

**T**OUT A COMMENCÉ il y a près de trois ans autour des mots : classer, évaluer, noter, palmarès, relations d'ordre... Les discussions étaient animées, mais il était difficile de dégager une ligne directrice.

L'évaluation des individus, des organismes est maintenant passée dans les mœurs. Pourquoi un tel succès ? Quand tout cela a-t-il commencé ? Les Jeux Olympiques ? déjà un bel exemple de palmarès. Une étape importante semble la mise en place des concours et des examens après la Révolution française, en vue de dégager une Élite Républicaine (page 29). Maintenant on peut dire que personne n'a été ou n'est épargné par cette pratique. Qui n'a pas été noté dans sa vie, à l'école ou au travail ? Habitué depuis votre enfance (page 35), maintenant vous pouvez même noter votre frigidaire. Noter pour se simplifier la vie, parce que c'est pratique (pour faire des moyennes par exemple...). Et pourtant, quantifier, n'est-ce pas déléguer la décision à l'ordre des nombres ?

Nous avons comparé plusieurs processus de notation chiffrée. Faire des catégories, comparer, classer, noter... Qu'est-ce qui fait valeur ? Noter, est-ce mesurer ? Classer pour noter ou noter pour classer ? Un art bien difficile à cerner (pages 7 et 23). Alain Desrosières, dans un exposé au groupe de travail le 24 avril 2001, soulignait la confusion opérée en trois siècles entre objectivité d'une mesure standardisée et objectivité d'un jugement, rendant commensurable ce qui ne l'est pas. Certains outils de mesure, le Q.I. pour n'en citer qu'un, sont même parvenus à créer une réalité. L'intelligence ? C'est ce que mesure le Q.I. La courbe de Gauss a justifié des expressions comme « l'élève normal », « l'individu moyen » (mais normal !). Tout est bien. Notre société suit la loi normale, l'ordre des choses. Son élite a ses raisons d'être.

« C'est injuste » pense un élève en recevant sa copie. Pourtant, la note « normale » en philo, moyenne de celles de 762 correcteurs serait-elle la note juste ? *Une vraie fausse note* (page 3) rend compte du contenu de nos réflexions à ce propos : de quoi "détraumatiser" un certain nombre d'élèves et les inciter à demander des comptes. Face à ces incertitudes, on pourrait penser que noter c'est tout un art, que ça s'apprend. À quand les experts notateurs pour arriver à une note légitime ? Le pouvoir des juges... Noter, c'est exprimer un jugement de valeur, mais gare aux échelles construites dans un « système de préférence manichéen » (page 37). Notez donc, en toute immoralité !

Nous n'avons pas trouvé de réponse définitive et globale. Il y a des problèmes de notation, de classement, d'évaluation aux réponses complexes, discutables et relatives. Côté anglo-saxon : concurrence, tests, palmarès, côté « latin » : culture civique, méritocratie et élitisme républicain. Et ailleurs ?

Il ne faudrait pas conclure que la note sous ses apparences de simplicité trompeuses, et sous prétexte d'imperfection, est à bannir de nos sociétés. Juger est souvent une nécessité qui sous-entend responsabilité et esprit critique, que le résultat soit nombre ou non, suivant les circonstances. Tenter d'y échapper, c'est risquer de voir s'installer des formes plus sournoises d'évaluation (QCM de philosophie...) ou de laisser faire le hasard ou l'arbitraire.

Justesse de la note, justice de la note ? Telle est la question au bout de ces trois ans de réflexion collective.

**Anne Le Bihan**

TOUTS LES MATINS je courais jusqu'à la colonne Morris pour voir les spectacles qu'elle annonçait. Rien n'était plus désintéressé et plus heureux que les rêves offerts à mon imagination par chaque pièce annoncée et qui étaient conditionnés à la fois par les images inséparables des mots qui en composaient le titre et aussi de la couleur des affiches encore humides et boursouflées de colle sur lesquelles il se détachait. Si ce n'est une de ces œuvres étranges comme *Le Testament de César Girodot* et *Œdipe-Roi* lesquelles s'inscrivaient, non sur l'affiche verte de l'Opéra-Comique, mais sur l'affiche lie-de-vin de la Comédie-Française, rien ne me paraissait plus différent de l'aigrette étincelante et blanche des *Diamants de la Couronne* que le satin lisse et mystérieux du *Domino noir*, et, mes parents m'ayant dit que quand j'irais pour la première fois au théâtre j'aurais à choisir entre ces deux pièces, cherchant à approfondir successivement le titre de l'une et le titre de l'autre, puisque c'était tout ce que je connaissais d'elles, pour tâcher de saisir en chacun le plaisir qu'il me promettait et de le comparer à celui que recélait l'autre, j'arrivais à me représenter avec tant de force, d'une part une pièce éblouissante et fière, de l'autre une pièce douce et veloutée, que j'étais aussi incapable de décider laquelle aurait ma préférence, que si, pour le dessert, on m'avait donné à opter entre du riz à l'Impératrice et de la crème au chocolat.

Toutes mes conversations avec mes camarades portaient sur ces acteurs dont l'art, bien qu'il me fût encore inconnu, était la première forme, entre toutes celles qu'il revêt, sous laquelle se laissait pressentir par moi, l'Art. Entre la manière que l'un ou l'autre avait de débiter, de nuancer une tirade, les différences les plus minimes me semblaient avoir une importance incalculable. Et, d'après ce que l'on m'avait dit d'eux, je les classais par ordre de talent, dans des listes que je me récitais toute la journée, et qui avaient fini par durcir dans mon cerveau et par le gêner de leur inamovibilité.

Plus tard, quand je fus au collège, chaque fois que pendant les classes, je correspondais, aussitôt que le professeur avait la tête tournée, avec un nouvel ami, ma première question était toujours pour lui demander s'il était déjà allé au théâtre et s'il trouvait que le plus grand acteur était bien Got, le second Delaunay, etc. Et si, à son avis, Febvre ne venait qu'après Thiron, ou Delaunay qu'après Coquelin, la soudaine motilité que Coquelin, perdant la rigidité de la pierre, contractait dans mon esprit pour y passer au deuxième rang, et l'agilité miraculeuse, la féconde animation dont se voyait doué Delaunay pour reculer au quatrième, rendait la sensation du fleurissement et de la vie à mon cerveau assoupli et fertilisé.

Mais si les acteurs me préoccupaient ainsi, si la vue de Maubant sortant un après-midi du Théâtre-Français m'avait causé le saisissement et les souffrances de l'amour, combien le nom d'une étoile flamboyant à la porte d'un théâtre, combien, à la glace d'un coupé qui passait dans la rue avec ses chevaux fleuris de roses au frontail, la vue du visage d'une femme que je



pensais être peut-être une actrice, laissait en moi un trouble plus prolongé, un effort impuissant et douloureux pour me représenter sa vie. Je classais par ordre de talent les plus illustres, Sarah Bernhardt, la Berma, Bartet, Madeleine Brohan, Jeanne Samary, mais toutes m'intéressaient.



Marcel Proust, *Du côté de chez Swann*.

## VRAIE FAUSSE NOTE

**A**PPROCHEZ, mesdames et messieurs, entrez, entrez... Vous allez voir ce que vous allez voir, ce que vous n'avez jamais vu ! Une VRAIE note... Allez ! vous, et vous aussi, et encore vous... 759, 760, 761, et voilà, 762 ! Nous y sommes. Vous êtes ici sept cent soixante-deux. Eh bien, je donne à chacun de vous le même devoir de philo réalisé par un élève, et chacun de vous le note. Allez-y, allez-y... Quoi ? Comment noter ? Mais je m'en fiche, moi ; vous notez, et puis voilà. Entre zéro et vingt, vous savez bien... Vous avez été à l'école oui ou non ?..... Et maintenant, chacun me donne sa note, mon assistant que voici que voilà calcule la moyenne des sept cent soixante-deux notes, et là, écoutez-moi : oui, mesdames et messieurs, la VRAIE note de ce devoir de philo est ...

Vous pensez que je vous emmène à la foire, à l'Uni-Parc avec les philosophes du Palace de la Rigolade ? Pas tant que ça... regardez :

Pour le Palace de la Rigolade, voir le premier chapitre de *Pierrot mon ami*, Raymond Queneau.

« Quand on sait qu'il faut 762 correcteurs d'une copie de philo et 95 d'une copie de physique pour obtenir une note identique, on se demande pourquoi l'organisation du bac ne serait pas confiée à l'expert en docimologie qu'est la Société Française des jeux ? »

C'est de la plume d'un inspecteur de l'Éducation nationale (Jean Desoli, *Libération*, Rebonds, 04-09-01), ça ne peut qu'être du sérieux...

Donc, il paraît qu'on sait qu'il faut 762 correcteurs... Ah bon ? Moi, je ne savais pas. Et j'ai essayé de comprendre d'où sortaient ces chiffres.

### L'enquête Carnegie

En 1936, la commission française de cette enquête soumet des copies du bac à une expérience de multi-corrrection : pour chaque copie, on dispose de six notes (sur 20) : celle du bac, et celles de cinq autres correcteurs mobilisés pour l'occasion. Les écarts de notes pour une même copie sont frappants :

Écart maximum :	13 points en composition française, 9 points en mathématiques, 12 points en philosophie, 8 points en physique...
Écart moyen :	3,3 points en composition française, 2,1 points en mathématiques, 3,4 points en philosophie, 1,9 point en physique...

En 1931, l'Université américaine Columbia lance une enquête internationale sur « les conceptions, les méthodes, la technique et la portée pédagogique et sociale des examens et concours ». La Carnegie Corporation de New York fournit les fonds nécessaires, et l'enquête sera nommée par la suite « Enquête Carnegie ». Des commissions allemande, américaine, anglaise, écossaise, finlandaise, française et suisse sont créées à Eastbourne en 1931. Seules les commissions anglaise et française continueront leurs travaux jusqu'en 1938.

En 1936, la commission anglaise a publié un rapport : « An examination of examination », et la commission française un autre : « La correction des épreuves écrites dans les examens ».

« H.Laugier et D.Weinberg en avaient conclu que pour obtenir la « note vraie », il fallait recourir à la moyenne de 13 correcteurs en mathématiques, 78 en composition française, 127 en philosophie... Les aléas de la correction sont donc considérables et la recherche d'une « note

H.Laugier et D.Weinberg, Recherche sur la solidarité et l'interdépendance des aptitudes intellectuelles d'après les notes des examens écrits du baccalauréat. Paris, Chantenay, 1938.

vraie » quelque peu utopique. » (Pierre Merle, Sociologie de l'évaluation scolaire, PUF, 1998, page 9)

« Devant de tels faits, Laugier et Weinberg ont cherché à calculer statistiquement le nombre minimum d'examineurs compétents auxquels il faudrait faire appel pour obtenir une moyenne des notes mises par eux qui ne variât plus sensiblement : ils ont trouvé pour la dissertation philosophique : 127 ; pour la composition française : 78 ; pour l'anglais : 28 ; pour la version latine : 19 ; pour la physique : 16 ; pour les mathématiques : 13. »

(A. de Perette, Controverses en éducation, Hachette Éducation, 1993, page 354).

Cette fois, les examinateurs doivent être compétents... Ce n'était peut-être pas le cas de mes 762 philosophes...

« Une première piste, proposée par Laugier et Weinberg (1936), considérait d'une part les copies à évaluer comme des objets physiques, dotés de propriétés physiques mesurables et, d'autre part, les divergences inter-notateurs comme des erreurs de mesure. Cette perspective reposait sur deux postulats positivistes : une copie serait assimilée à un objet physique qui aurait une « vraie masse » et les notations seraient des « estimations » de cette masse ; les divergences de notation seraient dues au hasard. Il convenait donc, à l'instar de la physique, de multiplier le nombre de mesures selon une procédure d'estimation statistique. En d'autres termes, il convenait de déterminer le nombre nécessaire de correcteurs pour obtenir une « vraie valeur » à chaque copie (la moyenne des notes ne variant plus, même avec des notes supplémentaires). C'est ainsi, après calcul, que ces auteurs estiment qu'il faudrait 78 correcteurs en français, 19 en latin, 28 en anglais, 13 en mathématiques, 127 en philosophie et 16 en physique. » (R. Amigues, M.T. Zerbato-Poulos, Les pratiques scolaires d'apprentissage et d'évaluation, Dunod, 1996, page 135).

### La formule de Spearman-Brown

« Comment obtenir la valeur réelle d'un devoir, quand on sait les divergences d'appréciations entre correcteurs, d'une part, et chez le même correcteur, d'autre part ?

« Pour répondre à cette question, H.Laugier et D.Weinberg (1936) ont étendu le principe de la double correction à autant d'évaluateurs que nécessaire pour obtenir la note d'une copie au degré de précision désiré. Ce nombre est donné par la formule de Spearman-Brown :

$$R = \frac{nr}{1+(n-1)r} \quad \text{soit} \quad n = \frac{R(r-1)}{r(R-1)}$$

The true-score-and-error model of a test score was presented by Spearman in 1904.

« Où R est le coefficient de fidélité  
r est le coefficient moyen de corrélation  
n est le nombre de mesures (ici le nombre de correcteurs). »

(On est rassuré, il y a une formule, ce n'est pas du flan.)

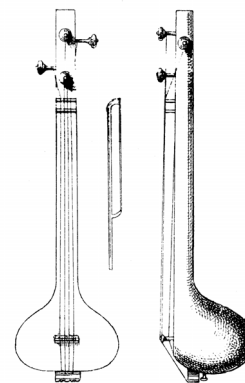
**साङ्गमपुधान साङ्गमपुधान साङ्गमपुधानसा**

sa, ri, ga, ma, pa, dha, ni. sa, ri, ga, ma, pa, dha, ni. sa, ri, ga, ma, pa, dha, ni sa.

« (...)En fixant le coefficient de fidélité à 0.99, les auteurs cités ont obtenu les chiffres suivants :

Mathématiques	13 correcteurs
Physique	16 correcteurs
Version latine	19 correcteurs
Anglais	8 correcteurs
Composition française	78 correcteurs
Dissertation philosophique	127 correcteurs »

(Yvan Abernot, Les méthodes d'évaluation scolaire, Dunod, 1996, pages 22-23).



Chikara à 4 cordes.

Mais est-ce que cette formule nous éclaire pour déterminer la « valeur » d'une copie ?

D'abord, je me place dans un cas extrême : 126 correcteurs ont mis 20 à un devoir de philo, et le 127<sup>e</sup>, une peau de vache patentée, ou un des philosophes de Queneau, met 0 au même devoir. Ça ne fera passer la moyenne que de 20 à  $20 \times 126/127$ , soit une variation d'à peine 0,16 point. Pas besoin de Spearman, ni de Brown pour comprendre ça.

Ensuite, je crois que ce que la formule de Spearman-Brown vise, est d'améliorer la fiabilité d'un instrument de mesure. Elle permet de savoir, par exemple, le nombre de pesées qu'il faut réaliser avec une même balance pour obtenir (en faisant la moyenne des résultats) toujours la même valeur pour un même objet. Il ne s'agit pas de savoir si la balance est juste ou fautive, mais si elle est fidèle. Elle peut très bien donner 10 kg de plus que le poids « réel », ça n'est pas la question.

De plus, une des hypothèses qui permettent d'établir cette formule est que les « erreurs » de mesure (différence entre une hypothétique valeur « vraie » et une valeur observée) sont dues au hasard, et donc que l'erreur moyenne tend vers zéro quand le nombre de mesures augmente indéfiniment.

Est-ce qu'on peut considérer que 127 correcteurs remplacent 127 corrections réalisées par un même correcteur (le pauvre...), que les variations entre les différentes corrections sont dues au hasard, et surtout est-ce qu'on peut considérer que parce qu'un correcteur a recorrecté 127 fois la même copie, la moyenne des résultats peut en quelque façon que ce soit être considérée comme une « vraie note » ?

« En fait, rien ne prouve que cette note soit plus « vraie » que celle donnée par un seul évaluateur. Si le nombre fait ici loi, il ne fait pas nécessairement vérité. Tel enseignant peut très bien avoir attribué une excellente note à une copie et la justifier très rigoureusement ; tel autre peut avoir attribué une très mauvaise note et la justifier tout autant. En revanche, de nombreux correcteurs seraient bien en peine d'expliquer le détail de leur système de notation. » (Y.Abernot, *op. cit.* page 24)

### **De 127 à 762**

Vous vous rappelez Monsieur Desoli ? Cela fait un moment que je vous raconte mes 127 philosophes, mais pour lui, il en fallait 762... Et en physique, ce n'était pas 13, mais 95. Inflation verbale ?

Que nenni. Là aussi il y aurait du sérieux, du scientifique :

J.J. Bonniol, Déterminants et mécanismes des comportements d'évaluation d'épreuves scolaires. Thèse. Université de Bordeaux II, 1981.

**Note** : Le calcul de Laugier et Weinberg, essayez de suivre un peu !

« On doit à J.J. Bonniol (1976, 1984) d'avoir dénoncé le mythe de la vraie note. En effet, en dépit de leur calcul (voir note), parmi les 78 correcteurs de français, 55% de ces évaluateurs se situent en dehors des limites statistiques de confiance, à la place des 5% prévus ; les divergences constatées dépassent systématiquement et largement celles dues au hasard. Les calculs ainsi proposés par J.J. Bonniol montrent qu'il faudrait multiplier par 6 le nombre de correcteurs proposé par Laugier et Weinberg. Ce qui donnerait 78 correcteurs en mathématiques et 762 en philosophie ! »

(R. Amigues, M.T. Zerbato-Poulos, *op. cit.* page 135).

Les voilà les 762 : il n'y a pas de doute :  $127 \times 6 = 762$ ...(et  $16 \times 6 = 96$  et non 95... il y a du flottement dans les multiplications à l'Inspection !)

Mais il faut lire jusqu'au bout :

« Selon Bonniol, la notation de copies ne peut s'apprécier à l'aune de la mesure physique, car une copie n'est pas un objet physique, au sens strict, mais un objet construit et multidimensionnel. Or, pour estimer, au sens de la mesure, un tel objet, il faudrait que chacune de ses dimensions soit isolable, pourvue d'une échelle métrique, combinable avec les autres selon une règle de composition, et que la nouvelle dimension obtenue soit munie d'une échelle numérique. Or cette conception métrologique ne peut s'appliquer à des copies, de français par exemple, comprenant plusieurs dimensions comme en témoignent les barèmes analytiques du type : plan, style, précision de la pensée, élégance, etc. Selon Bonniol, il est possible d'évaluer une copie, non pas au sens de la mesurer, mais de l'apprécier en référence à une échelle de valeur. En d'autres termes, la note serait un moyen de résumer des appréciations de nature différente afin de les communiquer à un élève. »

(R. Amigues, M.T. Zerbato-Poulos, *op. cit.* page 135).

Enfin un peu de bon sens... Pourtant il semble que J.J. Bonniol n'ait pas été vraiment écouté, puisque ne restent dans le débat public que les chiffres, alors que lui ne les avait construits que pour mieux les abattre.

On n'a pas encore trouvé le moyen de déterminer la « vraie » valeur d'un devoir d'élève. On aurait pu rêver, peut-être, qu'en mesurant la chose, elle se définirait. Theodore M. Porter raconte bien qu'au XVIII<sup>e</sup> siècle, c'est la mesure des températures qui a permis d'éclaircir et d'unifier des notions confuses et diverses jusqu'alors : la température de l'atmosphère, la température d'un corps humain...

Theodore M. Porter, Making things quantitative, in : Power M. (ed.), Accounting and Science, Cambridge, 1994, pp. 36-56.

Rien de tel pour les copies. Dommage, car « la fiabilité incertaine de la correction des copies (...) pose la question de la pertinence de l'expertise professorale concernant les décisions d'admission et d'ajournement ». (Pierre Merle, *op. cit.* page 9).

En tous cas, si vous voyez un jour passer 762 philosophes, renvoyez-les vite au Palace de la Rigolade.

**Françoise Dixmier**

## NOTER ET/OU CLASSER

*En cherchant à savoir [...] comment se passe l'attribution des notes [...] une perception bien particulière du jugement [...] émerge, et au sens large, l'expression d'arrangement évaluatif est celle qui convient le mieux.*

D'après Pierre MERLE, L'évaluation des élèves, PUF.

*Noter* : action de mettre une note.

*Note (chiffrée)* : appréciation donnée selon un barème préalablement choisi (Robert) ; appréciation [...] de quelqu'un, de son travail, d'un devoir... (Encyclopédie Larousse)

*Classer* : mettre un certain ordre.

On retiendra surtout dans ces extraits de dictionnaires que *noter* c'est en quelque sorte *apprécier*, ce qui peut nous laisser sur notre faim pour une approche un peu scientifique de cette affaire. L'action de *noter*, de *classer* ne peut être sortie de son inscription sociale. On ne note pas, on ne classe pas sans but plus ou moins explicite. Et dans la procédure de notation, de classement, à côté des procédures techniques il y a des questions de pertinence intellectuelle : pourquoi une faute de grammaire est-elle plus grave qu'une faute d'orthographe ? Pourquoi l'ergonomie d'un caméscope compte-elle autant que la qualité de son objectif ?... Même s'il peut paraître illusoire de vouloir isoler cet aspect, nous essayerons dans ce texte de nous intéresser aux procédures techniques de notation, de classement comme tentatives de *mathématisation du réel* dont l'ambition est de nous sortir d'une simple perception subjective par une démarche qui se veut plus rationnelle.

### **Noter est-ce mesurer ?**

*Mesurer* sous-entend qu'on ait défini *au préalable* la notion de *grandeur* à mesurer. Dans un ensemble d'objets de même nature (hôpitaux, copies, fonctionnaires, magnétoscopes,...). Introduire la notion de grandeur consiste à définir ce que sont des objets équivalents. Ce n'est évidemment pas suffisant : on peut très bien regrouper des objets ou des individus en classes d'objets équivalents (les femmes et les hommes, ou encore, comme le faisaient les théoriciens du racisme, les aryens, les juifs, les slaves, les roms, sans parler des noirs, des jaunes, des métis,...) sans qu'on puisse passer ensuite à une mesure ou à une notation. On peut parler ici de simples *classes nominales* : on peut *nommer* les classes (les hommes, les femmes, les noirs, les aryens, les juifs, ...). On peut aussi décider de les *classer* : les hommes devant les femmes, les aryens devant les juifs, ...

Au delà d'un simple classement on peut introduire une notion *d'écart* plus ou moins grand entre deux grandeurs classées en les *repérant* sur une échelle numérique établie à partir de valeurs attribuées arbitrairement à des grandeurs stables et répétables (par exemple, la température d'un corps par rapport à celle de la glace fondante et de l'eau bouillante). On parlera ici d'un *repérage numérique* des grandeurs.

Enfin, au sens plein du terme, la *mesure* demande qu'on sache définir la somme de deux grandeurs de même espèce. De telles grandeurs seront *additives*. On choisit alors une certaine grandeur comme unité (sa mesure sera par définition 1) et on mesure les autres grandeurs par le rapport entre la grandeur à *mesurer* et la grandeur unité, selon une procédure de mathématisation (passage de l'addition au rapport) que nous ne précisons pas ici.

Le processus de mesure le plus simple est le *dénombrement* appliqué à un groupe d'objets ou à une succession de phénomènes tenus pour identiques (par exemple les gouttes d'un liquide, le nombre de décès dans un hôpital, le nombre d'élèves présentés au bac).

Dans les sciences d'observation, tout processus de mesure est tributaire des instruments de mesure et des techniques mises en œuvre. En outre, il peut être troublé par des phénomènes parasites qu'il n'est pas toujours possible d'éliminer. On ne peut donc attribuer une signification objective à une mesure qu'à condition de définir le degré de précision et de fidélité de l'instrument de mesure ainsi que la marge globale des erreurs et incertitudes liées intrinsèquement à toute mesure.

En *docimologie* (le discours sur la notation), nous sommes loin de réunir toutes ces conditions. Sommes-nous en présence de grandeurs ? C'est-à-dire avons-nous défini *au préalable* ce qu'étaient des objets (hôpitaux, copies, fonctionnaires, magnétoscopes, ...) équivalents, meilleurs, moins bons ? NON ! Regardons ensuite l'échelle de notation, par exemple la plus courante en France, celle qui consiste à noter de 0 à 20. La mise en place d'une telle échelle exige plusieurs conditions. En particulier de définir la *grandeur nulle*, c'est-à-dire pour chaque catégorie (services hospitaliers, lycées, copies, fonctionnaires, magnétoscopes, ...) l'ensemble des nuls, et la *grandeur parfaite*, c'est-à-dire l'ensemble des *parfaits* de chaque catégorie. Puis de définir les échelons de l'échelle (pour les températures on décide qu'elles sont proportionnelles à la dilatation du mercure). Rien de tel pour une échelle de notes ! On fait même exactement l'inverse : on note d'abord et on définit ensuite, par l'intermédiaire de la note, ce que sont deux hôpitaux, copies, fonctionnaires, magnétoscopes, ... équivalents, distants d'un point, de deux points. Disons que l'usage intensif de ces procédures de notation finit par instaurer sur nos objets une notion de grandeur repérable, même si la démarche est viciée dès le départ (et sans parler ici de la pertinence des critères de notation).

Une échelle de notes, telle que nous venons de la décrire, n'est en réalité qu'un repérage et n'est pas, par essence, munie d'une structure additive. On ne devrait alors faire ni sommes, ni moyennes de telles notes. Mais c'est là qu'enseignants, notateurs, testeurs, évaluateurs, succombent presque tous à la magie du nombre, au charme du résultat global chiffré, à l'illusion du mesurable. On ne peut évidemment pas aller, sans contorsions multiples, jusqu'à la grandeur mesurable pour la notation. Et pourtant on y va bien souvent, c'est-à-dire qu'on va s'autoriser à faire des moyennes (pondérées ou non), de la même façon qu'on fait de l'eau tiède avec de l'eau chaude et de l'eau froide (on reste alors sur la même échelle de notes, par exemple de 0 à 20) ou à faire des sommes pondérées de notes (on se retrouve alors sur une autre échelle, de 0 à  $x$ , où  $x$  est la somme pondérée des « 20 »). Dans une démarche scientifique on mesure, on compare, on additionne des grandeurs de même nature. Peut-on additionner, ou faire la moyenne, d'une note d'anglais et d'une note de mathématiques d'un élève, des notes des différents services d'un hôpital, sous le simple prétexte que les notes (et



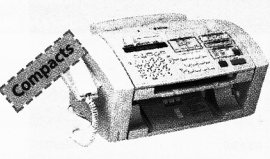
# Comparatif à la carte

**Suivez le guide**

Nous avons évalué les tout-en-un à l'aide de 5 critères : les caractéristiques générales, l'ergonomie, le prix, la qualité et la rapidité d'impression puis la qualité et la rapidité de numérisation et de copie. Nous avons associé aux trois premiers les coefficients respectifs suivants : 3, 5 et 4. Pour les deux derniers critères, les coefficients ont varié suivant nos catégories de produits : «laser» (4,3), «photocopieur» (3,4) et «compact» (3,3).

Caractéristiques	Les coeff. INFO PC	Les notes INFO PC	Votre coefficient	Votre note
Caractéristiques	3	3	X 4	= 12
Ergonomie	5	3	X 2	= 6
Prix	4	5	X 5	= 25
Vitesse/qualité impression	4	5	X 7	= 5
Vitesse/qualité numéri./copies	3	2	X 4	= 8
<b>Exemple : les lasers</b>		<b>Notre total</b>		<b>Votre total</b>
		70		56

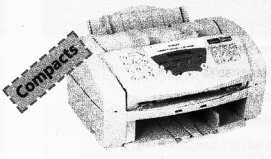
**SANITIZER**  
**MFC 740**



● Sa mémoire importante, une haute résolution et un chargeur de 200 feuilles en font un outil adapté aux travaux bureautiques. Seul point faible, l'installation est un peu complexe.

Les notes INFO PC	Votre coefficient	Votre note
5	X	=
3	X	=
5	X	=
4	X	=
4	X	=
<b>Notre total</b>		<b>Votre total</b>
74		

**COMPACT**  
**Multipass C50**



● Un prix élevé pour des performances tout à fait moyenne, on attendait mieux de ce constructeur. Le seul point fort de cette machine est son fax intégré fonctionnant à merveille.

Les notes INFO PC	Votre coefficient	Votre note
3	X	=
3	X	=
2	X	=
3	X	=
3	X	=
<b>Notre total</b>		<b>Votre total</b>
50		

Extrait de *Info PC*, n° 167, février 2000

non les grandeurs notées) sont des nombres « de même nature » ? Cela a-t-il un sens intrinsèque ou cela a-t-il pris du sens parce que nous le faisons ? La réponse à cette question est loin d'être simple. Nous pouvons trouver une telle démarche totalement aberrante ou, à l'opposé, trouver qu'une moyenne représente un indicateur pertinent d'une « valeur moyenne » de l'élève, de l'hôpital, du caméscope, à partir d'une notation de différentes composantes.

## **Avec un bon barème ce sera plus juste !**

Admettons qu'on note quand même, malgré tout et parce que dans la société actuelle c'est pratiquement incontournable. Remarquons qu'il y a alors plusieurs démarches possibles (ici on ne s'intéressera pas à la pertinence des critères, mais uniquement à la stratégie de notation). On peut noter intrinsèquement, globalement, sans barème, en s'appuyant sur une compétence et/ou une expérience plus ou moins avérées : telle dissertation vaut 11/20, l'ergonomie de tel magnétoscope vaut 14/20, tel chirurgien vaut 16/20, tel Côtes-du-Rhône 13/20. Les docimologues appellent une telle note une *note d'estime*. Souvent on y ajoute un petit commentaire « pour justifier la note ». Mais parfois le notateur est lui-même soumis à des contraintes : un fonctionnaire n'est pas noté intrinsèquement, mais de telle façon qu'il se trouve dans une zone de notes qui lui assure tel avantage et lui barre tel autre en matière de promotion, d'avancement, de mutation. On pourrait parler de *notation indirecte* ou de *notation rétrograde* (non, ce n'est pas un mauvais jeu de mot !). Cette idée de notation rétrograde, c'est-

Pierre : à l'ananas... mais j'aime mieux les nougats chinois. Les nougats chinois, je leur mets 7, ou 8... plutôt 8.

à-dire d'une notation fabriquée pour donner les résultats désirés, est présente dans la plupart des processus de notation, ne serait-ce qu'au nom de la pertinence de la notation envisagée. Parfois cette idée est simplement diffuse, implicite (on pense réellement ne « mesurer » *que* la performance d'un élève, sans avoir *fabriqué* une épreuve *ad hoc* dont on sait d'avance qu'elle va distribuer les notes « comme il faut »), parfois elle est explicite (quelles épreuves, quels barèmes fabriquer pour avoir 80 % de reçus au baccalauréat ?).

Souvent, pour « être plus juste », on se donne un barème. *Plus juste* : l'ambiguïté de cette formule est tout à fait remarquable parce qu'elle articule deux sentiments forts du notateur, du noté et de la société environnante : justice et précision scientifique. Il y a plusieurs types de barèmes : tant de points pour une faute de grammaire, tant de points pour une faute d'orthographe. Tant de points si on saute 1m20, tant de points si on saute 1m30, etc. De tels barèmes peuvent se discuter quant à leur pertinence pédagogique, mais ont ensuite le mérite d'être indiscutables (ou presque) dans leur application (on parle de *note formalisée*). Nous sommes là dans la même situation que lorsqu'on repère la température grâce à la dilatation du mercure dans le tube du thermomètre. Plus utilisée est la pratique de barèmes tels que

- le style sera noté sur 5, le plan sur 8, l'argumentation sur 7 ;
- la première question sur 3, la seconde sur 5, etc.

Mais de tels « barèmes » ne sont que des démultiplications du processus de *notation d'estime* et produisent souvent des dérives importantes par effet cumulatif des habitudes individuelles des notateurs.

### ***T'as quoi comme moyenne, toi ?***

Dans le milieu scolaire l'usage de *moyennes* à des fins diverses (bulletins, passage de classe, examen, concours, classement trimestriel pour parents et administration avides de ce genre d'information...) est monnaie courante. Cela dit, les moyennes ne se rencontrent pas qu'en milieu scolaire : un classement comparatif de services hospitaliers se fait à travers une moyenne d'indicateurs hétérogènes de par la nature des objets notés (notoriété, mortalité, soins ambulatoires, durée de séjour...) et de par la nature des indicateurs chiffrés utilisés (notes entre 0 et 3, pourcentages, nombre de jours, bascule « oui/non »,...) et celui des caméscopes itou à partir de la note d'ergonomie, de celle du zoom, de celle de la mécanique... Les gens aiment les classements : faisons des moyennes pour pouvoir classer facilement.

Pour fixer les idées, nous allons nous placer dans le contexte scolaire. Que veut-on faire, que fait-on en général ?

- comparer des séries de notes (notes des élèves de différentes classes, notes des élèves d'une classe dans les différentes disciplines, notes des copies de différents correcteurs) ;
- comparer plus globalement, plus synthétiquement un ensemble de résultats par l'intermédiaire de moyennes de plusieurs séries de notes, dans le but, par exemple de valider, justifier telle ou telle décision. On peut utiliser des moyennes dans plusieurs directions : faire la moyenne dans une discipline de l'ensemble des notes obtenues par les

## 11

différents élèves ; faire la moyenne de chaque élève sur un ensemble de disciplines ; faire une moyenne à travers le temps des notes obtenues au sein d'une discipline par un élève donné.

Pour comparer des séries, on se contente très souvent de ne regarder que la moyenne sans s'occuper de leur dispersion autour de cette moyenne.

Lorsqu'un notateur note une production d'élèves (prof dans sa classe, examinateur face à ses copies, interrogateur à un oral, ...) il produit une série N de notes que nous supposons, selon l'habitude la plus courante, prises sur l'échelle [0 ;20].

$$N = (x_1 ; x_2 ; x_3 ; \dots ; x_n)$$

(il y a ici n notes, c'est-à-dire n copies, n élèves...)

Une façon classique de décrire synthétiquement cette série N est d'en donner une caractéristique centrale (moyenne ou médiane sont les plus utilisées dans la pratique) et une caractéristique de la dispersion (écart-type ou quartiles sont en général respectivement associés à moyenne et à médiane). Restons-en à la moyenne  $\mu(N)$  et à l'écart-type  $\sigma(N)$ . Remplacer N par  $\mu(N)$  et  $\sigma(N)$  a les vertus et les défauts de toute synthèse : il y a simplification et amélioration de la lisibilité, que l'on paye par une certaine perte d'information.

Dans le quotidien du monde enseignant on s'occupe beaucoup des moyennes :

- la classe A est « meilleure » que la classe B, sous prétexte que

$$\mu(A) > \mu(B)$$

- le prof X est plus « vache » que le prof Y, parce que pour un même paquet de copies

$$\mu(X) < \mu(Y)$$

- l'élève E passera en Première et l'élève F redoublera, sous prétexte que

$$\mu(F) < 10 < \mu(E)$$

On se préoccupe beaucoup moins de la dispersion des notes qui joue pourtant un rôle important. Est-on bien conscient, dans un conseil de classe par exemple, que lorsqu'on fait la moyenne, pour chaque élève d'une même classe, de la note d'EPS élément de la suite N(EPS) de moyenne  $\mu(EPS)$  et d'écart-type  $\sigma(EPS)$ , affectée du *coefficient 1*, et de la note de français, élément de la suite N(F) de moyenne  $\mu(F)$  et d'écart-type  $\sigma(F)$ , affectée du *coefficient 3*, le classement des élèves sera bien plus gouverné par la note d'EPS que par celle de français, malgré le jeu des coefficients, dès que  $\sigma(F)$  est « nettement inférieur » à  $\sigma(EPS)$ .

Soyons conscients que si le poids « technique » d'une discipline, (il y a bien d'autres composantes dans le poids d'une discipline, ne serait-ce que la personnalité même du prof) est évidemment fonction du coefficient dont il dispose, il est aussi fonction de la dispersion de ses notes, dispersion qu'il peut influencer pour une bonne part.

Pour les moyennes il faut être conscient d'un effet mécanique : Faire des moyennes « rabote » la dispersion ! On est là aussi en présence d'un phénomène souvent mal perçu dans les conseils de classe ou dans des jurys d'examen : la contraction de la dispersion des notes lorsqu'on fait des moyennes. Pour bien comprendre le phénomène considérons une situation tout à fait particulière, mais dans la réalité le phénomène de contraction est du même ordre de grandeur. Soit donc dans une classe de  $n$  élèves, les notes des  $k$  différentes disciplines :

**Trop facile, l'examen est annulé**

À la faculté de droit de Nantes (Loire-Atlantique), on n'avait jamais vu ça. Pour la première fois, des étudiants de licence ont obtenu la note maximale à l'épreuve de droit fiscal : 20 sur 20. Le jury, alerté par cette suspecte excellence, a enquêté. En fait, le professeur, qui avait fait le choix du libellé de l'épreuve, avait choisi un mode d'évaluation peu orthodoxe, sous forme d'un question-réponse binaire « oui » ou « non ». Le doyen de la faculté a estimé qu'il s'agissait « d'un examen pas sérieux dont les résultats exceptionnels ne correspondent pas à une évaluation sérieuse de connaissances ». Le jury a donc annulé l'épreuve et la moitié des 323 candidats devra la repasser (la deuxième moitié avait passé l'examen avec un autre professeur). Les étudiants concernés sont convoqués le 18 juin.

$N(F) = (x_1 ; x_2 ; x_3 ; \dots ; x_n)$  les notes de français

$N(HG) = (y_1 ; y_2 ; y_3 ; \dots ; y_n)$  les notes d'histoire-géographie

:

$N(M) = (z_1 ; z_2 ; z_3 ; \dots ; z_n)$  les notes de musique

Et supposons, *idéalement*, que les moyennes et les écarts-type de la classe soient les mêmes dans chaque discipline. Soit  $\mu$  et  $\sigma$  cette moyenne et cet écart-type communs (le proviseur est très content d'un tel consensus...). Supposons que la sacro-sainte moyenne de chaque élève soit calculée en appliquant le même coefficient à chaque discipline. L'élève « 1 » aura donc pour moyenne :

$$m_1 = \frac{x_1 + y_1 + \dots + z_1}{k}$$

et ainsi de suite, ce qui produira la série des fameuses moyennes de chaque élève si déterminantes pour les décisions mettant en jeu leur scolarité ultérieure :

$$M = (m_1 ; m_2 ; m_3 ; \dots ; m_n)$$

Dans notre cas idéal on aura évidemment :

$$\mu(M) = \mu$$

mais combien de professeurs, combien de chefs d'établissement seront conscients que

$$\sigma(M) = \frac{\sigma}{\sqrt{k}}$$

c'est-à-dire qu'avec 9 disciplines l'écart-type est divisé par 3 ! et qu'avec 4 disciplines, l'écart-type est déjà divisé par 2 ! On m'objectera que je sous-entends que les séries de notes des différentes disciplines sont indépendantes les unes des autres et qu'ainsi je nie la notion de bon élève, d'élève moyen, de mauvais élève, c'est-à-dire que par-delà les disciplines il y a une certaine cohérence, une certaine corrélation entre les notes (les bons sont plus ou moins bons partout, les mauvais...). Certes cela est plus ou moins vrai et cela atténuera donc le phénomène de « rabotage », mais sauf à imaginer que chaque élève ait la même note dans chaque discipline, il y aura peu ou prou du rabotage. On peut ainsi avancer l'idée qu'en général le *contrôle continu* (note décisive comme moyenne de plusieurs notes de contrôles) sera moins discriminant en termes de dispersion des notes que l'épreuve unique. Cela ne préjuge en rien les avantages et inconvénients de l'une ou l'autre pratique.

### **Est-il normal que la distribution des notes soit normale ?**

Autre question : est-il naturel, normal, souhaitable, indispensable que la dispersion des notes (sous réserve qu'il y en ait une certaine quantité) ait l'allure de la célèbre courbe de Gauss (courbe en cloche) ?

La distribution des notes selon une courbe de Gauss (pas forcément centrée sur la moyenne 10, comme pourrait le faire croire la figure) est-elle une loi naturelle ou un modèle social porté par la culture du notateur qui consciemment ou non fabrique épreuves et barèmes pour y aboutir ? Il faut dire que le modèle a tout pour séduire : symétrie harmonieuse, cohérence avec les idées dominantes (petite élite de « forts », masse de « moyens », minorité de « faibles »). La courbe de Gauss serait justifiée si les performances des élèves face à un travail et la notation qui en est faite suivaient « objectivement » une *loi normale* (au sens statistique et mathématique de cette expression).

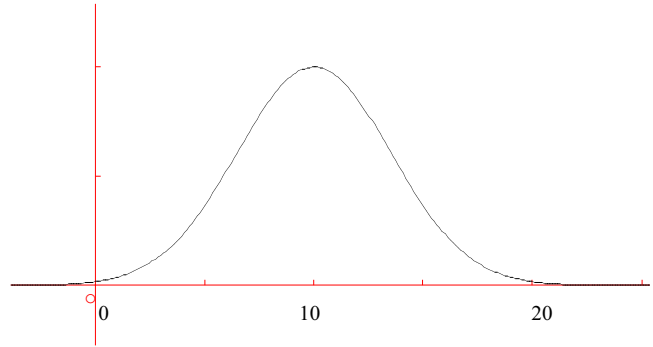


Figure 1

Une telle loi a une définition précise et décrit des phénomènes aléatoires fréquents dans la nature. Mais nous n'avons aucune certitude quant à l'adéquation de la situation scolaire décrite précédemment à une telle loi.

On peut se demander si la référence gaussienne, invoquée pour rassurer les acteurs scolaires (élèves, professeurs, parents, jurys, institution) n'agit pas en fait de façon perverse. La littérature consacrée à la docimologie rapporte souvent l'expérience suivante : on fait corriger par plusieurs correcteurs (afin de lisser les travers individuels) un même lot de copies. On aboutit en général à une distribution des notes suivant une belle courbe de Gauss. On garde ensuite de ce lot le quart des copies les plus faibles et le quart des copies les meilleures, qu'on réunit en un seul lot qu'on fait corriger à nouveau (par d'autres correcteurs !). La logique voudrait qu'on obtienne une distribution bimodale (courbe à deux bosses dont l'allure serait proche de la figure 2). Dans la plupart des cas il n'en est rien, les correcteurs recréent une distribution qui suit une courbe de Gauss. Au mieux elle sera plus étalée (écart-type plus grand) que la courbe initiale.

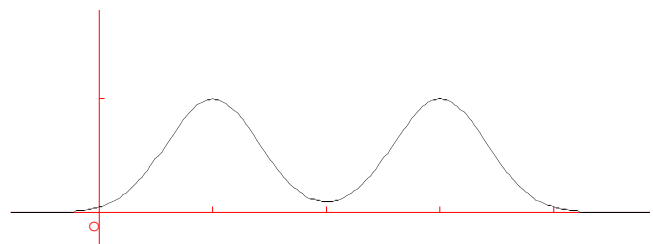


Figure 2

D'autres expériences de docimologie confirment cette tendance naturelle à produire des notes distribuées de façon gaussienne : derrière une loi statistique se cache sans doute aussi un biais comportemental de nature individuelle, sociale ou institutionnelle.

Et d'autres expériences permettent de mettre en évidence d'autres biais comportementaux des correcteurs, très souvent induits par l'attente de l'institution, des parents, des élèves.

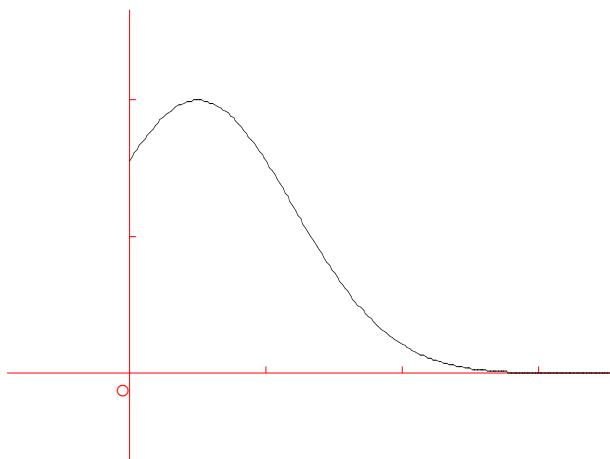


Figure 3

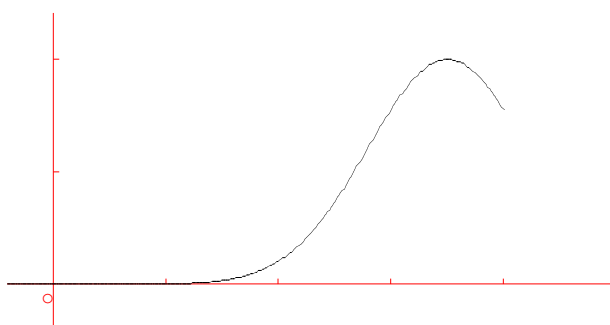


Figure 4

Les notes distribuées selon une courbe de Gauss répondent en général à une fonction plus ou moins explicite : séparer le bon grain de l'ivraie (ou plutôt *définir* le bon grain et l'ivraie) tout en gérant une masse de « moyens ». Lorsqu'on se propose simplement d'extraire d'une population une petite élite (certains concours par exemple), épreuves et corrections s'adaptent pour produire plutôt une distribution dont la courbe aura l'allure en « I » de celle de la figure 3.

Au contraire lorsqu'il s'agit d'un examen largement accordé (le brevet des collèges par exemple) on trouvera plutôt des distributions dont la courbe aura une allure en « j » (figure 4).

Et c'est bien la fonction (de l'épreuve à noter) qui crée en quelque sorte l'organe (du notateur).

***Classez, classez, il en restera toujours quelque chose !***

Derrière l'action de classer, il y a l'idée fondamentale d'ordre, de hiérarchie, mais au delà de cela il y a des aspects, des procédures, des utilisations bien divers.

On peut « classer » (le mot convient-il dans ce cas limite ?) une population en deux catégories : les bons et les mauvais, les reçus et les refusés, les gagnants et les perdants..., c'est-à-dire établir deux catégories *ordonnées* l'une par rapport à l'autre, sans que pour autant il soit nécessaire que les individus soient ordonnés entre eux au sein de chaque catégorie. Certes on peut construire ces deux catégories à partir d'un classement individuel complet, en fixant une *barre* ; c'est ce qui se passe dans un concours pour fabriquer la liste des reçus et des refusés. Rien n'empêche ensuite d'oublier au sein de ces deux catégories le classement individuel et de publier les deux listes par exemple par ordre alphabétique. Remarquons aussi que séparer une population en deux catégories ne sous-entend pas toujours un classement : séparer les hommes et les femmes, les blancs et les noirs n'induit pas (ne devrait pas induire) une hiérarchie.

Bien évidemment, au lieu de deux catégories, on peut « ordonner » une population en plusieurs catégories (par tranches de revenus, par calibres pour des fruits, par admis, admis à un oral de contrôle, refusés, pour les candidats au bac, ...).

À l'opposé de ces classements en simples catégories ordonnées, on peut avoir affaire à des classements individuels : chaque individu de la population considérée est affecté d'un rang précis. Encore que ce cas parfait soit assez rare : souvent la notion d'ex æquo induit un mélange de classement individuel et de catégorisation.

Il faut regarder de plus près le processus de « mise en ordre » sous-jacent. Les mathématiciens définissent avec précision ce qu'est une relation d'ordre et les propriétés qu'on

en exige. Sans aller jusqu'à une telle formalisation technique, il faut quand même souligner les caractéristiques communes de relations comme :

- x est plus grand que y
- x est plus beau que y
- x est meilleur en anglais que y
- x est un aïeul de y
- x est un multiple de y

Par analogie avec l'écriture  $x > y$  utilisée pour les nombres (qui sont naturellement ordonnés), nous noterons de façon générale les relations du type précédent par

$$x \gg y$$

Une première propriété nécessaire au bon fonctionnement de l'ordre est la *transitivité*, c'est-à-dire que chaque fois que  $x \gg y$  et que  $y \gg z$  on peut en déduire que  $x \gg z$ .

Une autre propriété serait bien utile, c'est celle d'interdire à deux individus distincts x et y de vérifier à la fois  $x \gg y$  et  $y \gg x$

Dans la pratique courante des classements on se libère de cette contrainte en créant les fameux *ex æquo* : si  $x \gg y$  et  $y \gg x$  on dira que x et y sont *ex æquo* et on peut donc ainsi, au sein d'une population à classer individuellement, trouver éventuellement quelques îlots d'*ex æquo*.

Considérons à présent deux individus x et y de notre population. Quelles sont les situations possibles ?

- a) x et y sont *ex æquo* (c'est-à-dire  $x \gg y$  et  $y \gg x$ )
- b)  $x \gg y$
- c)  $y \gg x$
- d) On n'a ni  $x \gg y$  ni  $y \gg x$  (on dit que x et y ne sont pas comparables)

C'est la dernière situation (d) qui va distinguer deux types d'ordre. Si tous les éléments sont comparables entre eux, on parlera d'*ordre total*. C'est le cas de la plupart des classements que nous rencontrons : classement à l'issue d'un concours, classement des joueurs de tennis professionnels, classement des équipes de football du championnat de France de 1<sup>e</sup> division. C'est aussi ce que la plupart des médias extraient d'études souvent beaucoup plus complexes et fines lorsqu'ils nous proposent/imposent le classement des hôpitaux, des lycées, des plus grosses fortunes... On peut mettre les individus *dans l'ordre*, i.e. leur affecter un rang : 1<sup>er</sup>, 2<sup>e</sup>, 3<sup>e</sup>... avec éventuellement des *ex æquo*.

Si par contre il existe dans la population à classer des individus non comparables, on parlera d'*ordre partiel*. Deux clubs jouant dans deux poules différentes d'une même division ne sont pas comparables, deux perdants des quarts de finale d'un tournoi ne sont pas comparables, deux personnes d'une même famille ne sont pas forcément aïeul l'un de l'autre (deux frères, un oncle et son neveu...), 12 est multiple de 3, 15 est multiple de 3, mais de ce point de vue (être multiple de), je ne peux pas comparer 12 et 15. Les évaluateurs, testeurs, notateurs... adorent les ordres totaux et font donc tout (et parfois n'importe quoi) pour transformer un ordre partiel

en ordre total. Il est facile de dire qu'une grosse Bentley est plus confortable qu'une Clio ou une Peugeot 206, mais entre ces deux dernières... ? Il est peut-être évident que l'hôpital A est meilleur que la clinique chirurgicale B et que la maternité C, mais entre B et C... ? Un des moyens les plus utilisés est d'attribuer à chaque individu, à chaque voiture, à chaque hôpital, une note chiffrée (avec tous les problèmes posés par la notation) et de faire ensuite le classement via ces notes. Ainsi noter résout tous les problèmes de classement, mais il n'est pas toujours facile de noter (même de façon non pertinente !).

Deux types de problèmes se présentent souvent.

## TOPAZE

### Acte I, scène 12 (extrait)

*(On entend chanter la musique. Topaze ne bronche pas.)*

**Topaze** : Prenons des exemples dans la réalité quotidienne. Voyons.

*(il cherche un nom sur son carnet)*

Élève Tronche-Bobine

*(L'élève Tronche-Bobine se lève, il est emmitoufflé de cache-nez ; il a des bas à grosses côtes, et un sweater de laine sous sa blouse.)*

Pour réussir dans la vie, c'est-à-dire pour y occuper une situation qui corresponde à votre mérite, que faut-il faire ?

**L'élève Tronche réfléchit fortement** : Il faut faire attention.

**Topaze** : Si vous voulez. Il faut faire... attention à quoi ?

**L'élève Tronche, décisif** : Aux courants d'air.  
*(Toute la classe rit.)*

**Topaze, il frappe à petits coups rapides sur son bureau pour rétablir le silence** : Élève Tronche, ce que vous dites n'est pas entièrement absurde, puisque vous répétez un conseil que vous a donné madame votre mère, mais vous ne touchez pas au fond même de la question. Pour réussir dans la vie, il faut être... Il faut être ?...

*(L'élève Tronche sue horriblement, plusieurs élèves lèvent le doigt pour répondre en disant « M'sieur... M'sieur... ».)* Topaze repousse ces avances.)

Laissez répondre celui que j'interroge. Élève Tronche, votre dernière note fut un zéro. Essayez de l'améliorer... Il faut être ho... ho...

*(Toute la classe attend la réponse de l'élève Tronche. Topaze se penche vers lui.)*

**L'élève Tronche, perdu** : Horrible !

*(Éclat de rire général accompagné d'une ritournelle de boîte à musique.)*

Un *premier problème* consiste en la production d'un classement d'une population lorsqu'on connaît de cette population plusieurs classements selon des critères distincts. Exemples : le classement général des élèves d'une classe à partir des classements dans les différentes disciplines, le classement général des coureurs du Tour de France à partir de leurs résultats dans chaque étape, le classement général des hôpitaux à partir des classements de leurs services de chirurgie, de gynécologie, de néphrologie, ..., le classement des meilleurs conducteurs de formule 1 à partir des classements obtenus aux différents Grands Prix. Chacun des exemples cités illustre des stratégies particulières et la liste est loin d'être exhaustive.

Le cas des élèves est le plus simple car les classements dans chaque discipline sont en général obtenus à partir de notes. On revient donc aux notes, on fait des moyennes (pondérées ou non) et on produit le classement général à partir de ces moyennes.

Pour le classement des meilleurs coureurs de formule 1 on est en présence d'une situation un peu analogue : chaque place obtenue dans chacun des Grands Prix rapporte un nombre donné de points (notes). On fait la somme pour aboutir ainsi au classement général.

Au Tour de France, sponsoring oblige, on procède de même, mais avec plusieurs systèmes de « notes » : un premier classement (maillot jaune) est établi en additionnant les temps mis par chacun pour parcourir les différentes étapes (système favorable aux grandes échappées), un second classement (maillot vert) est

établi en attribuant un nombre de points donné à chacune des places obtenues dans chaque étape et en faisant pour chaque coureur la somme des points obtenus (système favorable aux sprinters), un troisième classement (maillot à pois) classe les meilleurs grimpeurs par l'addition des points obtenus selon le classement au sommet des cols et la difficulté de chaque col.

c'est très sucré ; plus on en prend, plus on a soif...



Plus délicat est de réaliser un classement général des hôpitaux. On peut évidemment transformer le classement de chacun des services en notes (parfois d'ailleurs les classements ont été obtenus à partir de notes), puis faire la somme ou la moyenne pour chacun des hôpitaux. Mais tous les hôpitaux n'ont pas forcément un service de néphrologie. Si on somme, les gros CHU, avec tous leurs services, l'emporteront facilement sur de petits établissements. Si on fait la moyenne des notes des services existants chez chacun, la petite clinique avec juste deux ou trois (très bons) services l'emportera sur les gros établissements où il y aura ici ou là un service un peu plus faible.

*L'autre problème* se présente lorsqu'on a classé séparément deux sous-populations d'une population donnée et qu'on veut classer ensuite l'ensemble de la population. Exemples : établir le classement tous sexes confondus de l'ensemble des acteurs de cinéma à partir du classement des actrices et de celui des acteurs, établir un classement général (pour la montée en division supérieure) de différentes équipes ayant disputé un championnat dans des poules différentes d'une même division, établir un classement de la popularité des « célébrités » à partir d'un classement de la popularité des hommes politiques et un autre des personnalités du monde médiatique. Dans tous ces exemples la solution « facile » est de disposer (ou de fabriquer à partir des classements initiaux) des « notes ». Encore faut-il que ces notes soient comparables !

Ainsi si Lionel Jospin a 42% d'opinions favorables et se situe à la troisième place des politiques et que Béatrice Schoenberg avec 36% d'opinions favorables se situe à la seconde place des médiatiques, que peut-on faire ? Comparer des pourcentages non comparables ? Comparer des places dans deux classements non comparables ? Comparer des personnes non comparables ?

Les sportifs n'ont en général trouvé d'autres solutions que de confronter entre eux les deux ou trois premiers de chaque poule pour sortir un classement « inter-poules » afin de trier ceux qui accéderont à la division supérieure.

On peut d'ailleurs remarquer que nous pouvons beaucoup apprendre au sujet des classements dans le monde sportif. C'est dans ce secteur qu'on trouve la plus grande richesse de stratégies d'obtention de classements, dans la mesure où le classement est le nœud gordien du sport de compétition. On peut à ce sujet étudier le processus de classement des joueurs de tennis professionnels, véritable bijou dans son genre, dans la mesure où la compétition reine au tennis est le *tournoi* (qui ne produit qu'un ordre partiel et à chacun desquels ne participent pas tous les joueurs concernés) alors qu'on désire établir le classement, c'est-à-dire un ordre total, sur l'ensemble des joueurs professionnels.

## Conclusion

Classer / noter est un fait social, un besoin, une envie d'ordre socioculturel. Essayer d'objectiviser une telle procédure, de se donner les instruments pour réaliser cette objectivisation peut paraître une démarche naturelle, d'autant plus qu'intuitivement on a envie de rapprocher une procédure de classement / notation de procédures plus scientifiques, plus mathématiques : *mesurer / compter / ordonner*. Nous avons regardé dans ce texte quelques-uns des principes qui gouvernent cette modélisation, mais nous avons aussi essayé d'en

montrer les limites et la relativité des démarches utilisées. Tout cela devrait inciter notateurs, classificateurs, sondeurs, enquêteurs, médiatisateurs... à beaucoup de prudence. D'autant plus que nous n'avons pas du tout abordé un autre facteur qui devrait inciter à encore plus de prudence, c'est tout ce qui concerne la pertinence même des critères de jugement, la neutralité et la stabilité de jugement de l'évaluateur, l'usage de synthèses parfois bien réductrices, voire biaisées.

Tout cela me rend bien sceptique et je vais me consoler en buvant un coup de rouge qui a eu 14/20 dans le guide X et 17/20 chez Y, sur la terrasse d'un café « trois étoiles » d'une ville qui a 15/20 pour sa qualité de vie : le bonheur !

Daniel Reisz

**TEST QUE CHOISIR LABORATOIRE margarines et dérivés**

★★★ TRÈS BON  
 ★★ BON  
 ★ MOYEN  
 ■ MÉDIOCRE  
 ■■ MAUVAIS

NOS ANALYSES	PRIMEVÈRE spécial cuisine 80 % MG	AUCHAN tournesol 70 % MG	FRUIT D'OR tartines et cuisson 70 % MG	CARREFOUR au tournesol 80 % MG	CHAMPION au tournesol 80 % MG	CASINO au tournesol 80 % MG	BOUTON D'OR 1,2,3 soleil à tartiner 70 % MG	MONOPRIX cuisine et pâtisserie <sup>(1)</sup> 80 % MG	ASTRA à tartiner et à cuire 70 % MG
	produits contenant plus de 60 % de matières grasses								
Poids net (g)	250	500	250	500	500	250	250	500	500
Prix (€/F)	1,20/7,9	1,26/8,25	1,06/6,95	1,26/8,25	1,30/8,5	0,80/5,25	0,75/4,95	0,89/5,85	1,27/8,35
Prix au kg (€/F)	4,82/31	2,52/16,5	4,24/28	2,52/16,5	2,59/17	3,20/21	3,02/20	1,78/12	2,55/17
Valeur calorique (kcal/100g)	728/★	638/★	640/★★	727/★	724/★	731/★	637/★	725/★	640/★
Sel (%)	0,3/★★	0,3/★★	0,3/★★	0,5/★	0,5/★	0,3/★★	0,5/★	0,8/■■	0,6/■
Vitamine A (µg/100g)	25/■	1 250/★★	821/★★	760/★★	750/★★	900/★★	910/★★	< 2/■	< 2/■
Vitamine E (mg/100g)	33/★	39/★	16/■	51/★★	38/★★	40/★	26/★	< 3/■■	< 3/■■
Conservateurs	★★	★★	★	★★★	★★★	★	★	★★★	★★★
Acides gras saturés	★★	★★	★★	★★	★★	★★	★★	■	■■
mono-insaturés	★★	★★	★★	★★	★★	★★	★★	★★	★★
polyinsaturés	★★	■	★	■	■	■	■	★★	★★★
trans	★★★	★★★	★★★	★★★	★★★	★★★	★★★	■	★★★
Microbiologie	★★★	★★★	★★★	★★★	★★★	★★★	★★★	★★★	★★★
Tartinabilité à 8 °C <sup>(2)</sup>	★★	★★	★★	★★	★★	★★	★★	★	★
Étiquetage	★★★	★★★	★★★	★★★	★★★	★★	★★★	★★	■
JUGEMENT note sur 20	16	15,5	15,5	15,5	15,5	15	15	12	8
GLOBAL appréciation	★★	★★	★★	★★	★★	★★	★★	★	■

(1) Formule modifiée en juillet 2001 – uniquement 5 huiles végétales. (2) La tartinabilité n'a pas été prise en compte pour les graisses destinées uniquement à la cuisson.

Que Choisir, N°389 bis/6, Janvier 2002

N.D.L.R. : le coefficient de Spearman-Brown pour la tartinabilité est omis par cette revue pourtant sérieuse.

le deuxième, c'est que je ne les aime pas trop...

## Le classement des joueurs de tennis professionnels

Un bon exemple d'une procédure de classement est celle en vigueur pour les joueurs professionnels de tennis (classement ATP). Il est actuellement remis à zéro au début de chaque saison et ne tient donc pas compte du passé (saisons précédentes). Remarquons qu'il n'en a pas toujours été ainsi et qu'il y a encore quelques années, les points accumulés les saisons précédentes étaient repris en compte, avec simplement un coefficient de vieillissement qui faisait qu'un joueur sans résultats reculait progressivement au classement, mais qui malgré tout établissait un classement assez (trop) stable parmi eux. Trop stable aux yeux des sponsors qui, pour corser les choses et introduire plus de suspense et donc plus d'attractivité, ont imposé un classement remis à zéro à chaque saison, obligeant chaque joueur à refaire ses preuves, sans prise en compte de ses performances passées.

Le tennis se joue par tournoi, c'est-à-dire qu'à chaque tour les perdants sont éliminés et les gagnants s'affrontent deux par deux au tour suivant. Pour des raisons évidentes, un tournoi regroupe 128, 64, 32, 16, ... joueurs et se termine donc par des quarts de finale (8 joueurs), puis demi-finales (4 joueurs), puis finale (2 joueurs). Un tel tournoi ne produit pas un classement, puisque, si on peut comparer le vainqueur et le finaliste, il n'est pas possible de comparer les deux perdants des demi-finales, ni ceux des quarts de finale, etc.

Les exigences sportives, médiatiques et aussi commerciales exigent un classement. Il se fait donc à travers le cumul de points attribués à l'issue de chaque tournoi. Le nombre de points est à la fois fonction de la performance du joueur et de l'importance du tournoi. Ainsi, pour l'un des quatre tournois du « Grand Chelem » (Roland Garros, Wimbledon, U.S. Open et Open d'Australie) :

Vainqueur :	200 points
Finaliste :	140 points
Perdants des deux demi-finales :	90 points
Perdants des quarts de finale :	50 points
Perdants des huitièmes de finale :	30 points
Perdants du 3 <sup>e</sup> tour :	15 points
Perdants du 2 <sup>e</sup> tour :	7 points
Perdants du 1 <sup>er</sup> tour :	4 points ou 1 point selon que le joueur est issu d'épreuves de qualification ou non.

Pour un tournoi ATP officiel, mais bas de gamme, les points attribués sont les suivants :

Vainqueur :	35 points
Finaliste :	24 points
Perdants des demi-finales :	15 points
Perdants des quarts de finale :	8 points
Points pour les deux tours précédents :	3 et 1 points

Il y a ainsi, parmi la petite centaine de tournois reconnus par l'ATP, 9 catégories, chacune avec son barème, entre les quatre tournois du Grand Chelem et les tournois « bas de gamme ». Il y a donc chez chaque joueur, outre l'envie de gagner, toute une stratégie à mettre en œuvre pour choisir les « bons » tournois en fonction de son niveau, afin de marquer un maximum de points.

D.R.

## Championnat de France de football

Il y a quelques années de cela, les dirigeants du football professionnel ont introduit un changement dans l'attribution des points dont le but était de favoriser le jeu offensif. En effet,

1	NANTES	NANTES
2	LYON	LYON
3	LILLE	LILLE
4	BORDEAUX	BORDEAUX
5	SEDAN	SEDAN
6	RENNES	TROYES
7	TROYES	RENNES
8	BASTIA	GUINGAMP
9	PSG	BASTIA
10	GUINGAMP	PSG
11	MONACO	METZ
12	METZ	MONACO
13	AUXERRE	LENS
14	LENS	AUXERRE
15	MARSEILLE	MARSEILLE
16	St-ÉTIENNE	St-ÉTIENNE
17	TOULOUSE	TOULOUSE
18	STRASBOURG	STRASBOURG

jusqu'alors on attribuait 2 points pour une victoire, 1 point pour un match nul et 0 point pour une défaite. Afin d'éviter que les équipes, surtout en déplacement, ne se contentent trop facilement d'un match nul et ne se réfugient dans un jeu entièrement défensif peu spectaculaire, il fut décidé d'attribuer 3 points en cas de victoire, 1 point pour un nul et 0 point pour une défaite, c'est-à-dire de surpayer les victoires vis-à-vis des matchs nuls. En prenant le championnat de France 2000-2001 comme exemple et en faisant les deux décomptes, on obtient les deux classements ci-contre.

Que peut-on en conclure ? Tout d'abord que le passage de l'un des décomptes à l'autre ne bouleverse pas vraiment la hiérarchie générale des équipes. Seules quelques petites interversions se font jour, ici ou là, mais aucune équipe ne change significativement

de place au classement de fin de saison. Ceci est un constat technique objectif.

Peut-on aller plus loin ? On aurait alors envie de dire que l'objectif de la réforme n'a pas été atteint de façon significative puisque les deux barèmes fournissent à peu de choses près la même hiérarchie.

1	MONACO	MONACO
2	NANTES	NANTES
3	STRASBOURG	STRASBOURG
4	MARSEILLE	MARSEILLE
5	BASTIA	BASTIA
6	NANCY	NANCY
7	St-ÉTIENNE	St-ÉTIENNE
8	NICE	NICE
9	SOCHAUX	SOCHAUX
10	LAVAL	LAVAL
11	PSG	PSG
12	METZ	METZ
13	NÎMES	NÎMES
14	BORDEAUX	VALENCIENNES
15	LYON	REIMS
16	VALENCIENNES	BORDEAUX
17	REIMS	LYON
18	LENS	LENS
19	TROYES	TROYES
20	ROUEN	ROUEN

Plus intéressant est de constater que les trois lignes qui précèdent constituent une belle faute de raisonnement, que d'aucuns n'ont pas manqué de faire. En effet, que les deux barèmes n'introduisent pas de différence notable *après coup* ne signifie pas que le style de jeu n'a pas changé. Cela ne signifie d'ailleurs pas non plus le contraire ! On peut par exemple penser que les équipes, conscientes du nouveau barème, avaient eu un jeu plus offensif, donc plus spectaculaire, mais que cela n'a pas affecté la hiérarchie des équipes. Une autre idée est de se demander l'effet que cette modification aurait eu sur un championnat antérieur. Voir ci-contre par exemple les deux classements obtenus pour la saison 78/79. Quelques légères modifications entre la 14<sup>e</sup> et la 17<sup>e</sup> place. Là encore, rétrospectivement, la hiérarchie des clubs n'était pas bouleversée ! Là encore il est difficile d'en tirer des conclusions.

D.R.

## Quotient intellectuel (Q.I.)

*Peu important les tests pourvu qu'ils soient nombreux. A. Binet*

Qu'est-ce que l'intelligence ? Est-ce seulement une boutade que de répondre que c'est ce que mesure le Q.I. (quotient intellectuel) ? Sans rentrer dans un débat de fond qui dépasserait nettement le cadre de ce court texte, essayons de « mesurer » les conséquences du succès de cette tentative de *mesurer l'intelligence*.

Le principe du Q.I. tel qu'il fut mis au point par Alfred Binet entre les années 1904 et 1911 consiste en une batterie de tests progressifs, étalonnés par une échelle *d'âges mentaux*. Le sujet *normal* saura réussir tous les tests correspondants à des âges mentaux inférieurs ou égaux à son âge réel et échouera à ceux qui correspondent à des âges mentaux supérieurs à son âge réel. C'est dans le souci de pouvoir comparer les performances de sujets d'âges différents que le psychologue allemand Stern proposa de diviser la « note » obtenue aux tests par l'âge réel du sujet : le *quotient* intellectuel était né !

La mise au point par Binet de ces tests répondait à une commande du ministre de l'Instruction de l'époque qui désirait mettre à la disposition des instituteurs un outil simple permettant de dépister parmi les élèves en difficulté scolaire ceux qui relèveraient d'un enseignement spécialisé. Contrairement à l'usage qui en a été fait ultérieurement, A. Binet s'interdisait d'assimiler Q.I. et intelligence et était parfaitement conscient qu'un outil aussi grossier ne devait en aucun cas donner naissance à des classements individuels, mais rester un simple indicateur sans grande précision.

Alfred Binet s'interdisait aussi de considérer le Q.I. comme pratiquement invariant chez un individu donné. Il croyait profondément aux vertus de l'éducation et de nombreuses expériences attestent que le Q.I. d'une personne peut progresser sous l'effet d'une éducation sérieuse.

Ce qui nous amène à nous intéresser au Q.I. dans le cadre de cette étude sur les procédures de notation et de classement est le caractère exemplaire des (mauvais) usages du Q.I. Une première dérive a consisté à passer d'un indicateur grossier de dépistage à un instrument de mesure précis en soi : cela mesure avec une totale précision quelque chose qu'on est bien en peine de définir ! En réalité le Q.I. est un unique nombre obtenu par sommation de repérages de compétences très complexes, difficiles à isoler de tout contexte socioculturel et psychologique.

Une seconde dérive a consisté à passer de la détection d'enfants relevant d'une éducation spécialisée à une échelle couvrant toute l'étendue des niveaux de performances. C'est comme si, pour mesurer la distance entre Paris et Nice, on se servait d'un double décimètre conçu pour mesurer les longueurs d'une figure de géométrie sur une feuille de papier. Conscients de ce travers, certains psychologues ont d'ailleurs mis en place des tests distincts selon le niveau présumé des sujets à tester.

Une troisième dérive, conséquence directe des deux premières, consiste à faire du Q.I. un instrument de *classement* entre les individus ou entre des populations (le Q.I. moyen des Noirs, des femmes, des délinquants sexuels, ...) ou de *catégorisation* (âge mental inférieur à 3 ans : idiot, âge mental entre 3 et 7 ans : imbécile, ... jusqu'aux *doués* et aux *surdoués*, en passant

par les *moyens* et les *normaux*, ces derniers étant d'ailleurs supérieurs aux *moyens*). Ces catégories étant évidemment définies par des bornes précises, alors que la moindre des choses serait d'envisager des zones un peu floues.

Dans cette perspective on ne peut pas ne pas citer un extrait d'un discours tenu en 1922 par le président de la Colgate University : « *Nous ne pouvons concevoir pire forme de chaos qu'une démocratie véritable dans une population dont l'âge mental dépasse à peine les treize ans.* » (il s'agissait évidemment de la population des États-Unis)

Tout enseignant, tout évaluateur ne peut qu'être interpellé par les effets pervers d'une fausse scientificité introduite par l'usage d'un instrument technique censé mesurer avec précision quelque chose qui est loin d'être clairement défini.

Bibliographie : Stephen Jay Gould, La mal-mesure de l'homme, Éditions O. Jacob, 1997.

D.R.

TABLEAU D'HONNEUR $\frac{1}{5}$ sur $\frac{2}{5}$ Absences				Philosophie	
FACULTÉS	VALEUR 0 à 40	PLACES	NOMBRE des ÉLÈVES	M	
Philosophie.....			38	Franglais M = elle	Trimestre d'avant. Semble n'être
Composition franç <sup>e</sup> , littérature.....	26	10 <sup>e</sup>		Latin M = vide	mal adaptée à la classe. L'oral est souvent vraiment avec toutefois une certaine tendance à la dissipation] mais l'écrit est soigné. (P)
Français, Grammaire.....	14	18 <sup>e</sup>		Grec	
Diction.....	30	15 <sup>e</sup>		M	
Latin-thème.....	15 <sup>e</sup>	28 <sup>e</sup>		Histoire et Géographie	
Latin-version.....				M = Ager	Ensemble honorable à l'écrit. Attention en classe irrefutable - f#
Grec-thème.....				Langue vivante 1	français vivants, bon oral, assez un certain raffi de. mais écrit décevant.
Grec-version.....				M = Priv	
Histoire.....	25	10 <sup>e</sup>		Langue vivante 2	
Géographie.....	23	13 <sup>e</sup>		M	
Anglais.....				Mathématiques	Je laisse vivre - ne donne pas tout ce qu'on peut attendre. Doit beaucoup mieux (ai - résultats assez bons JP
Allemand.....	25	21 <sup>e</sup>		Sciences Physiques et Naturelles	Très bon travail EP
Italien.....				M = Aulo	
Espagnol.....					
Mathématiques.....	24	16 <sup>e</sup>			
Chimie.....					
Physique.....					
Sciences naturelles.....	27	20 <sup>e</sup>			
Musique.....	34	11 <sup>e</sup>			
Dessin.....	29,3	13 <sup>e</sup>			
Couture.....	8	27 <sup>e</sup>			
Gymnastique.....	18	11 <sup>e</sup>	195		
				APPRECIATION de la DIRECTRICE	
				Admise dans la classe supérieure sous condition. Voir lettre jointe	

Pierre : ah, les nems, je leur mets 10, sans hésiter !

## NOTER POUR EVALUER

L'ARTICLE de Daniel Reisz part du constat que, dans la notation, on manie des notions bien repérées en mathématiques : ordre, équivalence, mesure. Mais il remarque aussi que la pratique ne vérifie pas toujours bien les propriétés que les mathématiques ont ainsi dégagées : elle fait de la théorie une application confuse et contestable.

Faut-il dire qu'à cette approche critique, je vais tenter d'opposer une justification ? Sans doute pas. Je partirai cependant d'un autre point de vue. Non pas celui des qualités d'une (bonne) notation : celui du besoin auquel elle s'efforce de répondre.



Flûte à 7 trous.

De quoi s'agit-il ? Souvent, nous avons besoin ou envie de juger : apprécier une situation, une personne, une machine, etc. afin de prendre une décision. Besoin, envie ? pas même : juger des gens et des choses est dans la nature de notre activité éveillée. Constamment, nous estimons que le café est trop chaud, que Jules est sympathique, qu'il risque de pleuvoir, que nous préférons voter pour X plutôt que pour Y, que le trajet en métro est plus sûr en durée que celui en autobus, et ainsi de suite. La plupart du temps, nous nous contentons d'une évaluation intuitive. Pas même explicite, souvent : le geste implique le jugement sans que celui-ci ait été formulé. Mais parfois, nous doutons de la pertinence de cette intuition. Nous voudrions asseoir notre jugement sur des éléments tangibles ; ou bien, nous n'avons pas confiance dans l'appréciation portée par quelqu'un d'autre et nous voudrions la contrôler en nous assurant qu'il a bien pris en compte ce qui doit l'être. Ou encore, il s'agit de comparer des gens ou des choses qui ont été séparément évaluées par des juges différents. Alors, on va commencer à formaliser.

En fait, il y a là deux finalités distinctes ;

- d'une part, on essaie d'*objectiver*, c'est-à-dire faire échapper le jugement à des circonstances qui ne devraient pas jouer ou éviter au contraire d'en méconnaître d'autres qui importent. Se laisser influencer par les premiers serait une blâmable discrimination ou à tout le moins une perte de pertinence : la consonance du nom, la couleur de l'emballage, la qualité de notre digestion ce jour là, ... Négliger les seconds aurait des conséquences aussi fâcheuses ;
- d'autre part, il s'agit de trouver *un langage*, un moyen de communiquer notre jugement à d'autres.

Ces deux préoccupations sont indépendantes. Je puis, pour choisir une voiture, m'astreindre à regarder le prix, le poids, la consommation, la boîte de vitesses, le volume du coffre, le nombre de sièges et leur confort, etc. mais j'élabore cela pour moi-même et n'ai pas besoin de le transmettre. À l'inverse, l'usage de notations s'étant répandu dans la société, on entend parfois dire que l'on met à Untel une note de 15/20, comme ça, à l'estime, pour son aptitude ou son esthétique ou autre chose : sans qu'il soit besoin de détailler les critères pris en compte. C'est un langage imagé et l'on se fait très bien comprendre (du moins dans le contexte donné). Mais souvent aussi, on visera les deux objectifs à la fois : encadrer l'évaluation et la faire communiquer entre évaluateurs ou décideurs.

### **Désubjectiver le jugement**

D'abord, considérons un seul juge : nous-même, par exemple. Nous voulons asseoir notre jugement sur des critères exprès, car d'une certaine façon nous nous méfions de notre impulsion. Définir une notation contrôlée suppose trois opérations : l'une est d'arrêter la liste, la nomenclature des critères à examiner ; la deuxième est de doter d'une échelle chacun des critères ; la dernière est de décider d'une manière de les combiner.

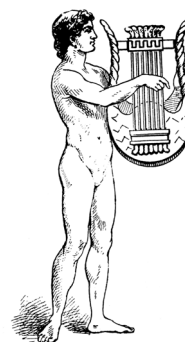
1. liste de critères. Par exemple, on dira que des candidats doivent faire une dictée, nager 100 m et subir un test psychométrique. Ou encore, dans une copie d'examen : le plan, le style, les connaissances et l'orthographe.

2. cotation. On définira pour chaque épreuve une échelle de cotation : peut-être à deux catégories (admis/refusé), peut-être à plusieurs niveaux (A, B, C, ...) en principe ordonnés, ou enfin une échelle quasi-continue telle qu'une note chiffrée. Cette note élémentaire donnée ici peut rester tout à fait subjective : simplement, on s'astreint à faire porter la subjectivité sur des critères bien explicités. Mais on peut aussi décider qu'on prendra une mesure objective : nombre de secondes du parcours, nombre de fautes à la dictée, etc.

À partir de ce schéma, et avant d'en venir à la synthèse, se situera la cotation proprement dite, qui consiste à examiner l'objet ou la personne au regard de chaque critère séparément et à la classer dans une catégorie ou lui attribuer une note.

3. synthèse (ou, agrégation). On dispose de notes séparées pour les différents critères et l'on veut les ramener à une note unique. Souvent, si l'on a des notes numériques, on en fait la somme ou la moyenne ; et il peut s'agir de la somme ou de la moyenne simple ou bien avec des pondérations qui donnent plus d'importance à certains critères qu'à d'autres. Cela permet une compensation entre bons et mauvais scores. Mais il y a d'autres façons de faire : on peut par exemple exiger un niveau minimum à toutes les épreuves. Ou, des formules plus compliquées ou hybrides.

Pour cette synthèse, souvent, il est commode d'avoir des notes partielles qui soient des nombres : car cela permet des opérations telles que des moyennes (pondérées) ou simplement parce que les nombres sont ordonnés. On peut en effet seulement s'intéresser au rang : à l'ordre dans lequel les différents objets ou personnes apparaissent au regard du critère en question. Le nombre n'est ici convoqué que pour la relation d'ordre, de même que les maisons sont numérotées dans une rue sans qu'on en déduise que la distance entre les numéros 2 et 10 serait la même qu'entre le 12 et le 20.



Lyre à 7 cordes.



Je ne crois pas essentiel - et, pour tout dire, il est un peu illusoire - de prétendre « échapper » à une supposée dictature des nombres en adoptant des barèmes de lettres ou de couleurs, par exemple. J'ai connu des appels d'offres pour l'attribution de crédits de recherche ou des marchés de fourniture, où l'on demandait de classer en trois catégories : les très bons (A), les bons (B) et les mauvais (C) : eh bien ! on voyait apparaître des A<sup>-</sup> et des B<sup>+</sup>, quand ce n'étaient pas des A<sup>+++</sup>. Si c'est pour en venir là, autant noter de 0 à 20...

Un point important est que c'est souvent le besoin de synthétiser plusieurs critères qui conduit, pour formaliser la synthèse, à utiliser des notes chiffrées : non parce que c'est nécessaire, mais parce que c'est commode. On est conduit à « numériser » ce qui n'est que qualitatif. Autrement dit, c'est le besoin d'une appréciation multidimensionnelle (ou, multicritère) qui conduit d'abord à plaquer un critère qualitatif sur une échelle de nombres : de 1 à 5, de 0 à 20 ou autre. Par exemple, on vous dit « donnez votre préférence entre 1 et 5 ». On soulignera, pour s'en indigner ou s'en émerveiller, qu'on a alors toute l'apparence d'une mesure comme une longueur ou un poids, alors qu'il s'agit du plaisir ou du degré de confiance que l'on éprouve. Qu'on s'en indigne, soit ! Mais aussi, qu'on considère que cette formalisation répond au but de contrôler, d'endiguer l'impétuosité de jugements qui, sinon, seraient autrement affectifs ou fantasques.

« Ceux qui jugent d'un ouvrage sans règle sont à l'égard des autres comme ceux qui ont une montre à l'égard des autres. L'un dit : il y a deux heures ; l'autre dit : il n'y a que trois quarts d'heure. Je regarde ma montre et je dis à l'un : vous vous ennuyez et à l'autre : le temps ne vous dure guère, car il y a une heure et demie et je me moque de ceux qui disent que le temps me dure à moi et que j'en juge par fantaisie.

Car il ne savent pas que j'en juge par ma montre. »

Pascal Pensées

### ***Intersubjectiver les jugements***

Ici, le problème n'est peut-être pas tant d'avoir une appréciation « objective » que de s'assurer d'une transitivité des appréciations entre des juges différents : que la même personne ou le même objet soit apprécié de la même façon par tous ; ou, ce qui revient au même, que l'appréciation par l'un vaille pour l'autre. Il s'agit que le jugement soit communicable. À la limite, peu importe qu'il reste subjectif (intuitif, affectif,...) pourvu que ce soit de la même façon<sup>1</sup> : d'où l'adjectif « intersubjectif ». Ceci peut se présenter de deux manières, que nous pourrions appeler verticale et horizontale.

1. Cette identité du ressenti, chez des personnes différentes, est sans doute illusoire. Mais elle est posée comme un idéal : hors d'atteinte, certes, mais dont on s'efforce de s'approcher.

La transmission - ou, transitivité - verticale est le cas d'un jugement *délégué* : si je n'ai pas le temps ou la compétence pour porter un jugement, je vais m'en remettre à vous. Et, je ferai mienne votre appréciation. Par exemple, je souhaite recruter un formateur en informatique : je n'ai pas le temps de recevoir moi-même tous les candidats, parce que mes fonctions m'occupent totalement par ailleurs. Ou bien, je pourrais recevoir les candidats, mais je ne connais rien à l'informatique et je ne sais pas reconnaître un bon pédagogue. Je demande donc à plus compétent et plus disponible que moi de faire cette sélection. Mais le juge, l'examineur, l'arbitre, l'expert, etc. auquel je m'en remets doit préparer une décision que je prendrai et qui soit ajustée compte tenu de ce que j'attends de la personne recherchée. Donc, ce juge que je mandate doit s'assurer de ce que je cherche : il me demande d'abord quels sont mes critères. Je lui fournis la première des trois étapes évoquées plus haut.

(Ou, nous discutons pour l'établir d'un commun accord.) Puis il accomplira les deux autres étapes. Mais le résultat me reviendra.

La transmission - ou, transitivité - horizontale est le cas où les différents candidats (personnes, équipements, procédés, etc.) mis en concurrence ne sont pas examinés par un juge unique mais par plusieurs. Il faut alors s'assurer qu'un même candidat serait jugé de la même manière par les différents juges (ou jurys). À la question de la pertinence des critères, comme précédemment, s'ajoute ici celle de l'équité. Il faut donc donner aux différents juges une liste commune de critères de jugement, une échelle commune pour chaque critère et un même schéma de synthèse (minima, moyennes, pondérations, etc.). Il faut aussi que les juges soient à peu près interchangeable, c'est-à-dire également compétents. Mais il restera sans doute une part de subjectivité dans leurs appréciations. On peut tomber sur un examinateur plus sévère ou qui se laisse séduire par une forme de présentation, etc. Cependant, le fait de s'être donné un ensemble de critères et de barèmes permet de réduire les différences d'appréciations entre les juges. Ceci demande un certain soin *a priori* et aussi une vigilance durant tout le processus. Par exemple, on peut croire qu'on a donné une juste pondération aux différents critères alors que (D. Reisz le mentionne) le fait qu'un examinateur note dans une large fourchette et un autre avec très peu de variation introduit une pondération implicite qui modifie le poids en principe affecté à leurs notes respectives.

### **Objectiver en fonction de l'objectif**

Une formalisation du processus de jugement ou notation permet d'échapper à l'inconvénient d'impulsions ou à la relativité des notateurs. (Permet de, ou en tout cas vise à.) Les notions



Double flûte.

mobilisées par cette formalisation ressemblent à celles des mathématiques : nomenclatures et relations d'équivalence, échelles et relations d'ordre, chiffrage et mesurage. On n'atteint pas toujours à la pureté des notions mathématiques en question, mais on se trouve quelque part à mi-chemin de la subjectivité incontrôlée et de la mesure objective. On peut donc se plaindre que le verre reste à demi-vide ; mais on peut être heureux qu'il soit déjà à demi-plein. On utilise en quelque sorte une mathématique dégradée, une pseudo-mathématique. Mais après tout, n'en est-il pas souvent ainsi dans la pratique ? Lorsqu'un jardinier dessine un massif circulaire, est-on bien sûr que le pourtour en est un cercle à la fraction de millimètre près ? Et, a-t-on vraiment besoin qu'il en soit ainsi ? Le tout est de savoir si l'approximation est suffisante pour qu'on s'appuie sur quelques propriétés intéressantes des outils mathématiques utilisés (on

dirait presque : imités). Et, ceci, compte tenu du but que l'on se donne. D'où le jeu de mots de mon dernier intertitre : échapper suffisamment à la subjectivité en fonction de l'objectif que l'on se donne.

Cette utilisation pragmatique de notions ressemblant à celles des mathématiques (des nombres, mais aussi, on l'a vu des notions non numériques : ordres, équivalences...) est

commode et tout à fait licite. À condition toutefois qu'on observe deux règles : être clair sur ce qu'on fait et ne pas tirer l'interprétation trop loin.

Si l'intérêt de cette formalisation du jugement est que l'on explicite un certain nombre de choses (critères d'évaluation, pondérations, etc.) il faut justement que cette explicitation soit accessible. Il faut donner la recette que l'on applique. C'est classiquement ce qu'on fait pour les examens scolaires, domaine où la notation est devenue paradigmatique : liste des matières, programmes des épreuves, barèmes de notes, statistiques ultérieures (et rétrospectives), ... Mais c'est aussi ce qui est fait dans d'autres cas. Une revue de consommateurs qui établit une comparaison de moulins à légumes ou de téléviseurs donne les divers critères qu'elle a examinés. Si elle propose une note de synthèse (un « meilleur choix ») elle vous donne aussi le détail et vous laisse libre de préférer d'autres pondérations. Elle a jugé pour vous, par délégation, mais vous êtes en état de vérifier si ses critères de jugement s'accordent aux vôtres.

La seconde précaution consiste à ne pas se piéger soi-même ou à ne pas surinterpréter des notes d'apparence rigoureuse. Il faut garder à l'esprit que ce n'est là qu'une amélioration par rapport à ce qu'un jugement brut aurait pu être. Le fait que ça s'exprime par des chiffres ne veut pas dire que c'est « scientifique ». Encore que : ne pas non plus donner à cet adjectif le sens qu'on lui donne souvent à tort. Un résultat n'est pas scientifique : c'est la démarche qui y conduit qui peut l'être... ou non. Or, la science est

(étymologiquement : scinder) l'art de séparer le vrai du faux. Et, la science a depuis longtemps repéré qu'il n'y a pas une telle distinction tranchée entre le vrai et le faux. Le vrai n'est souvent qu'approximation : si ma montre marque 12h03 et que je dis « il est midi », est-ce faux ? Parfois aussi, le vrai est seulement probable : si un médecin dit « ce malade n'en a plus que pour une semaine », comprenons-nous qu'il mourra certainement dans 168 heures ? En revanche, ce qui est scientifique, c'est la manière de contrôler le doute, la manière d'arriver à une affirmation approximative et probable ; et, de donner en même temps une indication de son degré d'approximation et de probabilité.

Or, qu'a fait notre processus de notation ? Il a caractérisé une ou des propriétés d'un objet ou d'une personne et l'a fait en vue d'une certaine décision (achat, recrutement, ...). On espère qu'il permet à cette décision de n'être pas trop mauvaise. Mais doit-on s'autoriser du résultat intermédiaire que la note constitue pour en tirer sur le produit ou la personne en cause d'autres conclusions ? Pas forcément. La notation a consisté à objectiver le ressenti, à donner une expression quantitative de ce qui n'était pas quantitatif, une mesure de ce qui n'est pas métrique. Elle l'a fait par approximation en vue d'un certain but. Elle n'a pas eu pour effet de mesurer le non mesurable.

En revanche, le fait d'utiliser le langage des nomenclatures et des nombres permet diverses opérations sur les résultats intermédiaires et finals : des indicateurs de l'écart entre notateurs, de la stabilité ou de l'homogénéité des jugements, etc. Ce qui permet une certaine appréciation de la qualité des appréciations portées.



Lyre à 7 cordes, double flûte.

Si l'on veut bien alors ne demander à la notation que ce pour quoi on l'a établie, ne pas lui faire dire ce qu'il n'a pas été prévu qu'elle dise, alors la procédure consistant à enserrer ainsi le jugement dans des grilles, des barèmes et des règles est acceptable : car elle est utile.

Mais il faut voir aussi que l'abus n'est pas seulement du côté de l'interprétation après coup. Il peut se situer dans le processus même de la notation ou chez ceux mêmes qui l'instituent : si certains manipulent le processus pour lui faire fournir le résultat qui les arrange (ou, l'empêcher de produire ce qui les gêne), ou encore s'ils instrumentalisent le processus lui-même pour s'arroger un pouvoir ou pour accréditer par avance une signification hors de propos ou hors d'atteinte.

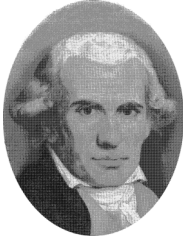
René Padieu



Violoncelle. Contrebasse à quatre cordes.  
(xviii<sup>e</sup> siècle.)

## CARNET DE NOTES IMAGINAIRES DE GASPARD MONGE

Alès, 1<sup>er</sup> Février 1789



UN VENT GLACÉ balayait les rues ce matin, mais le ciel était clair comme jamais. Une belle lumière d'hiver dessinait la ville comme une épure. J'ai retrouvé Alès, familière, pour la quatrième fois en trois ans. Le collège n'a pas changé. Les pensionnaires de l'école de marine, que je viens interroger, se mêlent sans heurt aux autres écoliers.

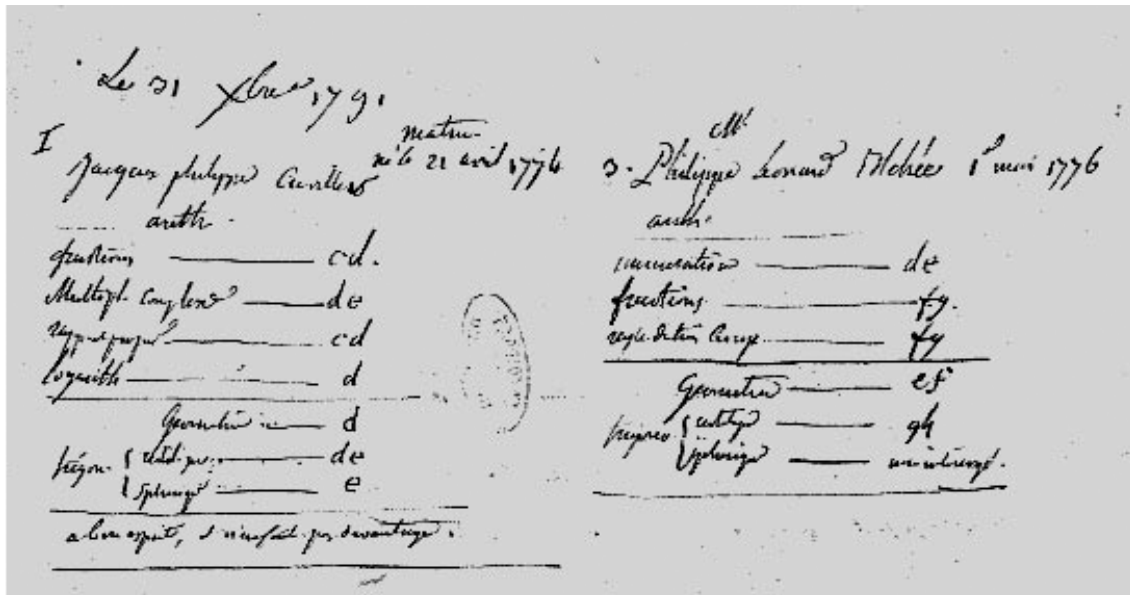
Des enfants encore, treize ou quatorze ans, parfois moins, comme ceux de Vannes, où je serai en mars, si tout va bien. Sur la liste qui m'a été remise avant mon départ, j'ai noté un petit Artur de la Villarmois qui n'a pas même douze ans... Saura-t-il déjà maîtriser l'arithmétique, la géométrie, la navigation, et même la trigonométrie rectiligne et sphérique, au programme pour la première fois ?

Je sais que j'aurai toujours autant de plaisir à questionner, à écouter, à évaluer ces jeunes intelligences en éveil. Pourtant, cette année, ma tâche m'inquiète : Monsieur de la Luzerne ne veut pas plus de quarante élèves officiers, nous en avons gardé trop dans les deux dernières promotions. Je vais devoir faire un tri sévère entre tous ces jeunes gens. Il ne s'agit plus de dire pour chacun : celui-ci oui, cet autre non. Ce n'était déjà pas facile d'être juste en interrogeant autant : plus de cent l'an passé entre Vannes et Alès. J'avais déjà dû prendre force notes pour ne pas laisser ma mémoire trahir ces enfants, et pour assurer à ma conscience une équité rigoureuse. Mais cette fois, être juste, c'est être sûr de choisir les quarante meilleurs. Et s'il s'en trouve cinquante qui auraient l'étoffe de bons officiers du roi, ce sera à moi de dire : ces dix-là ne valent pas ces quarante autres.

Je me suis fabriqué un petit carnet où j'ai inscrit le nom de chaque concurrent, et je compte, pour chacune des cinq disciplines où je dois vérifier leur instruction, en noter le degré par une lettre de l'alphabet. J'espère ainsi avoir à la fin une vision plus synthétique et plus juste de leurs

Le 25 octobre 1783, Gaspard Monge est nommé examinateur de la marine par le roi, sur proposition du maréchal de Castries, ministre de la Marine. Ses tournées l'amènent à Brest, Rochefort et Toulon, où se trouvent les écoles d'officiers de marine. En 1786, deux collèges navals sont créés, l'un à Vannes et l'autre à Alès, pour préparer aux écoles d'officiers. À partir de 1787, Monge doit alors se rendre dans ces deux collèges lors de ses tournées d'inspection, pour y choisir les jeunes gens qui pourront intégrer

César-Henri, comte de la Luzerne, ministre de la Marine de 1787 à 1790.



qualités respectives. C'est le professeur de Mathématiques de Vannes, Monsieur de la Farge, qui m'en a inspiré l'idée : l'année dernière, il m'avait fourni, avant l'examen, une liste classée et annotée de ses élèves, les meilleurs étant désignés par a, b, ou c, ceux qui lui paraissaient plus douteux, par d, et ceux qui n'avaient pas l'instruction nécessaire, par n. Je compte affiner ce système en utilisant davantage de lettres, et en notant chaque épreuve.

Vannes, 23 mars 1789

Je suis fatigué. Le ciel est bas. Un crachin silencieux mouille de tristesse les bâtiments encore inachevés de l'école de marine.

J'ai fait partir aujourd'hui pour Paris la liste des noms qu'on attend là-bas.

Pendant ces deux mois, mon esprit échauffé par le travail n'a pas douté de parvenir à un jugement sûr.

Mais à présent, je ne sais plus rien. Tant de questions me viennent, ai-je vraiment mené ma mission à bien ? Je croyais avoir trop de brillants sujets à départager, mais je n'en ai pas même trouvé quarante : en cherchant à trancher entre le quarantième et le quarante et unième, j'ai trouvé ces deux-là également médiocres, et même aussi deux autres que j'avais précédemment placés devant. J'ai classé les quatre *ex aequo*, en recommandant qu'on ne les garde pas. Il n'en reste donc finalement que trente-sept. La crainte d'en choisir trop m'a-t-elle rendu plus sévère ? Mon système de lettres y est-il pour quelque chose ? En parcourant mes carnets, je me rends compte que pas une fois je n'ai voulu donner de a... Mes questions de trigonométrie ont-elles été au-delà de ce qu'on peut attendre d'esprits encore si neufs ? La statique sera au programme de l'examen l'année prochaine. Trouverai-je encore moins de têtes

En 1791, les collèges de Vannes et Alès sont remplacés par douze écoles gratuites et publiques de mathématiques et d'hydrographie. C'est Monge qui est chargé de leur création, et du recrutement des élèves. Il sera lui-même ministre de la Marine d'août 1792 à avril 1793.

assez bien faites ? Je rêve à l'immense réserve d'intelligence à laquelle nous aurions accès en supprimant la règle qui exige des concurrents quatre quartiers de noblesse paternelle.

Je pars demain pour Lorient, où l'on instruit les troupes coloniales. Puis les élèves officiers de Brest m'attendent pour la fin d'avril. Cette tournée est la plus longue et la plus fatigante de

toutes. J'ai hâte de retrouver Paris, et ma chère famille, pour un été de repos...

Dijon, le 28 janvier 2002

Ces quelques lignes m'ont été inspirées par :

\*l'article de Dominique Julia : « Gaspard Monge, examinateur » (*Histoire de l'éducation*, n°46, mai 1990)

\*le fonds Gaspard Monge, aux Archives départementales de la Côte d'Or (Série 21F 13 à 20)

\*la biographie écrite par François Pairault : « Gaspard Monge, le fondateur de Polytechnique » (Tallandier, 2000)

Tous ces documents m'ont donné l'illusion d'avoir rencontré Gaspard dans une autre vie. C'est pourquoi j'ai eu l'audace de prendre la plume en son nom. Les détails historiques sont exacts. Quant à la psychologie que je lui attribue, il ne faut y voir que le reflet de mon imagination.

**Françoise Dixmier**

**Extrait du carnet de notes (authentique) de G.Monge (Vannes, mars 1789)**

	Arithmétique	Géométrie	Trigonométrie rectiligne	Trigonométrie Sphérique	Navigation
Huon de Kermadec (4 <sup>e</sup> )	c	b	b	c	c
Aubert de Trégomain (5 <sup>e</sup> )	b	b	c	c	c
de Mélient (7 <sup>e</sup> )	f	d	c	b	b
de Poix (9 <sup>e</sup> )	b	b	b	b	b
Artur de la Villarmois (10 <sup>e</sup> )	c	c	c	c	c
Majance de Camiran (11 <sup>e</sup> )	b	b	c	c	f
Le Roux de la Routière (12 <sup>e</sup> )	c	b	b	b	c
de Berset (13 <sup>e</sup> )	b	c	c	c	c
Sauvage (15 <sup>e</sup> )	c-d	c	c-d	c-d	c-d
de Carn (16 <sup>e</sup> )	c	c	c-d	d-e	d-e
d'Amphernet (17 <sup>e</sup> )	d	d	c-d	c-d	d-e
Lantivy de Trédion (18 <sup>e</sup> )	c	d	d	d-e	e-f
Achard de Bonvouloir (20 <sup>e</sup> )	d-e	d	d	e	d
Le Large (24 <sup>e</sup> )	e-f/2	e	f	f	g-h
de Vaufleury de la Durandière (25 <sup>e</sup> )	c	e	e	e	e
Augier (26 <sup>e</sup> )	d-e	d-e	d-e	d-e	e-f
Tardieu de Maleyssie (28 <sup>e</sup> )	c	d	d	d	f-g
de Verteuil (33 <sup>e</sup> )	e	e	e	e	e
Charette de la Colinière (34 <sup>e</sup> )	e	f	f	f	f
Garnier de Boisgrollier (36 <sup>e</sup> )	d	e	f	f	*
de Gilart de l'Archantel (38 <sup>e</sup> )	d	e	e	f-g	f-g
Huchet de Ceintré (38 <sup>e</sup> )	e	e	f	f	*
d'Aubenton (38 <sup>e</sup> )	f	f	f	f	g-h
Perrin de Cabriller de Laugary (38 <sup>e</sup> )	e	d	f	f	*

\*: « il ne la sait pas »

Gaspard Monge complétait ces notes par des remarques :

de Mélient : « Je l'avais mal jugé, il n'a que treize ans, et il sait fort bien lorsqu'il sait. Sa lenteur et sa timidité m'avaient trompé. »

Le Large : « Intelligent et propre aux sciences, mais trop peu instruit. »

d'Aubenton : « Intelligence bien faible, et cependant il n'est pas borné. »

Il mettait parfois une « note d'intelligence » : de Berset : c-d , Huchet de Ceintré : c , Perrin de Cabriller : d.

Il a réexaminé certains candidats une deuxième fois :

Huon de Kermadec : « sera examiné de nouveau sur les logarithmes ; assez intelligent mais lent. »

Sauvage : « Le 26, il a donné la statique jusqu'à la fin du 3<sup>e</sup> chapitre. »

Lantivy de Trédion : « Le 26, il a donné les mappemondes. »

## NOTES HISTORIQUES : PROSE MINISTÉRIELLE

### Arrêté du 5 juillet 1890 (extraits)

(Cet arrêté réforme l'enseignement secondaire, des programmes aux conditions d'enseignement. Le ministre qui le signe est Léon Bourgeois.)

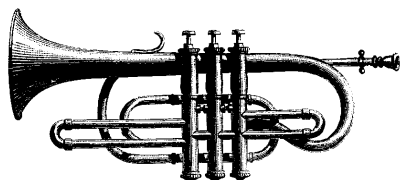
Eh bien voici, selon nous, des conditions qui peuvent faire de la note un moyen de discipline très efficace.

Tout d'abord, il faut qu'elle soit discrète, afin de garder sa valeur : aucune sanction ne peut se passer de cette qualité. On n'en trouvera pas qui, par elle-même, ait assez de vertu pour pouvoir être impunément appliquée à tort et à travers. Il faut donc renoncer à tout noter. Non ! avec les enfants, il ne faut pas tout noter, parce que tout n'est pas grave. Le premier avantage de la note, c'est que, devant être motivée dès qu'elle sort de l'ordinaire, elle ne pourra signaler que ce qui a de l'importance. Puis elle n'est ni immuable, ni irrévocable. Elle s'assouplit aux incidents de la journée. On peut sans bruit l'effacer, l'abaisser, la relever, admettre au bout de la journée, de la semaine, d'intelligentes compensations.

(...)

De toutes les manières dont un professeur consciencieux peut perdre son temps et sa peine, la plus évidente n'est-elle pas de passer des heures à relire des copies d'enfants en s'ingéniant à trouver des degrés où il n'y en a point, à mettre en balance, comme s'il s'agissait d'une affaire d'état (c'en est une en effet pour les rivaux et quelquefois pour leurs familles), des mérites qui souvent sont d'ordre différent, et par suite sans commune mesure ? Combien ce temps serait mieux employé en lectures et en travaux personnels par lesquels le professeur renouvellerait sa provision d'idées ; combien même il serait plus utilement donné au repos, source de bonne humeur et de fraîcheur d'esprit !

(...)



Cornet à pistons.

Article 3. Les seules punitions autorisées sont les suivantes :

- a. La mauvaise note ;
- b. La leçon à réapprendre en totalité ou en partie ;
- c. Le devoir à refaire en totalité ou en partie ;
- d. Le devoir extraordinaire ;
- e. La retenue du jeudi ou du dimanche ;
- f. La privation de sortie ;
- g. L'exclusion de la classe ou de l'étude ;
- h. L'exclusion temporaire ou définitive de l'établissement.

Article 16. Les notes des bulletins trimestriels ne seront arrêtées qu'après une réunion dans laquelle tous les maîtres à qui ont affaire les mêmes élèves auront échangé sur chacun d'eux leurs impressions. Ces notes exprimeront l'opinion propre de chaque maître, de façon à renseigner exactement les familles. Elles seront accompagnées d'une note générale rédigée par le proviseur.

Article 17. Les notes arrêtées à la fin de chaque trimestre, dans les conditions prévues à l'article 16 pour être communiquées aux familles, seront lues et commentées par le proviseur aux élèves réunis par classes dans la salle des actes. Ce compte rendu trimestriel devra être pour chaque élève la punition de ses défaillances ou la récompense de ses efforts. Il offrira au chef de la maison la meilleure occasion d'adresser à tous et à chacun les exhortations et les avis nécessaires.



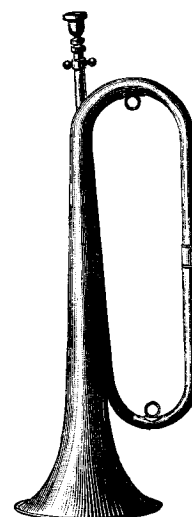
Article 21. Dans les compositions, chaque copie aura sa note chiffrée de 0 à 20.

L'attention des élèves sera appelée sur la note plus que sur la place.

Le classement linéaire comportera autant d'*ex æquo* que le professeur le jugera nécessaire.

#### **Circulaire de février 1942** (signée de Jérôme Carcopino)

Au moment où j'ai décidé d'accorder des bourses à tous les candidats reçus à l'examen sans qu'un pourcentage fixé auparavant vienne limiter arbitrairement le nombre des boursiers, mon intention très nette est d'exiger davantage de l'élève qui reçoit une aide pécuniaire de l'État que de celui qui, côtoyant tout juste la moyenne, réussit tout juste à passer dans la classe supérieure. J'entends que l'échelle des notes soit intégralement utilisée de 0 à 20 et non pas entre des limites plus resserrées particulières à chaque correcteur. C'est ainsi que la moyenne exigée de 12 points sur 20 prendra toute sa signification. Je ne veux, ni d'une indulgence excessive, fille de la crainte des responsabilités, ni d'une sévérité inadmissible ignorante des possibilités des candidats. J'entends que les notes se répartissent normalement et je ne veux plus avoir à constater, comme lors de l'examen du 4 décembre dernier, que, dans tel département, sur 80 candidats ayant subi l'examen, aucun d'eux n'a obtenu une moyenne inférieure à 9 et supérieure à 12. Les présidents des commissions d'examen s'assureront, avec le plus grand soin, que mes instructions sont exécutées et leur responsabilité sera engagée sur ce point. Je procéderai à des sondages en faisant revoir les corrections de certaines commissions. Ce contrôle établira si les notes données marquent bien les différences présentées par les diverses copies et il me sera rendu compte des vérifications ainsi effectuées.



Clairon

#### **Circulaire du 6 janvier 1969** (signée d'Edgar Faure)

(extrait)

Ainsi les trois notions essentielles de notre système d'appréciation des résultats scolaires, notions de composition, de note, de classement, doivent faire l'objet d'une triple révision, de trois réformes indissolublement liées, celle des procédés de notation servant d'instrument aux deux autres.

Les travaux scolaires les plus formateurs sont ceux où la préoccupation de la note s'efface : maître et élèves avancent ensemble dans la découverte d'un texte, d'un raisonnement, d'une expérience scientifique, d'une activité sportive, d'une donnée de géographie humaine etc., et ce n'est qu'à regret que le fil est interrompu pour permettre les contrôles cependant nécessaires. Une pédagogie véritablement active réussit d'ailleurs, sans difficultés, à inclure le contrôle dans le champ même de l'élaboration des connaissances. Le contrôle permet en effet au maître d'orienter de manière plus efficace les directions de son action. Sans doute l'élève a-t-il besoin de voir son travail apprécié, ses efforts motivés et sa progression jalonnée. Sans doute les parents comme les autorités scolaires ont-ils le besoin d'informations précises. Il faut cependant éluder l'obsession de la note, presque aussi pernicieuse que l'obsession de la « place », comme l'ont observé depuis longtemps bien des maîtres expérimentés.

À cet effet, il est bon d'abord de prendre conscience de la relativité de la note, et par suite d'écartier les procédés dont la précision apparente est trompeuse. La notation chiffrée de 0 à 20

peut être abandonnée sans regret. Une échelle convenue d'appréciation, libérée d'une minutie excessive, sera moins prétentieuse. En indiquant la zone dans laquelle l'élève se situe, on cerne déjà la réalité d'assez près, on évite de multiplier systématiquement les différences qui ne seraient pas confirmées par d'autres correcteurs, ni par le même correcteur à une autre époque. Des appréciations globales telles que « très satisfaisant », « satisfaisant », « moyen », « insuffisant », « très insuffisant », auxquelles on peut faire correspondre, si on le juge bon, les symboles A, B, C, D, E, ou 1, 2, 3, 4, 5, constituent donc un système non pas plus rudimentaire que le système traditionnel, mais plus rationnel et plus adapté aux données.

Il sera bien entendu utile à l'élève que cette appréciation globale s'accompagne d'annotations plus détaillées, (concernant par exemple l'orthographe, l'ordre, le vocabulaire, la syntaxe, la précision, l'habileté, les facultés de raisonnement, l'invention, le sens artistique, etc.)

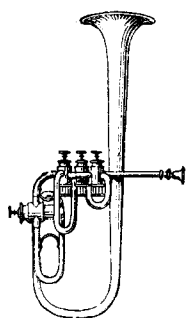
Ces modalités techniques de notation –pour lesquelles il n'existe pas de formule unique qu'il conviendrait d'imposer– ne doivent pas faire oublier la préoccupation essentielle : ramener la note à son rôle utile sans avoir à le payer par trop d'inconvénients. Il faut en particulier éviter de la « dramatiser ». Or, lorsque les notes attribuées par chaque maître à des exercices isolés au cours du trimestre ou du semestre –les compositions traditionnelles– figurent distinctement sur le bulletin de fin de trimestre ou d'année, la dramatisation est inévitable.

Aussi bien se préoccupe-t-on depuis longtemps des difficultés nécessairement rencontrées par les conseils de classe si le sort de l'élève doit être lié à un résultat aussi aléatoire que la « moyenne générale » des notes de « composition ». Les maîtres savent bien qu'il n'est pas raisonnable de prétendre apprécier le travail de toute une année en se fondant sur les notes de trois exercices écrits seulement.

L'article 3 de l'arrêté du 30 décembre 1964 (qui concerne les élèves de première) a cherché à remédier à l'insuffisance de la « moyenne générale » en invitant les conseils de classe à tenir compte « des autres éléments d'appréciation concernant le travail, les aptitudes et le comportement général de l'élève en classe de première, ainsi que l'évolution de sa scolarité depuis son entrée en seconde ». Une circulaire du 5 mars 1965 devait ensuite insister sur cette idée, une autre, du 6 juin 1967, préciser fermement « qu'en aucun cas le sort d'un élève ne doit dépendre de la note globale obtenue en faisant la moyenne des notes obtenues aux compositions ». Ce principe, posé pour les élèves de première, est valable pour tous.

Dès maintenant, il est recommandé aux chefs d'établissement et aux enseignants, professeurs et instituteurs :

- 1) de substituer à la notion de composition traditionnelle celle d'exercices de contrôle divers, faits en classe, en un temps limité et présentant les caractères ci-dessus décrits ;
- 2) de substituer à l'échelle de notation traditionnelle de 0 à 20 une échelle simplifiée d'appréciation globale du type ci-dessus défini, ou d'un type analogue ;
- 3) d'exclure en général les classements par rang, établis et annoncés par le maître.



Cornet à pistons perfectionné par Sax.

## SOUVENIRS, SOUVENIRS

LISANT la circulaire d'Edgar Faure, sur l'impropriété d'une notation inutilement (et, illusoirement) précise et l'opportunité des classements, il me revient que, lorsque j'étais en 11<sup>e</sup>, au Lycée Henry IV, il y a bien longtemps (en ce temps-là le lycée couvrait tout le primaire et le secondaire : du cours préparatoire aux classes préparatoires), à l'issue de la première composition, la maîtresse nous a fait mettre en rang par ordre de classement. Elle a appelé le ou la première, qui se tenait donc devant la classe encore assise ; puis le ou la seconde, qui s'est placée à côté. Elle les a fait tourner d'un quart de tour, puis a appelé le ou la troisième, qui s'est placée derrière les deux précédents ; et ainsi de suite jusqu'au dernier. Matérialisant de la sorte pour nous le classement.

J'entends déjà les cris d'orfraie. Mais, j'atteste que cette mise en scène n'a traumatisé apparemment personne à l'époque. Plus tard, au long de notre scolarité, nous avons été répétitivement notés et classés. Autant que je m'en souviens cette matérialisation n'a eu lieu qu'une fois.

Je me vois ensuite, pas mal d'années après, assistant à une réunion de parents d'élèves à L'École Alsacienne (alors à la pointe de l'innovation pédagogique : c'était l'expérimentation notamment des « maths modernes » dans trois établissements : L'E.A., Vitruve et un troisième). La sous-directrice, Mlle Æschmann nous expliquait que dorénavant, on ne noterait plus. Sur trente élèves, par exemple, le groupe central des 12 moyens serait noté "C" les 6 bons et les 6 passables, de part et d'autre, respectivement "B" et "D" enfin les 3 de tête "A" et les 3 de queue "E". Ainsi, ceux-ci recevaient un sérieux avertissement, mais aucun n'était stigmatisé comme étant LE dernier. Et, elle nous disait « vous savez, à cet âge-là, être le dernier, c'est dur ! ». (Symétriquement, côté tête, on ne distinguait pas entre l'or, l'argent et le bronze, de sorte que personne ne pouvait s'enorgueillir d'être le meilleur.)

Quelques années plus tard encore, je vois mon petit-fils très attentif à ses notes. Il nous les raconte (surtout les meilleures) et recalcule régulièrement sa moyenne. Comme il n'est ni petit génie ni cancre, il manie ça dans une calme normalité, sans gloriole ni angoisse.

Alors, je me dis qu'être noté et classé doit répondre à une normalité sociale, un usage accepté. Un usage qui, soit répond à un besoin, soit s'actualise en un besoin. Un peu comme de se vêtir pour sortir dans la rue. Sans doute, comme tout usage et comme toute norme, il connaît ses excès et ses perversions. Faut-il, pour les prévenir, supprimer la chose en totalité ?

Je me dis aussi que l'effet négatif ou pervers peut résulter d'une application trop abondante ou systématique. La mise en rang de ma 11<sup>e</sup> m'a frappé : puisque je m'en souviens soixante ans plus tard. Mais je ne crois pas que cette mise en scène se soit reproduite, une seconde fois. Unique, elle a pu avoir sa vertu pédagogique tout en évitant d'indurer la procédure et ses répercussions.

Je me dis enfin que les élèves - et, par delà l'expérience scolaire, les hommes - sont hétérogènes. Il n'y a pas de raisons pour qu'une façon de faire soit appropriée pour tous, ni funeste pour tous. Il doit y avoir des sujets pour qui être noté (et, classé) est bénéfique ;

d'autres, pour qui c'est néfaste. Pourquoi priver les uns de ce bénéfice au motif que ce serait traumatisant pour les autres ? Pourquoi imposer à ces derniers ce maléfice, au motif que les premiers en profitent ? Faut-il donc imposer un système universel ? ou, le prohiber pour tous ? On construirait alors ce tableau à 4 cases :

<i>ceux qui réagissent</i>	<i>bonne note</i>	<i>mauvaise note</i>
<i>positivement</i>	se sentent gratifiés et persistent	sont vexés et veulent se rattraper
<i>négativement</i>	s'endorment sur leurs lauriers	sont découragés et abandonnent

Évidemment, cette vision des choses laisse deux difficultés à résoudre :

\* il n'y a sans doute pas deux catégories (les positifs et les négatifs) mais davantage. En particulier, il doit y avoir beaucoup de personnes qui sont sur la diagonale de ce tableau : flattées par leur succès, elles s'améliorent là où elles ont réussi *et* convaincues par leur échec qu'elles « ne sont pas bonnes en maths/langues/gymnastique/... » font tout pour confirmer cette anti-vocation ;

\* une fois dressée une typologie acceptable, comment distinguer les individus selon leur type de réaction, de façon à leur appliquer ou à ne pas leur appliquer tel système de notation ? Et, cela étant supposé fait, comment faire coexister des systèmes différents, avec des gens qu'on note et d'autres qu'on « traite » autrement ?

**René Padieu**



## NOTES EN PHILO OU NOTES EN FOLIE ?

QUI A NOTÉ LE PREMIER ? Quand et où ? Comment l'idée lui en est-elle venue ? Faute d'une histoire encore à écrire, les mots ont-ils gardé la mémoire de la première note ?

### Le palmarès des dictionnaires

Partons donc du Petit Robert comme le propose D. Reisz. « L'appréciation chiffrée selon un barème préalablement choisi » vient en tout dernier lieu dans la liste des différentes significations de *note*. Et cela ne nous éclairerait pas beaucoup si nous n'avions pas été notés dès le premier âge, appris ce qu'est la note en étant noté, comme à compter en comptant (et en étant noté dans le même temps).

Revenons au début de la notice. *Note*, n.f ; (XII<sup>e</sup> ; lat. *nota*). Bon alors, c'est du latin de cuisine...

Passage au Gaffiot. *Nota*, æ, f. 1) signe, marque. Un signe matériel donc, mais vite avec un sens figuré péjoratif, *notæ ac vestigia flagitiorum* (Cicéron) : marques et traces d'infamie. Puis viennent des significations tournant toujours plus ou moins autour de la notion de signe, de la ponctuation des copistes aux signes musicaux. 3) marque sur le corps, tache naturelle, tatouage, au propre (sic) d'où, au figuré, encore flétrissure, honte : *quæ nota domesticæ turpitudinis non inusta vitæ tuæ est ?* (Cicéron) Est-il une flétrissure provenant de scandales domestiques qui ne soit imprimée sur ta vie ? 5) étiquette (mise sur les amphores pour rappeler l'année du vin, le bon cru), d'où marque, sorte, qualité : *mel secundæ notæ* (Columelle) miel de seconde qualité ; *de meliore nota commandare aliquem* (Curius) recommander quelqu'un de la meilleure sorte. Et puis vient le 6) annotation, remarque. Ce serait donc notre note de bas de page ? Pas tout à fait : *notam apponere ad malum versum* (Cicéron) : noter un mauvais vers. Pas étonnant de trouver alors 7) note du censeur, blâme (motivé, inscrit à côté du nom) : *ensoriæ severitatis nota inuri* (Cicéron), « par jeu de mots avec le sens du § 3 » précise Gaffiot pour les jeunes latinistes qui n'auraient pas encore fait l'amère expérience d'un zéro pointé, être marqué d'infamie par la rigueur des censeurs. D'où sans doute l'expression « une vraie tache, ce mec... ».

Donc finalement, les romains notaient à tour de bras.

Mais ce *nota* de Cicéron est-il le bon ? Retour chez Monsieur Alain Rey. La *note* comme appréciation chiffrée arrive après les deux significations principales à en croire la hiérarchie classificatoire du Petit Robert. Le I, c'est la note de musique. Le II, c'est la note écrite en marge d'un document ou pour garder la trace d'un discours ou d'un événement. Deux sens éloignés en apparence. Et hiérarchiquement toujours, notre note (celle du professeur) vient dans le II, après par exemple la *note* comme glose ou scolie (avec cette citation d'Alain que tous les élèves mal notés du pays devraient avoir en tête : « la note, c'est le médiocre qui s'attache au beau »), et, enfin, juste après la note d'hôtel, détail écrit d'un compte, mais qui au restaurant laisse place à

26205



Or mat  
2 roses, 1 saphir  
Fr. 23.25

l'addition. Une autre *note* chiffrée après la *nota* des amphores, curieux cette proximité de la note chiffrée et de la consommation de boissons diverses...

Le Petit Robert date la *note* comme « brève appréciation donnée par un supérieur sur le travail, la conduite de quelqu'un » de 1845. Notre groupe de recherches historiques et pénombresques sur la notation semble avoir trouvé des mentions plus anciennes de cette note (de gueule). La *note* comme élément textuel (signe de la médiocrité pour Alain) remonterait quant à elle à 1630 et l'usage, vieilli, de la *note* comme marque pour se souvenir de l'emplacement de quelque chose serait attesté en 1530. Bon, mais il n'y avait plus trop de romains à cette époque...



La *note* musicale est d'abord un signe marquant la hauteur et la durée du son, puis le son lui-même. Le PR rattache à ce sens, l'usage figuré de *note* : une note juste, une note gaie, forcer la note, être dans la note... Enfin, des notes moins stigmatisantes...

Encore un détour par la *notation* avant de donner une deux ou trois étoiles au Petit Robert. À la *note* donnée par le supérieur hiérarchique correspond la *notation* comme « action de noter ». L'exemple est donné bien sûr de la notation des devoirs par le professeur. Nulle référence à ce point cependant à la correction des devoirs (présupposé de la note ?) qui s'apparente plus à l'action de ramener à la règle comme dans la « correction des mineurs délinquants » ou la « correction typographique » des erreurs d'imprimerie. Pour le PR, la notation comme action de donner une note ne paraît être finalement qu'une signification sans lien avec le sens principal : la *notation* est l'action qui met en œuvre des symboles pour représenter. La première *notation* est alors celle des nombres, la notation numérique, puis vient la notation chimique... Le Petit Robert ne s'étend pas, la chimie, ce n'est pas vraiment son affaire. Il préfère ensuite nous parler de la notation musicale. Et surtout de la notation en peinture, « des notations vivement colorées », et cette citation de Francastel, l'impressionnisme « substitue au dessin classique la notation des ombres et des reflets », voilà qui devrait pouvoir resservir dans un conseil de classe ou sur un livret scolaire !

Finalement, le PR ne donne guère d'indication sur le lien entre la note comme signe ou comme petit texte et la note (chiffrée ou non) de nos professeurs. La *nota* latine semble plus proche des pratiques « pédagogiques » de la notation. Faut-il remonter si loin pour comprendre ?

### **La note portugaise**

Un coup d'œil rapide sur le Nouveau Larousse Illustré, sixième volume, édition non datée mais probablement du début du XX<sup>e</sup>, à lire les biographies et chronologies qu'il contient. L'entrée *notation* comprend un long développement sur la notation chimique (mais il ne s'agit pas encore de la notation atomique) et un autre sur la notation musicale. En fin de rubrique, *notation chiffrée* renvoie à musique chiffrée. Une utilisation oubliée peut-être quand même : dans le commerce, « la notation est la manière adoptée par les marchands pour indiquer sur leurs marchandises des prix de vente dont ils veulent dérober la connaissance à leurs clients ». Les chiffres sont alors transformés en



lettres. On n'a donc rien inventé en passant de la note chiffrée à la note avec des lettres ? Recette qui peut servir encore : « un marchand a choisi le mot *portugaise*, dont les lettres figureront pour lui les chiffres 1,2,3,4,5,6,7,8,9,0 ; s'il veut indiquer qu'il veut vendre une pièce de drap 14 fr. 85 c. le mètre, il marquera sur cette pièce les lettres *ptiu* ». Sur une copie, au lieu de 12,5 (sur 20), on pourra mettre alors *pou*. Ingénieux, n'est-ce pas ?

La rubrique *note* (lat. *nota*) nous déçoit. Déjà, pas la moindre trace de la note chiffrée du professeur. Peut-être cela vient-t-il de l'orientation encyclopédique du Larousse. Dans le désordre, sont évoqués les usages du mot *note* dans les diverses disciplines et l'on s'y perd un peu. La *note* est donc une marque pour se rappeler, un exposé succinct, la mention particulière d'une personne distinguée (Vx.), *personnage de note* ou personnage de marque. Vient alors l'usage moderne de ce sens : « observation concise par laquelle on apprécie la conduite ou le caractère de quelqu'un, particulièrement d'un militaire ou d'un élève : *avoir de bonnes notes*. » Donc les notes peuvent être bonnes sans être chiffrées, elles apprécient autant le caractère que la conduite.



Il n'est pas dit si ces notes peuvent être communiquées sous forme de *notes tironiennes*, selon une écriture abrégée qui « permettait d'écrire aussi vite que la parole » et surtout secrète pour authentifier les actes de façon sûre. Ce serait alors une note chiffrée pour éviter qu'elle ne soit trop facilement comprise... Signification que l'on retrouvait anciennement chez les notaires pour désigner les abréviations qu'ils utilisaient pour rédiger leurs minutes, et par extension leurs minutes elles-mêmes. D'où cette ancienne appellation du notaire, le *garde-note*. L'appellation étant tombée en désuétude, on pourrait la proposer pour d'autres utilisateurs intensifs de notes...

Illustré le Larousse l'est, et la note par la décomposition figurée des notes musicales. Une ronde vaut deux blanches ou 4 noires ou 8 croches ou 16 doubl. cr. ou 32 tripl. cr. ou 64 quadrup. croches. La ronde est en haut et les quadruples croches, en bas. La note musicale donne lieu à des utilisations dérivées comme « attaquer la note » ou « attaquer la note vivement » ou figurées, comme dans cette citation de Lamartine :

« La nature a deux chants, de bonheur, de tristesse ;  
de l'une à l'autre note elle passe sans cesse. »

On retrouve quand même au passage certaines utilisations de *note* ou *noter* qui peuvent bien provenir de la *nota* décrite par le Gaffiot. Même si cela est considéré comme un usage ancien : une *note infamante* ou d'infamie qui en droit désignerait la déclaration légale et juridique qui frappe une personne et l'attaque dans son caractère ou sa conduite. Et puis une personne *notée*, qui serait une personne mise en suspicion et dont on a reconnu qu'il fallait se méfier.

### **Le chaînon manquant**

Toutes ces définitions ignorent un usage savant plus ancien de la note, mais sans remonter à Cicéron pour autant. On le trouve mentionné dans les dictionnaires philosophiques. Celui des *Notions philosophiques* (PUF) comporte une entrée *nota* (un peu recopiée du vocabulaire de

Lalande en fait) à laquelle renvoie directement l'entrée *note* et qui sera citée ici assez longuement de peur d'en trahir l'aspect très pédagogique :

« NOTA

« Subst. lat. signifiant note ou marque.

« Dans la philosophie scolastique, attribut d'un sujet caractère d'un concept. De là vient la formulation en compréhension du *dictum de omni* : la note de la note est la note de la chose même, c'est-à-dire le prédicat du prédicat est le prédicat du sujet<sup>1</sup>. Ce caractère peut également

1. En V.O. : *Nota notæ est nota rei ipsius = prædicatum prædicati est prædicatum subjecti.*

note dans la note : "Dans toute énonciation où l'on peut distinguer ce dont on parle et ce que l'on en affirme ou nie, le premier terme est appelé sujet et le second prédicat" (Lalande). Espérons que la note scolaire soit attribuée en faisant

être conçu dans sa fonction de reconnaissance et de distinction de l'objet ; pour Leibniz, une *notion distincte* correspond aux choses dont nous avons une *définition nominale*, laquelle n'est rien d'autre qu'une énumération des marques suffisantes à leur reconnaissance (par ex. la notion que les essayeurs ont de l'or). Mais pour assurer cette fonction, il n'est pas nécessaire que la note ait un rapport intrinsèque et naturel à son objet (je souligne). C'est ainsi que pour Hobbes « une marque est un objet sensible qu'un homme érige pour lui-même volontairement, afin de s'en

servir pour se rappeler un fait passé, lorsque cet objet se présentera de nouveau à ses sens ». Elle correspond à la nature du langage dans son usage intérieur : « Le premier usage des dénominations est de servir de marques ou de notes en vue de la réminiscence ».

### ***Si la faculté de noter n'existait pas, il faudrait l'inventer***

Donc la note, est bien depuis longtemps, depuis les romains jusqu'à nos jours, l'expression d'un jugement sur le caractère des choses. Noter relève de la faculté de juger<sup>2</sup>. Nous pouvons

2. E. Kant, Critique de la raison pure, livre II. Introduction. Du jugement transcendantal en général.

« Si l'on définit l'entendement en général le pouvoir des règles, le jugement sera le pouvoir de subsumer sous des règles, c'est-à-dire de décider si une chose est ou n'est pas soumise à une règle donnée. La logique générale ne renferme aucun précepte pour le jugement et n'en peut pas renfermer...

« ...Aussi le jugement est-il la marque spécifique de ce qu'on nomme le bon sens et au manque de quoi aucun enseignement ne peut suppléer ».

L'auteur ajoute peu après cette note : « Le manque de jugement est proprement ce que l'on appelle stupidité, et à ce vice il n'y a pas de remède. Une tête obtuse ou bornée en laquelle il ne manque que le degré d'entendement convenable et de concepts qui lui soit propres, peut fort bien arriver par l'instruction jusqu'à l'érudition. Mais, comme alors, le plus souvent, ce défaut accompagne aussi l'autre, il n'est pas rare de trouver des hommes très instruits qui laissent incessamment apercevoir dans l'usage qu'ils font de leur science ce vice irrémédiable ».

« noter » comme nous pouvons juger. Y compris d'ailleurs par des jugements quantitatifs, la grandeur étant l'un des caractères (une catégorie) que nous savons attribuer au sujet.

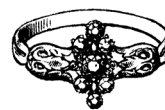
Notre faculté de noter n'a pas besoin pour s'exercer qu'une grandeur ait été définie au préalable et que la note résulte d'une opération de mesure de cette grandeur. Nous savons noter (d'infamie ou de dignité, de bonheur ou de tristesse, de bon cru ou de médiocre production) les objets soumis à notre jugement.

L'appel croissant fait à des nombres pour exprimer un jugement ne peut donc se comprendre que par notre maîtrise (*a priori*) d'un « instrument » de quantification sans grandeur définie au préalable. Une capacité psychologique à noter et même à noter de chiffres. Innée ou acquise ? Les neurobiologistes sont à la recherche du gène du nombre. Certains pensent l'avoir trouvé, les animaux possèderaient les rudiments



nécessaires à l'évaluation chiffrée comme le petit d'homme qui sur cette base innée pourra apprendre à compter. Mais peu importe, adultes nous sommes supposés savoir dire notre opinion sur les gens, les choses, les événements et ceci de plus en plus précisément. Si l'on nous demande « d'affecter une note » à une chose ou une personne, nous mettons notre tuttimètre en route, et la réponse vient. Au sondé, on ne demande en général que de dire s'il est tout à fait favorable, plutôt favorable, plutôt en désaccord ou tout à fait en désaccord avec une opinion stimulus. Arrive un moment où on lui demande de mettre une note et ceci ne fait qu'imposer une forme particulière à l'expression de ses préférences. Et dans un premier temps, ce n'est pas trop compliqué. Ayant mis une certaine note pour une chaîne de télévision ou un responsable politique, la préférence des suivants sera marquée en faisant varier la note.

28308



Ciselé or mat  
6 roses, 1 rubis  
Fr. 46.50

### **Notes colorées, notes claires ou obscures, blanches ou noires**

Car noter avec des chiffres, ce n'est pas autre chose que l'expression d'un système de préférence. Les propriétés de cet instrument propre à noter toutes choses sont-elles bien connues ? Dans le domaine de la consommation de biens matériels, les économistes pensent normal d'attribuer à notre système de préférence et à la formation de nos choix des propriétés formelles stables et permettant une mathématisation sans trop de difficultés. D'où les bonnes propriétés d'ordre et de mesurabilité qu'auraient les préférences du consommateur.

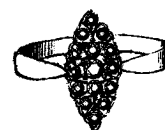
28231



10 roses  
Fr. 47.50

Si je préfère avoir deux pommes plutôt qu'une, et trois plutôt que deux, alors je suis supposé préférer avoir trois pommes plutôt qu'une. Admettons. Mais cette transitivité est-elle toujours assurée ? Que se passe-t-il si l'on me demande quelle est ma couleur préférée ? Si je dis préférer le rouge au jaune, et le jaune au bleu, pourquoi devrais-je préférer le rouge au bleu ? Mettez donc les couleurs sur un cercle et voyez ce qu'il en est de cette « relation d'ordre » ! Je peux tout à fait préférer le bleu au rouge et de nouveau le rouge au jaune en parcourant le cercle... Ou alors il faut avoir en tête la disposition des feux tricolores pour transformer trois couleurs en système d'expression de préférences garantissant le bon ordre, par exemple pour évaluer des élèves en cours primaire : vert, c'est bien tu peux continuer, orange c'est moyen, il faut faire une pause, rouge, c'est mauvais, tu t'arrêtes et tu ne repars pas. Dans ce cas, vert est mieux que orange, orange mieux que rouge, et vert est bien mieux que rouge.

28009



Platine pavé roses  
Fr. 59.

Cet usage scolaire des couleurs comme signe d'évaluation (que mes enfants ont expérimenté) n'est possible qu'après « rectification » du cercle des couleurs. L'opposition du bien et du mauvais le permet et l'impose. En allant vers le bien (vert) et au-delà, on ne peut revenir au mal (rouge).

28635



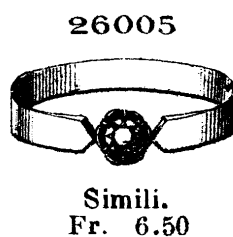
Patiné  
1 rose et 2 saphirs  
Fr. 60.

On peut alors construire toutes les échelles que l'on veut (échelle lettre comme le marchand de tissu cité par mon Larousse ou note chiffrée) cela ne résulte pas d'une propriété objective de ce qui est noté, mais de la mise en œuvre d'un système de préférence manichéen.

Dans le monde des couleurs, dont la notation n'est pas absente selon l'histoire des mots, ce système cohabite avec le « cercle » des teintes : c'est l'opposition du clair et de l'obscur.

L'obscur ne peut rejoindre le lumineux (au moins pour des occidentaux). Dans le monde des notateurs, le bien ne peut être confondu avec le mal. La note moderne ne commence-t-elle pas avec les représentations (les marques) mettant le paradis en haut, là où règne la lumière, et l'enfer en bas, monde des ténèbres ?

**Bruno Aubusson de Cavarlay**



### Illustrations

Pages 2, 4, 5, 23, 24, 26 à 28, 32 à 34, 36 :

À partir des gravures de *l'Histoire de la musique depuis les temps anciens jusqu'à nos jours* par Félix Clément, Paris, Hachette, 1885.

Page 9 :

À partir de *Info PC*, n° 167, février 2000.

Page 18 :

À partir de *Que Choisir*, N°389 bis/6, janvier 2002.

Page 29 :

Portrait de Gaspard Monge.

Carnet de notes de Gaspard Monge (octobre 1791), A.D. Côte d'Or, cliché Françoise Dixmier.

Pages 37 à 39, 41, 42 :

À partir du catalogue de la Bijouterie Savard et C<sup>ie</sup> à Paris (sans date).

## Table des matières

Ce volume contient :

Inscrivez votre note	Présentation, par Anne Le Bihan	page 1
A ! B ! C !	Première sortie, par Marcel Proust	page 2
A ! B ! C !	Vraie fausse note, par Françoise Dixmier	page 3
A ! B ! C !	Noter et/ou classer, par Daniel Reisz	page 7
A ! B ! C !	Noter pour évaluer, par René Padieu	page 23
A ! B ! C !	Le carnet de notes imaginaires de Gaspard Monge	page 29
A ! B ! C !	Notes historiques, par divers ministres	page 32
A ! B ! C !	Souvenirs, souvenirs, par René Padieu	page 35
A ! B ! C !	Notes en philo ou notes en folie ? par Bruno A. de Cavarlay	page 37
A ! B ! C !		
Analyse de votre score page		

## Le résultat de votre test

Votre score	Votre profil
9 A	C'est bien. Continuez à progresser. Si vous n'avez pas encore écrit de texte pour Pénombre, il faudrait y songer...
De 5 à 8 A et moins de 3 C	Vous êtes sans doute à jour de votre cotisation. Sinon, ne tardez pas trop.
Plus de 7 B	Êtes-vous sûr(e) d'avoir lu tous les textes attentivement ?
Moins de 3 A et de 5 à 8 C	Si vous n'aimez ni les maths, ni la musique, ni la littérature, peut-être nos pages historiques vous auront-elles plu !
9 C	Définitivement marqué(e) par les notes que vous avez subies ou qui vous ont sur-valorisé(e), vous avez des difficultés à apprécier Pénombre. Avez-vous essayé Péhaimécy-roi, la Lettre grise sur l'usage du nombre en santé mentale ?

Ce numéro de la Lettre grise est le fruit des rencontres régulières d'un groupe de travail de Pénombre dit « relation d'ordre » ou RDO entre 1999 et 2002. Ce groupe a été animé par Karin van Effenterre.

© Pénombre 2002.