« Qui fera l'histoire […] de l'« examen », de ses rituels, de ses méthodes, de ses personnages et de leur rôle, de ses jeux de questions et de réponses, de ses systèmes de notation et de classement ? Car dans cette mince technique se trouvent engagés tout un domaine de savoir, tout un type de pouvoir. » Michel Foucault.

Nous reproduisons ici l'article de Grégory Chambat initialement publié dans le n° 25 de la revue *N'Autre école* (dossier "Évaluation"

Une histoire de la note et de sa contestation

La mise en perspective historique des modes de notation ayant cours au sein de notre système éducatif se révèle une approche stimulante mais ardue, tant les recherches sur le sujet sont rares. Quelques études offrent des pistes, nous livrons ici une synthèse de ce qu'il est possible de lire sur la question, en nous appuyant en particulier sur les travaux d'Olivier Maulini [[1. « Qui a eu cette idée folle, un jour, d'inventer (les notes à) l'école ? Petite histoire de l'évaluation chiffrée à l'usage de celles et ceux qui désirent s'en passer (et des autres) », Olivier Maulini et « L'école de la mesure – Rangs, notes et classements dans l'histoire de l'enseignement (L'éclairage de l'Histoire) » in *L'Éducateur*, n° spécial « Un siècle d'évaluation en Suisse Roamnde, mars 2003. Deux textes en lien sur le site de la revue *N'Autre école*.]].

La note, invention des Temps modernes [[2. Titre d'un article de Pierre-Philippe Bugnard, en lien sur le site de la revue.]]

Nombre de nos collègues s'étonneront probablement d'apprendre que la notation est une idée neuve. Ce qui semble, aujourd'hui, une évidence admise par tous, est en réalité le résultat d'un long processus. L'idée même de hiérarchiser les élèves n'est apparue qu'assez récemment. Cette naissance tardive a de quoi interroger puisqu'elle nous laisse entendre que non seulement l'école a longtemps su se passer des notes, mais elle invite également à se demander si ce ne serait pas « la nature sélective du système scolaire qui entraînerait la hiérarchisation des performances des élèves, elle-même nécessitant l'élaboration d'un système d'évaluation ad hoc : la note scolaire et le calcul des moyennes ». Et Olivier Maulini de poursuivre en posant la question de manière directe : « Est-ce la sélection scolaire qui inventa la note ou la note qui provoqua la sélection ? ». De la réponse à cette question dépendront forcément les modalités de notre lutte contre la notation telle qu'elle est pratiquée aujourd'hui.

Au commencement était la marchandisation de l'école

C'est au Moyen-Âge, dès l'époque carolingienne, que se tisse le premier réseau scolaire. Durant cette période, l'école, en particulier dans les campagnes, fonctionne sur le mode du service. Un maître offre son enseignement contre une rétribution, parfois versée directement en nature sous forme de nourriture. Un fonctionnement qui répond aux attentes d'une société féodale où la naissance définit à l'avance la trajectoire sociale des individus et où l'école n'est investie d'aucune fonction de sélection des élèves-clients « Les élèves se préoccupaient d'étudier – ou de vivre ensemble – plus que de réussir, les maîtres plus d'enseigner que d'évaluer. (…) À l'élève et à sa famille revenait la responsabilité de décider s'il y avait quelque intérêt spirituel ou quelque profit matériel à suivre un enseignement, et à quel moment il convenait de l'abandonner, peut-être pour chercher un autre maître. L'auto-évaluation l'emportait sur l'évaluation. (…) Les premiers maîtres d'école offraient leurs services sur le "marché scolaire", les premiers élèves agissaient comme des consommateurs ; comme tels il leur appartenait de savoir quels étaient leurs besoins et dans quelle mesure l'enseignement reçu leur donnait satisfaction. (Perrenoud, 1984, p. 84). Dans cette logique, chaque objectif atteint par l'élève est tarifé, mais si l'élève n'apprend pas, le maître n'est pas payé…

Les Jésuites : former contre

Si, dans certaines campagnes, ce système perdura jusqu'au milieu du xix e siècle, c'est aux jésuites que nous devons la remise en cause de ce principe marchand. Dans le cadre de la Contre-réforme, cet ordre entend sélectionner une nouvelle élite. Ignace de Loyola et ses disciples dotent leurs collèges d'un code d'enseignement, le *Ratio studiorum*, qui pose comme principe que l'enseignant se doit de « favoriser une honnête émulation » qui fera effet de « grand aiguillon pour l'étude » (comprendre : qui sera aussi efficace, sinon plus, que la pratique des châtiments corporels). Ce code définit un système particulièrement complexe de concours, de prix, de devoirs écrits… suivant un rythme mensuel, trimestriel, annuel. Il encourage également une concurrence à double détente, entre les élèves eux-mêmes et entre les groupes qu'ils constituent selon un modèle inspiré de l'organisation des légions romaines. Chacun est le rival de l'autre, le sens de l'honneur servant de ressort à cette compétition aux accents guerriers.

L'âge du classement

C'est un collège jésuite, au Portugal, qui invente, en 1558, la première distribution des prix. L'idée s'étend à tout le continent et, à la fin du siècle, chaque établissement verse aux plus méritants des prix en argent, puis des médailles. Dès lors, il convient d'établir un classement : le maître compte les fautes et ordonne les compositions. Les résultats sont envoyés aux familles, avec un commentaire précisant le rang de l'élève.
Ces classements vont être progressivement remplacés par des indications chiffrées (au collège jésuite de Caen, on adoptera une échelle à 4 niveaux: 1 = bien ; 2 = assez bien ; 3 = médiocre ; 0 = mal.) À la fin de l'année, les « optimi » passent dans la classe supérieure (cette répartition en classe de niveaux n'existait pas auparavant, le maître allant d'un élève à l'autre). Les « inepti » sont recalés, les « dubii » (les « incertains ») seront admis dans la classe suivante, mais à l'essai. En cas de problème, ils redescendront dans leur classe de départ. Enfin, les parents seront priés de « retirer les bornés et les cancres ».

Discipline, répétition et concurrence définissent la pédagogie des jésuites, une pédagogie souvent caricaturée. Le codage des résultats est variable selon les époques et les lieux, mais il repose toujours sur le même principe : ce qui permet de « juger » un élève, c'est son classement dans la hiérarchie des groupes au sein de l'établissement.
Trop « méritocratique » pour l'aristocratie jalouse de ses privilèges, le système sera combattu et les écoles seront fermées jusqu'à ce que la Révolution française, bien qu'opposer à l'idéologie des jésuites, s'inspire et perfectionne ce modèle pour organiser son propre système secondaire.

« Je ne dois pas battre la campagne… »

L'évolution est bien plus lente dans les campagnes. Il faut attendre par exemple 1822 pour assister au premier prix dans le primaire. On y descerne des « prix absolus » et des « prix de progrès », puis des « prix de bonnes notes » (quand son travail est bon, l'élève reçoit des billets. Qui lui permettent, à la fin de la semaine, du mois ou du trimestre, de recevoir – déduction faite des mauvaises notes – de menus objets: couteaux, sifflets ou cravates pour les garçons ; ciseaux, dés à coudre ou peignes pour les filles – Mutzenberg, 1974). « Le classement des élèves en groupes hiérarchisés va aboutir à la "notation" de chacune des cohortes. Qu'elles se présentent sous forme de billets palpables ou de simples écritures, de lettres ou de chiffres, qu'elles se situent sur une échelle graduée de 0 à 20 ou de 0 à 6, ces « notes » découleront toutes du découpage imaginé par les jésuites et leurs contemporains » (Maulini).

Peut toujours mieux faire !

La notation, de 0 à 20, des compositions est quant à elle officialisée en France en 1890. Si la République s'inspire des systèmes de classement antérieurs, elle entend les optimiser selon une idéologie qui lui est propre. Alors que le classement des jésuites ne valait que pour un lieu et un moment, la note prétend à une valeur universelle : « Ce que la note va progressivement signifier, c'est moins le rang de l'élève dans sa classe, que sa place sur une échelle universelle : l'échelle d'appréciation. Comme la monnaie pour le produit, la note apprécie le travail de l'élève, c'est-à-dire qu'elle lui donne son prix. […] Ce que dit la note, ce n'est pas seulement le rang que l'on mesure ("5 e sur 25"), mais la valeur du "travail" ou de la "conduite" que l'on évalue. Peu importe, à la limite, que Jules soit 3 e ou 15 e sur 25. Si sa note est une bonne note, c'est que son travail est un bon travail. » (Maulini).
Le système permet de dépasser le traditionnel enseignement individuel des petites écoles au profit d'autres méthodes jugées économiquement plus rentables. Avec l'enseignement simultané le maître s'adresse directement à l'ensemble des élèves regroupés de façon homogène. La méthode permet d'augmenter significativement les effectifs. Mais, pour avoir des « classes », il faut… classer : d'où l'apparition de concepts nouveaux, tels ceux de « moyenne » ou de « redoublement ».

Le versant obscur

Il existe une seconde piste pour étudier l'histoire de la notation. C'est celle qui, marchant dans les traces de Michel Foucault, s'interroge sur l'évolution des systèmes de « contrôle », dans toute l'acception du terme. On connaît la formule « surveiller et punir » qui marque le passage à notre époque moderne. Progressivement – mais il faudra quand même, en France, attendre 1834 et surtout 1918 – le châtiment physique, comme méthode pédagogique disparaît. Il avait été concurrencé puis supplanté par « divers procédés d'examen, punitions morales ou psychologiques, surveillance disciplinaire récompensant les bons/flétrissant les mauvais élèves, leur portant des marques visibles, les classant » (*La Note, invention des temps modernes*) « L'examen combine les techniques de la hiérarchie qui surveille et celles de la sanction qui normalise. Il est un regard normalisateur, une surveillance qui permet de qualifier, de classer et de punir. Il établit sur les individus une visibilité à travers laquelle on les différencie et on les sanctionne. C'est pourquoi, dans tous les dispositifs de discipline, l'examen est hautement ritualisé. » (Foucault, 1975, p. 187). La note, comme « punition-signe » va se substituer à la punition-expiation (qui, visant le corps, faisait expier dans et par la douleur) et s'articuler avec la punition-exercice qui vise à dresser le corps (typologie des punitions de Prairat *Éduquer et punir. Généalogie du discours psychologique*).

Vers la disparition des notes ?

Dès son instauration dans sa version « moderne », le système de notation ne va cesser de susciter des critiques.
Critiques « externes », qu'un florilège de citations (de Claparède à Freinet, en passant par Ferrer, Freire et bien d'autres) ne saurait complètement épuiser (ainsi, Claparède, dès 1920, défendant « une école sur mesure et non de la mesure » écrit : « L'école actuelle veut toujours hiérarchiser ; ce qui importe avant tout, c'est de différencier. Cette idée fixe de hiérarchie provient de l'emploi des divers systèmes usités pour aiguillonner les écoliers: bonnes ou mauvaises notes, rangs, punitions, concours, prix… Mais il est entendu que, dans l'école de demain, tous ces expédients seront mis au rancart, ou n'auront en tout cas plus l'importance d'antan. L'intérêt, tel sera le grand levier qui dispensera des autres. » (Claparède, 1920, p. 30). La docimologie, l'étude scientifique de la fiabilité de la notation et de l'évaluation, apparaît également dans ces mêmes années 20 avec les travaux d'Henri Piéron (1922). Aujourd'hui, les travaux sur la constante macabre (André Antibi, 2003) sont significativement encouragés par le ministère de l'Éducation nationale.

C'est que la critique « interne » de la notation a également une longue histoire. Pour Olivier Mauline « Ce que montre l'histoire de la notation, c'est une lente érosion de la quantification. En 1910, les

enseignants genevois calculaient et répertoriaient environ 10 moyennes mensuelles pour chacune des 12 disciplines scolaires et pour la conduite, soit 130 informations chiffrées pour un seul bulletin scolaire. En 1966, on ne comptait plus que 5 moyennes bimestrielles pour 12 disciplines, la conduite et l'application, soit 70 notes. En 1980, 3 moyennes pour 10 disciplines, soit 30 notes. En 1992, 3 moyennes pour 5 disciplines, soit 15 notes. Demain, ou après-demain, seules devraient rester 3 "notes globales" de fin d'école primaire. Le mouvement est régulier, et il semble irréversible : l'école obligatoire réduit les classements ».
Comment le comprendre ? Les réponses sont multiples. La mutation des procédures de sélection, mises à jour par Bourdieu, est une piste. La cause essentielle de l'apparition de la note, le classement, ayant disparu au profit de systèmes de distinction plus subtiles et plus feutrés, la note perd sa première raison d'être. Autre piste féconde, les travaux sur *Le Nouvel Esprit du capitalism*e (Boltanski et Chiapello), dont Jacques Guigou propose une intéressante application aux questions scolaires[[3. « Retouche pour une histoire de l'évaluation », Jacques Guigou. En ligne sur le site de la revue.]]. Il est facile de voir dans l'acte d'évaluation l'opérateur essentiel de la naissance du capitalisme. Jusque dans les années 1960, le système va chercher à perfectionner techniquement les méthodes d'évaluation pour répondre au plus près aux besoins de la société. Il faut mesurer, compter, rationaliser, scientifiser l'évaluation et donc perfectionner la notation dans un souci « tayloriste ». La rupture est sans conteste les événements de 68 (Guigou, de Peretti) et le mouvement généralisé de boycott des examens. Une période définie ainsi par Guigou : «1968 : fin de l'examen final (et de la « lutte finale ») et début du contrôle continu » (Retouche pour une histoire de l'évaluation, Jacques Guigou). De manière symptômatique, l'apparation de la distinction entre évaluation-sommative et évaluation-formative date justement de 1967 (Scriven).
Pourtant, il faut se défier d'une vision « progressiste » et « scientiste » de cette rupture « selon laquelle on serait passé des tests et des mesures – réducteurs et indifférenciés – des origines de la docimologie, aux dispositifs élargis, autonomisants, différenciés et régulateurs des évaluations-formatives et des évaluations-recherches d'aujourd'hui ». 1968, c'est la naisssance d'une école où les « éduquants » gèrent leurs ressources humaines et les éduqués leur capital humain.
« L'évaluation-formative contribue intensivement à l'identification de ce qui peut-être mis en "unités de valeur capitalisables", en "crédits", en "validations". Longtemps rebelle à cette vaste computation, le qualitatif en fait désormais entièrement partie » (Guigou). Cette critique connaît un nouveau souffle avec les analyses sur la marchandisation de l'école (N. Hirrt et la critique de l'évaluation par compétences). Ironie de l'histoire, Philippe Perrenoud souligne malicieusement que les pratiques contemporaines de « consommation » scolaire ne sont pas sans rappeler l'éducation marchande du Moyen Âge. Tant et si bien que les familles qui sélectionnent aujourd'hui les établissements en fonction d'attentes et de stratégies personnelles pourraient selon lui amorcer une forme de retour aux sources.

Y a pire qu'une mauvaise note ?

« Suffit-il de jeter le thermomètre au vide-ordures pour que la fièvre (sans parler du virus) disparaisse ? » (Maulini). Si l'abolition de la note peut s'imaginer dans un système éducatif inégalitaire et sélectif, c'est bien que c'est avant tout à l'école du tri social qu'il faut s'attaquer en premier lieu. ■

La docimologie ?

« Étude des épreuves», la docimologie (dokime, épreuve en grec) entend analyser les ressorts cachées de l'évaluation. À l'origine, cette nouvelle discipline, fondée par Henri Piéron en 1922, entendait explorer les arguments scientifiques permettant la remise en cause des schémas traditionnels d'enseignement. Elle centre son étude sur l'influence de l'évaluateur, l'influence de l'énoncé, les objectifs de l'évaluation et ses risques.

En 1930, le professeur Laugier effectue une expérience de multicorrection de copies d'agrégation d'histoire puisées dans les archives : 166 copies ont été corrigées par 2 professeurs travaillant séparément, sans connaître leurs appréciations respectives. Les résultats furent surprenants. La moyenne des notes du premier correcteur dépassait de près de deux points celle du second. Le candidat classé avant-dernier par l'un était classé second par l'autre. Les écarts de notes allaient jusqu'à 9 points. Le premier correcteur donnait un 5 à 21 copies cotées entre 2 et 14 par le second ; le second donnait un 7 à 20 copies cotées entre 2 et 11,5 par le premier. La moitié des candidats reçus par un correcteur était refusée par l'autre.
En 1932, la Commission Carnégie propose une expérience de multicorrection en prélevant, au hasard, cent copies dans les archives du baccalauréat à Paris. Ces copies furent distribuées à 6 groupes de 5 examinateurs. Les disciplines concernées étaient : le français, la philosophie, le latin, les mathématiques et la physique. Aucune copie ne reçut deux fois la même note. Une copie de français est notée 3 et 16 ; en philosophie et en latin l'écart maximum est de 12 points. Les mathématiques et la physique, réputées pour des sciences exactes, ne sont pas épargnées : l'écart maximum est respectivement de 9 et 8 points.
En 1975, l'Institut de recherche sur l'enseignement des mathématiques de Grenoble (Irem) entreprend une expérience analogue de multicorrection. Un échantillon de 6 copies photocopiées de mathématiques (niveau BEPC) est soumis à 64 correcteurs, avec un barème, très précis, sur 40 points. Les résultats confirment ceux de l'enquête précédente, effectuée 43 ans plus tôt. La dispersion des notes atteint près de 20 points.
Laugier et Weinberg ont montré que la double correction est illusoire. En effet, selon eux, il faudrait, pour obtenir une note « exacte » (une note « exacte » étant une moyenne de notes telle que l'adjonction d'une autre note ne modifie pas sensiblement cette moyenne) : 127 correcteurs en philosophie, 78 en composition française, 28 en anglais, 19 en version latine, 16 en physique, 13 en mathématiques.
Avec des enseignants de l'Irem de Reims, M. Jacques Nimier a fait corriger 20 copies de BEPC par 30 professeurs de mathématiques après l'établissement d'un barème, tandis que lui-même corrigeait également ces copie, mais sans les lire, en regardant l'écriture et la présentation. Le maximum de l'écart pour une même copie était de 11 points sur 20, sa note n'étant jamais à une extrémité !
À la demande d'enquêteurs anglais, un professeur de physiologie de la Faculté des sciences accepta 37 copies – dactylographiées et anonymes – qu'il avait corrigées trois ans et demi auparavant. Dans 7 cas seulement, il remit la même note au même devoir. Dans les
30 autres cas, il y eut des divergences comprises entre 1 et 10 points. L'admissibilité, avec ses nouvelles notations, aurait été modifiée ; la moitié des précédents admissibles aurait été refusée et la moitié des refusés déclarée admissible.

Source principale : site personnel de Jacques Nimier

Artice extrait de N'Autre école n° 25

N'Autre école n°25 / dossier évaluation

Lire plus de publications sur Calaméo