

Rapport sur une première expérimentation d'évaluation par compétences

Le présent document décrit les motivations et la mise en œuvre d'une notation par compétences, dans quatre classes de seconde.

1. Critique de la notation traditionnelle

Voici, parmi d'autres, les reproches que l'on peut formuler à l'égard de la notation « traditionnelle » consistant à attribuer un certain nombre de points à chaque question posée dans un devoir.

- **manque de lisibilité instantanée** : face à sa copie notée, disons 8/20, l'élève a du mal à savoir ce qui a été maîtrisé, compris ou ignoré ; et même si une analyse de la copie rend cela possible, bien peu d'élèves la font en recevant leur copie ; cette tâche est encore plus difficile après coup.
- **manque de lisibilité cumulée** : Difficile, en voyant évoluer sa moyenne, de savoir dans quels domaines on a progressé ou, au contraire, là où les difficultés persistent.

Le professeur rencontre également ce problème lorsqu'il essaie de sélectionner les élèves pour une activité de soutien : il doit agir de mémoire pour associer élèves et difficultés.

- **classement plutôt qu'évaluation** : que maîtrise un élève ayant 10/20 de moyenne ? On est tenté de répondre « environ la moitié du programme ». Cependant plusieurs études indiquent qu'en fait, cela montre davantage que l'élève se trouve en milieu de classement.

Ceci signifie qu'un élève ayant toutes les compétences requises mais cohabitant avec de meilleurs élèves que lui est condamné à ne jamais recevoir de très bonnes notes. Réciproquement, et ce n'est pas moins grave, un élève ne comprenant pas grand chose mais placé dans une classe de niveau suffisamment faible sera sur-noté.

En résumé, *la notation traditionnelle a tendance à classer les élèves plutôt qu'à les évaluer en termes de savoirs.*

- **vitesse d'acquisition opposée à l'acquisition** : Imaginons un cours, suivi de trois évaluations entrecoupées de révision/soutien.
 - L'élève qui a compris de suite reçoit 20/20 à chaque interrogation : sa moyenne est 20/20
 - L'élève qui comprend de mieux en mieux reçoit 0, 10 et 20 : moyenne 10/20
 - L'élève qui ne comprend qu'à la fin reçoit 0, 0, 20 : moyenne 7/20

Pourtant tous les élèves ont compris ! La notation traditionnelle tient donc compte de la vitesse d'acquisition des connaissances d'une manière significative : est-ce réellement le but visé ?

- **rattrapage impossible** : la notation traditionnelle ne permet que difficilement le rattrapage : un élève qui comprend une notion tardivement a déjà été sanctionné et ne peut se rattraper que si cette notion est encore évaluée ; même dans ce cas, le rattrapage ne pourra être que partiel (point précédent)

Il en découle la volonté de développer une notation qui puisse :

- permettre une lisibilité en termes de savoirs, pour les élèves comme les professeurs ;
- évaluer l'acquisition des savoirs et non classer ;
- évaluer l'acquisition *in fine*, en favorisant la progression, sans sanctionner la lenteur du processus ;
- permettre le rattrapage, notamment en cas d'« accidents de parcours » ;
- ne pas demander un surcroît de travail déraisonnable au professeur.

2. Mise en œuvre de la notation par compétences

La notation par compétence permet de répondre de façon satisfaisante aux desiderata précédents. On en décrit ici le principe et la mise en œuvre testée au lycée Béhal, en classe de seconde.

Au début de l'année chaque élève reçoit :

- **le référentiel des compétences** : c'est la liste, codée et commentée, de l'ensemble des compétences que l'élève est censé maîtriser à la fin de l'année de seconde. C'est une version « fragmentée » du

programme officiel qui dicte le contenu des enseignements de seconde en mathématiques. Dans cette liste, chaque compétence à un code, une description et éventuellement un commentaire.

Tout au long de l'année, les codes de compétences sont indiqués pendant le cours et les exercices, pour signaler que telle ou telle compétence est travaillée. Ainsi l'élève sait exactement à quelles parties du cours et à quels exercices une certaine compétence fait référence.

- o **la grille d'évaluation** : elle permet à chaque élève de garder une trace du résultat des différentes évaluations pour chaque compétence. Il doit la tenir à jour à chaque fois qu'un devoir noté lui est rendu.

Dans le système des compétences, il n'y a plus directement de note chiffrée. Sur chaque énoncé de devoir figure la liste des compétences qui y sont testées et cette grille est complétée par le professeur puis rendue lors de la distribution des copies. C'est cette grille qui constitue le résultat au devoir, avec la codification suivante : + : compétence maîtrisée ; n+ : compétence acquise, avec quelques détails perfectibles ; n : compétence en cours d'acquisition ; n- : éveil à la compétence ; - : compétence totalement ignorée, ou non traitée.

Ainsi, au cours de l'année, chaque compétence est évaluée un certain nombre de fois, avec à chaque fois un des cinq résultats précédents.

Le professeur possède sa propre version de cette grille, qu'il complète pour chaque élève à chaque correction de devoir : c'est la nouvelle « saisie des notes ».

L'évaluation par compétences aurait pu s'arrêter là, mais il nous a semblé nécessaire d'avoir une note chiffrée ; pour le bulletin d'une part, pour effectuer une synthèse d'autre part et enfin pour ne pas casser trop brutalement les habitudes des élèves.

Il est important de comprendre que le calcul de cette note ne saurait être une simple moyenne des différentes évaluations, sous peine de retomber dans les travers de la notation traditionnelle.

D'abord, un pourcentage d'acquisition est calculé pour chaque compétence, en se basant sur la suite des évaluations qui lui correspondent. La méthode de calcul de ce pourcentage est un point central dans l'ensemble de la notation, et est détaillé au paragraphe suivant.

La note finale de l'élève est alors la moyenne de tous ces pourcentages, ramenée sur 20. À ceci vient s'ajouter la possibilité de pondérer par un coefficient chaque devoir ainsi que chaque compétence si l'on en juge certaines plus importantes que d'autres.

Pour être cohérents jusqu'au bout, nous avons décidé de ne pas faire de « coupure trimestrielle » dans la notation : toutes les évaluations sont prises en compte jusqu'à la fin de l'année dans le calcul de cette note. Ainsi, toute compétence validée reste acquise, mais toute compétence laissée invalidée handicape l'élève jusqu'à la fin de l'année.

Calcul du pourcentage d'acquisition d'une compétence

La méthode essaie de privilégier la progression dans la maîtrise d'une compétence donnée et notamment de faciliter le rattrapage de compétences. On commence par transformer les codes en pourcentages : de '-' à '+' on attribue les valeurs 0, 25, 50, 75 et 100. Pour une compétence donnée, on obtient donc les notes : $u_1, \dots, u_n$, dans l'ordre où les n évaluations ont été effectuées en classe.

On définit ensuite pour chaque compétence sa « valuation » : c'est un entier qui en quelque sorte représente le nombre minimal d'évaluations nécessaires pour considérer que la réussite (%) à une compétence puisse être chiffrée. Par défaut, la valuation v est $\lceil \frac{n}{2} \rceil$, c'est-à-dire le plus petit entier supérieur ou égal à la moitié des évaluations, mais on pourrait également choisir un nombre fixe à la place si on le désire (par exemple 2).

Le pourcentage de réussite pour une compétence est alors donné par (sans tenir compte ici des coefficients)

$$p = \sup_{1 \leq j \leq v} \frac{1}{n - j + 1} \sum_{i=j}^n u_i.$$

En clair, on prend la plus grande valeur parmi les moyennes des $v, v + 1, \dots, n$ dernières évaluations.

De cette manière, si les évaluations se terminent par de bons résultats, les moyennes partielles des dernières évaluations sont les meilleures et les premières évaluations ne sont plus prises en compte. Par exemple, si un élève obtient, dans l'ordre, -- ++, alors sa compétence est validée à 100 % puisqu'ici $v = 2$, donc la plus grande moyenne sera celle des deux derniers résultats, à savoir 100 %. À l'opposé, un élève dont les évaluations décroissent n'est pas pénalisé outre mesure, par exemple quelqu'un ayant ++ -- aura une évaluation de 50 % (moyenne des 4 évaluations) : on retrouve alors le comportement de la notation traditionnelle.

Tous ces calculs sont faits automatiquement par des macros sur tableur.

La note finale est communiquée aux élèves une fois par mois. Elle résume leurs résultats sur l'ensemble des évaluations subies depuis le début de l'année. Le détail des évaluations en pourcentage pour chaque compétence leur est quant à lui communiqué à la demande, et à défaut une fois par trimestre, ce qui semble suffisant car la grille d'évaluation leur permet déjà suffisamment de savoir là où il faut encore progresser.

Le fait que l'élève ne soit pas lui-même en mesure de calculer sa note, et le côté un peu obscur de cette opération, ne semblent pas avoir posé de problème.

Une possibilité à l'avenir serait de rendre la grille d'évaluation consultable par les élèves « en temps réel » via internet.

Avantages attendus de la notation par compétence

Celle-ci embrasse bien l'ensemble de la problématique décrite ci-avant :

- elle augmente considérablement la lisibilité des résultats ; c'est pour le professeur un avantage considérable pour bien cibler les élèves que l'on choisit en AI pour retravailler telle ou telle compétence ;
- elle se focalise sur l'obtention finale d'un savoir en négligeant considérablement la vitesse d'acquisition ;
- elle ouvre ainsi théoriquement la possibilité pour tout le monde d'avoir une très bonne note (ou une très mauvaise) ;
- elle permet le rattrapage à tout moment de l'année. Les élèves ont la possibilité de travailler telle ou telle compétence, soit par eux-mêmes, soit en AI, puis de venir volontairement en AI pour se faire évaluer. Ils deviennent ainsi bien plus responsabilisés face à leurs résultats, et conservent à tout moment la possibilité de les améliorer.
- la surcharge de travail pour le professeur reste très raisonnable et se retrouve essentiellement dans la saisie des notes (grille d'évaluation sur le tableur), le temps de correction d'une copie restant à peu près inchangé.

3. Bilan et perspectives

Quelques remarques pèle mèle ; ce qui a marché ou moins bien marché :

- la lisibilité apportée par les codes de compétences, à la fois dans le cours, les exercices, la description du contenu d'un devoir, *etc.*, a beaucoup intéressé les élèves faibles mais studieux ;
- après une petite période d'adaptation, les élèves se sont habitués sans problème particulier à ce nouveau type de notation.
- pour les possibilités de rattrapage en AI, le bilan est plus contrasté : elles n'ont pas été exploitées autant qu'elles auraient pu l'être — bien des fois, des élèves pris en AI et comprenant bien une compétence n'ont pas pris la peine de revenir à un autre moment pour la valider : même la carotte a ses limites !

Dans l'ensemble, ce sont plutôt les bons élèves qui ont cherché à rattraper leurs compétences que ceux qui en auraient réellement eu besoin. Cela étant dit, je nuance personnellement cet aspect des choses car le qualificatif « bon » ou « mauvais » doit être pris relativement à l'environnement du lycée Béhal, où un « bon » élève est simplement un élève qui arrive à suivre et fournit un travail raisonnable – les « mauvais » élèves, quant à eux, croulent sous les lacunes depuis plusieurs années et ne fournissent en général aucun travail personnel, fut-ce simplement de réviser avant de venir (pourtant volontairement !) valider une compétence.

- Le choix de la liste des compétences ainsi que sa « granularité » jouent un rôle central dans l'ensemble du processus : les compétences trop macroscopiques sont difficiles à évaluer, tandis que celles trop microscopiques demandent des exercices simplistes pour pouvoir être évaluées.
- Dans un exercice qui enchaîne les compétences, il est parfois difficile de savoir si une compétence n'a pas été traitée parce qu'elle n'est pas comprise ou simplement par manque de temps, ou parce qu'elle dépend des réponses à d'autres questions préalables, *etc.*

Une certaine attention doit être de mise lors de la conception des sujets pour éviter les questions trop dépendantes les unes des autres et ainsi faciliter l'évaluation à venir.

- Il est apparu dès le début que pour pouvoir donner un sens à l'évaluation d'une compétence et permettre pleinement l'évaluation de la progression d'un élève vis à vis d'elle, il fallait évaluer chaque compétence au minimum 3 fois. Or ceci est difficile à réaliser car très coûteux en temps.

En procédant à une mini-interrogation de 15 minutes (2 ou 3 compétences) chaque semaine plus un devoir d'une heure chaque mois, cela suffit à peine. On pourrait y ajouter les devoirs à la maison mais nous n'avons pas retenu cette possibilité pour deux raisons : d'une part, les élèves recopient trop les uns sur les autres et faussent ainsi l'évaluation, d'autre part, les devoirs à la maison sont l'occasion d'aborder des exercices plus complexes et plus ouverts qui sont assez difficiles à évaluer en compétences.

Ce problème du nombre d'évaluation par compétence n'est pas un problème intrinsèque à ce mode de notation : il a simplement le défaut de mettre cela excessivement en exergue, et donne en retour sérieusement à réfléchir quant à la signification des notes que l'on attribue aux élèves dans un bulletin.

L'ampleur des programmes associé à la faiblesse des élèves ne laisse pas le temps suffisant à l'évaluation et la maturation des concepts et rend peut-être veine la recherche d'une évaluation « juste ».

Conclusion

Pour ce premier galop d'essai, l'évaluation par compétence me paraît fournir plus d'avantages que d'inconvénients et mérite donc d'être poursuivie, en modifiant légèrement certaines des compétences dont l'évaluation s'est avérée délicate ou sans intérêt particulier.

Son principal défaut, au final, est peut-être de morceler à outrance le programme, de trop le techniciser. Avec des classes de faible niveau en effet, la validation des compétences occupe presque toute l'année et ceci se traduit par une proportion trop importante d'exercices répétitifs et techniques laissant croire que l'activité mathématique se résume à savoir appliquer un nombre important de « recettes de cuisine ». Ce genre d'enseignement ne déplaît pas forcément aux élèves car il les rassure, et on peut ainsi prédire sans risque un grand succès à la notation par compétences dans la filière SMS, par exemple. On aurait cependant souhaité pouvoir valider les compétences à travers la pratique de problèmes plus complexes et plus ouverts, ou au moins avoir le temps de pratiquer ceux-ci après que les compétences de bases aient été assurées. Mais cela ne semble possible qu'avec les meilleurs élèves, c'est-à-dire précisément ceux auxquels toute cette démarche ne s'adresse pas de prime abord.

Une solution est peut-être simplement de réduire le nombre de compétences dans le but non pas de couvrir le programme mais de cibler l'essentiel : *une sorte de « socle commun » exprimé et noté en compétences évaluées de nombreuses fois au cours de l'année, et accompagné d'activités plus riches, « ouvertes », « expérimentales », et notées traditionnellement.*

Ce problème d'équilibre entre les activités techniques et répétitives et les activités plus globales est central dans l'enseignement des mathématiques. La notation par compétence telle qu'elle est présentée ici se prête particulièrement bien aux premières, nettement moins aux autres.

A l'heure actuelle, faiblesse des horaires aidant, si l'on veut pouvoir fournir aux élèves à la fois un minimum des bases techniques liés à des savoir-faire clairs — imposés en quelque sorte par le format du BAC — mais aussi une activité mathématique plus diversifiée — sûrement formatrice à long terme, mais dont l'évaluation reste délicate — la seule solution consiste à les inciter à fréquenter un inexistant « club de maths » au lycée, où l'on prendrait enfin le temps simplement de réfléchir.

C'est là sûrement, à l'échelle modeste du professeur, que se situe une des clés pour un enseignement mathématique de qualité.