

Les statistiques : de leur utilisation dans le domaine public en remontant vers leurs sources.

Exemple tiré du Débat National sur l'école

René Mulet-Marquis(*)

Le « Débat National » sur l'école s'est déroulé en fin d'année 2003 et en début d'année 2004. Je souhaite, en partant d'un exemple, réfléchir avec vous, sur l'utilisation et l'interprétation de données chiffrées dans les débats publics. Je vous propose une réflexion en trois temps. Tout d'abord nous rencontrerons un exemple d'interprétation de données proposées au grand public, puis ferons un voyage en remontant jusque aux sources où elles ont été produites et enfin nous nous intéresserons à ces sources.

1. Le cadre d'apparition des données dans le domaine public

Le « Débat National » a été organisé à la demande des autorités politiques par Claude Thélot, spécialiste de l'évaluation. Il s'est entouré de personnalités et d'experts. Les chiffres que nous allons consulter sont extraits du rapport « Éléments pour un diagnostic sur l'école »⁽¹⁾ (nous le nommerons dans la suite EDE) dont les auteurs sont quatre de ces experts⁽²⁾. Ce rapport a été relu par le Haut Conseil de l'évaluation de l'école⁽³⁾ qui en a tiré un avis. L'ensemble est consultable sur le site internet du Débat.

Lisons maintenant dans le cadre ci-dessous un paragraphe d'EDE. Il est cité intégralement.

2.3 Les disparités s'accroissent lors du passage de l'école primaire au collège

L'écart moyen des performances entre enfants de cadres et enfants d'ouvriers est considérable en CE2. Ces performances sont en français 79,8% et 67,5 % respectivement ; elles ne sont que 60,3% pour les enfants d'inactifs. Le résultat des performances en mathématiques est semblable 73,8% et 63,4% (54,4% pour les enfants d'inactifs).

L'écart est encore plus grand à l'entrée en sixième : 78% et 63% en français (59,2% pour les enfants d'inactifs) ; en mathématiques 74,9% et 59,1% (53,7% pour les enfants d'inactifs).

(*) IREM de Lyon. Professeur au collège de Corbas.

(1) Jean-Claude Hardouin, André Hussenot, Georges Septours, Norberto Bottani, *Éléments pour un diagnostic sur l'école*, Document préparatoire au débat national sur l'école.

(2) Jean-Claude Hardouin universitaire, André Hussenot Inspecteur général de l'éducation nationale, Georges Septours Inspecteur général de l'éducation nationale, Norberto Bottani Directeur du service de la recherche en éducation (Genève).

(3) Le Haut conseil à l'évaluation de l'école est suivant la présentation faite sur son site internet une instance indépendante. Présidé à sa création en 2000 par Claude Thélot, il est dirigé depuis avril 2003 par Yves Mariani.

Le message contenu dans le titre: « **Les disparités s'accroissent lors du passage de l'école primaire au collège** » est développé dans le texte : « L'écart moyen des performances entre enfants de cadres et enfants d'ouvriers est considérable en CE2... L'écart est encore plus grand à l'entrée en sixième ». Une série de douze pourcentages vient appuyer ce message. Pour bien comprendre le cheminement des experts, il faut d'abord mettre de côté les chiffres concernant les enfants d'inactifs puisque ce sont les performances des enfants de cadres et d'ouvriers qui font l'objet du commentaire, puis rassembler les données concernant d'une part le français et d'autre part les mathématiques. C'est ce qui est fait dans les deux tableaux ci-dessous.

Français

	CE2	Sixième
Enfants de cadres	79,8	78
Enfants d'ouvriers	67,5	63

Les experts ont constaté que, en français, l'écart entre les performances des enfants de cadres et d'ouvriers est de 12,3% au CE2 et de 15% en sixième. Ils concluent à l'accroissement des écarts.

Mathématiques

	CE2	Sixième
Enfants de cadres	73,8	74,9
Enfants d'ouvriers	63,4	59,1

Les experts ont constaté que, en mathématiques, l'écart entre les performances des enfants de cadres et d'ouvriers est de 10,4% au CE2 et de 15,8% en sixième. Ils concluent à l'accroissement des écarts.

Nous avons maintenant compris comment l'affirmation du titre « les disparités s'accroissent lors du passage de l'école primaire au collège » s'articule avec les pourcentages fournis. Armés de cette expertise, nous pouvons maintenant l'utiliser pour comparer l'évolution des écarts de performances, par exemple en français, entre les enfants de cadres et d'inactifs, toujours à partir des données qui nous sont fournies (le constat que nous allons faire étant sensiblement identique en mathématiques, nous prenons le français comme exemple prototypique) :

Français

	CE2	Sixième
Enfants de cadres	79,8	78
Enfants d'inactifs	60,3	59,2

L'écart en français entre les performances des enfants de cadres et d'inactifs est de 19,5% au CE2 et de 18,8% en sixième. **Il y a une stabilité des écarts.**

Poussons jusqu'au bout la logique des experts en comparant l'évolution des écarts des performances en français entre les enfants d'ouvriers et les enfants d'inactifs.

Français

	CE2	Sixième
Enfants d'ouvriers	67,5	63
Enfants d'inactifs	60,3	59,2

L'écart en français entre les performances des enfants d'ouvriers et les enfants d'inactifs est de 7,2% au CE2 et de 3,8% en sixième. **Il y a une baisse des écarts.**

Que venons-nous d'établir ?

À partir des chiffres fournis par les experts et en adoptant la même logique de comparaison, suivant les catégories que l'on choisit de comparer, on peut conclure à l'augmentation, à la stabilité ou à la baisse des écarts de performances. En utilisant les chiffres et les méthodes des experts il est totalement déraisonnable de conclure quoi que ce soit de général.

2. Origine des données chiffrées

Les pourcentages cités dans EDE sont extraits d'un Rapport⁽⁴⁾ du Conseil de l'emploi des revenus et de la cohésion sociale (CERC)⁽⁵⁾. On les trouve page 13 de ce rapport, dans un grand tableau où figurent bien d'autres catégories que les trois (enfants de cadres, d'ouvriers, d'inactifs) retenues par les auteurs de EDE. Dans le rapport du CERC, le titre qui précède le tableau est : « **Les disparités s'accumulent pendant le premier degré** » ; dans Éléments pour un diagnostic sur l'école, il est devenu « **Les disparités s'accroissent lors du passage de l'école primaire au collège** ». Or les performances dont il est question étant mesurées à l'entrée en sixième, l'enseignement au collège *n'est pas concerné*. Le choix d'un titre n'est pas neutre : il influe sur la suite de la lecture et sur l'interprétation que le lecteur en fera. Introduire le mot collège dans le titre c'est engager le lecteur sur une fausse piste. Il y a là une très grande ... maladresse. Selon qu'on titre : « Les disparités de salaire s'accumulent pendant le gouvernement Dupont » ou bien : « Les disparités de salaire s'accroissent lors du passage du gouvernement Dupont au gouvernement Durand », on n'envoie pas le même message. De plus dans le rapport du CERC on trouve : « Les différences de scores obtenus par les élèves entrant en sixième sont légèrement plus marquées que celles des élèves entrant en deuxième année de cours élémentaire » quand il est question dans EDE d'« écart encore plus grand à l'entrée en sixième ». Passer de « différences légèrement plus marquées » à « l'écart est encore plus grand », c'est opérer un glissement. Choisir de remplacer le mot score par le terme performance, c'est, dans un cas faire référence au passage d'une épreuve, ce qui limite la portée des nombres, dans l'autre cas assimiler le pourcentage à une mesure des performances des élèves (voire à une mesure de performance du système).

(4) Rapport numéro 3 du Conseil de l'emploi des revenus et de la cohésion sociale (CERC) de Mars 2003.

(5) Le CERC est présidé par Jacques Delors, on trouve, entre autres, parmi ses membres Xavier Emanuelli (Fondateur du SAMU social), Jean Marc Espalieu (Président d'ACCOR).

Poursuivons notre voyage. Le CERC n'a pas produit ces chiffres, mais il fournit la référence permettant de remonter à leur source – il faut plutôt écrire leurs sources, *car ce sont deux documents séparés* –. L'un concerne des évaluations faites en CE2 en 2000, l'autre celles faites en début de sixième, également en 2000⁽⁶⁾. *Il ne s'agit pas d'une cohorte d'élèves suivie sur plusieurs années, mais d'élèves différents* ! Il n'est pas question dans ces documents de comparer les résultats des évaluations faites la même année en CE2 et en sixième par des élèves différents ! Des précautions sont prises, par exemple : « les protocoles ne recouvrent pas tous les apprentissages prévus par les textes ». On est très loin de l'idée de la mesure absolue d'une performance en français ou en mathématiques.

Méthodologiquement, annoncer dans EDE un accroissement des disparités alors que l'on manipule des chiffres portant non seulement sur des épreuves différentes mais sur des populations différentes pose quelques problèmes ... Le BO⁽⁷⁾ pour dénoncer par avance toute utilisation abusive utilise des mots très clairs : « **Aussi, ces résultats ne peuvent-ils en aucun cas être utilisés à des fins de comparaisons d'une année sur l'autre et détournés de leur objet pédagogique** ».

Nous sommes arrivés aux sources, cette partie du voyage s'achève. Résumons-la, dans l'ordre chronologique.

Les données de départ sont les résultats de deux évaluations diagnostiques destinées à détecter certaines difficultés et certaines réussites individuelles des élèves en début de CE2 et de sixième. Elles transitent par un document du Conseil de l'emploi des revenus et de la cohésion sociale qui les regroupe, tout en prenant encore quelques précautions avant d'arriver dans *Éléments* pour un diagnostic sur l'école. Les auteurs de ce dernier document choisissent certaines données et les interprètent en dehors du cadre et des objectifs pour lesquels elles sont produites. Ils comparent des pourcentages qui ne peuvent pas l'être. Ils transforment des scores à des tests concernant l'école primaire en indices de performances, et enfin les présentent sous un titre qui peut laisser supposer au lecteur que ce sont des indices concernant le collège.

3. Au sujet des évaluations

Les évaluations de début de sixième⁽⁸⁾ sont conçues pour repérer les difficultés et les réussites des élèves. À partir d'un cahier des charges fourni par un groupe de pilotage

(6) Andrieux, Brézillon, Chollet, Remvikos. *Les élèves en début de CE2. Évaluation de septembre 2000*. Note d'information numéro 01.35 juillet. Direction de la programmation et du développement

Andrieux, Dupé, Robin. *Les élèves en début de sixième. Évaluation de septembre 2000*. Note d'information numéro 01.36 juillet. Direction de la programmation et du développement.

(7) Citons un extrait du BO N° 25 du 29 Juin 2000 : « Il faut rappeler en effet que les évaluations nationales CE2 et sixième, tout comme celles d'entrée en seconde de lycée, n'ont de valeur qu'annuelle puisque les supports des évaluations et les objectifs évalués diffèrent chaque année. Aussi, ces résultats ne peuvent-ils en aucun cas être utilisés à des fins de comparaisons d'une année sur l'autre et détournés de leur objet pédagogique ».

(8) René Mulet-Marquis. « *Point de vue* ». Repères Irem n° 52. Ce qui suit dans la partie 3 est inspiré de ce point de vue.

national, les questions sont produites par des groupes académiques. Elles sont testées un an à l'avance dans des classes témoins suffisamment variées pour qu'on ait une bonne connaissance de leur difficulté réelle.

Le groupe de pilotage national dispose donc des informations nécessaires pour placer le curseur là où il le souhaite : si demain on veut 30% ou 50% ou 80% de réussite aux évaluations, il est facile de les obtenir. Disons-le clairement : le pourcentage obtenu à l'issue des évaluations sixièmes relève (avec une certaine marge d'erreur) du **choix conscient** du groupe de pilotage. Avec un choix de questions légèrement différentes, sur environ 90 questions posées, il suffit d'en remplacer 5 dont on sait qu'elles sont généralement ratées par les élèves en difficulté, par 5 dont on sait qu'elles sont généralement réussies pour que l'on ait des résultats qui sembleraient infirmer les affirmations citées ci-dessus. Bien entendu les conclusions que l'on pourrait être tenté d'en tirer en suivant la logique de comparaison des experts ne seraient pas davantage fondées. Voilà qui relativise les résultats bruts.

Pour avoir une idée de l'évolution des élèves, il faudrait faire passer, par exemple en début et en fin d'année, **à la même cohorte d'élèves, les mêmes épreuves**. On saurait alors si, sur cet ensemble de questions, et pas plus !, le taux moyen de réussite a augmenté ou diminué. De là à conclure en terme de régression ou de progression des élèves, il y a encore un grand pas qu'il faut franchir avec précaution.

J'ai essayé de faire ce travail deux fois dans des classes de sixième de consolidation à effectifs réduits constituées d'élèves en difficulté. Dans les deux cas, après avoir fait passer en fin d'année les mêmes évaluations qu'en début d'année, j'ai observé des trajectoires très variées. Les résultats de certains élèves avaient stagné, voire légèrement régressé, la plupart s'étaient améliorés dont quelques-uns, peu nombreux, de manière spectaculaire.

Même dans ce cas de répétition de tests identiques, le passage du taux moyen de réussite à l'idée d'une progression ou d'une régression **des** élèves ou même de certains groupes d'élèves est risqué. Le taux moyen ne nous dit rien du parcours personnel de chaque élève. On peut tout à fait avoir un taux moyen en progression en même temps qu'un nombre non négligeable d'élèves a décroché. L'un des grands défis auquel est confronté le collège c'est la scolarisation d'élèves en grande difficulté. L'observation d'un taux moyen ne permet pas de savoir ce qu'il advient de ces élèves. Le taux moyen de réussite, comme toute moyenne, donne une information assez pauvre. Entre ce taux moyen et des affirmations telles que : les élèves ont progressé ou régressé, ou même, les disparités entre telle et telle catégories d'élèves se sont accrues, il y a un abîme !

En conclusion

Le travail ci-dessus porte sur un paragraphe d'un rapport comportant de nombreuses pages. Je ne tirerai donc aucune conclusion sur l'ensemble de ce rapport. En revanche les mécanismes mis au jour me semblent avoir une portée très générale. Chaque fois que l'on présente, utilise ou consulte des données chiffrées, il faut faire preuve d'une extrême vigilance. En particulier lorsque les données sont utilisées en dehors et parfois très loin de leur contexte de production, les interprétations qui en sont faites ont plus de chance de traduire les conceptions des rédacteurs sur le sujet que ce qui peut être raisonnablement déduit des chiffres.

Il serait souhaitable d'assurer une meilleur « traçabilité » des informations : pour éviter les déformations successives, il faut donner leur source première. Dans un autre document consultable sur le site du Débat : « *Ouvrir le débat national sur l'avenir de l'éducation* », on trouve quelques pages de tableaux de nombres et graphiques divers. Les sources sont indiquées globalement de la manière suivante : MJENR, DEP, OCDE, CERI, INSEE. C'est dire qu'il est très difficile, voire impossible de les consulter. Or les informations données sont très loin d'être transparentes. Par exemple que penser de ces chiffres : les dépenses pour l'éducation en 2002 s'élèvent à 108,1 milliards d'euros et le budget du MEN est de 61,4 milliards d'euros. Que recouvrent ces deux chiffres ? Autre exemple : on apprend en lisant un graphique que le nombre moyen d'élèves par enseignant en 2002 est de 11,8 et qu'il a régulièrement baissé depuis 20 ans. Doit-on comprendre que les élèves sont en général dans des classes d'une douzaine (c'est contraire à l'expérience commune !). Donnés sans aucun éclairage, ce chiffre ne peut que laisser le lecteur très perplexe ou l'entraîner dans des interprétations « sauvages ».

De même qu'il faut prendre avec prudence les données présentées hors de leur contexte de production, il faut accueillir avec précaution le titre d'expert : dans son domaine, il dispose d'une compétence particulière ; en dehors de son domaine d'expertise, son jugement n'est pas meilleur que celui du commun des mortels.

Le Débat sur l'école doit déboucher sur une réforme, en particulier celle du collège. On peut trouver dans des déclarations des ministres ou de leurs proches les grandes lignes de ce que pourrait être cette réforme : enseignants plus polyvalents, travail plus orienté vers le pluridisciplinaire, autonomie des établissements, etc. Rien ne permet d'affirmer à ce jour que ces réformes sont susceptibles d'apporter une amélioration pour les élèves ou les catégories sociales les plus en difficulté. Nous ne disposons pas d'études rigoureuses qui iraient dans un sens ou dans l'autre. Tout ce qui est dit à ce sujet relève de l'opinion. Il faut être prudent avec les opinions. Pendant très longtemps on pensait qu'il fallait saigner les patients pour les soigner : c'était une opinion de « bon sens ». Le jour où des méthodes plus sérieuses ont été disponibles, on a cessé de le faire : c'était contraire au « bon sens », et pourtant quantité de malades ont été sauvés.

Nous risquons de nous trouver dans la situation suivante : à partir d'un diagnostic mal posé, qui amène à des conclusions erronées, vont être proposés des remèdes dont on ignore les effets, et en particulier s'ils sont une réponse aux problèmes dont souffre le collège à l'heure actuelle.