique, recommandations pédagogiques.

▪ **2.3.1.2. Le développement des évaluations à initiative académique : influence du « modèle lillois » et nouvelles pistes**

➤ *Les évaluations à initiative académique calquées sur le dispositif lillois*

On assiste dans la deuxième moitié des années 1990 à une floraison d'initiatives académiques en rapport avec l'évaluation des EPLE qui, peu ou prou, s'inspirent du « modèle » lillois sans en présenter la même systématisme. En 1996, l'académie de Rouen développe les inspections globales d'établissement, en 1997 l'inspecteur d'académie du département de l'Ariège – précédemment en poste dans le département du Pas-de-Calais – lance des « évaluations de politiques d'établissement »³² ; le collège des IA-IPR de l'académie de Clermont-Ferrand initie des inspections croisées en 1998. Enfin, une seconde vague d'évaluation d'établissements touche les academies de Bordeaux, Besançon, Montpellier, Paris et Toulouse entre 1999 et 2003.

La démarche, les objectifs, les protocoles, les thématiques privilégiées et le déroulement des opérations présentent nombre de similitudes avec ce qui s'est fait à Lille. L'évaluation s'appuie à la fois sur des documents statistiques et administratifs, des inspections individuelles, des observations d'activités transversales, des réunions de conseils d'enseignement et sur une série d'entretiens avec le chef d'établissement, des représentants des personnels, des parents et des élèves. L'opération se conclut par la rédaction d'un rapport de synthèse communiqué, au minimum, au recteur, à l'inspecteur d'académie et au chef d'établissement.

Cependant ces dispositifs sont loin d'avoir l'envergure des audits lillois de la première génération. A l'exception de l'académie de Besançon où l'on a programmé – dans le cadre de la contractualisation avec la Desco – l'audit de tous les collèges en quatre ans, le nombre d'établissements retenus est bien plus modeste (entre 3 à 12 selon les academies et/ou les années) et leur choix se fait selon des modalités variables, qui vont du volontariat à la désignation d'établissements en difficulté par le recteur, en passant par des propositions plus ciblées pédagogiquement par les corps d'inspection territoriaux. Dans nombre d'academies, l'accent est mis sur l'analyse des pratiques pédagogiques, dans leurs dimensions disciplinaires et transversales, au détriment d'une approche plus globale de l'établissement. Les protocoles sont de qualité très inégale : certains sont très (trop ?) détaillés tandis que d'autres se réduisent à un simple vade-mecum. Surtout, les évaluateurs, généralement recrutés sur la base du volontariat, n'ont pas toujours bénéficié d'une formation aux pratiques de l'audit ou de l'évaluation institutionnelle.

L'évolution des appellations et des pratiques signale aussi des tâtonnements. Dans l'académie de Clermont-Ferrand, on change de dénomination en cours de route, passant des inspections

³² Les évaluations sont systématiques et récurrentes.

croisées aux inspections concertées pour « éviter que ne disparaisse le regard disciplinaire » : au lieu de privilégier l'observation simultanée de situations de classe par plusieurs inspecteurs, l'accent est désormais mis sur la concertation avec les chefs d'établissements. Dans l'académie de Besançon, on semble renoncer pour la campagne d'audits à venir à faire figurer des recommandations dans les rapports, de manière à préserver l'autonomie des établissements alors que dans le même temps le protocole des « nouveaux » audits lillois préconise de rendre les recommandations plus opérationnelles et plus précises.

Certaines académies mettent aussi en avant des dimensions qui avaient été négligées dans les audits de Lille : par exemple, l'académie de Besançon a élaboré une charte des droits et devoirs des auditeurs et des audités, Bordeaux place au premier plan l'analyse des résultats, Toulouse accorde une attention toute particulière aux « demandes enregistrées au cours de l'audit ».

➤ *Des évaluations à initiative académique plus ciblées*

Après des expérimentations d'inspections croisées en 1997 puis d'audits de réseaux d'établissement au sein d'un territoire en 1999, restées sans lendemain, Poitiers s'est lancé dans une démarche d'évaluation de tous les lycées à partir de l'année scolaire 2000-2001 en liaison avec le développement de la contractualisation de l'académie avec ses établissements. L'évaluation s'est focalisée sur la mise en œuvre des projets d'établissement à mi-parcours avec l'objectif de « mesurer les écarts entre les volontés affichées par l'établissement et les réalisations en cours ». La démarche vise à sensibiliser les acteurs à des problématiques nouvelles : promouvoir le pilotage par les objectifs et les résultats, développer l'auto-diagnostic et l'auto-évaluation. Cette approche a été complétée par l'élaboration et la mise à disposition des lycées d'un guide d'auto-évaluation, précisant les dimensions pédagogiques et organisationnelles à évaluer et proposant une série d'indicateurs à construire par les établissements pour compléter ceux qui existent aux niveaux national et académique.

A Créteil, le recteur a mis en place en 2002 des « interventions coordonnées » des inspecteurs territoriaux dans 12 lycées de l'académie en ciblant plus particulièrement les observations sur « les facteurs pédagogiques susceptibles de rendre compte des performances observées ». Cette opération s'est prolongée cette année par des visites coordonnées dans douze collèges appartenant aux mêmes districts que les lycées visités l'année passée, de manière à explorer les problèmes d'articulation entre différents degrés d'enseignement. Des visites coordonnées de réseaux d'éducation prioritaire sont également prévues. Ces interventions visent essentiellement à mieux identifier les actions éducatives et pédagogiques les plus efficaces³³.

³³ Cette démarche s'inspire d'une recommandation du rapport des inspections générales sur l'évaluation de l'enseignement dans l'académie de Créteil.

L'évaluation de la microstructure expérimentale « Clisthène » située dans une annexe du collège du Grand Parc à Bordeaux, dont l'objectif est de lutter contre l'échec scolaire en mettant en place des formes innovantes d'apprentissage, d'évaluation et d'organisation scolaire, mérite une attention toute particulière. L'action des inspecteurs a consisté à prendre les acteurs du collège au mot, sur leurs intentions, en évaluant, en parfait accord avec les enseignants, les résultats plus que des activités. L'appréciation des progrès des acquisitions des élèves dans trois disciplines s'est effectuée en mesurant les compétences et savoirs des élèves en fin de période et en début de période par comparaison avec un échantillon témoin d'élèves appartenant à des collèges extérieurs à l'expérimentation. La dimension « vie scolaire et apprentissage de la prise de parole » a fait également l'objet d'une notation par des binômes d'observateurs dans le contexte du cours, dans les « groupes de tutorat », lors des conseils de classe et dans les différentes structures de vie scolaire. Bon exemple d'évaluation fine des acquisitions scolaires et des compétences sociales des élèves au sein d'une petite unité, l'opération est cependant difficilement reproductible à grande échelle compte tenu de l'investissement lourd en temps et en évaluateurs qu'elle suppose.

2.3.2. Des processus d'évaluation tirillés entre des objectifs et des attentes contradictoires

▪ 2.3.2.1. L'ambiguïté des objectifs

Il est remarquable d'observer que la plupart des académies n'ont pas choisi de nommer « évaluation » les opérations qu'elles initiaient, comme si on considérait, à tort ou à raison, que le terme était connoté trop négativement au sein de la communauté enseignante. De fait, dans les documents de présentation, on conseille aux auditeurs de ne pas porter de jugements de valeur mais simplement d'en rester aux faits et de renvoyer aux établissements une « photographie », « une image globale », un « reflet » aussi fidèle que possible du fonctionnement de l'établissement. Tout se passe comme si les évaluations n'osaient pas s'afficher pour ce qu'elles sont, comme si le fait de porter des jugements de valeur sur l'efficacité des établissements scolaires était en quelque sorte, inconvenant, déplacé, illégitime. Ce défaut d'affichage accrédite d'une certaine manière le fait que l'évaluation institutionnelle, lors même qu'elle est prévue explicitement par la loi d'orientation et les statuts des corps d'inspection, reste encore en France un sujet largement tabou. Cette présentation excessivement précautionneuse a souvent été la source de malentendus et des tensions parfois observées sur le terrain. Elle ne peut que faire naître incompréhensions et défiances auprès des professionnels de l'évaluation que sont les enseignants.

L'image de ces opérations d'audit ou d'inspections plus ou moins globales s'est trouvée d'autant plus brouillée que le *modus operandi* choisi n'a pas toujours contribué à en clarifier les objectifs officiels ou masqués. Ainsi dès lors que l'opération ne porte que sur quelques

établissements, trois ou quatre dans l'année par exemple, on ne peut guère s'étonner que les chefs d'établissement et les personnels enseignants s'interrogent sur les critères qui ont présidé à ces choix, souvent présentés comme le résultat d'un tri sur critères dans le cadre d'un échantillonnage rigoureux, alors que l'on peut supputer que les établissements ont été désignés en raison de difficultés particulières qu'ils rencontraient. Par de telles pratiques, on risque d'introduire une confusion entre opération d'évaluation et opération de pacification et, du coup, de manquer l'un et l'autre objectif.

▪ 2.3.2.2. Des « flottements » dans le déroulement des visites en établissement

On observe aussi un écart important entre les ambitions fortes affichées au début du processus d'évaluation qui touchent à la qualité de l'enseignement, au travail en équipe, à la participation aux nouveaux dispositifs transversaux, à l'implication dans le projet d'établissement, à l'analyse des besoins de formation et les conditions effectives de déroulement des investigations au sein des établissements dont, ni la nature, ni la durée, ne peuvent manifestement répondre aux attentes suscitées par l'effet d'annonce initial.

Les inspections individuelles dans les classes apparaissent trop souvent déconnectées de l'évaluation globale, en même temps qu'elles ne respectent pas toujours le programme initialement annoncé à l'établissement, de sorte que les enseignants ne font pas nécessairement le lien entre les deux opérations. Programmées avant l'évaluation globale, ces visites se déroulent parfois après ; tantôt elles donnent lieu à notation individuelle, tantôt il s'agit de simples observations non prises en compte pour la mise à jour de la note pédagogique ; dans certaines disciplines, tous les professeurs sont inspectés, dans d'autres seulement quelques-uns, voire aucun, la rédaction de la synthèse disciplinaire se faisant alors à partir des échanges entre l'inspecteur et les professeurs lors de la réunion du conseil d'enseignement.

La qualité du dialogue avec les équipes pédagogiques est, par ailleurs, inégale d'une discipline à l'autre : si dans certains cas, les professeurs se félicitent de la capacité d'écoute des inspecteurs, dans d'autres, il apparaît que l'on reste dans une logique de contrôle et de prescription qui s'accorde mal avec ce qui est présenté par les pilotes au début de l'opération. La logique de questionnement n'est pas toujours explicitée et la conduite d'entretien est plus ou moins bien contextualisée ; un ton critique et péremptoire est parfois adopté. Plus fondamentalement comme le souligne le doyen des IA-IPR de l'académie de Rouen « il n'est pas toujours facile de concilier l'objectif d'aide et de conseil avec des inspections individuelles éventuellement négatives et donc mal vécues par les enseignants ». Et cela est d'autant plus vrai que les inspecteurs sont peu formés à la démarche d'audit ou d'évaluation globale et plus ou moins convaincus de son intérêt.

▪ 2.3.2.3. Des modalités de restitution et de suivi très variables

La phase de restitution est d'autant plus délicate que le même propos peut avoir des résonances très différentes auprès de catégories d'acteurs qui n'occupent pas les mêmes positions institutionnelles et qui ne partagent pas nécessairement les mêmes valeurs. Voilà qui explique en partie l'extrême variabilité des pratiques en ce domaine. Ainsi, à Clermont-Ferrand, la restitution orale s'est effectuée, les premières années, en réunion plénière devant l'équipe de direction et tous les enseignants. Puis, l'expérience ayant montré les difficultés à gérer à chaud ce genre d'échanges, surtout lorsque certaines conclusions sont sévères pour l'établissement, on a par la suite différé dans le temps ces rencontres. Dans la plupart des académies, la restitution orale n'est d'ailleurs effectuée que devant le chef d'établissement ou l'équipe de direction, de sorte que les autres catégories d'acteurs (enseignants, parents, notamment) ne peuvent pas toujours faire part de leurs observations. La qualité de la restitution elle-même semble inégale, parfois plus proche de la semonce que du dialogue ouvert recommandé par les protocoles.

La pratique qui consiste à adresser le projet de rapport au chef d'établissement pour en corriger les éventuelles erreurs factuelles paraît également ambiguë. Ou bien il s'agit d'un simple contrôle formel, ou bien il s'agit d'une véritable procédure contradictoire qui doit logiquement aboutir soit à la réintégration des observations formulées dans le corps du rapport, soit à leur insertion *in extenso* en fin de rapport. Dans la pratique, les observations faites par les représentants de l'établissement sont encore trop rarement intégrées ou annexées dans le rapport final, alors même que l'on n'a pas toujours pris la peine de les réfuter.

Les suites données au rapport sont influencées à la fois par la manière dont il est diffusé au sein de l'établissement et par le suivi mis en place par l'autorité de tutelle. Dans certaines académies, le rapport est transmis au chef d'établissement et aux membres de la commission permanente ; dans d'autres, seul le chef d'établissement en est destinataire, avec le recteur et l'IA. Dans ce dernier cas de figure, son exploitation est tributaire de la diffusion qu'il veut bien lui assurer : ainsi dans tel collège de l'académie de Besançon, le rapport n'a pas du tout été diffusé ; dans tel autre, il l'a été à toutes les catégories d'acteurs (enseignants, parents, élèves) ; dans tel autre, encore, le rapport a juste fait l'objet d'une communication sans débats devant le conseil d'administration. La volonté du chef d'établissement de se saisir ou non du rapport afin de faire bouger les choses s'avère, dans la pratique, un paramètre difficilement contrôlable par les évaluateurs.

Le suivi de l'audit est également une préoccupation constante dans toutes les académies, même s'il est effectué avec une plus ou moins grande ténacité selon les cas. Dans les académies de Rouen, de Besançon et de Lille, une seconde visite d'inspection à échéance de deux ans est programmée pour observer les effets de l'audit. Est également prévu à Lille l'envoi par l'équipe de direction, dans l'année qui suit l'audit, d'un rapport indiquant dans

quelle mesure les recommandations faites ont été prises en compte ; en outre, le recteur réunit tous les auditeurs et les chefs d'établissement à la rentrée de septembre et présente publiquement les recommandations qu'il a retenues ; il adresse ensuite une lettre de mission aux chefs d'établissement pour en préciser la teneur³⁴. A Toulouse, le suivi peut prendre la forme d'une visite, en appui, par des inspecteurs à la rentrée suivante et d'une demande d'envoi des comptes rendus des conseils d'enseignement au pilote de l'audit pédagogique. Dans d'autres académies, le suivi est plus informel.

2.3.3. L'impact des évaluations reste limité en dehors des personnels d'encadrement

L'analyse de l'impact de ces évaluations peut se faire à partir de l'analyse du contenu des rapports, d'une part, et du jugement que les équipes d'évaluateurs et les personnels des établissements évalués portent sur le processus d'évaluation, d'autre part.

▪ 2.3.3.1. Les contenus des rapports intègrent peu la dimension pédagogique

Les rapports de synthèse présentent souvent une structure proche de ceux réalisés dans le cadre du suivi permanent des établissements par l'IGAENR, si l'on excepte la dimension financière et comptable ; finalement, ils ne diffèrent guère des rapports de visites d'établissements conduites dans le cadre des évaluations de l'enseignement en académie par les deux inspections générales. Ils traitent plutôt bien des caractéristiques du public scolaire et des enseignants, de l'environnement de l'établissement, de son équipement et de la qualité du pilotage. En revanche, certains points, comme les processus d'orientation, les modes d'évaluation, les effets de la formation, les caractéristiques des pratiques enseignantes, font généralement l'objet d'un examen trop rapide.

L'analyse fine des résultats par niveau de scolarisation et par discipline fait le plus souvent défaut. A cet égard, symptomatique est le fait que dans la quasi-totalité des protocoles académiques³⁵, l'analyse des résultats n'apparaît qu'en dernière partie des rapports, soit qu'on leur accorde peu d'importance, soit que l'on considère en quelque sorte qu'ils découlent naturellement des dimensions préalablement analysées.

Le plus souvent, on se borne à opérer de simples constats sans rechercher des explications. Encore trop peu de rapports se livrent à des croisements entre indicateurs pour en tirer quelques éléments probants, même si on relève, ici ou là, des tentatives de recoupement ou de

³⁴ Cette pratique fait courir un risque de double emploi ou de contradiction avec la lettre de mission adressée au chef d'établissement à la suite de la communication de son diagnostic dans le cadre de la procédure d'évaluation des personnels de direction.

³⁵ A l'exception de Bordeaux où l'analyse des performances constitue la première partie du rapport.

confrontation des données statistiques : par exemple, en collège, entre les évaluations de français et de mathématiques en 6^{ème}, les résultats ultérieurs aux épreuves ponctuelles ou au contrôle continu du brevet et les taux d'orientation en fin de seconde ; ou encore, en lycée, entre la sélectivité différentielle de l'orientation en fin de première vers les différentes séries du baccalauréat et les résultats ultérieurs à cet examen.

Les observations des pratiques d'enseignement font l'objet de fiches séparées qui sont rarement intégrées dans le rapport de synthèse, faute d'avoir été construites selon une grille harmonisée. Les propos sont plutôt convenus et déplorent généralement, quels que soient d'ailleurs les résultats, l'insuffisance du travail en équipe et le caractère traditionnel des pratiques pédagogiques sans pousser plus loin les investigations. En dépit de la priorité qui lui est officiellement accordée dans les protocoles académiques, l'évaluation de l'efficacité des pratiques pédagogiques bute sur la difficulté à objectiver ces pratiques et constitue encore le maillon faible des audits ainsi que des autres formes d'évaluation institutionnelle.

Au total, loin d'être organisés autour de quelques grandes problématiques, les rapports restent descriptifs et prennent souvent l'allure d'un compte rendu de faits ou d'opinions, privilégiant ainsi une vision statique au détriment de l'analyse des tensions et lignes de force qui sont susceptibles de favoriser ou, au contraire, de freiner le développement d'une dynamique d'établissement.

- **2.2.3.2. Des perceptions de l'évaluation contrastées entre évaluateurs et évalués**

- *Le point de vue des académies et de leurs évaluateurs*

Le développement de ces opérations a permis la mise en place de tableaux de bord académiques qui apportent des données et des outils souvent fort riches sur les établissements. La qualité des indicateurs mis à la disposition des évaluateurs prédétermine d'ailleurs, dans une large mesure, la qualité de l'analyse des résultats³⁶.

L'évaluation des EPLE a également permis le développement d'une nouvelle professionnalité au sein des corps d'inspection territoriaux. Elle a en quelque sorte joué le rôle d'une véritable formation sur le tas. Une meilleure compréhension entre inspecteurs et personnels de direction s'est établie et une culture commune d'encadrement prend corps, petit à petit, avec le risque aussi d'en rester là, comme le donne à entendre la remarque de Lise Demailly sur les audits de Lille : « un audit de cadres pour les cadres ».

³⁶ Toulouse compare les résultats du brevet à celle du baccalauréat et Lille a reconstruit des évaluations à l'entrée en seconde afin de permettre une mesure plus fine de la valeur ajoutée en lycée.

Si les audits de Lille ont pu contribuer à l'animation de la vie de l'académie, on n'a pas su capitaliser les connaissances acquises sur les établissements pour améliorer la qualité du pilotage académique. Dans les autres académies, *a fortiori*, le trop faible nombre d'établissements audités chaque année ne permet évidemment pas d'envisager une éventuelle utilisation de l'évaluation des EPLE comme instrument de pilotage académique.

➤ *Le point de vue des évalués*

Les effets de l'évaluation dans les établissements dépendent de facteurs nombreux et divers et il faut donc se garder de conclusions définitives qui ne seraient pas solidement étayées par des preuves. On peut cependant relever quelques grandes tendances. Dans l'enquête réalisée à Lille entre 1995 et 1997, une forte majorité de chefs d'établissement répondent que l'audit ne leur a pas apporté d'informations réellement nouvelles sur leur établissement, soit qu'ils connaissent déjà ce que révèle le rapport, soit qu'ils soient en désaccord avec ses conclusions. Plus généralement, les audits semblent avoir eu un faible impact sur la vie des établissements et sont rarement suivis d'une exploitation en interne. L'absence d'appropriation des conclusions de l'audit peut prendre d'ailleurs différentes dimensions : indifférence, rejet, détournement des conclusions du rapport pour obtenir davantage de moyens, voire détérioration du climat après l'opération. Dans certains cas, cependant, les chefs d'établissement ont exprimé la volonté de poursuivre l'évaluation par la mise en place d'un dispositif d'auto-évaluation.

Les enseignants, quant à eux, manifestent souvent une forte appréhension par rapport à ces nouvelles formes d'évaluation que seul un long travail d'explicitation pourra désarmer. Dans nombre de cas, ils affirment que l'évaluation n'a rien apporté qu'ils ne sachent déjà et qu'elle leur semble peu utilisable pour améliorer leur pédagogie. A Lille, comme ailleurs, on constate que l'audit est d'autant mieux reçu qu'il s'inscrit dans le cadre d'une dynamique préexistante dans l'établissement. C'est donc aux chefs d'établissement de faire vivre l'audit, ce qu'ils ne font pas toujours, comme on l'a vu précédemment.

Une amélioration de la qualité et des effets de l'audit semble moins devoir être recherchée du côté de l'élaboration des protocoles – généralement bien faits – qu'en direction d'une clarification des objectifs poursuivis et d'un effort de formation significatif des corps d'inspection et des personnels de direction.

3. L'évaluation de l'EPLÉ demain : perspectives méthodologiques et stratégiques

3.1. Revisiter l'évaluation de l'EPLÉ

Toute activité d'évaluation suppose un jugement de valeur porté sur l'action des personnes ou tout au moins sur les effets de leurs actions. A ce titre, l'évaluation de l'action publique est souvent perçue comme une menace. Elle suscite appréhension, méfiance et controverses. A tort ou à raison, on en conteste la rigueur, l'objectivité et parfois l'utilité. Cette crispation autour de l'évaluation est particulièrement marquée au sein de l'Éducation nationale où la frontière entre évaluation institutionnelle et évaluation des personnes n'a pas toujours été clairement posée. Pour asseoir sa crédibilité, l'évaluation des EPLÉ doit donc obéir à des règles déontologiques strictes et respecter en quelque sorte un code de bonnes pratiques, comme il en existe pour l'évaluation des politiques publiques, mais qui tienne compte de la spécificité de la mission éducative remplie par les établissements scolaires.

3.1.1. L'évaluation de l'EPLÉ : une démarche spécifique

▪ 3.1.1.1 Qu'est-ce qu'évaluer une action publique ?

L'évaluation de l'action publique est un genre hybride qui se différencie à la fois du contrôle ou de l'audit, d'une part, et des activités relevant de la recherche scientifique, d'autre part, tout en empruntant à l'un et l'autre genre.

Le contrôle consiste à vérifier que les actions sont conformes à des normes explicites qui s'imposent à tous les acteurs. Les écarts à la norme sont généralement considérés comme des anomalies, voire des irrégularités et les personnels concernés peuvent éventuellement être mis en cause et faire l'objet de sanctions. L'audit comporte également une dimension de contrôle par rapport à des normes juridiques ou comptables mais s'élargit, dans l'audit de performances, à la recherche d'une meilleure efficacité de l'organisation auditée. L'évaluation relève d'une démarche plus ouverte : l'appréciation sur laquelle elle débouche peut être adossée à des valeurs et à des normes externes à l'institution ou négociées avec les acteurs évalués qui peuvent être associés plus ou moins étroitement au processus d'évaluation lui-même.

L'évaluation de l'action publique emprunte également aux méthodes de recherche en sciences sociales en ce sens que l'évaluateur, comme le chercheur, doit expliciter ses hypothèses de départ et construire une démonstration rigoureuse qui permette d'en tester la validité. Comme dans une recherche scientifique, l'administration de la preuve constitue un temps essentiel du processus d'évaluation. De sa rigueur et de sa solidité dépend en effet la crédibilité de tout le processus.

L'évaluation se distingue cependant de l'activité de recherche dans la mesure où elle constitue elle-même une dimension de l'action publique ; en tant que telle, si elle vise à produire des effets de vérité, elle doit aussi se préoccuper de son utilité sociale, de sa capacité à modifier le cours des choses, à faire évoluer le comportement des acteurs. Bref, l'évaluation a partie liée avec la prise de décision. En conséquence, elle utilise des méthodes éclectiques et associe aux apports de la science, des compétences d'expertise, inséparables d'une connaissance intime des pratiques de terrain et des modes de prise de décisions au sein des organisations publiques. En outre, à la différence d'une recherche « désintéressée », l'évaluation doit mettre en relation son utilité sociale et son coût. D'où la nécessité d'opérer un arbitrage entre les moyens et le temps consacrés à l'évaluation³⁷ et l'utilité de ses apports. Ce choix est en grande partie politique et dépend de l'importance que les pouvoirs publics attachent à la fonction évaluation dans la régulation du système. Une comparaison rapide des moyens consacrés à l'évaluation de l'EPLÉ montre ainsi que dans nombre de pays, notamment en Angleterre, cette fonction y apparaît plus essentielle qu'en France.

▪ **3.1.1.2. Quelles spécificités présente l'évaluation des établissements scolaires ?**

L'évaluation de l'EPLÉ présente des caractéristiques propres qui la différencient des autres formes d'évaluation de l'action publique. Elle ne porte pas sur l'action d'une politique ou d'un programme ciblé pour une période de temps donnée mais sur le fonctionnement continu d'une organisation complexe aux objectifs pluriels et plus ou moins bien délimités. L'établissement scolaire regroupe en effet des catégories d'acteurs dont les intérêts, les valeurs, les attentes et les aspirations peuvent se révéler différents voire antagoniques, de sorte qu'il est parfois difficile d'imputer à l'établissement des objectifs d'action clairs et indiscutables. De fait, il existe souvent une grande distance entre la politique affichée dans le projet d'établissement et sa traduction réelle sur le terrain par des acteurs qui n'en partagent pas toujours les objectifs.

En outre, parmi les multiples variables qui interfèrent dans la réussite des élèves, nombre d'entre elles ne sont pas sous le contrôle direct de l'établissement. Surtout, l'action éducative ne peut être assimilée à la production d'un bien, fût-il collectif. C'est en quelque sorte une

³⁷ Nombre de personnes mobilisées multiplié par le nombre de journées consacrées à l'évaluation.

co-production entre des personnes que l'on a l'habitude de désigner comme les prestataires de services (personnels de direction, d'éducation et enseignants) et des « usagers » (les élèves).

Or, si on peut obliger l'élève à aller à l'école, on ne peut l'obliger à apprendre. D'où l'importance d'établir un climat de coopération entre les différentes catégories d'acteurs qui composent la population d'un établissement scolaire pour assurer la réussite de l'ensemble. La production d'effets d'apprentissage au sein de l'EPLE est donc en quelque sorte une fabrication conjointe dans laquelle l'élève a souvent une part aussi importante que celle du professeur.

De ces particularités naissent des problèmes spécifiques à l'évaluation de l'EPLE que l'on ne rencontre pas, tout au moins avec le même degré d'acuité, dans d'autres formes d'évaluation de l'action publique. Ils expliquent d'ailleurs en partie l'opposition à l'évaluation d'une frange importante du corps enseignant qui se refuse, avec de plus ou moins bons arguments, à voir les progrès des élèves assimilés à la production de la valeur ajoutée d'un bien économique. Il est vrai que cette inquiétude a pu être alimentée par certaines dérives anglo-saxonnes comme la publication de palmarès d'établissements ou les tentatives d'indexation des rémunérations des enseignants sur les résultats de leurs élèves. De fait, l'abus de tests et de classements et leur utilisation comme méthode de gestion des ressources humaines ne sont pas sans produire des effets pervers comme la réduction du contenu des curricula aux seuls contenus évalués, la recherche d'effets d'apprentissage à court terme au détriment de la construction des compétences à plus long terme, et même la tentation de certains enseignants d'aider leurs élèves lors de la passation des épreuves pour augmenter leur taux de succès.

Il convient donc, d'entrée de jeu, d'élaborer une posture d'évaluation compatible avec la culture professionnelle des acteurs du système éducatif en France. Plus qu'une technique, l'évaluation est d'abord un certain type de démarche, faite à la fois de regard critique mais aussi d'une attitude de compréhension et d'ouverture à l'autre, d'une volonté de faire partager par ceux qui sont évalués certains constats pour faire évoluer la qualité de la formation. La réussite de l'évaluation repose donc principalement sur la capacité des équipes d'évaluateurs à établir un climat de confiance au sein de l'établissement. Ce qui suppose d'abord que tous les acteurs concernés par l'évaluation soient clairement informés à la fois des objectifs, des méthodes et de l'utilisation qui en sera faite ultérieurement ; ce qui suppose également qu'ils soit associés étroitement au processus d'évaluation en ses différentes phases et qu'ils puissent faire part de leurs observations à chaque moment du processus ; ce qui implique, enfin, que soit clairement affirmé que l'évaluation de l'EPLE est une évaluation globale du fonctionnement d'une organisation et de ses résultats et qu'elle n'a pas pour objectif d'imputer tel ou tel effet constaté à la responsabilité de telle ou telle personne. De ce point de vue, l'évaluation institutionnelle doit être clairement distinguée d'autres formes d'intervention des inspections qui ressortissent, peu ou prou, au contrôle de conformité.

3.1.2. *Evaluation interne et évaluation externe : entre tensions et complémentarités*

▪ 3.1.2.1. Une opposition canonique

On oppose classiquement évaluation interne et évaluation externe³⁸ des établissements scolaires. Effectuée par des membres de la communauté scolaire, l'évaluation interne est censée mobiliser les personnels autour d'un projet d'amélioration de la qualité du service éducatif. Réalisée par des « experts » extérieurs à l'établissement, inspecteurs, chercheurs, ou membres de l'administration de tutelle, l'évaluation externe a pour principale finalité d'établir une analyse objective des performances des établissements. Alors que l'évaluation interne doit permettre aux acteurs de « se rendre compte » des effets de leurs actions, l'évaluation externe vise à en rendre compte à leur hiérarchie, à l'utilisateur, au citoyen. Evaluation interne et évaluation externe ne sont donc ni équivalentes ni substituables car leur finalité initiale diffère fondamentalement.

Par ailleurs, l'auto-évaluation n'est que l'une des modalités que peut prendre l'évaluation interne alors que les deux termes sont employés abusivement comme synonymes. Souvent donnée en exemple, portée par plusieurs projets de la Commission européenne, l'auto-évaluation a aujourd'hui « le vent en poupe ». Il s'agit en quelque sorte d'une auscultation de l'école par elle-même. Idéalement, elle suppose à la fois que chacun développe une capacité à avoir un regard critique sur les effets de son action au sein de l'école (dimension individuelle) et que l'ensemble de la communauté scolaire, personnels, parents, élèves, et représentants des collectivités territoriales, se mobilisent pour conduire une analyse réflexive sur la manière dont fonctionne l'établissement (dimension collective). Mais toute évaluation interne ne peut être assimilée à une auto-évaluation. L'évaluation interne résulte en effet, le plus souvent, de l'action, au sein d'un établissement, d'un évaluateur spécifique qui est, en bien des cas, le responsable hiérarchique. Ainsi, l'évaluation d'un enseignant par le chef d'établissement est bien une évaluation interne mais ne relève pas du registre de l'auto-évaluation. De même les diagnostics réalisés par les chefs d'établissement, lors de leur prise de fonction, ne peuvent être considérés comme de véritables auto-évaluations dans la mesure où ils n'impliquent pas toutes les parties prenantes et où ils ne sont pas portés à la connaissance des membres de la communauté scolaire mais seulement communiqués aux représentants de la hiérarchie (recteur et inspecteur d'académie).

³⁸ L'évaluation est ici appelée « externe » lorsqu'elle est effectuée par des agents extérieurs à l'organisation évaluée, qu'ils soient membres ou non de l'Éducation nationale. En matière d'évaluation des politiques publiques, l'adjectif « externe » prend un sens plus restrictif : on l'emploie pour désigner des évaluations réalisées par des instances extérieures à l'administration qui conduit la politique évaluée.

Les arguments en faveur du développement de l'auto-évaluation sont nombreux : les acteurs de terrain ont une connaissance plus intime de ce qui se passe au jour le jour dans l'établissement que des visiteurs occasionnels ; la participation à l'évaluation des personnes directement concernées par le bon fonctionnement de l'établissement semble être la meilleure garantie de leur mobilisation ultérieure au service d'une amélioration de la qualité du service éducatif ; enfin, les acteurs de terrain peuvent aborder en interne des questions vives qu'ils ne sauraient exposer devant des interlocuteurs extérieurs, perçus comme les représentants de la hiérarchie. L'auto-évaluation se voit ainsi parée de toutes les vertus. On ne saurait pour autant sous-estimer ses limites évidentes : les acteurs ont une propension inégale à se distancier de leurs pratiques, ils peuvent céder à l'autosatisfaction et faire preuve d'une certaine complaisance ; ils manquent généralement de la capacité d'expertise requise pour garantir une évaluation rigoureuse ; en outre, le « grand déballage » auquel donnent lieu certaines formes d'auto-évaluation peut faire naître ou réactiver des tensions ou des conflits difficilement contrôlables. Surtout, l'auto-évaluation ne semble pas pouvoir se suffire à elle-même, car l'établissement fait partie d'un ensemble plus large – le système éducatif – et son évaluation se conçoit avant tout de manière relative, par comparaison avec ce qui se fait dans d'autres établissements à profil comparable. Or pour comparer, il ne faut pas simplement disposer d'outils de comparaison mais avoir aussi l'expérience concrète de ce qui se fait ailleurs. Si un effort de transparence et de mutualisation d'expériences entre établissements peut aider fortement à la diffusion de « bonnes pratiques », encore faut-il qu'un arbitre extérieur les authentifie. Même si elle n'est pas exempte de critiques, l'évaluation externe reste donc incontournable, ne serait-ce que pour s'assurer de la qualité de l'auto-évaluation et en valider ou en infirmer les conclusions.

▪ 3.1.2.2. Une opposition en grande partie factice

Dans la pratique, ce que l'on observe, c'est en fait tout un dégradé de situations. Les évaluations combinent généralement différents éléments, externes et internes, selon des doses variables. De ce point de vue quatre dimensions méritent plus particulièrement d'être considérées :

- Qui est à l'origine de l'évaluation ? Est-ce une action spontanée de la communauté scolaire ou bien l'initiative provient-elle d'une impulsion extérieure ?
- L'évaluation se fait-elle en vase clos ou bien associe-t-elle des personnalités extérieures ?
- L'établissement utilise-t-il des indicateurs et instruments d'évaluation qu'il a conçus lui-même ou bien recourt-il à des outils fournis par des experts ?
- Quelle utilisation est faite de l'évaluation ? Sert-elle exclusivement au développement interne de l'établissement ou alimente-t-elle également la

connaissance que l'autorité de tutelle souhaite avoir du fonctionnement des établissements qui relèvent de sa compétence de gestion ?

Idéalement, l'auto-évaluation « pure » est celle qui serait effectuée par l'ensemble des membres de la communauté scolaire, à leur initiative et à partir d'indicateurs et d'instruments qu'ils auraient construits en fonction de leurs besoins propres et dont l'établissement serait seul destinataire des conclusions. Mais dans la pratique, ce cas de figure reste exceptionnel et nombre d'évaluations dites internes font intervenir, à un moment ou à un autre, l'action de personnes extérieures à l'établissement. Le degré de contrôle de l'évaluation par la communauté scolaire est donc généralement partiel : l'évaluation est bien réalisée par les membres de la communauté scolaire, mais souvent à l'initiative de l'autorité de tutelle, et à partir d'outils et de méthodes conçues à l'extérieur par des experts. Elle bénéficie parfois du regard extérieur d'un « ami critique »³⁹ qui favorise par ses questions une certaine distanciation par rapport aux préoccupations internes. Enfin, il arrive aussi que des établissements d'un même bassin de formation s'entendent pour effectuer des évaluations croisées de leurs pratiques.

▪ 3.1.2.3. Des formes d'évaluation complémentaires

De fait, auto-évaluation et évaluation externe entretiennent des rapports complexes faits de tensions et de complémentarités. Elles peuvent s'enrichir mutuellement par fertilisation croisée selon un processus en boucle : l'auto-évaluation peut servir de point de départ ou de point d'appui à l'évaluation externe ; l'évaluation externe peut favoriser la mise en place d'une auto-évaluation ou, quand elle existe, en améliorer la qualité par les effets d'apprentissage qu'elle induit. La solution la plus réaliste n'est-elle pas de faire cohabiter les deux selon des modalités qui restent à préciser ? A cet égard, l'utilisation d'un noyau commun d'indicateurs aux deux formes d'évaluation présente le double avantage de construire un langage commun à tous les échelons du système et de rendre commensurables évaluations externes et internes. Il ne faut pas se cacher, pour autant, la réticence des établissements à effectuer une évaluation interne en fonction de critères et d'indicateurs « prêts à l'emploi » dont ils ressentent bien qu'ils répondent davantage aux exigences de la tutelle qu'à leurs besoins propres. S'il apparaît souhaitable de fixer un cadre à l'auto-évaluation, il est tout aussi nécessaire de laisser aux établissements des marges de manœuvre significatives pour « créer » en complément des indicateurs et des instruments d'évaluation adaptés aux problèmes spécifiques qu'ils peuvent rencontrer dans leur environnement.

³⁹ Le « critical friend » est un personnage clé des auto-évaluations réalisées dans les pays anglo-saxons.

3.1.3. Construire une problématique de l'évaluation

▪ 3.1.3.1. Des thématiques aux questionnements

Une entrée dans l'évaluation par l'analyse de différents domaines ou thématiques que l'on peut inventorier au sein d'un établissement présente des limites évidentes. Une telle approche réduit le processus évaluatif à vérifier que l'établissement a bien mis en place l'ensemble des actions et des dispositifs figurant sur une *check-list* dont les items peuvent être allongés à l'infini. Elle débouche sur la tentation de l'exhaustivité alors qu'il est illusoire de penser pouvoir analyser toutes les dimensions d'un établissement et rendre compte de l'action de l'ensemble de ses diverses composantes. Elle conduit au mieux à une simple recension de ce qui se fait, ou ne se fait pas, sans s'interroger sur l'efficacité et l'impact des actions. A vouloir tout analyser, on risque, au pire, d'en rester à un simple compte rendu de ce que les responsables ont pu eux-mêmes écrire ou dire de leurs pratiques et, au mieux, de ne mettre en évidence que des traits de surface qui apporteront peu d'informations pertinentes.

Voilà pourquoi, il paraît préférable de construire l'évaluation à partir d'une série de questionnements ordonnés autour de quelques grandes problématiques. Déterminer ce qui est prioritaire et ce qui est secondaire doit être l'une des premières préoccupations d'une évaluation bien faite. Évaluer un établissement signifie donc faire des choix, sélectionner dans l'ensemble des données disponibles et qui peuvent être interrogées, certaines d'entre elles qui paraissent susceptibles de permettre à l'établissement de « faire la différence » ou qui révèlent, au contraire, des défaillances. L'approche par les résultats, à condition de s'entendre bien évidemment sur ce que cela signifie (cf. supra II), semble devoir être privilégiée, car la mission première d'un établissement scolaire est bien de faire réussir ses élèves. Le choix de tel mode d'organisation, la mise en œuvre de tel ou tel dispositif, qu'il soit dans l'air du temps ou qu'il satisfasse simplement une demande de l'autorité hiérarchique, n'est qu'un moyen au service de cette réussite.

Il ne s'agit pas seulement de mesurer des résultats mais d'examiner aussi comment et par quels mécanismes et modes de transmission la politique mise en œuvre au sein de l'établissement produit des effets plus ou moins positifs sur les acquisitions et les comportements des élèves. A cet égard, l'opposition entre une évaluation centrée sur les produits du système éducatif et une évaluation centrée sur les processus qui servent à les expliquer doit être dépassée. Ce qu'il faut, c'est articuler les deux : mesurer les acquisitions des élèves et les relier aux pratiques professionnelles – pédagogiques et organisationnelles – qui ont contribué à les faire advenir dans un contexte donné. Il ne s'agit pas tant d'ailleurs de repérer les différents facteurs ou variables qui influent sur la réussite des élèves, mais bien plutôt d'analyser comment ils interagissent entre eux pour constituer des configurations efficaces.

Enfin, évaluer revient donc à répondre à quelques grandes questions. Certaines sont communes à tous les établissements car elles correspondent aux grandes missions de notre système éducatif. D'autres sont en relation avec le positionnement de telle ou telle catégorie d'établissements et les problèmes particuliers que l'on peut y rencontrer. A cet égard, la mission dévolue aux collèges doit être clarifiée : l'objectif du collège est-il de faire acquérir par tous les élèves un socle de références communes (connaissances et compétences sociales) dans le cadre de la scolarité obligatoire, de préparer à la poursuite d'études en lycée général et technologique ou en lycée professionnel, ou encore de mener de front ces deux missions ? On comprendra que selon le choix opéré, le point focal de l'évaluation va se déplacer. L'évaluation des lycées professionnels conduit également à se poser des questions spécifiques à la nature des formations offertes : celles qui sont relatives à l'adaptation de l'offre de formation aux besoins de l'économie, à l'insertion professionnelle, aux relations avec les entreprises, à la continuité de l'offre et des parcours de formation, à l'adéquation des équipements aux référentiels professionnels, etc. En lycée d'enseignement général et technologique, les questions concernant la détermination des choix d'option en classe de seconde, celles relatives à l'orientation des élèves vers les différentes séries du baccalauréat et à leur devenir dans l'enseignement supérieur présentent également un intérêt tout particulier.

Au-delà, on peut et on doit considérer que chaque établissement présente des singularités, parfois une identité propre, qui ne le fait ressembler à aucun autre. Un équilibre sera donc à trouver entre recherches extensives et investigations intensives, entre le « prêt-à-porter » obéissant à des protocoles relativement standardisés et le « sur-mesure » qui suppose d'imaginer des modes d'investigation originaux susceptibles de rendre compte des différents profils d'établissement.

▪ 3.1.3.2. Evaluation et référentiels

Évaluer un établissement scolaire, c'est mesurer les effets observés au regard d'un référentiel qui fixe les objectifs à atteindre. L'EPL peut être évalué à partir d'un double référentiel, externe et interne. Les grandes finalités de l'action éducative sont fixées par la loi d'orientation de 1989 : développer la personnalité de l'élève ; transmettre des connaissances et faire acquérir des compétences ; former à la citoyenneté ; préparer à l'insertion sociale et professionnelle. Est également affirmée une finalité transversale : l'égalité des chances scolaires. Mais, en dehors de l'objectif de mener 80 % d'une classe d'âge au niveau du baccalauréat et de faire accéder tous les élèves au moins au niveau V de formation, la loi ne fixe pas de cibles chiffrées. C'est donc aux projets académiques que revient la tâche de décliner les grandes orientations nationales en déterminant des objectifs plus opérationnels. Mais la loi d'orientation fait également obligation à chaque établissement d'élaborer un projet qui lui permette de mettre en avant ses propres priorités en fonction du contexte dans lequel il

évolue et des problèmes concrets qu'il doit résoudre. C'est donc à l'aune de ce double référentiel que les résultats de l'EPLE devront être appréciés⁴⁰.

Dans la pratique, les résultats de l'EPLE peuvent être mesurés à partir de trois référents. L'évaluation des acquis des élèves peut se faire par rapport à un seuil minimum de compétences qui doit être atteint par tous les élèves ou, tout au moins, par une certaine proportion d'entre eux. L'évaluation se fait alors en termes absolus. Il n'existe pas en France de tels standards de compétences pour chaque niveau de cursus. En conséquence, sauf à ce que l'établissement s'en soit fixé un, l'évaluation sera relative. Les performances de l'établissement seront comparées à des moyennes nationales, académiques ou départementales ou aux performances des établissements ayant des caractéristiques comparables. On peut également suivre l'évolution des résultats au sein d'un même établissement à deux moments du temps. On déterminera alors dans quelle mesure il a atteint, ou pas, les objectifs de progression qu'il s'était lui-même fixés. Le croisement des deux types d'évaluation – dans l'espace et dans le temps – est en réalité nécessaire car l'établissement peut toujours se fixer des objectifs exagérément modestes ou, au contraire, irréalistes. Seule la comparaison des évolutions des résultats de l'établissement évalué par rapport à celles des autres établissements est susceptible de donner un éclairage pertinent sur la réalité des progrès accomplis au regard des évolutions générales.

▪ 3.1.3.3. Critères d'évaluation et indicateurs

En matière d'évaluation, s'il est relativement facile de s'accorder sur quelques grands objectifs, c'est parfois au prix de bien des malentendus. Dès lors qu'il s'agit de préciser ce que l'on met derrière les mots, peuvent apparaître des clivages jusque-là insoupçonnés. Une évaluation rigoureuse se doit donc de décomposer les objectifs en critères d'évaluation et d'assigner à chaque critère un ou plusieurs indicateurs permettant de mesurer le degré auquel il est atteint.

La distinction entre critères et indicateurs est souvent source de confusions et mérite donc clarification. Les critères servent à spécifier les différentes dimensions d'un objectif tandis que les indicateurs permettent de mesurer le degré de réalisation de chaque critère. Objectifs, critères et indicateurs se déclinent donc selon un axe qui va du plus général au plus détaillé. A titre d'exemple, l'amélioration du climat au sein de l'établissement pourra être considérée comme un objectif à évaluer ; il sera décliné en plusieurs critères qui permettront d'en cerner les différentes dimensions : l'un de ces critères pourra être une plus forte implication des

⁴⁰ Les directives de la commission européenne fixent des objectifs chiffrés pour 2010 qui devront également être pris en compte dans évaluations à venir.

parents dans la vie de l'établissement. La réalisation de ce critère se mesurera à partir d'indicateurs : par exemple, l'évolution du taux de participation des parents aux élections de délégués, au conseil d'administration, aux conseils de classe, l'évolution de la demande d'entretiens avec les professeurs, etc. Selon que les différents indicateurs donneront des visions concordantes ou discordantes du phénomène étudié, on pourra conclure avec une plus ou moins grande sûreté, à la réalisation du critère.

Tout indicateur est une représentation chiffrée d'un phénomène que l'on veut contrôler. Lorsque les données sont directement mesurables, il se présentera sous la forme d'un pourcentage, d'un ratio. Ainsi en est-il des indicateurs d'intrants, d'activité, de résultats ou de contexte. Mais un indicateur peut également servir à mesurer des éléments qualitatifs portant sur les processus : la qualité du pilotage, les formes d'organisation, les pratiques pédagogiques, les modes d'apprentissage. On attribuera alors aux pratiques observées un score sur une échelle de valeurs plus ou moins standardisée. Plusieurs pays ont défini de la sorte des scores de notation pour mesurer des processus qualitatifs qui vont de 1 (excellent) à 4 (médiocre), ou de 1 à 7 (Angleterre). En réduisant, somme toute, le qualitatif à du quantitatif, cette approche permet un accroissement de formalisation et rend commensurables des situations qui ne l'étaient pas *a priori*. Mais elle réduit finalement l'évaluation à une métrologie. Dans le cadre français où la culture de l'évaluation est encore mal assise, une telle évolution, à supposer d'ailleurs qu'elle soit souhaitable, risque de susciter la défiance d'une partie du corps enseignant.

De fait, l'opérationnalisation des objectifs sous forme d'indicateurs est à la fois réductrice et indispensable. Réductrice, car tout n'est pas quantifiable dans l'éducation ; et il n'est pas même évident que ce qui est quantifiable soit le plus important, le plus pertinent de ce qu'il faut évaluer. Indispensable, car sinon l'évaluation en reste au niveau de l'indescriptible, de l'invérifiable, finalement de simples procédés rhétoriques.

Si l'évaluation ne peut se passer d'indicateurs, il convient cependant de fixer quelques règles de précaution dans leur utilisation. Les indicateurs doivent servir à répondre à des questions que l'on juge pertinentes alors que trop souvent, c'est, au contraire, la disponibilité ou l'accessibilité des indicateurs qui déterminent les questions que l'on se pose. L'évaluation ne saurait donc se réduire à une simple compilation d'indicateurs, car un indicateur ne dit rien par lui-même. Il n'est qu'un instrument qui ne prend sens qu'interprété dans un contexte et croisé avec d'autres données statistiques ou d'autres observations.

Le risque existe pour tout système d'évaluation d'être emporté par une véritable « quantophrénie » et de tendre à considérer que tout peut se mesurer par des indicateurs et même que l'évaluation se réduit à ce que mesurent les indicateurs. Il convient donc de rappeler que les indicateurs sont simplement des signaux qui, lorsqu'ils sont bien utilisés,

permettent de repérer des tendances, des anomalies. Ils constituent, et ce n'est pas rien, une base objective de discussion avec les représentants de l'établissement évalué. Mais l'évolution d'un indicateur laisse généralement ouvertes plusieurs interprétations concurrentes, voire contradictoires. Ainsi la baisse d'un taux de redoublement peut aussi bien signifier une amélioration des acquisitions des élèves, un plus grand laxisme dans la politique d'orientation, l'obéissance à une injonction hiérarchique, la volonté « d'évacuer » les élèves dont on n'attend plus rien, etc. Seules des observations plus qualitatives et le croisement des informations permettront de trancher entre ces différentes options.

3.2. Améliorer l'évaluation de l'EPLÉ dans le cadre de l'existant

Il s'agit ici, non pas de proposer un protocole d'évaluation en tant que tel, mais seulement de fixer quelques grandes orientations qui pourront servir de cadre aux évaluations à venir.

3.2.1. Assurer une meilleure transparence à l'évaluation externe et associer davantage l'établissement à ses différentes phases

Une attention toute particulière doit être apportée au processus de l'évaluation lui-même, à l'évaluation « en train de se faire » car il est bien connu que la majeure partie des effets de l'évaluation se produit au cours de son déroulement et non après la remise du rapport. En conséquence la qualité de l'évaluation ne doit pas être seulement considérée en fonction de la qualité du contenu du rapport mais aussi à la lumière de la dynamique et des mobilisations que le processus même de l'évaluation aura su, ou non, enclencher.

Les évaluations actuellement pratiquées par les inspections générales ou par les équipes d'évaluateurs académiques distinguent clairement trois phases : la préparation de l'évaluation, son déroulement et sa restitution. L'enchaînement de ces trois phases gagnerait à être mieux articulé et devrait associer plus étroitement les membres de l'établissement à la détermination des principaux enjeux de l'évaluation. En particulier, une plus grande attention doit être portée à la phase amont qui détermine toute la qualité du déroulement de la suite de l'opération. Des améliorations peuvent être apportées à deux niveaux : celui de l'attention portée à l'analyse du dossier qui servira à l'élaboration d'un pré-diagnostic, d'une part, et celui de l'explication de la démarche d'évaluation aux représentants de l'établissement, d'autre part. La constitution du dossier documentaire doit être à la fois plus sélective et plus solide de façon à ce qu'il soit plus facile à exploiter. Rien ne sert de multiplier les demandes de documents, si on n'a pas le temps d'en faire une analyse approfondie, d'autant que les acteurs locaux peuvent manquer de disponibilité pour répondre à des demandes plus

spécifiques, compte tenu de la multiplicité des tâches quotidiennes qu'ils doivent remplir. Le dossier de base doit comprendre néanmoins le projet d'établissement et les différents projets pédagogiques, le diagnostic du chef d'établissement et sa lettre de mission, les procès verbaux des conseils d'administration des trois dernières années et les données statistiques fournies par la DEP, le rectorat et l'établissement.

L'analyse de ce dossier fera l'objet d'une première étude par le pilote de l'opération qui établira une sorte de pré-diagnostic mettant en évidence quelques grandes hypothèses de travail que le déroulement *in situ* de l'évaluation aura pour objet de confirmer ou d'invalidier. Ce premier diagnostic sera présenté en interne au sein du groupe des évaluateurs, discuté et éventuellement amendé. Les points principaux ainsi que les pistes de recherche qu'ils induisent seront communiqués au chef d'établissement qui pourra faire toutes remarques qu'il jugera utiles. Ce n'est qu'après la finalisation de cette première démarche contradictoire que seront arrêtés les domaines devant faire l'objet d'investigations plus poussées. Des documents complémentaires pourront alors être demandés à l'établissement en fonction des pistes d'évaluation privilégiées. Parallèlement les objectifs de l'évaluation, sa méthodologie et l'utilisation ultérieure qui en sera faite devront être clairement exposés aux différents interlocuteurs au sein de l'établissement (direction, représentants des personnels, des parents et des élèves) oralement ou par un courrier. La concertation entre l'équipe d'évaluateurs et les représentants de l'établissement doit être authentique et déboucher, dans la mesure du possible, sur un premier repérage partagé des forces et faiblesses de l'établissement et des questionnements qu'elles suscitent.

Le déroulement de l'évaluation à proprement parler comprend, au moins pour ce qui concerne les évaluations réalisées par des équipes académiques, à la fois des entretiens individuels ou avec des groupes, d'une part, et des observations de situations de pilotage, d'enseignement ou de vie scolaire, d'autre part. L'expérience montre qu'entretiens et observations dans les classes sont trop souvent déconnectés l'un de l'autre. Outre une mauvaise synchronisation dans le temps des visites d'inspection disciplinaires par rapport à la phase d'évaluation dite globale, il apparaît que ces observations de classe sont encore trop souvent calquées, quant à l'esprit et aux modalités de leur déroulement, sur le modèle des inspections individuelles. Il conviendrait que les inspecteurs qui interviennent au titre de leur seule discipline soient donc mieux associés à l'ensemble de l'opération, de manière à pouvoir articuler les observations de classes à la problématique globale d'évaluation. Cela suppose qu'ils aient eu, préalablement à leurs visites, connaissance du pré-diagnostic et des axes d'évaluation ; cela implique également que leurs observations se fassent à partir de grilles d'observation harmonisées entre disciplines pour permettre de capitaliser les états des disciplines.

A la fin du déroulement de l'opération, de premières conclusions orales devront être présentées au chef d'établissement et aux représentants de l'établissement pour qu'ils

puissent, de nouveau, faire part de leurs observations et de leurs désaccords éventuels. Ces remarques pourront être prises en compte dans la rédaction du rapport final si elles semblent pertinentes à l'équipe d'évaluateurs. Dans le cas où les appréciations portées sur les faits observés restent irréductiblement contradictoires, la contribution des représentants de l'établissement sera annexée au rapport. La version définitive du rapport, intégrant éventuellement cette annexe, sera alors communiquée officiellement au chef d'établissement, aux membres du conseil d'administration, au recteur et à l'inspecteur d'académie (et au ministre s'il s'agit d'un rapport réalisé par les inspections générales) dans un délai rapide qui devra être déterminé à l'avance dans le cahier des charges.

Le rapport devra comporter des recommandations à la fois réalistes et opérationnelles qui devront faire l'objet de la part de l'établissement d'un plan d'amélioration comportant des engagements précis ainsi qu'un échéancier. Au vu des conclusions du rapport, des mesures d'aide et de soutien pourront être prises par l'autorité de tutelle et un suivi systématique des évaluations sera opéré. On peut envisager, par ailleurs, de publier ultérieurement ces rapports – qui devraient à terme concerner tous les établissements – sur le site internet de chaque académie, à l'exemple de ce qui s'effectue dans certains pays étrangers.

3.2.2. Recentrer l'objet de l'évaluation sur la construction des performances au sein de l'EPL

L'évaluation peut conduire à explorer plus particulièrement telle ou telle dimension jugée essentielle dans le mode de fonctionnement de l'EPL. On ne saurait à cet égard les prédéterminer car elles dépendent de paramètres liés à la localisation, à la nature du public accueilli, au niveau et aux types d'enseignement dispensés. Il est cependant possible de fixer quelques axes communs à l'ensemble des établissements qui peuvent être regroupés sous forme de trois grands questionnements tournant tous autour de la construction des performances au sein de l'EPL : quelles sont les performances de l'établissement ? En quoi et comment le fonctionnement de l'établissement permet-il de les expliquer ? Quel est le niveau de maîtrise de l'établissement sur son environnement ?

▪ 3.2.2.1. Quelles sont les performances de l'EPL ?

Les performances de l'EPL peuvent être examinées à partir de trois perspectives qui se complètent : l'efficacité appréhendée à partir de la fluidité des parcours scolaires et des résultats des élèves, l'équité évaluée à travers sa capacité à réduire les inégalités de réussite entre élèves, et enfin, l'efficience qui jauge la capacité de l'établissement à utiliser de manière optimale ses ressources.

L'efficacité peut être définie comme la capacité de l'école à faire progresser les élèves dans les acquisitions scolaires et la maîtrise de compétences sociales, au cours d'une période de temps donné, au-delà de ce que l'on pouvait attendre compte tenu de leur niveau initial et de leurs caractéristiques sociales. Elle s'apprécie d'abord, à orientation et taux d'accès comparables, en termes d'acquisitions disciplinaires, telles qu'elles peuvent être approchées par les évaluations et les examens ; mais aussi en termes de compétences transversales – capacité à se documenter, à résoudre des problèmes, à communiquer, à faire preuve d'esprit critique – qui s'exercent plus particulièrement dans le cadre des nouveaux dispositifs pédagogiques (TPE, PPCP, IDD). Au-delà, il serait souhaitable de prendre également en compte les compétences dites socio-affectives comme le développement de la sociabilité, de la solidarité, l'adhésion aux valeurs démocratiques, l'estime de soi, la motivation pour les études. Plus difficilement opérationnalisables et mesurables en l'état actuel de notre système d'informations, ces dernières données ne peuvent être appréhendées que par des observations de situations de classe ou de vie scolaire ou à partir d'entretiens avec des élèves.

L'évaluation ne doit pas seulement permettre de mesurer le niveau de progression des élèves ou leur taux de réussite mais aussi la capacité de l'école à réduire les inégalités initiales par la détection et la prise en charge effective de la difficulté scolaire. Il convient donc de mesurer la capacité de l'établissement à réduire les inégalités de performances entre les élèves les plus faibles et les plus forts, d'une part, les plus favorisés socialement et les défavorisés⁴¹, d'autre part, tout en élevant le niveau moyen de chaque catégorie. Les évolutions des écarts des taux d'accès et de réussite au sein d'un établissement, appréhendées à partir d'indicateurs de dispersion ou de disparité sociale sont évidemment à comparer aux tendances académique ou nationale.

La mesure de l'efficacité doit conduire à mettre en relation les résultats des élèves avec le coût par élève. Un établissement sera jugé efficace s'il parvient pour un coût par élève, moyen ou faible, à obtenir de meilleurs résultats (exprimés en valeur ajoutée) que la moyenne des autres établissements ou encore s'il obtient des résultats bons ou moyens, pour un coût par élève inférieur à la moyenne. Cette question ne peut être abordée avec toute la rigueur souhaitée en l'état actuel de notre comptabilité qui ne permet pas encore – dans l'attente de la mise en œuvre de la LOLF – de calculer des coûts complets. Il est néanmoins possible d'apprécier la capacité de l'établissement à optimiser la gestion de ses moyens en mettant en relation la variété et la qualité de l'offre de formation (nombre d'options offertes, statut des enseignants, équipement informatique et documentaire, ratios d'encadrement) et les résultats obtenus par les élèves aux évaluations ou aux examens. La question mérite plus particulièrement d'être posée pour les établissements situés en ZEP pour lesquels on peut se

⁴¹ La catégorie « défavorisée » au sens de la direction de l'évaluation et de la prospective (DEP) correspond *grosso modo* à l'ensemble constitué par les enfants dont les parents sont ouvriers ou inactifs.

demander dans quelle mesure les moyens supplémentaires octroyés (en moyenne de l'ordre de 10%) ont contribué ou non à améliorer la réussite scolaire de leurs élèves.

▪ **3.2.2.2. Quelle est la contribution de l'établissement à la production des résultats et par quels processus son influence s'exerce-t-elle ?**

Les résultats des élèves sont le produit de l'action de l'école mais sont également influencés par des variables contextuelles. Il importe donc de discriminer dans l'évaluation ce qui correspond à un effet propre de l'école et ce qui s'explique par un effet induit exogène. Si un établissement ne peut être entièrement tenu responsable de mauvaises performances qui sont liées à un niveau initial faible des élèves ou aux difficultés sociales du public scolaire qu'il recrute, inversement, il ne peut davantage s'enorgueillir de résultats qui seraient produits en dehors de lui, dans et par le milieu familial, par exemple.

Par la prise en compte de la composition sociale du public et des retards d'âge scolaire, la mesure de la valeur ajoutée est censée neutraliser ces effets externes. Elle ne le fait cependant qu'imparfaitement car la DEP ne retient que les effets liés aux caractéristiques sociales individuelles des élèves alors qu'il existe une influence du contexte, notamment de la « tonalité sociale » de l'établissement ou de la classe, qui se surajoute au simple effet d'agrégation des caractéristiques individuelles des élèves. Cette influence s'explique par toute une série de mécanismes qui se renforcent mutuellement : le groupe social statistiquement dominant au sein d'un établissement ou d'une classe influe sur le niveau des aspirations et le rapport aux normes scolaires ; il marque de son empreinte le climat de l'établissement : atmosphère plus ou moins studieuse, présence ou absence de problèmes de discipline, degré de vigilance parentale ; un établissement favorisé attire généralement les professeurs les plus expérimentés et les plus qualifiés et vice-versa. Un cercle vertueux ou vicieux s'enclenche ainsi. Voilà qui explique d'ailleurs pourquoi la DEP constate généralement que les établissements les plus favorisés socialement produisent davantage de valeur ajoutée que les autres. Il s'agit là en fait d'un artefact produit par l'instrument de mesure statistique lui-même qui ne prend pas en compte les effets contextuels liés à la composition sociale des publics scolaires.

Une évaluation juste se doit donc d'examiner l'impact de la concentration plus ou moins grande dans l'établissement des catégories favorisées ou défavorisées ; elle doit aussi estimer de manière plus fine le degré de la difficulté sociale en intégrant des données sur les familles monoparentales, le niveau du chômage, la précarité sociale, la proportion d'élèves boursiers au niveau le plus élevé, etc. Ce n'est pas là céder à un fatalisme sociologique, mais tout simplement se donner les moyens de mesurer ce qui est vraiment imputable à l'action propre de l'établissement. Faute de prendre en compte de manière rigoureuse toutes ces variables, on risque de discréditer tout le processus d'évaluation lui-même. On s'empêche surtout de pouvoir prendre la mesure de l'ampleur et de la nature des améliorations et actions à mettre en

œuvres qui peuvent souvent déborder le cadre strict de l'établissement : développement de la mixité sociale au sein du bassin d'éducation, établissement de relations fortes avec les familles, etc. Il est bien évident que le constat de difficultés sociales particulières ne saurait pour autant servir de caution à des comportements de résignation face à la difficulté scolaire, comme on l'observe trop souvent.

Une fois identifié l'impact propre de l'établissement sur les résultats des élèves, il reste à examiner comment l'action de la communauté éducative, prise dans ses différentes composantes, contribue à produire les effets observés. En d'autres termes, il faut expliquer pourquoi l'effet « établissement » est plus que la simple addition de l'action de chacune de ses composantes. Se pose ainsi la question de l'articulation des actions menées aux différents niveaux du système et de leur synergie. On peut faire l'hypothèse d'un fonctionnement en termes de niveaux emboîtés. Le niveau « supérieur », la direction de l'établissement, cadre l'action des autres niveaux, (équipes disciplinaires, professeurs) et conditionne le degré de leur efficacité par son rôle de facilitation, de blocage ou encore d'inertie. Le pari peut être fait que l'instauration d'un climat de calme, de travail et de respect d'autrui au sein de l'établissement se répercutera nécessairement sur les conditions de fonctionnement des classes, et partant, sur les possibilités de progression des élèves.

L'appréciation de la qualité du pilotage n'appelle pas cependant de réponse unilatérale car le succès d'un mode de management ne dépend pas seulement de ce que font ceux qui l'exercent, mais aussi de la réceptivité de ceux auxquels il s'adresse. Autant dire qu'évaluer l'efficacité du pilotage suppose d'analyser l'ensemble des relations d'interdépendance entre acteurs (organigramme formel) et la nature des interactions qu'ils entretiennent effectivement (organigramme réel) et donc de confronter les propriétés objectives de l'organisation (structure, objectifs, styles de leadership) et les perceptions qu'en ont les agents en fonction de leurs systèmes de valeurs, de leurs conceptions des missions du service public et de leurs attentes. L'engagement des acteurs dans leur travail et leur implication dans la vie de l'établissement dépendent, en conséquence, autant des perceptions – plus ou moins positives – qu'ils ont du climat de l'établissement que des seules caractéristiques objectives de l'organisation.

Chacun admet aujourd'hui que les effets « classe » et les effets « maître » sont bien plus importants que l'effet « établissement » *stricto sensu*. On ne saurait donc concevoir une évaluation de l'EPL qui ferait l'impasse sur les représentations et les pratiques professionnelles des enseignants et sur les processus d'apprentissage des élèves. Le choix fait, dans la plupart des académies qui se sont livrées à des audits d'établissement, de faire procéder, parallèlement à l'évaluation globale de l'établissement, à des observations en classe ou dans d'autres situations éducatives paraît, à cet égard, un acquis incontournable. Il reste à

en redéfinir les modalités pour rendre les comptes rendus d'observation plus facilement utilisables dans le rapport de synthèse (cf. infra 3).

▪ **3.2.2.3. L'établissement maîtrise-t-il ou subit-il son environnement ?**

L'établissement scolaire est une entité ouverte sur son environnement dont il subit *a priori* les contraintes ou bénéficie des atouts. La capacité des établissements à endogénéiser les données environnementales par une action résolue en direction de leurs partenaires est extrêmement variable et constitue la troisième dimension dont doit rendre compte toute évaluation. Trois attitudes peuvent être repérées : l'établissement subit sans réagir un environnement qui favorise ou entrave son développement (comportement d'inertie) ; l'établissement adapte sa stratégie à l'évolution de son environnement (comportement réactif) ; l'établissement conduit une action de nature à transformer certaines contraintes extérieures en ressources (comportement proactif).

La capacité de l'établissement à agir sur son environnement devra être examinée de plusieurs points de vue. D'abord dans le rôle plus ou moins moteur qu'il peut jouer au sein du bassin d'éducation, ou du réseau d'éducation prioritaire, pour organiser des formations communes des personnels, mutualiser les ressources pédagogiques, adapter la carte des formations aux besoins sociaux⁴² ; ensuite, dans la nature et les modalités de sa relation – verticale et latérale – avec les autres EPLE.

Ce qui se passe dans un établissement ne peut en effet être isolé de ce qui se déroule en amont et en aval : pour le collège les relations avec les écoles de son secteur de recrutement et les lycées ; pour les lycées, celles avec les collèges viviers et l'enseignement supérieur ; pour le lycée professionnel avec ces mêmes collèges, les LEGT (classe de première d'adaptation, de STS) et, dans un autre registre, l'entreprise. Que fait l'établissement pour s'assurer que la nature et les niveaux des exigences des professeurs sont bien articulés d'un ordre d'enseignement à l'autre, que les attentes et les méthodes de travail sont bien en continuité ? Existe-t-il une analyse du devenir des élèves et des résultats aux différentes évaluations pour s'assurer qu'elles concordent et, dans le cas contraire, expliquer les variations constatées et y apporter des remédiations ?

Le rayonnement d'un établissement est aussi délimité par son attractivité relative par rapport aux établissements voisins et concurrents. De ce point de vue, les situations sont différentes en zone rurale où chaque établissement se trouve dans une situation de quasi-monopole et dans les agglomérations urbaines où il existe plusieurs établissements avec des réputations et des publics contrastés. L'image extérieure de l'établissement, et partant, sa réputation,

⁴² L'étude du microsystème constitué par l'ensemble des établissements du bassin d'éducation peut être, à cet égard, particulièrement éclairante.

peuvent alors être appréhendées à partir de sa zone d'attractivité et des demandes de dérogation en sa faveur ou pour l'éviter.

Enfin l'action menée en direction des parents d'élèves des collectivités territoriales et, pour les lycées professionnels, des partenaires économiques, constitue également un paramètre clé pour évaluer le degré de réactivité de l'établissement à son environnement et sa capacité à le mobiliser en faveur de son action éducative. Il faut bien reconnaître cependant que dans ce domaine les indicateurs qui pourraient permettre de se forger un jugement étayé sur des données comparatives solides restent encore à construire pour l'essentiel, même si des éléments tels que le nombre de conventions passées avec les entreprises, l'évolution du volume de la taxe d'apprentissage collectée, l'insertion des élèves dans une branche correspondant à leur diplôme peuvent fournir une première indication.

3.2.3. Mieux utiliser les indicateurs d'évaluation disponibles

Le système éducatif français dispose d'un nombre impressionnant d'indicateurs statistiques. Ils restent malheureusement encore méconnus, sous-utilisés et trop souvent mal interprétés. Comme le relevait un récent rapport des deux inspections générales⁴³, il y a d'ailleurs quelques paradoxes à se plaindre du manque d'indicateurs et, dans le même temps, à ne pas utiliser ceux qui existent. Soit on ne les connaît pas ou on ne sait pas comment les utiliser, soit on les juge inadaptés aux besoins de l'évaluation et, dans ce cas, il faut alors dire par quoi ils devraient être remplacés et quelles explications complémentaires sont nécessaires pour rendre leur utilisation plus aisée.

▪ 3.2.3.1. Prendre davantage appui sur les indicateurs de la DEP

L'évaluation des lycées⁴⁴ devrait tirer un meilleur parti des renseignements fournis par les indicateurs produits par la DEP qui restent insuffisamment mobilisés aussi bien dans les auto-évaluations qui accompagnent l'élaboration des projets d'établissements que dans les évaluations externes à initiative académique⁴⁵. Les indicateurs pour le pilotage des établissements secondaires (IPES) fournissent pourtant des informations précieuses sur les taux d'accès et de réussite au baccalauréat calculés à la fois en taux bruts et en valeur ajoutée par référence aux niveaux académique et national. Le croisement de ces deux types d'indicateurs permet de délimiter un premier profil des établissements en termes d'efficacité et de sélectivité.

⁴³ Rapport Duval, Saurat, op. cité.

⁴⁴ L'absence de la publication des taux de réussite au diplôme national du brevet dans les IPES oblige à ne considérer ici que les seuls lycées.

⁴⁵ Il n'y aurait d'ailleurs que des avantages à rendre les IPES directement accessibles à l'ensemble des personnels de l'EPLÉ voir à les rendre publics.

Typologie des établissements en fonction de leur degré de sélectivité
et de réussite au baccalauréat selon les indicateurs de la DEP

	Indice de valeur ajoutée du taux d'accès seconde/bac supérieur à 100	Indice de valeur ajoutée du taux d'accès seconde/bac inférieur à 100
Indice de valeur ajoutée du taux de réussite au bac supérieur à 100	Établissements les plus performants	Établissements sélectifs
Indice de valeur ajoutée du taux de réussite au bac inférieur à 100	Établissements accompagnateurs	Établissements les moins performants

Source : tableau constitué à partir de la note de la DEP du 04-03-2004

L'analyse des résultats peut ensuite être affinée à partir de la recension des notes obtenues par les élèves au baccalauréat. La DEP a en effet introduit depuis l'an dernier un nouvel « univers » statistique dans la Base centrale de pilotage (BCP) qui couvre de manière détaillée les résultats au baccalauréat : cette base de données permet de connaître les moyennes et la distribution des notes pour chaque discipline au sein d'un établissement par comparaison aux autres échelons géographiques. On peut dès lors mesurer le degré d'homogénéité ou d'hétérogénéité des résultats au baccalauréat entre les différentes disciplines et déterminer ce qui relève d'un effet d'établissement (homogénéité forte des résultats d'une discipline à l'autre), d'un effet « maître » (résultats atypiques de telle ou telle discipline) ou encore d'un effet « cru » (résultats instables d'une année à l'autre pour une même discipline). Il devient également possible d'effectuer un croisement de ces résultats avec les catégories socioprofessionnelles, le sexe, les retards d'âge scolaire et le nombre de fois où l'élève s'est présenté à l'examen. Un nouveau champ s'ouvre ainsi à l'évaluation des acquisitions cognitives au sein des disciplines représentées dans l'établissement.

Pour les collèges, les IPES ne donnent que les taux d'accès de la 6^{ème} en troisième, d'une part, et de troisième en seconde, d'autre part. La mise en relation des taux d'accès en seconde lors de l'année N et des taux d'orientation ou de réorientation à la fin de cette classe au cours de l'année N+1 peut cependant apporter de précieux renseignements sur la continuité des parcours scolaires. Ces données peuvent être utilement complétées par les informations publiées par les services statistiques rectoraux concernant les résultats au brevet (épreuves ponctuelles et contrôle continu) et, quand ils existent, les résultats aux évaluations de français et de mathématiques en 6^{ème} au cours de l'année N-4, même s'il faut garder à l'esprit que les épreuves sont de nature différente : sommatives dans le premier cas, diagnostiques dans le second.

▪ **3.2.3.2. Respecter quelques règles de méthode dans l'utilisation des indicateurs**

L'interprétation des indicateurs s'avère généralement problématique et il faut donc se garder d'explications toutes faites qui peuvent rater leur cible. Pour prendre un exemple simple, un écart élevé entre les résultats aux épreuves ponctuelles du brevet et en contrôle continu (classes de 4^{ème} et 3^{ème}) peut s'interpréter, selon le contexte, dans des directions tout à fait opposées. Il peut signifier un certain laisser-aller des professeurs en matière de notation, une volonté d'acheter la paix sociale par les notes, une perte de repères sur les niveaux d'exigence, ou encore, une politique volontariste et consciente de motivation des élèves par le système de notation. Le niveau ou la variation d'un indicateur ne fournit donc que des conjectures qui devront être confrontées, pour confirmation ou invalidation, à d'autres données d'observation plus qualitatives.

De manière générale, on observe dans les rapports d'évaluation d'EPLÉ une tendance à la surinterprétation des données qui a pour effet d'altérer le caractère probant de leurs conclusions. Il convient donc de rappeler ici quelques règles de prudence à respecter dans l'utilisation des statistiques :

- Des données synchroniques ne doivent pas conduire à tirer des conclusions diachroniques comme on l'observe trop souvent. En revanche, le suivi de cohortes – réelles ou fictives –, même sur une période courte, présente un intérêt tout particulier, surtout lorsqu'il porte sur des degrés d'enseignement différents (CM2/6^{ème} ; 3^{ème}/2^{nde} ; terminale/DEUG) car il permet d'apprécier la continuité des parcours scolaires des élèves par comparaison de leurs résultats d'un niveau d'enseignement à l'autre.
- Les écarts entre taux d'orientation, d'accès ou de réussite ne sont significatifs que lorsqu'ils concernent des effectifs d'élèves relativement nombreux, qu'ils sont suffisamment contrastés et qu'ils présentent une certaine stabilité dans le temps. Des écarts faibles peuvent être dus au simple hasard, d'autant que la DEP n'indique pas le degré de fiabilité de ses données. Pour les mêmes raisons, le classement des établissements en fonction de la valeur ajoutée ne présente une véritable signification que pour les cas extrêmes.
- Les évolutions des données de l'EPLÉ ne prennent tout leur sens que par comparaison avec celles observées dans des espaces plus larges (bassin, département, académie métropole) ; la baisse d'un taux de redoublement ou l'augmentation d'un taux de réussite aux examens peuvent en réalité cacher une détérioration de la position relative de l'établissement si ces variations sont inférieures à celles observées dans les autres établissements.

- Une relation de corrélation ne doit pas être interprétée comme une relation causale unilatérale. Ainsi la relation entre le climat de l'établissement et les résultats des élèves ne doit pas laisser penser que le premier élément est nécessairement la cause du second. La relation peut aussi être réversible : dans certains cas, c'est parce que les élèves ont d'excellents résultats que le climat de l'établissement est bon.

Par ailleurs, un indicateur isolé apporte rarement une information significative : d'où la nécessité d'opérer des croisements entre indicateurs dont la nature dépend des questions que l'on se pose ; et s'il n'existe pas à cet égard de formules « toutes faites », on peut cependant proposer quelques pistes potentiellement heuristiques :

- Mettre en relation les indicateurs de résultats avec les données de contexte pour affiner la mesure de la valeur ajoutée.
- Croiser les indicateurs de résultats, les décisions d'orientation et les vœux d'orientation pour déterminer le poids relatif des résultats scolaires dans les décisions d'orientation par comparaison avec celui des familles, et partant, le degré de sélectivité scolaire et sociale de l'établissement.
- Comparer les taux d'accès, les indicateurs de résultats et le devenir des élèves en fin de seconde et de première, pour repérer la pertinence des choix d'orientation.

3.3. Repenser a terme l'évaluation des EPLE

3.3.1. *Revoir l'articulation entre les différents niveaux d'évaluation : constituer une chaîne de l'évaluation*⁴⁶

▪ 3.3.1.1. Le cadre académique paraît devoir être privilégié

Dans la logique de la déconcentration, il paraît normal que le recteur d'académie qui a la compétence de gestion des EPLE soit à la fois le commanditaire et le premier destinataire du rapport d'évaluation des établissements de son académie. Ce choix s'impose également pour des raisons plus pratiques. En l'état actuel, les inspections générales sont dans l'incapacité d'assurer un suivi régulier des établissements scolaires (voir supra). Le suivi permanent des établissements réalisé par l'IGAENR ne touche guère qu'environ 150 établissements chaque année sur un total de plus de 11000 (privé compris) en 2002-2003⁴⁷. Dans ces conditions, un

⁴⁶ Expression empruntée à un document interne de l'IGEN vie scolaire (1989).

⁴⁷ Repères & références statistiques, DEP, 2003.

suivi régulier des évaluations d'EPL ne peut être réalisé qu'au niveau académique, si on veut parvenir à une évaluation régulière de tous les établissements ou tout au moins atteindre une certaine taille critique. Le choix du cadre académique s'impose encore plus si l'on souhaite intégrer aux évaluations globales les résultats des observations des pratiques enseignantes dans les classes. Dès lors deux scénarios peuvent être envisagés.

- Le premier reste dans la logique esquissée actuellement dans plusieurs académies, qu'il se contente de systématiser, en envisageant un plan d'évaluation de tous les établissements sur une période de quatre ou cinq ans. Les corps d'inspection territoriaux assistés par d'autres experts, cadres de l'administration, chefs d'établissement, éventuellement des enseignants et des représentants des collectivités territoriales, effectuent les opérations d'évaluation selon des protocoles améliorés. L'expérience lilloise montre que cela est possible, même dans une académie de grande taille. Mais elle révèle également le risque d'un essoufflement rapide des équipes d'évaluateurs qui éprouvent très vite les plus grandes difficultés à assumer parallèlement leurs autres missions, sauf à mener de front évaluation globale et inspection individuelle, ce qui introduit une confusion dommageable auprès des enseignants. La généralisation de ce scénario suppose donc à terme, soit une création massive de postes d'inspecteurs territoriaux, soit un recentrage de leurs missions autour de la fonction d'évaluation des EPL dont le corollaire serait vraisemblablement l'abandon des inspections individuelles et de leur rôle dans la notation des personnels et la gestion des carrières.
- Le second scénario nous semble plus porteur d'avenir et s'inscrit d'ailleurs dans la perspective des expérimentations initiées par la commission européenne. Il place l'auto-évaluation au premier plan en faisant le pari du développement d'une dynamique interne des établissements. Les corps d'inspection pédagogiques territoriaux sont alors chargés de mettre à la disposition des établissements et de leurs personnels des outils et méthodes d'évaluation, de former les personnels de direction à l'évaluation et de sensibiliser les autres catégories de personnels – notamment les enseignants – à ses enjeux. Les inspecteurs, correspondants de bassin, peuvent en outre apporter une aide aux établissements de leur ressort et les accompagner dans la mise en œuvre des auto-évaluations. Au-delà, les corps d'inspection territoriaux s'assurent, en aval, de la qualité des auto-évaluations menées par les établissements et veillent à ce que toutes les conséquences en soient tirées notamment par la mise en place d'un plan d'action dont les effets devraient eux-mêmes être évalués. Le schéma ici retenu est donc celui d'une auto-évaluation « accompagnée ». Il n'exclut pas cependant la possibilité que des évaluations externes ciblées puissent continuer à être réalisées à la demande du recteur.

▪ 3.3.1.2. Un nouveau positionnement des inspections générales

Dans ce nouveau cadre, le recteur pourrait confier la coordination de l'ensemble du dispositif d'évaluation à un responsable placé à un niveau hiérarchique élevé et distinct du directeur de la pédagogie.

Les inspections générales interviendraient alors en amont et en aval du processus d'évaluation. En amont, elles seraient chargées de veiller à la qualité globale du système d'informations statistiques (fiabilité, cohérence et pertinence des indicateurs) et des autres instruments d'évaluation, qu'ils soient produits au niveau de la centrale ou au niveau académique. Elles auraient également pour mission de concevoir, en liaison avec la direction de l'encadrement et l'ESEN⁴⁸, les formations des corps d'inspection territoriaux et des chefs d'établissement à l'évaluation des EPLE. En aval, elles procéderaient à des méta-évaluations qui porteraient principalement sur la capacité des académies à évaluer leurs établissements (premier scénario) ou à leur fournir des supports, aides et protocoles de qualité pour s'auto-évaluer (deuxième scénario).

Les inspections générales continuent, en outre, à procéder à des évaluations d'EPLE dans le cadre de commandes ministérielles ciblées sur des thématiques spécifiques : « l'établissement efficace », les collèges de zone rurale, les zones d'éducation prioritaires etc. Elles bénéficient, pour mener ces enquêtes, de l'appui logistique de la DEP pour la constitution des échantillons d'établissements et peuvent passer commande des travaux statistiques qu'elles jugent utiles pour l'accomplissement de leurs missions d'évaluation⁴⁹.

3.3.2. *Réformer notre système d'informations statistiques*

Si les données statistiques actuellement disponibles sont abondantes, elles ne forment pas système en raison d'un manque d'homogénéité entre les informations disponibles pour les collèges et celles qui le sont pour les lycées. A terme, notre système d'informations doit donc être amélioré sur plusieurs points.

▪ 3.3.2.1. Comblent certaines lacunes du système d'informations statistiques

Avec la mise en application de la LOLF, une réflexion d'ensemble doit être conduite sur les indicateurs, pour harmoniser les données disponibles en collège et en lycée. En l'état actuel, les seuls indicateurs disponibles en collège concernant le niveau de réussite des élèves sont les taux d'accès de sixième en troisième, d'une part, et de troisième en seconde, d'autre part. Il convient donc de s'assurer au plus vite d'une remontée systématique vers la centrale des

⁴⁸ Ecole supérieure de l'éducation nationale.

⁴⁹ L'évaluation des EPLE par les inspections générales peut également conserver toute sa place dans le cadre de l'autoformation des nouveaux inspecteurs généraux ou dans le cadre de visites ponctuelles.

résultats au brevet, de façon à faire figurer dans les IPES des données sur les taux de succès au brevet (taux bruts et valeur ajoutée) ainsi que des indications fiables sur le pourcentage d'élèves ayant obtenu la moyenne aux seules épreuves ponctuelles (plus significatives de leurs acquis cognitifs que les notes du contrôle continu). Devraient également figurer dans les IPES rénovés, les écarts entre les moyennes obtenues aux épreuves ponctuelles et celles du contrôle continu.

D'un autre côté, si l'on trouve dans les IPES des données sur le devenir des élèves de collège en fin de classe de seconde, assez étonnamment, aucune donnée sur le devenir des bacheliers n'y figure. De fait, on ignore le plus souvent dans chaque lycée les choix d'orientation des élèves après le baccalauréat, à l'exception de ceux qui s'orientent en classes préparatoires ou en STS. On ne dispose pas non plus d'informations, par établissement, sur la proportion d'élèves qui sont passés en seconde année de DEUG et sur ceux qui ont obtenu ce diplôme en deux ou trois ans. Le développement des systèmes d'inscriptions informatisés dans l'enseignement supérieur, comme Ravel en région parisienne, devrait conduire la DEP à produire, à échéance rapprochée, de telles données sans lesquelles on ne peut appréhender les effets différés des enseignements dispensés en lycée. Une telle évolution n'est cependant envisageable que si la CNIL autorise la DEP à suivre des cohortes d'élèves avec un cryptage des numéros d'identification.

Au-delà, les indicateurs de vie scolaire, font cruellement défaut au niveau national : les taux d'absentéisme des élèves – en différenciant les absences excusées et celles qui ne le sont pas – le nombre et la nature des actes de violence⁵⁰ et d'incivilité, d'une part, et des sanctions disciplinaires, d'autre part, ne font pas l'objet d'une recension systématique, comme dans les pays anglo-saxons, alors que leur évolution constitue pourtant de bons indices de l'amélioration ou de la détérioration du climat d'un établissement.

▪ 3.3.2.2. Améliorer la fiabilité des données

Les données produites par la DEP sont généralement de bonne qualité, surtout si on les compare à leurs équivalents étrangers. Il reste qu'elles sont perfectibles dans plusieurs de leurs dimensions. Il conviendrait notamment d'assurer la consolidation nationale des taux d'accès de 3^{ème} en seconde qui en l'état actuel minorent les performances des collèges « frontaliers » dont les élèves poursuivent leurs études dans les lycées d'académies voisines.

Les données des enquêtes IVA sur l'insertion des élèves dans la vie active présentent également un intérêt tout relatif dans la mesure où les taux de réponse varient dans des proportions importantes d'une académie à l'autre (du quart à plus de 60 %). Il conviendrait

⁵⁰ Les items du logiciel Signa restent encore trop inégalement renseignés par les établissements pour permettre des comparaisons fiables.

certainement de simplifier le questionnaire adressé aux anciens élèves, peu clair dans ses formulations et trop détaillé dans ses demandes, et de conduire une politique plus volontariste auprès des proviseurs de lycées professionnels pour obtenir un meilleur retour de cette enquête.

Dans un autre registre, l'identification des catégories socioprofessionnelles des parents d'élèves, qui joue actuellement un rôle important dans le calcul de la valeur ajoutée des lycées, souffre de plusieurs imperfections qui en altèrent considérablement la validité : l'enregistrement se fait sur une simple base déclarative alors que l'Insee reconstruit l'appartenance aux catégories socioprofessionnelles à partir d'une pluralité de questionnements ; les personnels chargés de l'enregistrement de ces données n'ont généralement pas bénéficié d'une formation à la codification des professions, de sorte que les classifications, pour une même profession déclarée, peuvent varier dans des proportions importantes d'un établissement à l'autre : aujourd'hui, la marge d'incertitude affectant ces opérations de codification rend toute comparaison d'un établissement à l'autre d'autant plus problématique que les effectifs concernés à l'échelle de l'EPLÉ sont généralement faibles et que l'on n'est pas assuré que les biais observés se compensent.

▪ **3.3.2.3. Assurer un meilleur chaînage des indicateurs**

On assiste depuis quelques années à une prolifération de données construites aux niveaux académique ou départemental. Certaines d'entre elles visent à combler des lacunes du système d'information de la DEP et n'auront plus leur raison d'être lorsque ces problèmes seront résolus. D'autres sont tout simplement redondantes avec celles qui sont publiées dans les IPES ou dans la BCP et n'ont donc pas de justification intrinsèque : comme elles ne sont pas toujours calculées à partir des mêmes bases que celles de la DEP, leur publication ne peut que contribuer à semer la confusion dans les esprits et à renforcer le risque d'illisibilité des opérations d'évaluation.

Il convient donc d'améliorer le chaînage des indicateurs produits aux différents niveaux du système en harmonisant définitions, champ, et modes de construction. Une meilleure concertation entre les différents acteurs qui interviennent dans la production des statistiques au sein des académies, d'une part, et entre les services statistiques rectoraux et la DEP, d'autre part, devrait aboutir au renforcement souhaité de notre système d'informations. Cette évolution est rendue encore plus urgente avec la mise en place de la LOLF à partir de 2006.

▪ **3.3.2.4. Redéfinir la notion de valeur ajoutée et en revoir le mode de calcul**

Le concept de valeur ajoutée actuellement utilisé par la DEP est daté. Novateur à l'époque où il a été mis en place, il est maintenant souvent dépassé par certains indicateurs construits dans

les pays anglo-saxons qui prennent en considération une pluralité de dimensions : le niveau initial des élèves, leurs caractéristiques sociales et les effets de contexte⁵¹.

En fait, plus qu'un véritable indicateur de valeur ajoutée, la DEP détermine des taux d'accès ou de réussite corrigés en fonction des catégories socioprofessionnelles et des retards d'âge scolaire. Outre le fait que ce mode de mesure de la valeur ajoutée présuppose une forme de déterminisme social des performances scolaires qui le disqualifie d'avance auprès de certains interlocuteurs, il n'est pas même en mesure d'assumer correctement sa fonction. Et ce pour plusieurs raisons : la catégorie défavorisée est à la fois trop large (43,5 % des élèves en collège) et trop disparate pour mesurer de manière fine le degré réel de la difficulté sociale. En outre le second critère utilisé, le retard d'âge scolaire, est un prédicteur très inégal du niveau initial des élèves dans la mesure où il peut être influencé par les politiques de redoublement pratiquées en amont dans les établissements viviers. Un lycée peut donc avoir une proportion élevée d'élèves en retard sans que cela dénote nécessairement des difficultés scolaires fortes (et vice-versa). A terme, la mesure de la valeur ajoutée devrait donc être calculée, conformément aux standards internationaux, par comparaison des niveaux de compétence en début et en fin de cycle selon des modes d'évaluation standardisés et comparables.

▪ **3.3.2.5. Produire et exploiter les données individuelles aux niveaux académique et infra-académique**

Le développement de l'informatisation des bulletins scolaires et leur exploitation par les logiciels actuellement disponibles devraient permettre aux établissements et aux académies d'établir des comparaisons plus fines des résultats scolaires à partir de données désagrégées. Il serait indiqué, par exemple, de mettre en relation les notes des différentes disciplines pour un même niveau de classe, ou encore, de mesurer le degré de corrélation des notes d'une même discipline à différents niveaux de cursus. De telles analyses permettraient, par-delà les variations des systèmes de notation de chaque professeur, de s'assurer de la cohérence d'ensemble des acquisitions des élèves d'une année d'étude à l'autre au sein d'un même établissement et entre degrés d'enseignement (CM2/6^{ème}, troisième/seconde) apportant ainsi aux professeurs un retour d'informations précieux pour harmoniser leurs pratiques pédagogiques et leurs modes d'évaluation. Les conseils scientifiques et pédagogiques qui doivent être installés dans les établissements pourraient alors devenir le lieu d'une réflexion mettant en relation les processus d'apprentissage et les résultats des élèves en prenant appui sur une analyse plus systématique de leurs notes aux examens et dans les évaluations intermédiaires.

⁵¹ Par exemple, l'académie de Toulouse a déjà élaboré un indicateur de « niveau ajouté » en lycée qui prend en compte, outre la PCS des parents et le retard, le niveau scolaire d'entrée des élèves, appréhendé à partir des résultats aux épreuves ponctuelles du brevet.

Au-delà, une formation de base de l'ensemble des enseignants à l'utilisation des indicateurs et leur libre accès sur internet semblent constituer un pré-requis nécessaire au développement d'une culture d'évaluation qui prenne appui sur des données tant soit peu objectivées.

3.3.3. Élargir et confronter divers modes d'investigation

Les évaluations d'EPLÉ menées par les inspections générales s'appuient principalement sur des documents statistiques, des dossiers, des résultats d'enquêtes réalisées par les établissements scolaires et des entretiens. Celles qui sont conduites dans le cadre des académies recourent en plus à des visites d'inspection.

Ces modes d'investigation sont susceptibles d'évoluer différemment selon que l'on privilégie l'hypothèse d'une évaluation externe des établissements ou celle de l'auto-évaluation. Mais dans tous les cas, une réflexion doit être conduite sur l'utilisation d'un spectre plus large de méthodes d'investigation.

▪ 3.3.3.1. Introduire des enquêtes par questionnaires

La distribution de questionnaires, en amont des entretiens, comme cela se pratique à l'étranger, permettrait de mieux cerner les perceptions et représentations du fonctionnement de l'établissement par les différentes catégories d'acteurs et de repérer plus facilement les principaux enjeux. L'utilisation de questionnaires, par le respect de l'anonymat qu'elle assure, peut également favoriser l'explicitation, par les enseignants les moins portés à s'exprimer en groupe, de leurs démarches pédagogiques sur des questions sensibles ou délicates et l'expression par les élèves des problèmes relationnels qu'ils peuvent rencontrer et de leurs difficultés d'apprentissage (auto-évaluation).

Le tri croisé des réponses aux différents questionnaires peut alors servir à affiner les protocoles utilisés lors des entretiens individuels et collectifs (évaluations externes) ou à mieux cibler les thématiques à aborder dans les groupes de discussion (auto-évaluation).

En recoupant ces divers types d'informations, on bénéficierait à la fois des apports de données extensives, exprimables sous la forme de statistiques, et des comptes rendus d'investigations qualitatives qui rendent mieux compréhensibles la construction des représentations des acteurs et les tensions ou dynamiques qu'elles peuvent engendrer.

▪ 3.3.3.2. Réévaluer le statut des documents fournis par les établissements

Les conditions d'analyse des documents remis par l'établissement méritent une attention toute particulière. Il existe en effet une tendance forte et, quelque peu naïve, à considérer que ce qui

est écrit a été fait et, qu'inversement, ce qui n'a pas été formalisé n'existe pas. C'est là confondre la réalité des choses avec les moyens d'attester leur existence, ce qui n'est pas sans susciter des incompréhensions au sein des établissements évalués, où cette approche « bureaucratique » du travail pédagogique entretient, non sans raison, les craintes des enseignants. De fait, il n'est pas rare qu'un beau projet sur le papier n'aboutisse à aucun résultat tangible en termes d'apprentissages et qu'inversement des choses intéressantes puissent être réalisées sans que les intéressés aient éprouvé ou éprouvent le besoin de les formaliser par écrit. Les documents ne doivent donc pas être considérés comme « faisant foi » ou ayant valeur probante, au sens administratif du terme, mais simplement comme point d'appui pour mener entretiens et observations dans le cadre de l'évaluation externe.

Il reste que si l'activité d'enseignement est par essence fugace et difficile à saisir, elle laisse cependant des traces (cahiers de textes, sujets d'évaluations, devoirs des élèves, progressions des professeurs, notes et appréciations sur les bulletins scolaires), qui en conservent certains effets, à condition cependant qu'évaluateurs et évalués soient au clair sur ce qu'on peut y trouver. En un grand nombre de cas, la démarche d'évaluation elle-même sera donc une invitation pour beaucoup d'acteurs, et notamment pour beaucoup de maîtres, à opérer un retour réflexif sur leurs démarches pour les objectiver.

- **3.3.3.3. Recentrer les grilles d'observation sur les activités des élèves et les processus d'apprentissage**

Comme on l'a déjà souligné (cf. supra 2), les observations effectuées en classe et les entretiens menés par les corps d'inspection territoriaux restent encore très marqués par le modèle de l'inspection individuelle. Les rapports qui en résultent ne permettent généralement pas de comparer les pratiques d'une discipline à l'autre ni de capitaliser les observations. Il conviendrait donc de donner une nouvelle orientation à ce type d'investigation de manière à produire des données exploitables sur les représentations du métier et sur les méthodes et pratiques d'enseignement : préparation des cours (construction des progressions, référence aux programmes et instructions, utilisation des manuels et des différentes ressources documentaires), pratiques pédagogiques en classe, modes et critères d'évaluation, relations avec les autres acteurs en dehors de la classe (élèves, parents, collègues, direction, personnels d'éducation, d'orientation). Les grilles d'observation des pratiques enseignantes devraient donc être homogénéisées entre disciplines, tout en faisant place à la spécificité de la construction des savoirs au sein de chacune d'entre elles. Les observations gagneraient également à être recentrées sur les comportements, activités et acquisitions des élèves de façon à apporter une attention plus systématique aux démarches de construction et d'évaluation de leurs acquis. La dimension collective de l'enseignement, disciplinaire et interdisciplinaire, devrait être davantage prise en compte. Les observations de classe pourraient être suivies d'entretiens avec des élèves de la classe de façon à se faire une idée plus précise de ce qui s'y est joué et de ce qu'ils ont effectivement appris au cours des séances

observées. Ces grilles d'observation des pratiques de classe peuvent aussi bien être utilisées par des observateurs extérieurs qu'entre pairs dans le cadre d'une auto-évaluation au sein des établissements.

▪ **3.3.3.4. Mieux articuler les différents types d'entretiens et la nature des informations qu'ils peuvent apporter**

Il est souhaitable que les entretiens soient ouverts à une très grande diversité d'acteurs, les parents évidemment, mais aussi les représentants des collectivités territoriales, sans oublier le monde de l'entreprise pour l'évaluation des lycées professionnels. A côté des porte-parole d'organisations syndicales ou associatives, des représentants au conseil d'administration, de professeurs exerçant des responsabilités de coordination, il paraît aussi indispensable d'interroger (évaluation externe) ou de faire participer (auto-évaluation) des personnes moins impliquées dans le pilotage de l'établissement et dont on peut attendre des discours moins convenus et conséquemment plus riches en informations brutes, spontanées, bref sociologiquement exploitables.

Le statut des différents entretiens devrait également être clarifié en distinguant plus clairement la nature des questions et des renseignements qu'elles peuvent livrer. Il convient notamment de faire le départ entre des questions qui visent à déterminer « qui fait quoi au sein de l'établissement ? » et celles qui portent sur les perceptions, les représentations que les acteurs ont du fonctionnement de l'établissement.

S'agissant des questions du premier genre, il importe de croiser systématiquement ce qui est dit à ce que l'observation *in situ* des pratiques de pilotage ou d'enseignement permet de constater. L'expérience montre en effet qu'il peut exister des écarts sensibles entre le « déclaré » et le « réalisé », des distorsions fortes entre ce que les acteurs disent de leurs pratiques et ce qu'ils font effectivement et ce pour au moins deux raisons : d'abord parce qu'en situation d'entretiens, les acteurs sont en représentation et ont intérêt à définir la situation à leur avantage ; ensuite parce qu'ils ont rarement conscience de leurs comportements réels.

S'agissant des questions du second genre, on peut considérer les représentations elles-mêmes comme des faits, au sens scientifique du terme, mais qui doivent être pris en compte d'une manière différente des constats opérés par ailleurs. L'analyse des représentations permettra surtout de déterminer le degré d'homogénéité des perceptions que les différents acteurs ou catégories d'acteurs ont de l'établissement. A partir de là, on pourra déterminer les zones de consensus, les points de divergence et les marges de manœuvre disponibles pour faire évoluer le fonctionnement de l'établissement. L'évaluation, en effet, ne saurait se réduire à un simple constat, ni même à la mise en évidence des atouts et des points faibles d'un établissement ; elle se doit aussi de repérer les lignes de force à partir desquelles une dynamique de

changement peut s'engager, sans pour autant ignorer les résistances qui devront être surmontées pour en assurer toutes ses potentialités. En d'autres termes, l'évaluation est toujours à la recherche d'un équilibre fragile entre ce qui est du domaine du souhaitable et ce qui est du domaine du possible.

Conclusion : l'évaluation, une ambition pour le service public

On aura compris que la définition de l'évaluation au sein du service public comme succédané du marché n'est pas celle ici retenue : le service public est à assurer, en effet, quoi qu'il adienne, et il n'a pas le choix entre réussir et autre chose. Il doit donc être clair que les résultats de l'évaluation ne posent pas la question du maintien ou non du service public. La question posée est bien en revanche celle de l'amélioration du service public, avec cette idée que la qualité ne se construit pas seule, qu'elle nécessite une stratégie continue et que cette stratégie passe d'abord par la connaissance, attitude qui devrait aller de soi dans un établissement scolaire d'ailleurs fondé sur le rapport à la connaissance : cette connaissance des effets des enseignements est-elle autre chose que pratiquer une sorte de didactique générale, partant du principe qu'aucun enseignement, dans son élaboration comme dans ses effets, n'est détachable du milieu où il se construit, et ce recul réflexif est-il autre chose que la pratique de la réflexion philosophique à l'égard de l'école ? L'évaluation de l'établissement, loin d'être réductrice, s'adresse en effet aux capacités les plus hautes de réflexion et de connaissance de ses acteurs et s'adresse en particulier aux maîtres comme à des intellectuels⁵².

Comme toute institution humaine, un établissement scolaire agit dans un environnement complexe au sein duquel il n'a pas nécessairement une conscience claire de ce qu'il fait (en particulier dans le domaine de l'évaluation des élèves, comme si l'évaluation était *a priori* réticente à l'idée d'être évaluée, le juge à celle d'être jugé). Ses acteurs, en effet, sont pris au quotidien dans une infinité de tâches exigeantes et difficiles. L'évaluation, est dès lors le moyen par lequel l'établissement peut en apprendre plus qu'il n'en sait sur son action et les effets de son action et donc peut par là percevoir en quoi son activité est conforme à ce qu'on peut attendre d'un service public.

⁵² Cf. Jean-Pierre OBIN, *Enseigner pour demain...*

On ne saurait également perdre de vue que l'amélioration et le progrès sont les seules justifications véritables de l'évaluation et de l'énergie qu'y investissent les agents : c'est ainsi que toute évaluation doit s'ouvrir à sa propre évaluation, que les changements intervenus à la suite d'une évaluation doivent à leur tour être évalués et que toute évaluation réussie contribue à construire et à renforcer la culture d'évaluation des acteurs. Plus généralement, l'évaluation ne peut se réduire à un protocole figé *ne varietur* et doit ménager, dès le stade de sa conception, une place à l'évolution de sa démarche qui prenne en compte le développement de l'expérience professionnelle des évaluateurs.

Jean ETIENNE

Roger-François GAUTHIER

Annexes

Annexe 1 : Définitions

Annexe 2 : L'évaluation des établissements dans le « débat national sur l'avenir de l'école » : des attentes qui ne parviennent pas à surmonter un grand flou conceptuel où s'abritent de nombreuses craintes

Annexe 3 : Résumés et propositions

Définitions

Certaines définitions sont à ce stade indispensables :

Contrôle	Vérification de l'application par un agent ou une organisation d'une norme ou d'un ensemble de normes fixées extérieurement à cette organisation
Audit	Regard détaillé porté sur une organisation sous la forme d'un contrôle de l'ensemble de ses fonctions
Contrôle de qualité	Vérification de l'application par une organisation d'un ensemble de normes entrant dans la définition de la qualité
Evaluation	Jugement porté sur l'activité d'un agent ou d'une organisation, en considération de ses résultats, proposant des explications du niveau d'atteinte de ces résultats dans le but de permettre à l'agent ou à l'organisation de les améliorer
Auto-évaluation	Évaluation conduite sur lui-même ou sur elle-même par un agent ou une organisation
Evaluation interne	Évaluation conduite sur une organisation par des agents désignés à cette fin faisant partie de cette organisation
Evaluation externe	Évaluation d'une organisation par des acteurs extérieurs à elle, que leur statut d'évaluateur soit occasionnel ou permanent qu'ils appartiennent ou non à l'éducation nationale
Indicateurs	représentation chiffrée d'un phénomène que l'on veut contrôler
Résultats	mesure de l'atteinte par un agent ou une organisation d'objectifs prédéterminés
Effets	conséquences, conscientes ou non, de l'activité d'un agent ou d'une organisation, sur cet agent, cette organisation ou l'environnement
Efficacité	partie de l'évaluation d'une organisation qui mesure si ses différentes activités atteignent les objectifs fixés
Équité	partie de l'évaluation d'une organisation qui mesure sa capacité à bénéficier de façon égale à ses différents usagers
Efficienc	partie de l'évaluation d'une organisation qui met en rapport les résultats de ses différentes activités avec les moyens (financiers, humains, etc.) qui leur sont consacrés

L'évaluation des établissements dans le « débat national sur l'avenir de l'école » : des attentes qui ne parviennent pas à surmonter un grand flou conceptuel où s'abritent de nombreuses craintes

Il a déjà été dit que les questions relatives à l'autonomie de l'établissement et à son évaluation ont occupé une part limitée du débat national sur l'école.

C'est ainsi que le « miroir du débat » :

- fait apparaître que les débats sur le thème de l'évaluation ont été assez flous « *Les opinions émises font apparaître ... des confusions entre évaluation des élèves, des personnels, des établissements* », un peu comme ceux relatifs à l'autonomie, « *tronqués ou partiellement, voire totalement hors sujet du fait des lacunes dans l'information des participants* » ;
- montre l'expression fréquente d'une opposition à l'autonomie et l'évaluation des établissements (« *l'Education nationale est un service public et, à ce titre, ne peut et ne doit être soumise à évaluation* » citation tirée du débat des lycées publics de l'arrondissement de Dôle) ;
- fournit quelques « explications » à cette opposition : refus de la gestion des établissements en fonction des résultats, méfiance à l'égard de la remise en cause de la carte scolaire, crainte que l'autonomie soit une étape vers une « privatisation », malgré la conscience de la nécessité de répondre au réel (« *on ne veut pas d'autonomie et d'évaluation à l'anglo-saxonne mais il faut des réponses à la diversité des situations* » citation du lycée de l'arrondissement d'Istres) ;
- montre que l'effet d'une conception ambiguë de l'évaluation, laissant en redouter les pires conséquences pour les acteurs, peut être de « *crisper* » le « *débat sur l'autonomie elle-même* » : en sens opposé le lien fait entre autonomie et décentralisation « politique » (compétences des collectivités territoriales) plutôt qu'entre autonomie et décentralisation « technique » vers l'établissement crispe le débat sur l'évaluation ;
- fait émerger l'idée de l'importance de l'évaluation des résultats (« *L'évaluation de l'établissement doit consister en une prise de conscience des résultats des élèves, de leur évolution* », citation des écoles publiques de l'arrondissement de Bourg-en-Bresse), à condition que cette évaluation n'ait pas de conséquence sur la situation des personnels (unanimité), ni sur les moyens accordés (très majoritaire) ;
- lance des idées sur ce que pourrait être cette évaluation (à noter que les évaluations existantes ne semblent pas connues) : « *strictement interne* » pour les uns, ou « *par un organisme neutre* », « *mixte* » ou « *paritaire* » pour les autres ; cette évaluation doit être une « *évaluation-conseil* », et « *comprendre une part d'auto-évaluation* » (collèges publics de l'arrondissement de Nîmes) ;
- montre que la question de l'évaluation est souvent perçue comme seconde par rapport à la notion de projet d'établissement : elle est parfois ramenée au bilan du projet d'établissement et de l'atteinte de ses objectifs ;

Résumé et propositions

Première partie : l'évaluation de l'EPL, avec quels enjeux ?

Alors que dans la plupart des pays européens le développement de l'autonomie des établissements scolaires s'est accompagné de la mise en place de procédures d'évaluation ou d'auto-évaluation, tant pour répondre aux attentes de l'institution qu'à celle des usagers, il est symptomatique **qu'en France le sujet n'ait pas été considéré comme une priorité politique.**

Après avoir noté cette réticence apparemment plus forte en France de l'École par rapport à l'évaluation, le rapport rappelle qu'il s'agit pourtant d'une obligation légale que la mise en place de la LOLF va rendre encore plus nécessaire et qu'en tout état de cause le développement de l'autonomie des établissements ne saurait être réel, sans la contrepartie de formes nouvelles de régulation se substituant à la simple conformation à une norme unique.

Plusieurs rapports, dont un émanant du Haut Conseil de l'évaluation de l'école, ont d'ailleurs regretté la faiblesse de l'évaluation des collèges et des lycées français et, de fait, plusieurs académies ont tenté de pallier cette carence par la mise en place de procédures d'évaluation des établissements.

Laisser les choses en l'état et se désintéresser d'un sujet en apparence difficile ne manquerait pas d'avoir deux conséquences :

1. priver les établissements de la maîtrise de leur destinée et de la capacité à exercer pleinement leur autonomie, au moment même où la difficulté des problèmes qu'ils ont à résoudre est accrue ;
2. laisser les écarts se creuser encore entre les différents établissements, dans l'indifférence apparente de l'institution.

Tout en jugeant le développement de l'évaluation des établissements indispensable à l'amélioration du terrain éducatif et des résultats de l'école, les auteurs du rapport ne méconnaissent pas les difficultés de l'exercice et tentent de répondre aux arguments qui tendraient à dissuader de s'engager dans une telle démarche :

- si les caractéristiques initiales des élèves et la qualité des enseignants expliquent, pour une grande part, le degré de la réussite scolaire, plusieurs recherches ont montré que dans certaines situations – très bons ou très mauvais établissements, difficultés des élèves – l'établissement peut faire la différence ;
- si l'interventionnisme croissant en matière pédagogique des administrations centrale et académique a pu faire douter de l'autonomie, des établissements, l'évaluation des collèges et lycées, loin d'être rendue inutile, est devenue d'autant plus nécessaire qu'il faut mesurer l'effet en contexte des orientations nationales ou académiques ;
- enfin, l'évaluation doit permettre de déterminer dans quelle mesure l'établissement s'est saisi des marges d'autonomie dont il dispose, notamment dans le domaine pédagogique, où l'autonomie des maîtres, distincte de celle qui a été récemment consentie, est ancienne et forte.

Les auteurs du rapport prennent aussi au sérieux les arguments par lesquels l'école tente d'échapper à l'évaluation :

- l'impression que l'école fait de l'évaluation sans le savoir, puisqu'elle évalue sans cesse ses élèves : l'argument ne provient que d'un effet de miroitement autour du mot « évaluation », puisque l'activité d'évaluation des élèves fait précisément partie de ce qui doit être évalué dans un ensemble pédagogique ;
- les résultats d'un établissement, les notes de ses élèves, seraient des données « indigènes » auxquelles il ne conviendrait pas de donner une signification à l'extérieur de l'établissement : ne convient-il pas pourtant de lire les résultats scolaires à une aune différente de celle de l'école qui les produit ?
- des exemples étrangers, notamment américains, ont pu à juste titre conduire à se méfier des aspects « appauvrissants » sur l'enseignement de certaines modalités d'évaluation par tests dont il faudrait dire clairement que ce n'est pas ce qui est visé ; il en va de même pour la crainte de la marchandisation et de la privatisation souvent invoquée : tout montre, au contraire, que le service public sera au contraire renforcé s'il se dote de modes de régulation de qualité.

Il convient aussi de ne pas méconnaître certains obstacles que, *volens nolens*, l'institution met elle-même au développement de l'évaluation des établissements :

- en ne veillant pas au respect des obligations d'auto-évaluation qui figurent dans les textes et à la qualité de leurs productions ;
- en mettant centralement l'accent sur la question du projet d'établissement, conçu trop souvent en faisant l'économie d'une évaluation véritable ;
- en maintenant un régime d'inspections individuelles qui, confiées souvent aux mêmes agents que l'évaluation d'établissements, mélange tout et dément l'importance par ailleurs proclamée de l'établissement ;
- en ne donnant pas le même régime aux écoles du primaire, ce qui ne favorise pas la promotion d'une conception unifiée de l'évaluation dans le système ;
- en fournissant aux établissements des indicateurs nombreux mais qui ne sont pas toujours les plus pertinents, avec en particulier une insistance exagérée sur l'impact des catégories socioprofessionnelles au détriment de l'étude du niveau des acquisitions des élèves.

Les auteurs du rapport considèrent en conséquence qu'une relance de l'évaluation des établissements constitue un enjeu important, mais qu'il convient de prendre un certain nombre de précautions « clés » :

- **appréhender l'établissement dans sa globalité**, et non comme une simple juxtaposition de fonctions à évaluer séparément ;
- **clarifier la démarche d'évaluation et la rendre crédible et utile**, notamment aux usagers, en répondant à des questions simples : l'évaluation pour quoi faire ? avec quels destinataires ? quelle périodicité ? quels protocoles ? quels acteurs ?
- **privilégier les résultats des élèves**, avec une conception élargie de ces résultats, qui visent non seulement les « acquisitions scolaires » ponctuelles, mais aussi le devenir des élèves à l'issue de leur départ de l'établissement, ainsi que des aspects plus difficiles à saisir, bien qu'importants, comme les apprentissages sociaux et la constitution progressive d'une culture.

Deuxième partie : le paysage de l'évaluation des établissements aujourd'hui est marqué par son caractère disparate

Le rapport procède à un état des lieux des évaluations d'établissements selon trois focales :

1. Une focale européenne

Il est observé que dans plusieurs pays l'évaluation des établissements a été jugée matière si importante qu'on en a chargé une instance indépendante du ministère de l'éducation nationale. A la différence de la situation française, il est fréquent que les évaluations soient obligatoires, avec une périodicité imposée, que les protocoles d'évaluation soient définis publiquement et s'imposent à ceux qui les mettent en œuvre et que les évaluations débouchent sur un suivi obligatoire. La publicité des rapports est également une différence significative, ainsi que l'insistance des autorités sur l'auto-évaluation, en limitant en ce cas l'évaluation externe à un contrôle qualité des auto-évaluations.

2. Une focale nationale

Si à l'origine, l'inspection générale était bien chargée d'un regard global sur les établissements du second degré, on a assisté par la suite à une différenciation entre des activités d'inspection « administrative » des établissements, qui se sont poursuivies, et d'inspection « pédagogique », qui ont délaissé l'approche d'établissement pour se centrer sur le contrôle individuel des maîtres. L'IGAENR, effectuée, ainsi que l'IGEN « établissements et vie scolaire », des visites systématiques d'établissements : elle a élargi son regard au fonctionnement systémique de l'établissement et élaboré des protocoles de qualité, mais :

- les protocoles ne sont ni publics ni toujours respectés ;
- les critères de choix des établissements évalués ne sont pas définis ;
- la distinction entre le commanditaire (le ministre) et le destinataire (le chef d'établissement et les autorités déconcentrées) est source de difficultés ;
- il n'existe pas de procédure de suivi ;
- les évaluations d'établissements sont peu visibles dans le système et généralement méconnues des usagers ;
- le nombre même des établissements ne permet pas à ces visites d'établissements d'avoir une périodicité crédible ;
- de par ses compétences, l'IGAENR ne s'intéresse pas centralement aux acquisitions des élèves, ainsi qu'à l'utilisation fine des moyens pédagogiques, parmi lesquels figurent les talents des maîtres.

3. Une focale académique

Le rapport aborde aussi la question des évaluations d'établissements scolaires mises en place dès la décennie 1990 dans plusieurs académies, sous des dénominations diverses.

Si le cas de l'académie de Lille se signale par le caractère systématique des évaluations qui y ont eu lieu dès 1991, d'autres démarches académiques souvent inspirées par l'expérience lilloise, mais portant sur un nombre d'établissements plus restreint, émergent depuis quelques années et tentent d'apporter un essai de réponse au besoin de disposer d'un dispositif éprouvé d'évaluation d'établissements. Ces tentatives rencontrent cependant des limites évidentes :

- les objectifs sont souvent ambigus : les évaluations n'osent pas s'afficher pour ce qu'elles sont et les critères du choix des établissements sont flous ;
- le déroulement des visites a souvent connu des flottements avec une prédominance d'une logique de contrôle et des inspections individuelles souvent déconnectées de l'ensemble ;
- la restitution et le suivi obéissent à des procédures et à des logiques fort disparates.

Au total, l'impact de ces évaluations reste limité en dehors des personnels d'encadrement :

- les contenus des rapports intègrent peu la dimension pédagogique : l'analyse des résultats, des processus d'orientation, des modes d'évaluation, des effets de la formation, ne dépassent qu'exceptionnellement le stade du constat pour rechercher des explications. L'objectivation des pratiques enseignantes est faible et les préconisations souvent conventionnelles ;
- les perceptions des évaluations sont contrastées entre évaluateurs et évalués : pour les premiers ces évaluations ont permis la mise en place d'indicateurs appréciés et le développement d'une culture commune aux chefs d'établissements et inspecteurs territoriaux. Pour les évalués, l'impact est très limité : seule une clarification des objectifs de l'évaluation ainsi qu'un effort significatif de formation des personnels d'inspection et de direction sembleraient susceptibles d'améliorer cet impact.

Troisième partie : l'évaluation des EPLE demain : perspectives méthodologiques et stratégiques

Le rapport fait des propositions dans trois directions :

1. « Revisiter l'évaluation de l'EPLE »

L'évaluation de l'action publique dans l'éducation nationale demande, plus encore qu'en d'autres domaines, que sa crédibilité soit assise sur des règles déontologiques strictes et adaptées à la mission d'éducation.

C'est ainsi que le rapport insiste sur la nécessité :

- **de recourir pour l'établissement scolaire à une démarche spécifique d'évaluation** : comme toute évaluation d'action publique l'évaluation d'établissement scolaire emprunte à la fois aux activités de contrôle ou d'audit, tout en étant multiforme, et aux activités de recherche scientifique, tout en se préoccupant davantage de son utilité sociale ; mais elle tient compte aussi du fait que l'activité éducative est une co-production entre prestataires de service (personnels de direction, d'éducation et enseignants) et usagers (les élèves) : elle se distingue d'autant plus clairement du contrôle des personnes qu'elle développe une attitude d'écoute compréhensive et associe les évalués à la méthode et aux constats de l'évaluation.
- **de rechercher la complémentarité entre évaluations interne** (dont la finalité est que les acteurs « se rendent compte » des effets de leur action) **et externe** (dont la finalité est de rendre compte à une hiérarchie ou à des usagers) : si l'auto-évaluation a l'avantage de mieux impliquer les acteurs de l'établissement, ses limites (difficulté des acteurs à se distancier de leur action et de l'établissement à

se situer dans un ensemble plus large) conduisent plutôt à rechercher une complémentarité entre les deux formes d'évaluation : l'auto-évaluation peut servir de point de départ à l'évaluation externe, ou être son aboutissement ; l'évaluation externe peut favoriser l'amélioration de la qualité de l'auto-évaluation par les effets d'apprentissage qu'elle produit ; les indicateurs externes peuvent se combiner avec des indicateurs conçus en interne plus adaptés à des problématiques spécifiques.

- **de construire une problématique adéquate de l'évaluation :**
 - en préférant la recherche de quelques grandes problématiques à celle de l'exhaustivité ;
 - en privilégiant, dans tous les cas, l'approche par les résultats articulée à une évaluation par des processus et des pratiques professionnelles – pédagogiques et organisationnelles – qui ont contribué à les faire advenir dans un contexte donné ;
 - en ayant recours à des protocoles uniformes mais aussi à des approches adaptées à la spécificité de chaque établissement ;
 - en faisant la part du recours aux référentiels national et académique externes et au référentiel interne que constitue le projet d'établissement ;
 - en approchant les résultats des élèves non de manière absolue (absence de référentiels de compétences d'élèves par niveau en France), mais de façon relative, par comparaison dans l'espace et dans le temps ;
 - en s'entendant, pour chaque objectif dont on veut mesurer l'atteinte, sur des critères et des indicateurs (chiffrés quand ils existent), tout en les considérant avec précautions et en les croisant.

2. « Améliorer l'évaluation dans le cadre de l'existant »

Le rapport conclut que dans le court terme il faudrait :

- **assurer une meilleure transparence à l'évaluation externe et associer davantage l'établissement à chacune de ses phases**

L'analyse du dossier initial devrait servir à l'élaboration d'un pré-diagnostic, communiqué pour avis au chef d'établissement. L'association des inspecteurs des différentes disciplines, travaillant sur des grilles d'évaluation harmonisées, devrait être plus forte, de manière à permettre d'articuler les observations de classes à la problématique d'évaluation.

- **recentrer l'objet de l'évaluation sur la construction des performances au sein de l'EPL**

Il s'agit d'abord de l'évaluation des performances de l'EPL lui-même envisagées sous l'angle de l'**efficacité** (capacité de l'école à faire progresser les élèves dans les acquisitions scolaires et la maîtrise de compétences sociales), de l'**équité** (capacité à réduire les inégalités de réussite entre les élèves) et de l'**efficience** (capacité de l'établissement à utiliser de manière optimale ses ressources).

Il s'agit ensuite de déterminer la contribution propre de l'établissement à la production des résultats : recherche de sa valeur ajoutée, au-delà des seules catégories socioprofessionnelles ; recherche de l'effet des différentes composantes emboîtées – de la classe à la direction – qui constituent l'établissement.

Il s'agit enfin de mesurer la capacité de l'établissement à maîtriser son environnement : l'établissement a-t-il un comportement d'inertie, ou un comportement réactif ou proactif par rapport à son environnement ? quel est son rôle au sein du bassin, dans ses relations verticale et latérale avec les autres EPL,

et avec l'ensemble de ses partenaires (collectivités, institutions et services, entreprises, associations) ?

- **mieux utiliser les indicateurs d'évaluation disponibles**

Les indicateurs statistiques disponibles sont nombreux mais mal utilisés, à commencer par les IPES produits par la DEP. Ils doivent être enrichis, au moins pour les lycées, par la recension, depuis peu disponible, des notes obtenues au baccalauréat.

Il convient, par ailleurs, d'utiliser ces indicateurs avec méthode, en distinguant le diachronique du synchronique, en comparant les indicateurs avec ceux des établissements voisins, en se gardant de prendre une relation de corrélation pour une relation causale unilatérale.

3. « Repenser à terme l'évaluation des EPLE »

Le rapport propose à terme de :

- **revoir l'articulation entre les différents niveaux d'évaluation pour constituer une « chaîne » :**

- **le cadre académique** doit être privilégié dans la logique de la déconcentration et pour répondre au problème soulevé par le nombre des établissements (11000, avec l'enseignement privé sous contrat). La généralisation des démarches d'évaluation initiées dans certaines académies, fait encourir un risque de surcharge des tâches des inspections territoriales ; on lui préférera, dans la logique des expérimentations initiées par la commission européenne, la mise en œuvre d'« auto-évaluations accompagnées », les corps territoriaux d'inspection étant alors chargés de la formation des chefs d'établissements et de la sensibilisation des enseignants aux démarches d'évaluation, ainsi que du contrôle de la qualité des auto-évaluations ;

- **l'organisation des rectorats et la fonction des inspections générales** dans l'évaluation des établissements devrait être redéfinie en conséquence. Une direction de l'évaluation devrait être créée dans chaque rectorat et confiée à un responsable de haut niveau dans des conditions où elle serait indépendante des directions opérationnelles. Quant aux inspections générales, même si elles continuent à visiter les établissements dans un cadre ciblé à la demande du ministre, elles auraient à la fois à veiller à la cohérence du système d'informations statistiques, à aider à la construction d'outils d'auto-évaluation de qualité, et à contrôler la qualité de la mise en œuvre académique de l'évaluation des établissements.

- **réformer notre système d'informations statistiques**

Il s'agirait de rendre plus homogènes les données disponibles, en :

- **comblant certaines lacunes du système d'information** : ce qui est en cause est l'hétérogénéité entre les données disponibles pour les collèges (très peu de renseignements sur les résultats de l'examen terminal, mais d'assez bonnes données sur les poursuites d'études après le collège) et les lycées (où c'est un peu l'inverse), ainsi que la faiblesse des indicateurs de vie scolaire ;

- **améliorant la fiabilité des données** : certaines enquêtes ont des taux de réponses insuffisants et les procédures d'identification des catégories socioprofessionnelles dans les établissements sont entachées de trop d'incertitudes ;

- **assurant un meilleur chaînage des indicateurs** produits aux différents niveaux du système, notamment académique et national ;

- **redéfinissant la notion de valeur ajoutée et en revoyant son mode de calcul** conformément aux standards internationaux, par comparaison des niveaux de compétence en début et en fin de cycle par des épreuves standardisées ;
- **réservant la production et l'exploitation de données individuelles aux niveaux académique et infra-académique** : pour profiter notamment des fonctions d'études des notes mises à la disposition des établissements avec les logiciels de saisie de notes dans les bulletins.
- **élargir et confronter divers modes d'investigation par :**
 - l'introduction d'enquêtes par questionnaires, afin de mieux saisir les perceptions des acteurs ;
 - la réévaluation du rôle et du statut des documents écrits dans l'évaluation en ne leur accordant une valeur probante que dans la mesure où ils sont corroborés par les autres formes d'investigations ;
 - un recentrage des grilles d'observation sur les activités des élèves et les processus d'apprentissage : une homogénéisation raisonnée des grilles d'évaluation des pratiques enseignantes entre disciplines est à rechercher, en particulier dans le cadre d'une auto-évaluation entre pairs, de même que leur ouverture aux aspects interdisciplinaires ; les observations de classes pourraient être suivies d'entretiens avec des élèves pour mieux approcher ce qui s'est passé ;
 - une meilleure articulation entre les types d'entretien et la nature des informations qu'ils peuvent apporter, en distinguant clairement les apports factuels des simples représentations d'acteurs.

Bien conduite, l'évaluation de l'EPLC constitue l'un des vecteurs susceptibles d'améliorer grandement la qualité du service public d'éducation en France en favorisant chez les acteurs une connaissance réflexive de leurs pratiques.