

QUELQUES REMARQUES A PROPOS DE L'OPERATION D'EVALUATION DECIDEE PAR LE MINISTERE POUR LA RENTREE 89-90 ET DESTINEE A L'ENSEMBLE DES ELEVES ENTRANT EN CLASSE DE SIXIEME

Antoine BODIN
IREM de Besançon

Bien que *petit x* n'ait rien d'un journal d'actualité, il était difficile de ne pas dire un mot, dans notre revue, sur l'opération "*Evaluation-Diagnostic-Formation-Remédiation*" mise en place par le Ministère de l'éducation Nationale à la rentrée 89/90

On sait que cette opération concerne tous les élèves de CE2 et de sixième de France, et qu'il s'agit essentiellement de donner aux enseignants la possibilité de repérer, dès les premiers jours de l'année scolaire, les difficultés des élèves en français et en mathématiques, puis de leur donner, par l'intermédiaire de formations adaptées, les moyens de prendre en charge et si possible d'éliminer ces difficultés. Il ne sera question, dans ce qui suit que de la partie de cette évaluation qui concerne les mathématiques et le niveau sixième.

Au moment où nous écrivons ces lignes, l'évaluation proprement dite a eu lieu, c'est-à-dire que 850 000 élèves ont rempli les quatre cahiers de tests qui leur étaient proposés (quatre fois 20 minutes), que les enseignants ont codé les résultats (entre 10 et 14 heures de travail de codage pour chaque classe !), que les résultats des établissements ont été envoyés aux services académiques ou rectoraux, et que les enseignants se sont entretenus avec les parents pour leur communiquer les résultats de leurs enfants. Simultanément, la formation se met en place, de façon très diverse selon les académies. Disons simplement que l'on compte beaucoup sur un effet de démultiplication des formations. Par exemple, deux jours de formation pour les formateurs académiques, eux-mêmes chargés de former en quatre jours un enseignant par collège, cet enseignant ayant à son tour pour mission de répercuter localement la formation dans des conditions qui restent mystérieuses. Il ne s'agit là, répétons-le que d'un cas de figure, d'autres scénarios coexistent qui semblent assez rarement satisfaisants.

L'évaluation elle-même s'est déroulée sans problèmes sérieux. Cela montre au moins que lorsque l'entreprise Education Nationale décide de réaliser quelque chose, lorsque l'administration est mobilisée pour atteindre un objectif précis, des miracles sont possibles. En effet, quoi que l'on puisse dire ensuite sur la qualité pédagogique ou didactique des documents produits, quel que soit l'effet final de cette opération, il faudra saluer comme une prouesse logistique le bouclage en quatre mois (dont la période des vacances d'été) d'une opération d'une telle ampleur.

Toutefois, prouesse logistique n'est pas synonyme de pertinence. Il n'est pas absolument certain que l'objectif ait été repéré avec suffisamment de finesse et que le matériel amené à pied d'oeuvre soit vraiment celui dont on ait besoin. Il n'est pas non plus certain, qu'après cet effort, l'intendance suive, que la formation soit distribuée sans parcimonie par des formateurs que l'on aura pris la peine de former ; des formateurs qui pourront réellement distribuer des outils d'aide à la remédiation. En réalité l'existence même de ces outils reste problématique et le temps de formation donné pour leur communication et dans certain cas leur création risque fort d'être très insuffisant.

Pour tout dire, nous pensons que la rapidité avec laquelle cette opération a été réalisée, n'a pas permis de la faire précéder des recherches et expérimentations indispensables, ni même de recenser les travaux antérieurs et d'en tirer tout le parti possible. Il faut toutefois souligner que bien que les modules de formation aient été conçus avec trop de hâte, ils témoignent de la qualité de la recherche en Didactique des mathématiques dans notre pays et de la qualité de certains travaux IREM, et donnent une bonne idée de ce qui serait réalisable dans de meilleures conditions.

Malgré les critiques qui précèdent, nous nous garderons d'être trop négatifs. Beaucoup d'enseignants se sont investis avec beaucoup d'énergie et d'espoir dans cette opération ; ils ont accepté de bonne grâce la charge énorme de codage qui leur était imposée, ils estiment souvent que cette opération leur apporte dès à présent une meilleure connaissance de leurs élèves. Nous voulons simplement rester vigilants. Nous pensons que les enseignants doivent garder la maîtrise et la responsabilité de ce qu'ils font, que s'il est bon de leur fournir des outils et de leur apprendre à s'en servir, il convient qu'ils gardent une totale liberté de jugement, ce qui ne semble pas être mis en cause pour l'instant, mais aussi une grande liberté en ce qui concerne leur utilisation. Des instruments, même de qualité, auront un bien piètre effet entre les mains d'enseignants qui se les seront vu imposer sans avoir la possibilité de les mettre en question, sans possibilité de choix. Des déviations auront d'autant plus de chance de se produire que les enseignants abandonneraient une part de leur responsabilité au profit de directives venues "d'en haut".

On nous dit maintenant que l'opération sera reconduite pendant plusieurs années et qu'il y aura toute possibilité d'amélioration et de prise en compte des difficultés et des manques rencontrés la première année. Dans ce cas, espérons en effet que les observations faites lors de cette *première* seront soigneusement analysées et utilisées pour l'amélioration du dispositif. En fait, il semble être sérieusement envisagé que les opérations ultérieures soient décentralisées et confiées aux Académies. Pour une opération d'évaluation de ce type, la centralisation est plutôt un atout. Elle permettrait de mettre les quelques moyens disponibles sur quelques équipes solides, pouvant vraiment effectuer le travail de recherche et de synthèse qui a manqué jusqu'ici. De plus, cette centralisation éviterait de glisser d'une évaluation diagnostique à une évaluation normative, puis petit à petit, par le jeu des surenchères, à une espèce de concours général local. De telles déviations se sont déjà produites dans le passé, et nous espérons pouvoir publier (en français) dans un prochain numéro une histoire (passionnante!) de la prise en compte et du traitement des erreurs aux Etats-Unis, histoire dont l'origine remonte à 1893.(*)¹

Il est donc trop tôt pour faire le bilan de l'opération 88/89. Dans ce qui suit, pour préparer l'avenir, on trouvera un texte écrit au moment du lancement de l'opération et qui à notre avis reste d'actualité, surtout si les opérations futures devaient être décentralisées.

¹BELANGER M., Errors in arithmetic computation: a century of american speculation - Actes de la 39^e CIEAEM - Université de Sherbrooke.

L'évaluation dont il est question présente quelque analogie avec ce que l'on nomme dans l'entreprise le "contrôle réception des marchandises". Dans ce dernier cas il existe toutefois des paramètres mesurables associés à des normes. De plus, le contrôle réception peut très bien refuser de prendre en charge le produit livré. Malgré l'existence de mesures et de normes, la décision est en général basée sur des méthodes statistiques qui incluent la notion de risque. Le "risque du vendeur" ou risque de première espèce (risque de refuser un produit bien qu'il satisfasse aux normes) étant antagoniste au risque de l'acheteur ou risque de seconde espèce (risque d'accepter un produit bien qu'il ne satisfasse pas aux normes). Antagoniste signifiant ici qu'il n'est pas possible de diminuer l'un des risques sans augmenter l'autre.

Certes, l'esprit qui anime initialement cette évaluation n'est sans doute pas, a priori, celui du contrôle réception. Dans son application, il pourrait toutefois le devenir, et cela constitue, à mon sens, un premier risque de déviation. Ne pouvant pas retourner le produit défectueux au fournisseur, il ne restera plus aux professeurs de collègues qu'à faire les réserves d'usage, réserves qui seront tout d'abord destinées à dégager en partie leur responsabilité en ce qui concerne l'évolution ultérieure des élèves. Réserves qui apparaîtront aussi comme une mise en accusation publique de l'Enseignement Élémentaire ou de telle de ses composantes, pouvant aller jusqu'à des mises en accusation individuelles. Contrairement à ce qui est souhaité, la liaison entre l'Ecole et le Collège risque fort d'être la première victime de ces opérations. Les adaptations qui en résulteront ne manqueront pas d'adapter les objectifs de l'école élémentaire à l'évaluation, risquant, en ce qui concerne les mathématiques, de renforcer l'aspect apprentissages des algorithmes au détriment d'autres objectifs dont la maîtrise est d'accès plus difficile, mais dont les probabilités de transfert sont plus grandes.

L'analogie faite ci-dessus entre l'élève et le produit brut livré à l'entreprise peut choquer, et sans doute le doit-elle. Aussi, il est probable qu'elle ne corresponde pas à l'esprit du projet. Nous avons simplement voulu signaler que cette correspondance risquait d'avoir lieu au niveau des effets.

L'analogie médicale est bien sûr plus proche de l'inspiration de cette évaluation : il y est question de diagnostic, de remédiation, de piqûre de rappel.... On peut dès lors s'étonner de voir toute une population, les bien-portants comme les mal-portants, soumise au même test diagnostique, sans prise en compte d'éventuels symptômes individuels. La démarche pour le moins simpliste consistant à déclarer que toute la population sera soumise à un test polyvalent de diagnostic de "*Maladie*" sans qu'une maladie particulière ne soit visée et avant même que le test ait été mis au point, n'est heureusement pas habituelle dans le monde médical. Jusqu'à présent, l'analogie médicale faisait référence à l'évolution de la pratique médicale en fonction des résultats de la recherche, elle ne concernait pas directement les malades. Cette analogie était donc relative à la pratique pédagogique et non aux acteurs eux-mêmes. Il y était question de recherches à mener pour parvenir à une meilleure connaissance des faits d'enseignement, recherches conditionnant la mise au point d'instruments d'enseignement plus efficaces ainsi qu'une formation des enseignants plus adaptée aux difficultés de leur tâche. Il n'était pas, jusqu'à présent, question de porter l'analogie au niveau des élèves, ce qui conduirait à les traiter comme des malades.

Un effet bien connu de l'évaluation a priori est de transformer la prédiction en réalité, c'est l'effet Pygmalion: prévenir un enseignant de l'insuffisance d'un élève en début de formation renforce les risques de voir cet élève parvenir à un niveau insuffisant en fin de formation. Bien des enseignants ont conscience de ce phénomène : ils renoncent

à se faire une opinion a priori sur le niveau global d'un élève, attendant pour cela d'avoir rassemblé par eux-mêmes suffisamment d'informations sur ses compétences avant d'aller consulter un éventuel livret scolaire. Cet attentisme n'étant pas absolument satisfaisant, une pratique d'évaluation de positionnement s'est développée en particulier autour des IREM, elle consiste à faire le point sur les PREACQUIS des élèves avant d'aborder une question nouvelle. La différence essentielle avec l'évaluation uniforme qui nous est proposée (imposée?) est que ces évaluations ne cherchent pas à établir un niveau absolu des élèves et que l'enseignant garde la maîtrise de l'utilisation et de l'interprétation de ces évaluations. Elles ne sont faites que pour lui permettre de mieux adapter la suite de la formation à l'état de savoir de ses élèves et peuvent bien sûr le conduire à davantage différencier, voir individualiser sa pédagogie.

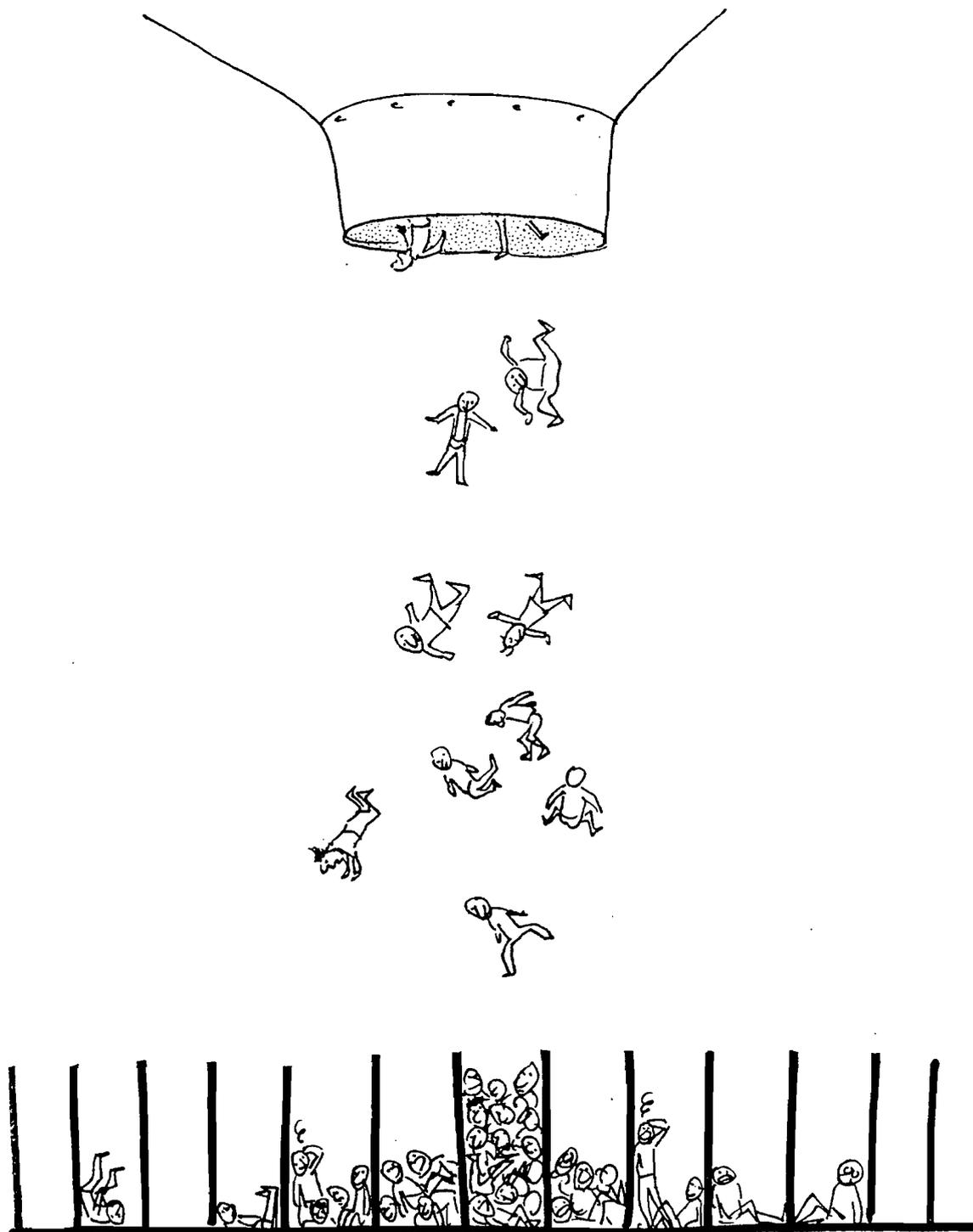
L'évaluation mise en place est fortement liée au souci de mettre corrélativement en place une formation des enseignants leur permettant de mieux prendre en compte les difficultés des élèves. L'intention est louable, mais les effets de système, dont certains seulement sont évoqués ci-dessus, se soucient peu des intentions. L'évaluation devrait donc servir de catalyseur de formation. Mais connaît-on bien la nature de la réaction qui sera initialisée par cette catalyse? Le risque n'est pas négligeable de voir ces formations concourir à la rigidification des idées reçues concernant le faible niveau des élèves arrivant au collège, les carences de l'Ecole Élémentaire, l'impossibilité du "Collège pour tous", la nécessité de mettre en place des filières... Par quoi et comment va-t-on nourrir cette formation? Les éléments de solution qui existent, et qui concernent essentiellement l'analyse et la prise en compte des erreurs, accompagnés d'une instrumentation permettant une certaine différenciation de la pédagogie, ont peu à gagner à être proposés aux enseignants dans le contexte d'une évaluation massive et publique. Bien au contraire, cette évaluation risque de faire obstacle à la diffusion des éléments évoqués.

Longtemps, le corps enseignant a considéré avec beaucoup de suspicion les différentes tentatives d'évaluation globale des effets de l'enseignement. La façon dont sont maintenant accueillies les enquêtes de fin de cycle du SPRESE ou de la DEP, et plus encore, dans un genre différent, celles de l'APMEP, montre que les conceptions ont considérablement évoluées. L'une des règles, essentielle, qui régissent ces évaluations est toutefois qu'il ne puisse y avoir aucune ambiguïté sur les raisons de leur mise en place, ni sur leur caractère réellement global. Ainsi, dans ces évaluations, toutes garanties sont données pour qu'aucun résultat concernant un élève, un enseignant, ou même un établissement particulier ne puisse être communiqué à qui que ce soit. Toute transgression de cette règle aurait pour premier effet d'affaiblir considérablement la fiabilité de l'évaluation. Elle risquerait aussi de ramener les enseignants à leur état antérieur en ce qui concerne l'évaluation en général et de casser la dynamique enclenchée depuis quelques années, au moins chez les professeurs de mathématiques des collèges.

Il doit être clair que le paragraphe qui précède ne met pas en cause la nécessité d'une évaluation individuelle des élèves, des enseignants ou des établissements. Il veut simplement rappeler qu'il n'est pas possible de poursuivre ces divers buts simultanément et en utilisant les mêmes instruments.

Enfin, il y a dans notre pays des Instituts de Recherche sur l'Enseignement des Mathématiques, il y a dans ces IREM et à l'Université une recherche permanente en ce qui concerne l'enseignement des mathématiques, et, de façon plus fondamentale, la Didactique des Mathématiques. Il y a aussi l'Association des Professeurs de Mathématiques (APMEP) qui a accumulé des réflexions et des propositions visant à l'amélioration de l'enseignement des Mathématiques. Il y a encore des commissions

officielles telles que la COPREM, le GREM, la Commission présidée par M. DA CUNHA CASTELLE. Il ne semble pas que beaucoup de ces organismes aient été associés au projet. N'est ce pas faire preuve de beaucoup de mépris pour le travail des uns et des autres, à moins qu'il ne s'agisse d'ignorance, que de penser qu'il serait possible, par simple décision administrative, de régler de façon simple des questions dont chacun s'accorde à reconnaître la complexité.



Jean Pierre Petit