

# Évaluer autrement

André Gagneux

*André GAGNEUX, professeur de mathématiques ayant exercé de 1973 à 1992 dans un collège de ZEP, est désormais responsable du site de Bourges de l'IUFM d'Orléans-Tours. Lors des journées d'Orléans en octobre 2004, il a animé l'atelier intitulé « Un environnement de l'enseignement : évaluer autrement ». Ce thème a retenu toute notre attention et nous remercions André Gagneux d'avoir accepté que le compte rendu se transforme en article pour notre rubrique « La parole à... ». Voici donc une lecture qui vient compléter celles proposées sur ce même thème dans nos précédents numéros.*

**I - Décrivons ce qui nous est demandé par les Instructions Officielles (IO), et ce que la société exige de nous de plus en plus précisément.**

**- un contexte historique et institutionnel**

La loi d'orientation de 1989 affirme pour la première fois l'obligation faite aux enseignants d'évaluer les copies des élèves - cela peut surprendre car l'idée de mettre des notes pour trier et classer les élèves vient des jésuites et a plus de 250 ans ! Elle ne précise pas sous quelle forme, — et en particulier on ne dit pas qu'il faut mettre une note sur 20 — 0 mais affirme que l'évaluation fait partie du travail de l'enseignant. Un autre choix aurait pu être fait, celui **de séparer les deux activités**, comme cela se passe lors des examens ou concours.

**- des pratiques courantes traditionnelles**

Même s'il y a de nombreuses formes d'évaluation (les anciennes compositions trimestrielles, le contrôle continu, etc.), la tradition toujours en vigueur veut que ce soit une notation chiffrée qui soit majoritairement employée. Après quelques tentatives d'avoir une autre réflexion (docimologie, lettres, appréciations) l'idée se renforce qu'« on impose de mettre des notes », par exemple dans le dossier 4<sup>e</sup> / 3<sup>e</sup> qui accompagne l'évaluation continue au

moment du Brevet, ou qu'il faut un certain nombre de contrôles par trimestre pour en faire la moyenne.

**- un environnement professionnel en mutation**

Dans tous les métiers, on définit ce qu'on appelle un **référentiel de métier** qui décrit, par un certain nombre de phrases, les compétences que les professionnels exerçant ce métier doivent posséder. Depuis 1994, il y a aussi un référentiel du métier de professeur des écoles et, depuis 1995, le référentiel du professeur de lycée-collège, puis celui du chef d'établissement. C'est d'ailleurs ce qui permet à l'IUFM de pouvoir finaliser ses formations. On peut ainsi, à partir de ces documents, attester, et valider ces compétences, y compris dans la « valorisation des acquis de l'expérience », lors des recherches d'emploi, ou des entrées en formation continue. Une procédure d'unités capitalisables, que les collègues de LP ou de SEGPA connaissent aussi, permet de valider des « morceaux » d'acquisition pour, tout au long de la vie, pouvoir progresser par la formation continue.

**- une demande sociale et institutionnelle plus précise**

Le rapport THELOT introduit la notion de « socle » dont l'évaluation « doit porter

André Gagneux est auteur de « Évaluer autrement les élèves » Ed. PUF, collection Éducation et Formation. Vous pouvez lire la critique de cet ouvrage, écrite par Henri Bareil, dans le BV n°456.

aussi bien sur les connaissances que sur les compétences, y compris les compétences transversales aux disciplines ». À l'inauguration de nos journées, M. Moisan, le doyen de l'Inspection Générale, disait : « le socle, ça ne doit pas être le programme, ça doit être des compétences ». Dans la société — et notre Institution traduit dans les textes cette évolution — pour valider une formation ce qu'il faut pouvoir dire, c'est la liste des compétences qui sont atteintes ou qui ne sont pas atteintes.

**Accompagnant les programmes, dans la plupart des disciplines, il existe des listes de compétences à atteindre, notamment en fin de cycle.**

En mathématiques, nous sommes d'ailleurs favorisés, car c'est depuis 1985 qu'il y a cette cohérence entre la demande de la société, et les textes officiels, et que, jusqu'en seconde, les programmes sont présentés sur 3 colonnes dont l'une s'appelle « compétences exigibles ».

*Exemples :*

- l'examen d'ouvrier professionnel de lingerie : c'est une *liste de compétences* qui sont rangées, pour lesquelles on se pose la question « *le formé est-il capable de ?* ». Et l'attribution de la qualification est don-

née en fonction des croix mises dans des cases, d'éléments indiquant si la compétence est possédée ou non.

- à l'école maternelle il existe, depuis 14 ans, des carnets d'évaluation « *pour informer les familles des progrès de l'enfant et servir d'outils de liaison entre les enseignants* ». On conçoit assez bien qu'il n'y ait pas de notes en maternelle : on n'a pas cherché à en mettre, donc on fait autrement : deux fois dans l'année, le maître marque sur ce carnet d'évaluation les atteintes aux objectifs. Conformément à la demande de la société, ce carnet est communiqué aux parents.

**- des aides institutionnelles pour changer les pratiques**

Les évaluations nationales (GS, CP, CE2, 6<sup>e</sup>, 5<sup>e</sup>, 2<sup>de</sup>) ont apporté une originalité : au contraire des évaluations qui certifient les compétences et que l'on fait donc **après l'apprentissage** pour savoir si les connaissances sont acquises, ces évaluations doivent se faire **en début d'année**, et il est **interdit de les noter**.

Une aussi grosse opération, sur le plan national, doit coûter très cher depuis quinze ans ! C'est donc qu'elle doit porter un message et participer d'un changement des pratiques selon **trois principes** :

- on peut évaluer **sans forcément noter**,
- on peut, on doit même, **évaluer au début**, et pas seulement à la fin,
- dans une évaluation initiale, on est en situation de continuer à enseigner. Il y a donc un **contrat**, contrat moral pour le moins, de tenir compte des résultats de cette évaluation pour faire acquérir les compétences manquantes et, en quelque sorte, « boucher les trous ».



## II – Le système traditionnel permet-il ces évolutions ?

### - une notation peu fiable

Les travaux sur l'injustice, la non justesse et la non répétition des notes sont anciens (vers 1925 Pieron a créé la docimologie) et bien connus.

Premier exemple : l'expérience *Carnegie* montre qu'il faut 127 professeurs de philo-

sophie, pour que la nouvelle note mise

par un nouveau correcteur

ne fasse plus changer

la moyenne des

notes déjà mises

à une copie don-

née. Et même en

mathématiques où

on se dit qu'il y a des

barèmes qui peuvent cadrer

les réponses, il faut tout de

même 13 correcteurs.

Deuxième exemple : vous savez bien que votre 12 ne vaut pas celui d'un autre collègue ! Avec 12, l'élève est premier avec certains professeurs et dans la moyenne avec d'autres. Pour les parents qui passent de professeurs en professeurs, dans la classe, et d'année en année qu'est-ce que ça signifie que d'avoir 12 ? On ne sait pas très bien...

Troisième exemple : des études montrent les effets de halo (corriger une copie après une autre copie change la façon dont cette dernière va être notée). Si vous corrigez en deux jours, vous n'allez pas noter de la même façon la première copie et la dernière copie et il y a même des travaux récents qui affirment que les filles ne sont pas notées comme les garçons...

### - une notation réductrice

Quand vous rendez vos copies, vous avez effectué un travail considérable, souvent long et fastidieux, et en corrigeant vous

avez remarqué beaucoup de détails. Même si vous avez annoté ces détails sur la copie (des études montrent que 80% des copies sont annotées), ce travail reste inutile :

- pour vous, parce que vous n'en gardez pas trace : seule la note globalisante est généralement reportée sur votre carnet ou vos fiches ;

- pour les élèves, car des études montrent qu'il n'y a que 50 % des élèves dans les collèges dits de centre ville, qui disent que ces annotations leur servent à quelque chose, et qu'il y a 75% des élèves dans les ZEP, qui affirment ne jamais lire les annotations, et ne regarder que la note !

On pourrait d'ailleurs se demander pourquoi utiliser autant de notes, puisque, pour l'élève et son avenir, ce qui change tout c'est d'avoir plus de 10 ou moins de 10 ! C'est en tout cas ce qu'ont intégré les élèves et leurs parents !

- pour la transmission des informations utiles et précises, (aux parents, aux collègues, à l'établissement) que la société attend, puisque cette globalisation de la note ne permet plus d'attester des compétences acquises, mais seulement de situer l'élève dans les « bons », les « moyens » ou les « faibles » !

Ces études de docimologie montrent qu'il n'y a :

- rien de précis, de fiable, de reproductible, rien de ce qui, scientifiquement, pourrait rassurer et donner une valeur à la note (car évaluer c'est bien donner une valeur à la copie).

Pourquoi tant de variation ? Tout simplement parce que le produit attendu n'est pas défini de manière claire. Le professeur a seulement une image de ce que serait la bonne copie, celle qui vaudrait 20 !

- rien d'utile pédagogiquement pour continuer à progresser.

Quand l'élève a 15, il est content, et il ne



va pas chercher à retravailler ses lacunes. Pourtant, s'il n'a pas 20, c'est qu'il a « perdu des points » quelque part ! Et nous rencontrons ainsi des élèves qui ont toujours « eu la moyenne » mais qui seront incapables d'aller en « S » parce qu'ils ne sont pas au niveau en géométrie, par exemple.

### **III - Il existe des théories et des outils pour éviter ces effets pervers.**

Robert F. MAGER (*Comment mesurer les résultats de l'enseignement* - Bordas 1986) a finalisé la notion « **d'objectif** » : c'est une compétence précise, sans ambiguïté et observable. La phrase qui traduit l'objectif doit comporter un **verbe d'action qui décrit ce que l'on attend de l'élève**, (ex : tracer une perpendiculaire), les conditions matérielles de réalisation (avec une règle et un compas) et les contraintes de la passation du test (documents, matériel, temps...). Elle doit fixer clairement les critères de l'évaluation (degré de réussite, fréquence, tolérance...), c'est-à-dire du jugement de la production de l'élève, d'une manière binaire : atteint ou non atteint.

**Ni les programmes, ni l'enseignement ne se réduisent aux « objectifs », bien au contraire, les activités mathématiques sont riches et donnent du sens, par contre il n'y a pas d'évaluation utile et juste autrement que par « objectifs ».**

**Bien sûr, à chaque fois que cela nous est demandé par l'institution, il nous faut traduire par des notes ces renseignements binaires. Et cela nous ouvre une autre piste de réflexion sur les inconvénients de la notation traditionnelle.**

#### **- un effet pervers amplifié**

Non seulement nous avons déjà des notes trop globalisantes pour bien renseigner,

mais l'usage traditionnel de la moyenne aggrave la situation en décourageant les progrès.

#### **Faire la moyenne c'est mettre en place une pédagogie de l'échec.**

Ainsi, dans le meilleur des cas - celui où nous avons l'occasion de re-tester exactement les objectifs sur lesquels l'élève a échoué - un échec lors du premier contrôle dévalorise la réussite au 2<sup>e</sup>. En effet, si la première production est notée « zéro » parce que l'élève n'a pas réussi, et si la deuxième production, cette fois-ci réussie, est notée 20 sur la même compétence, la moyenne sera de 10, niant ainsi l'acquisition complète de cette compétence.

Pour que l'élève progresse, pour qu'il ait envie de travailler, il faut qu'il puisse constater que ce travail a été vraiment utile.

Dans une moyenne, toutes les notes « comptent » : on amalgame donc les échecs et les réussites pour une même compétence.

**Ainsi, aucune évaluation ne peut plus être considérée comme « formative » si elle « compte dans la moyenne ».**

#### **- un nouveau regard sur l'apprentissage**

Les collègues reconnaissent que, souvent, ils n'ont pas pris le temps de refaire un contrôle sur les mêmes objectifs. Les élèves, dont nous attendons que, même faibles, ils aient acquis « quelque chose » à la fin de l'année ou du cycle ne peuvent donc pas valoriser ces acquisitions, y compris après une « leçon particulière ». Nous favorisons ainsi la rapidité d'acquisition.

Cette absence de visibilité des « progrès » ne nous choque pas assez. En effet, les résultats obtenus ainsi correspondent à la représentation que nous avons de la répar-

tition « normale » des notes dans une classe : une approximative courbe de Gauss, qui correspond grosso-modo à la courbe de répartition de ce qu'on pourrait appeler « l'intelligence générale ».

**Cette représentation est implicitement influencée par l'image du concours et du classement** (dans cette logique, un élève qui ne comprend pas immédiatement et n'assimile pas rapidement est un élève moyen voire faible).

**Elle va à l'encontre de la certification des compétences.** En effet, si les Instructions Officielles qualifient les compétences d'« exigibles », c'est qu'elles doivent être **toutes acquises, et par tous** les élèves.

Le « bon professeur », au contraire de l'imaginaire collectif qui assimile à tort « exigence » et « mauvaises notes », est donc celui qui obtient la réussite de la plupart de ses élèves. La courbe de répartition de ses notes devrait correspondre à la partie gauche de la « courbe de Gauss » qu'on appelle « courbe d'apprentissage » : peu échouent, et presque tous réussissent. L'autre partie de la courbe ne correspond qu'à la sélection (concours ou « prédiction » servant à une orientation élitiste).

### - des principes concrets

1°) Lorsque la liste des objectifs (une certaine maximum par année) est établie, grâce aux compléments des programmes – c'est le verbe d'action employé qui nous permet de savoir à quel « niveau » de difficulté la compétence doit être acquise – un système de grille (quelle qu'en soit la forme) avec des croix dans des cases doit suffire à visualiser de manière binaire quels sont les objectifs qui sont atteints - et à quel niveau -, et ceux qui ne le sont pas. À chaque instant, l'élève aussi bien que l'enseignant sait où il en est.

2°) Pour que le « test » soit fiable, sans

trop multiplier les passations, et mesure bien l'acquisition de la compétence, il faut qu'il ait lieu à une certaine distance de l'apprentissage et de l'entraînement.

3°) Il faut se donner les moyens de proposer au moins un deuxième test pour les compétences non acquises (la réussite vient alors effacer l'échec précédent). Bien sûr le deuxième test est différencié selon les élèves, en fonction des objectifs à re-tester.

4°) À tout moment, pour les nécessités de communication avec les parents ou l'administration, une note peut traduire l'état de la grille : dès le collège les élèves savent très vite la calculer eux-mêmes par le simple rapport entre le nombre de croix obtenu et le nombre de croix « attendu » à ce moment de l'année, c'est-à-dire le quotient du nombre de croix obtenu par le nombre d'objectifs testés, éventuellement pondérés par leur niveau.

Un modèle idéal parce que la note y augmente forcément, mais peu réaliste, car on partirait de « 0 », consisterait à faire le quotient du nombre de croix obtenu par le nombre total d'objectifs à valider donnés par les IO.

**En conclusion, nous devrions garder à l'esprit qu'il faut faire évoluer l'environnement pour faire évoluer l'enseignement, c'est-à-dire entrer dans l'évaluation pour « enseigner autrement ». Les deux principaux leviers de transformation pour plus d'efficacité de l'évaluation seraient de :**

- **globaliser les renseignements le plus tard possible,**

- **effacer les notes au fur et à mesure des progrès pour que la note communiquée reflète l'état actuel des compétences, et non pas les aléas de l'apprentissage.**

**Ainsi les concepts d'évaluation formative et sommative seraient mis en œuvre.**