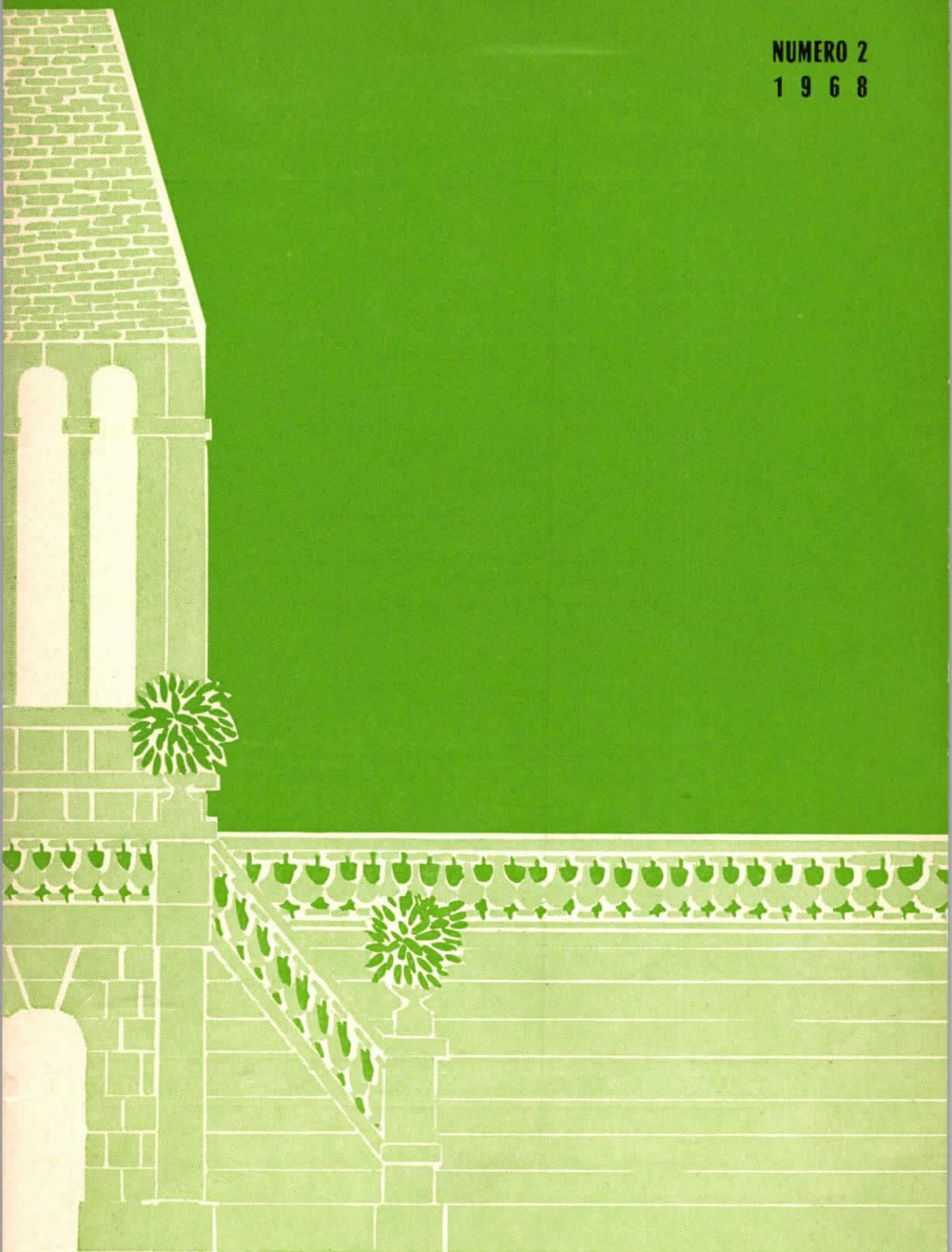


NUMERO 2
1 9 6 8



LES AMIS DE SÈVRES

ET LES CHEVAUX TREMPAIENT LEUR COU DANS L'AVENIR
POUR DEMEURER VIVANTS ET TOUJOURS AVANCER...

————— JULES SUPERVIELLE —————

ASSOCIATION DES AMIS DE SÈVRES

CENTRE INTERNATIONAL D'ÉTUDES PÉDAGOGIQUES

BUREAU DE L'ASSOCIATION

PRESIDENTE - FONDATRICE D'HONNEUR :

Edmée HATINGUAIS

Inspectrice Générale honoraire de l'Instruction
Publique.

PRESIDENT : **Jean AUBA**

Inspecteur général de l'Instruction Publique
Directeur du C.I.E.P.

VICE-PRESIDENT : **Jacques QUIGNARD**, proviseur
directeur des Etudes Pédagogiques au C.I.E.P.

SECRETAIRES : **Paule ARMIER**, professeur de
Lettres.

Marcel HIGNETTE, agrégé de
l'Université.

TRESORIERE : **Madeleine LANSON**.

TRESORIERE ADJOINTE : **Renée LESCALIE**.

1, AVENUE LÉON-JOURNAULT - 92 - SÈVRES - TÉL. 027.08.00

MEMBRES BIENFAITEURS 20 F MEMBRES ADHÉRENTS 10 F C.C.P. PARIS 6959-99 LES AMIS DE SÈVRES

DOCIMOLOGIE

Sommaire

- 3 EDITORIAL
- 5 **PRISE DE CONTACT**, par Marcel SIRE, Inspecteur Général de l'Instruction Publique.
I. PRISE DE CONSCIENCE.
- 14 II. NOTES ET MOYENNES.
- 19 III. MESURER, REPERER, NOTER.
- 23 **LES TESTS DE CONNAISSANCE**, par Jacqueline PELNARD-CONSIDERE, chargée de recherches à l'INOP.
- 33 **LA DOCIMOLOGIE, effort d'explicitation**, par Maurice REUCLIN, Directeur de l'INOP.
- 41 **BACCALAUREAT ET DOCIMOLOGIE**, par Jean PASQUIER, Directeur de l'Office du baccalauréat.
- 47 **COMMENT AMELIORER LA NOTATION**, par Louis LEGRAND, Chef du département de la recherche pédagogique à l'I.P.N.
- 53 BIBLIOGRAPHIE
- 55 THEMES D'EXPERIENCES
- 57 LA VIE DE SEVRES

N° 2
LES AMIS DE SÈVRES

DOCIMOLOGIE

« **D**ocimologie »... ce mot barbare a de quoi faire frémir nos lecteurs : recouvre-t-il une science mystérieuse, une technique pour initiés ? Ils seront peut-être déçus de constater qu'il s'agit tout simplement d'un art qu'ils pratiquent tous les jours, l'art de noter.

Pourtant ce numéro n'est pas un petit manuel de notation, c'est un cri d'alarme ! Dès l'ouverture du premier stage de Sèvres consacré à la docimologie en mai 1967, M. l'inspecteur général SIRE créa un choc en tirant, selon ses propres termes, « un obus explosif... »

La prise de conscience fut saisissante et brutale : ainsi nous commettions, en toute bonne foi, des erreurs aussi lourdes ! Les preuves qu'en apportait M. SIRE étaient irréfutables : le même devoir recevait de différents correcteurs des notes différentes, et de très loin ; bien plus, le même professeur, notant le même devoir à quelques mois d'intervalle, se déjugait totalement... Et nous avons, presque tous, bonne conscience ! Chacun de nous se rapporte inconsciemment à un système de références tout personnel qu'il a le tort de croire universel. En toute honnêteté, des gens parfaitement informés, de bons pédagogues, admettent implicitement qu'ils sont infaillibles.

Certains de nos collègues ont bien voulu se prêter à des expériences contrôlées. M. SIRE en présente ici le bilan en trois articles extraits d'un ensemble plus important :

— le premier, par son caractère expérimental, pose le problème de la notation traditionnelle, son titre est significatif : « Prise de conscience » ; il peut suggérer aux chefs d'établissements et aux Inspecteurs de conduire des expériences semblables dans leur établissement ou dans leurs circonscriptions et départements afin de provoquer chez les notateurs l'inquiétude lucide que nous souhaitons ;

— le second fait le procès des notations traditionnelles et de tout ce qui s'y rattache, en particulier, les moyennes générales ;

— le troisième demande au lecteur de réfléchir sur trois opérations qui conduisent à des nombres : mesurer, repérer, noter.

Cette prise de contact assez rude avec la réalité du problème remet

en question la sacro-sainte note chiffrée. Il reste à trouver des formules neuves mettant en jeu des critères nouveaux de jugement et d'appréciation des résultats :

Les tests de connaissance? Mme PELNARD-CONSIDERE fait pour nous le tour du problème. Par leur fidélité et leur validité, sinon par leur valeur prédictive, ils méritent que les pédagogues s'y intéressent et collaborent avec les psychologues dans le cadre d'une orientation mieux adaptée à l'évolution actuelle.

La science de la docimologie a progressé en cherchant à expliciter ses propres démarches. Cet effort d'explicitation que décrit M. REUCHLIN, permet d'améliorer dès maintenant les procédures d'appréciation. Mais il amène aussi à définir plus étroitement les objectifs assignés à l'enseignement et à poser la question fondamentale : quels sont les buts de l'éducation?

Problème particulièrement préoccupant si l'on songe aux examens. M. PASQUIER, sous le titre Baccalauréat et docimologie touche un certain nombre de points sensibles. Un effort a été accompli pour l'information des professeurs, pour l'homogénéisation des notes par l'intermédiaire des présidents de jury, pour la prise en considération du dossier scolaire, pour la définition du niveau requis. Le baccalauréat perdant son caractère redoutable de « verdict » peut devenir un instrument de mesure et d'orientation.

Dans une synthèse positive, M. LEGRAND se demande comment améliorer la notation? Des systèmes, comme celui des questions à choix multiple, permettent de « désobjectiver » la notation, et par suite la comparaison et le classement. Mais ne risquent-ils pas d'infléchir le contenu même de notre enseignement et de privilégier l'encyclopédisme au détriment de l'activité intellectuelle d'organisation et de synthèse? La notation en catégorie, et surtout le dossier scolaire, permettent d'exclure le verdict brutal et aveugle de la note et de quantifier les aptitudes de façon complexe et nuancée. Chacun des procédés a sa valeur et ils ne devraient pas s'exclure, mais trouver dans le dossier scolaire un point de rencontre et de synthèse.

Ainsi, « l'examen de conscience » auquel la docimologie nous convie ne doit pas nous induire à raffiner sur la précision des notes. « Notre pays est atteint d'un véritable délire de notation, disait récemment le Professeur LICHNEROVITZ au Colloque d'AMIENS, et nous pouvons nous demander si, en tant que maîtres, nous sommes des formateurs ou des notateurs. » La forme actuelle de nos examens nous fait un devoir de savoir noter, et les apports de la docimologie nous aident à le faire le mieux possible. Mais nous pensons que d'autres voies doivent être trouvées, qui pourraient bien être celles que les Lycées-pilotes ont depuis longtemps explorées. Au moment où la Commission de Renovation Pédagogique intègre nos expériences dans un vaste mouvement de recherche, nous pouvons espérer la diffusion de méthodes d'appréciation qui soient cohérentes avec les fins mêmes de notre enseignement. Ennemis de la sélection brutale qui stérilise les efforts, nous estimons que toute pédagogie véritable est une pédagogie de l'encouragement. Comme le dit encore le professeur LICHNEROVITZ, « Un maître doit être un éducateur bien plutôt qu'un juge... on ne dialogue pas avec un juge ».

Jean Auba

PRISE DE CONTACT

I. prise de conscience

Le 25 mai 1967 soixante-seize chefs d'établissement et professeurs de l'enseignement de second degré se réunissaient à Sèvres, au Centre International d'études pédagogiques, à l'occasion d'un stage consacré à la docimologie... le premier dans l'histoire de Sèvres.

Le groupe comprenait trente-quatre chefs d'établissement et quarante-deux professeurs de Lettres et de Sciences. Chacun des participants avait reçu dans les trois semaines qui précédaient le stage :

— la lettre circulaire ci-jointe leur demandant de vouloir bien corriger un devoir et d'en faire corriger un second par un collègue ;

— le texte ronéotypé d'une rédaction d'un élève d'une classe de 3^e ;

— le texte ronéotypé d'un problème de géométrie rédigé par un autre élève de 3^e.

Pour le français, nous avons reçu cinquante-huit notes ; pour le calcul, cinquante-quatre. Quelques correcteurs ont noté à un demi-point près ; nous avons arrondi leurs notes au nombre entier inférieur.

MINISTERE
DE
L'EDUCATION NATIONALE

INSPECTION GENERALE

Vous avez été désigné pour participer, le 25 mai, au Centre International d'Etudes Pédagogiques de SEVRES, à un stage consacré à la docimologie.

En tant que responsable de l'organisation de ce stage j'aimerais que vous corrigiez l'une des deux épreuves ci-jointes et que vous demandiez à l'un de vos collègues d'une autre discipline de corriger l'autre.

Ces deux devoirs ont été exécutés par des élèves de la classe de 3^e.

Afin de procéder au dépouillement des résultats avant le 25 mai, nous vous serions obligés d'adresser les deux notes (en distinguant nettement dans votre envoi le Français des Mathématiques) au C.I.E.P. de SEVRES (1, rue Léon-Journault).

Avec mes remerciements, veuillez agréer, M. l'expression de mes sentiments très dévoués.

L'Inspecteur Général
de l'Instruction Publique.

Signé : H. SIRE.

REDACTION (B.E.P.C.)

Note :

Observations :

SUJET :

« Objets inanimés, avez-vous donc une âme qui s'attache à notre âme et la force d'aimer. »

Avez-vous, vous aussi déjà ressenti l'impression que certains objets paraissent avoir une âme à laquelle nous sommes attaché? Raconter.

Objets qui vous promenez dans les vitrines, on vous voit, on vous achète ; et puis un beau jour on vous range ou on vous jette, sans regrets.

Mais, il y a d'autres objets qui tout en étant moins beaux et moins bien faits vous plaisent davantage. Et chaque fois que vous passez devant, une sorte de petit dialogue s'entame. Car ceux-là, c'est vous qui les avez fait.

D'un morceau de bois trouvé, au fond d'une boîte, après de nombreuses heures de travail, vous avez tiré, une forme gauche. Puis elle s'est anoblie et dans chaque coup de couteau que vous avez donné, une partie de vous-même est partie. Ensuite vous l'avez poncé, et des dizaines de fois vous avez regardé à la lumière pour dépister une petite anomalie et vos yeux ont admiré cette forme imparfaite, bien sûr, mais qui était vôtre.

Quand enfin, cet objet, un oiseau pour ma part, a été fini. Vous l'avez fait voir à votre famille, à des amis, à des camarades qui vous ont complimentés sûrement. Mais qui ne pouvaient comprendre pourquoi, chaque fois que vous passiez devant, chaque fois que vous le regardiez votre visage s'illuminait.

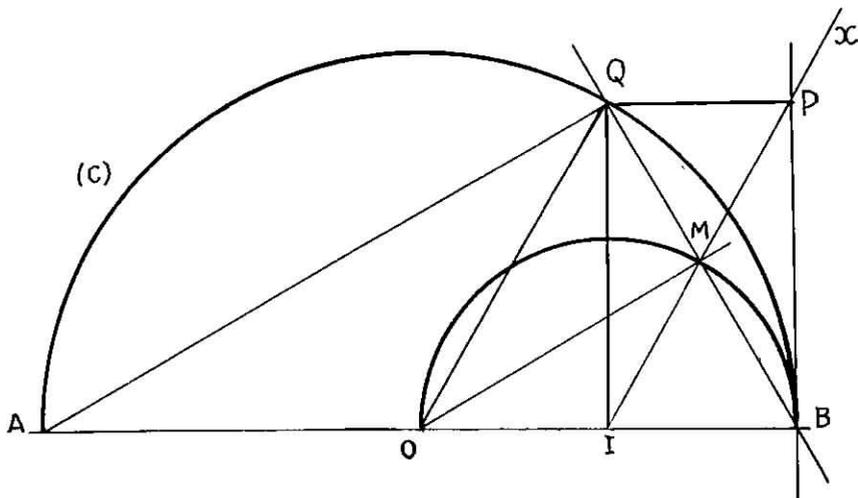
On peut avoir l'impression que certains objets ont une âme. Mais ce sont des objets que vous avez faits ou qui vous ont coûtés et quand vous les regardez, ils vous rappellent une époque, un souvenir. C'est pour cette raison comme le fait si bien comprendre Lamartine, que ce ne sont pas, les grandes et belles choses qui peuvent avoir une âme car tout le monde se les attribue, mais des régions plus retirées ou des objets plus simples qui emportent avec eux une partie de l'âme de celui ou de ceux qui les ont aimés.

Note : Observations :

Problème :

On considère le demi-cercle (C) de centre O, de diamètre $AB = 2R$ et le point I, milieu de OB. La demi-droite IX faisant avec IB un angle $BIX = 60^\circ$ coupe en M le demi-cercle de centre I de et rayon $\frac{R}{2}$ – situé du même côté que (C) par rapport à AB et en P la tangente en B à (C). La demi-droite BM coupe (C) en Q.

- 1) Evaluer les angles \widehat{OBM} et \widehat{BOM} .
- 2) Montrer que le triangle \widehat{BIQ} est rectangle en I.
- 3) Montrer que les triangles \widehat{OBM} et \widehat{BIP} sont égaux.
- 4) Calculer BM, OM, BQ, AQ, IP et IQ.
- 5) Montrer que le quadrilatère OIPQ est un parallélogramme.



1. — IM et IB sont rayons du cercle de centre I : ils sont donc égaux.

Le triangle IMB est équilatéral parce qu'il a un angle \widehat{BIM} de 60° compris entre 2 côtés égaux. Donc $\widehat{OBM} = 60^\circ$.

Le triangle OBM est rectangle en M parce qu'il a pour hypoténuse le diamètre du cercle de centre I et de rayon $\frac{R}{2}$. Donc \widehat{BOM} est égal à 180° moins $(\widehat{OMB} + \widehat{OBM}) = 180^\circ - (90^\circ + 60^\circ) = 30^\circ$.

2. — Lorsque dans un triangle, la médiane est égale à la moitié de l'hypoténuse, le triangle est rectangle. Comme MIB est équilatéral, QMI est isocèle. Donc, $MQ = MI = MB$. Cela entraîne donc que le triangle **IBQ est rectangle en I**.

3. — Dans le triangle OBM, $\widehat{M} = 1d$; dans le triangle BIP, $\widehat{B} = 1d$ parce que B est le pied de la tangente à (C). Donc $\widehat{M} = \widehat{B} = 1d$.

Dans le triangle OBM, $\widehat{B} = 60^\circ$ parce que MIB est un triangle équilatéral; dans le triangle BIP, $\widehat{I} = 60^\circ$ par construction. Donc $\widehat{B} = \widehat{I} = 60^\circ$. Le triangle MIB étant équilatéral, $IB = MB$.

Les triangles OBM et BIP sont égaux (1^{er} cas d'égalité).

$$4. \text{ — } BM = BI = \boxed{\frac{R}{2}}$$

$$\begin{aligned} \text{— } OM^2 &= OB^2 - BM^2 \\ &= R^2 - \frac{R^2}{4} \text{ ou } \frac{4R^2 - R^2}{4} \end{aligned}$$

$$OM = \sqrt{\frac{3R^2}{4}} \text{ ou } \boxed{\frac{R\sqrt{3}}{2}}$$

$$\begin{aligned} \text{— } BQ &= 2 BM \text{ ou } 2 IB \\ &= R \end{aligned}$$

$$= \boxed{R}$$

$$\begin{aligned} \text{— } AQ^2 &= AB^2 - BQ^2 \\ &= 4R^2 - R^2 \end{aligned}$$

$$AQ = \sqrt{3R^2} \text{ ou } \boxed{R\sqrt{3}}$$

$$\text{— } IP = OB \text{ car } OMB = IBP$$

$$IP = \boxed{R}$$

$$\begin{aligned}
 - IQ^2 &= BQ^2 - IB^2 \\
 &= R^2 - \frac{R^2}{4} \quad \text{ou} \quad \frac{4R^2 - R^2}{4}
 \end{aligned}$$

$$IQ = \sqrt{\frac{3R^2}{4}} \quad \text{ou} \quad \boxed{\frac{R\sqrt{3}}{2}}$$

5. — $\widehat{QIB} - \widehat{BIM} = 30^\circ$, donc $\widehat{MIQ} = 30^\circ$ et égal à \widehat{MIO} , \widehat{MPB} , \widehat{MBP} et \widehat{IQO} .
 $\widehat{QIO} = 1d = \widehat{PQI}$ parce PQ est parallèle à AB.

Donc $\widehat{PQO} = \widehat{PIO} = 240^\circ$

Dans le triangle QPI, \widehat{P} est égal à $\widehat{Q} - \widehat{I} = 60^\circ$

$\widehat{QIO} = \widehat{QIB}$, donc $\widehat{O} = \widehat{B} = 60^\circ$

Donc $\widehat{QOI} + \widehat{QPI} = 120^\circ$

Le triangle QOI est égal PIB, donc QO est parallèle à PI.

OQPI est un parallélogramme car il a ses côtés parallèles deux à deux et ses angles opposés égaux deux à deux.

Les tableaux suivants donnent l'ensemble des résultats.

FRANÇAIS

Note sur 20 = p	Nbre de notes = fréquence = f	Produit p × f
4	1	4
5	1	5
6	9	54
7	9	63
8	8	64
9	9	81
10	7	70
11	5	55
12	3	36
13	3	39
14	2	28
15	1	15
Totaux	58	514

p = valeur de la variable
f = fréquence

$$\text{moyenne pondérée} = \frac{514}{58} = 8,86$$

CALCUL

p	f	p × f
4	1	4
5	—	—
6	2	12
7	—	—
8	1	8
9	2	18
10	2	20
11	14	154
12	9	108
13	9	117
14	8	112
15	3	45
16	1	16
17	2	34
Totaux		648
54		

moyenne pondérée = $\frac{648}{54} = 12,00$

Deux conclusions se sont imposées immédiatement aux participants au stage :

- Les notes sont dispersées :
- en français, elles s'échelonnent entre 4 et 15°;
- en géométrie, entre 4 et 17.

Cette dispersion étonne les correcteurs ; certes ils s'attendaient à des variations ; a priori, ils pensaient qu'elles seraient plus fortes en français qu'en géométrie... mais, ils étaient loin de s'imaginer qu'elles auraient pareille ampleur :

- 11 en français sur 20 !
- 13 en géométrie sur 20 !

La prise de contact est rude ; c'est ce que nous souhaitons ; la tranquillité dans laquelle vivaient jusqu'alors les notateurs a fait place à une anxiété certaine ; on sourit, on rit même de constater qu'on a perdu la vérité qu'on croyait posséder depuis de nombreuses années ; qui désormais la possède cette vérité ? moi ou le collègue ? mais quel collègue ? y a-t-il même une vérité dans la notation ? peut-on noter ? la note n'est-elle pas une fausse monnaie que la loi ne punit pas mais que la tradition maintient ; un correcteur n'est-il pas un faux-monnayeur ?

— Tentons de trouver un refuge dans la **moyenne** ; si chacun a tort, la collectivité a peut-être raison ? Calculons la moyenne arithmétique de toutes les notes de français ; elle est de 8,86 sur 20 ; neuf notateurs sur cinquante-huit qui ont noté 9 sont contents ; vingt-huit sont en-dessous ; vingt-et-un sont au-dessus.

La moyenne en géométrie est 12 ; neuf notateurs ont mis cette note (sur cinquante-quatre) ; vingt-deux sont en dessous ; vingt-trois sont au-dessus ; quatorze avaient noté 11 ; la majorité semble avoir eu tort (ce ne serait pas la première fois !). Voici une question que nous laissons sans réponse :

ces 14 qui ont noté 11

ont-ils raison devant 9 qui ont noté 12 ? On a envie de dire « oui » ; ont-ils raison devant 54 notateurs dont la moyenne des notes est 12 ; on a envie de dire « non ».

— Une troisième conclusion aurait pu être établie si le temps ne nous avait pas manqué pour dépouiller les appréciations des notateurs ; ces appréciations sont parfois contradictoires ; nous avons relevé les principales en français autour de deux idées :

— le sujet était-il compris ? neuf notateurs disent « oui » ; dix disent « non ».

— l'élève avait-il ou manquait-il de sensibilité ?
deux disent qu'il en manquait,
six affirment qu'il en avait.

Le dépouillement montre que certains notateurs attachent plus d'importance que d'autres

à la longueur du devoir (ici il était court),
à l'orthographe,
à l'originalité des idées,
à l'aisance du style, etc...

Bref, les points de vue des notateurs diffèrent profondément les uns des autres.

Ci-joint les appréciations que nous avons relevées :

COMPRIS OU PAS

— Sujet mal compris	1/2
+ Devoir qui contient une idée bien personnelle ; le 3 ^e et le 5 ^e paragraphes expriment avec une certaine finesse une expérience qui semble vraie	7 1/2

— Devoir superficiel	9
+ Sujet bien compris	10
— Fond sec et pauvre	6
+ Vous avez semble-t-il compris le sujet	10
— Le développement ne s'appuie sur aucun élément ou souvenir personnel précis	11
+ Sujet bien compris	12 1/2
+ Sujet bien compris	14
— Devoir très médiocre	7
— Les idées demeurent le plus souvent insuffisantes	8
— Le sujet n'est pas très bien compris	6
+ Devoir intéressant	9 1/2
+ Bien	15
— Sujet mal compris	7
+ Sujet compris	11
+ Le sujet est traité	10
+ Sujet assez bien compris	10 1/2
+ Excellent début montrant une maturité d'esprit supérieure à la maturité normale des élèves de 3 ^e	13
— Vrai sujet non compris, non traité	6
— L'ensemble est nettement insuffisant	6

SENSIBILITE

— Des qualités de composition, mais non d'imagination et de sensibilité	9
+ De la finesse d'esprit et de la sensibilité	10
+ Développement court mais révélant une assez grande sensibilité et des qualités littéraires	13
+ Le devoir ne manque ni d'originalité ni de sensibilité	11
+ Ce devoir n'est pas sans montrer une sensibilité certaine	7
+ De la sensibilité, des idées originales	11 1/2
+ Travail qui révèle une certaine sensibilité	8
— Absence de sensibilité	6

II. notes et moyennes

Au moment où l'on se penche avec tant d'intérêt, et même d'inquiétude, sur une orientation scolaire continue tout au long du premier cycle et à son issue, orientation qui repose, pour une très large part, sur les appréciations que les maîtres portent sur chacun de leurs élèves ; où il devient nécessaire d'analyser les résultats obtenus par un élève dans des disciplines différentes ; où il faut suivre l'évolution de ces résultats, tout au moins dans les disciplines dites fondamentales, pendant les quatre années du premier cycle ; où il faut comparer les résultats d'un élève à ceux d'un élève « normalisé » dont aucune définition objective ne semble possible (quelle est, par exemple, la norme d'un élève de 3^e B apte à suivre en seconde C ?), on doit se demander si les échelles de valeurs utilisées par les professeurs sont concordantes et stables. Si la réponse à cette question est négative nous aurons alors à craindre des erreurs dans les conclusions des maîtres et les décisions des conseils ; et nous connaissons les mêmes craintes en ce qui concerne les examens et les concours.

La tradition en France est de considérer la note chiffrée comme le premier élément, sinon l'élément fondamental, de toute appréciation. Nous avons la manie de la note chiffrée ; trop de maîtres croient à leurs notes chiffrées ; ils ont l'impression, et même la certitude, que leurs échelles de notation sont comparables, qu'un 13/20 a une valeur, disons, nationale ; que s'ils notent 13 1/2 au lieu de 13 c'est que ce demi-point est nécessaire pour caractériser des différences réelles, appréciées parce qu'appréciables, dans la chose jugée. Nous avons déjà montré que cette confiance est absolument illusoire.

La majorité de nos élèves travaillent pour la note ; je dirai même que le classement leur importe avant tout et qu'ils sont animés du plus fâcheux esprit de compétition qui consiste à vouloir se classer par rapport à leurs camarades et non pas par rapport à eux-mêmes.

Et puisqu'il faut des points pour être bien jugés ou pour « passer » tous les moyens sont bons pour les obtenir ; ainsi s'explique, en partie, l'état de fraude devant lequel nous demeurons impuissants.

Nous apparaissions comme les premiers responsables de ce fâcheux état d'esprit :

— nous ne décidons du sort final d'un élève qu'en le soumettant à des examens et concours, c'est-à-dire à un nombre limité d'épreuves différentes ; leur correction aboutit à des notes ; des coefficients interviennent ; une addition termine les opérations... une délibération plus ou moins brève... et la décision est prise... l'addition est une commodité ; elle est aussi une tranquillité : « Madame, votre fils n'a pas atteint la moyenne : il est donc collé ». Madame s'incline : une addition ne saurait se discuter ; et d'ailleurs, le jury n'est-il pas souverain ? Tout ceci est une façon de se réfugier dans l'irresponsabilité et de pratiquer le fétichisme à l'égard du diplôme.

M. ZAMANSKY, doyen de la Faculté des Sciences de Paris, écrit : « L'enseignement du second degré semble s'organiser sur l'idée fixe que, quand leur enfant a 10 ans, les familles ont déjà l'obsession du concours de recrutement auquel il devra se présenter huit ans plus tard. Dès lors, on organise les études en vue d'assurer le succès au concours. »

A certaines époques de l'année toute la France est en transe ; tous les enseignants et une bonne partie des enseignés jonglent avec les chiffres, les coefficients, les totaux, les moyennes... et l'on criera aux injustices, surtout sur celles dont seront victimes les meilleurs élèves sans qu'on s'attarde par trop sur celles dont pourront bénéficier les mauvais.

— tout au long de la scolarité, nous plaçons des compositions trimestrielles ; véritables accidents qui, trop souvent encore, conduisent au calcul — sans qu'on sache trop comment la calculer — d'une moyenne générale annuelle qui, elle aussi, est notre commodité et notre tranquillité ;

— abusivement, le redoublement d'une bourse nationale est accroché à la moyenne générale annuelle des compositions ;

— le passage d'un élève de Première dans la classe terminale dépend largement de la moyenne annuelle des compositions sans, toutefois, que joue automatiquement cette moyenne ; celle-ci peut défavoriser les élèves brillants dans quelques disciplines et faibles dans d'autres ; un élève moyen en tout ne court aucun risque ; c'est nier les aptitudes et les goûts personnels.

Certes, le dossier scolaire est destiné à corriger, s'il y a lieu, l'automatisme de la moyenne ; mais que renferme-t-il ? reconnaissons que c'est plutôt un livret scolaire qui ne comprend que les notes et les places dans les compositions, assorties de quelques appréciations rapides, rarement, explicatives des professeurs.

Comment, dans ces conditions, les élèves et leurs parents n'attacheraient-ils pas une importance majeure à la question des notes chiffrées puisque nous-mêmes, professeurs et administrateurs, montrons cet attachement ?

Remarquons qu'en tant que fonctionnaires, nous sommes notés de 0 à 20... et je sais qu'on tient à la note... Mais, fort heureusement, la sagesse s'est emparée des notateurs car, pratiquement, l'échelle 0-20 s'est singulièrement rétrécie : la note administrative 20 est rare ; la note 17 et les notes inférieures le sont également. Il en résulte que presque toutes les notes sont des 18 et 19... 19 quand on est très bien ou bien... 18 quand on est, disons, moins bien.

Ainsi, l'échelle prévue de 0 à 20 est devenue, à l'usage, une échelle binaire encadrée de deux nuances, ce qui fait finalement quatre catégories :

celle des « exceptionnels »	20
celle des « très bons-bons »	19
celle des « assez bons-moyens »	18
celle des « insuffisants »	17 et moins

Avant d'aborder la très difficile, sinon impossible, question de la normalisation des notes, il est nécessaire que tous les enseignants, professeurs et examinateurs, prennent conscience, mais une conscience claire, de l'importance du problème ; je dirais même de son caractère parfois dramatique.

Nous connaissons bien, mais nous n'osons pas toujours le reconnaître, la part de subjectivité qui intervient dans l'appréciation d'un résultat scolaire ; jouent :

— le tempérament du correcteur, naturellement sévère ou indulgent,

- les rapports qu'il entretient avec l'élève, rapports de sympathie, d'antipathie, de voisinage, d'amitié avec les parents,
- l'aspect du devoir auquel il attache le plus d'importance :
 - l'exactitude du résultat ou la justesse, ou même l'élégance du raisonnement,
 - la concision du style, sa souplesse, son coloris, son charme,
 - la clarté, la précision de la pensée, ou, au contraire, ses ondulations, ses ramifications, voire ses hésitations,
 - l'orthographe, parfois l'écriture, l'imagination que manifeste l'élève ou son sens de l'observation,
 - la longueur du devoir ou la brièveté.

Il est bien certain que le correcteur d'un devoir de français ou de philosophie ne peut s'empêcher de filtrer les idées de l'élève à travers ses propres conceptions ; et ces filtres sont nécessairement à mailles plus ou moins serrées.

— joue également le manque de stabilité du correcteur suivant les moments de la journée, le nombre et la quantité des devoirs corrigés antérieurement — et l'on a mis nettement en évidence les « effets de contraste » ;

— enfin, ajoutons que le jour de l'épreuve, l'élève a pu avoir de la « chance » ou, au contraire, en manquer.

LA QUESTION DE LA MOYENNE GENERALE ET DU CLASSEMENT

— Quand le professeur de mathématiques Dupont corrige les devoirs d'algèbre de ses élèves de 3^e A², et que **ces devoirs sont nombreux**, on peut, à la rigueur, admettre que sa stabilité est suffisante pour que la moyenne arithmétique des notes mises à l'élève Marcel, et celle des notes mises à l'élève René aient une valeur relative ; si la moyenne de Marcel est supérieure à celle de René on peut considérer que la valeur de Marcel en algèbre est supérieure à celle de René

$$\begin{array}{ccc}
 \text{moyenne Marcel} & & \text{moyenne René} \\
 \text{en algèbre} & > & \text{en algèbre} \\
 = 12,3 & & = 10,1
 \end{array}$$

— On peut dire la même chose pour chacune des disciplines que pratiquent Marcel et René.

— Nous admettrons donc que les moyennes par discipline peuvent être retenues aux conditions suivantes :

1) elles doivent porter sur un nombre suffisant de notes (ce qui condamne la composition trimestrielle qui ne donne qu'une note) ;

2) la discipline doit être une discipline élémentaire et non pas globale :

— les mathématiques sont une discipline globale ; algèbre et géométrie sont des disciplines élémentaires,

— le français est une discipline globale ; orthographe, récitation, dissertation (et encore !) sont des disciplines élémentaires ; etc.

Marcel et René peuvent être caractérisés par les moyennes arithmétiques obtenues dans chacune des disciplines élémentaires ; mais ces moyennes caractérisent aussi les professeurs qui ont noté, de telle sorte que si Marcel est supérieur en dissertation à René avec le professeur Duval, il lui serait peut-être inférieur avec le professeur Blanc.

Cette remarque suffirait à condamner la moyenne générale qui consiste à faire la moyenne arithmétique des moyennes élémentaires... Dans cette moyenne arithmétique générale

— faut-il faire intervenir **toutes** les disciplines ? (par exemple, l'éducation physique ? et pourquoi pas !),

— et avec le même poids ? certains chefs d'établissement affectent les moyennes élémentaires d'un coefficient égal au nombre d'heures hebdomadaires de la discipline élémentaire ; c'est sans doute mieux que le coefficient 1, bien que ce soit tout à fait arbitraire.

Bon sens et logique condamnent absolument cette moyenne générale qui est l'image même de la confusion ; partir de moyennes élémentaires qui peuvent avoir chacune une certaine signification pour arriver à une note c'est partir d'une macédoine de légumes — qu'on excuse cette vilaine comparaison — dans laquelle on peut reconnaître les constituants, petits pois, carottes, etc., pour arriver à une bouillie où l'on aurait broyé tous les constituants ; je saurais, à la rigueur comparer la macédoine Marcel à la macédoine René la première étant plus riche, par exemple, en petits pois que la seconde ; mais comment comparer la bouillie Marcel à la bouillie René ?

La moyenne générale est donc inconcevable ; il en est de même des classements que l'on établit à partir d'elle.

III. mesurer, repérer, noter

Qu'est-ce que noter ?

Est-ce mesurer ?

Est-ce repérer ?

Ce n'est ni l'un ni l'autre.

I. — Le physicien **mesure** des grandeurs mesurables. Mesurer une grandeur, pour lui, c'est la comparer à une grandeur de même espèce choisie comme unité.

Cette comparaison se traduit par un nombre qui est le résultat de la mesure.

a) Pour mesurer la longueur d'une salle, je choisis une unité de mesure : le mètre, par exemple ;

b) je compare la longueur de la salle à celle du mètre : pour ce faire, je porte bout à bout **plusieurs mètres** jusqu'à ce que j'aie réalisé une longueur **égale** à celle de la salle. Je suis donc conduit à faire la **somme** de plusieurs longueurs égales à un mètre, puis à apprécier **l'égalité** entre la longueur ainsi obtenue et celle de la salle.

Ainsi, très généralement, la mesure d'une grandeur exige qu'on sache définir la **somme** et **l'égalité** des grandeurs de cette espèce.

Sont mesurables les longueurs, les surfaces, les volumes, les masses, les forces, les pressions, etc. Ne sont pas mesurables les températures, les valeurs des devoirs de nos élèves.

— Tout résultat d'une mesure comporte nécessairement une erreur, une incertitude qui tient aux erreurs commises lors de la mesure :

- comment assurer un « bout à bout » des mètres avec une précision absolue ?

- comment apprécier la longueur de la salle au centimètre près si le mètre n'est gradué qu'en décimètres ?

- comment lever toute indétermination dans la définition de la grandeur à mesurer : dire que la salle a une longueur de huit mètres c'est sous-entendre qu'elle est rectangulaire ; si le maçon a tenu à lui donner une forme rectangulaire, il a réussi l'opération sans doute à 2, 3, 4, 5 centimètres près ; il n'a certainement pas atteint la précision du centimètre.

— De toute évidence un devoir d'élève n'est pas une grandeur mesurable ; on ne saurait définir l'égalité de deux devoirs (sauf dans le cas d'un copiage absolu), ni leur somme ; et nous ne pouvons imaginer une unité de mesure.

Un notateur n'est pas un physicien ; il n'a en commun avec lui que l'incertitude — et quelle incertitude !

II. — Le physicien **repère**.

Lorsque deux corps sont côte-à-côte, leurs températures tendent à devenir égales. Mais peut-on dire qu'on ajoute une température à une autre et qu'un litre d'eau à 50°C ajouté à un autre à 50°C donne deux litres d'eau à 100°C ?

La température, dont on ne peut définir la somme, n'est pas une grandeur mesurable. Le thermomètre est un appareil qui porte des **repères** dont l'ensemble constitue sa graduation ; cette graduation résulte d'un **étalonnage** à partir de repères thermométriques fixes, ce qui permet d'avoir des thermomètres comparables entre eux.

III. — Le correcteur ne mesure pas, ne repère pas.

Dans les conditions habituelles de la notation des exercices scolaires traditionnels, peut-on utiliser une échelle de notation, comparable à une échelle thermométrique, étalonnée à partir de repères fixes, qui permettrait aux notateurs d'un même exercice scolaire de lui affecter la même note-repère ? Non. Noter un exercice scolaire n'est donc pas repérer.

Là encore le notateur n'est pas un physicien. Faire, pour un élève donné, dans un exercice scolaire déterminé, le graphique de ses notes mises toujours par le même correcteur est certes comparable au tracé du graphique des températures d'un malade, à cette différence près que si les degrés de température peuvent être correctement repérés puisqu'on dispose d'échelles de repères thermométriques stables, il n'en est pas de même des notes pour lesquelles aucun étalonnage n'est possible.

Une note ne peut être, comme une pesée, ni juste, ni fidèle ; elle ne peut davantage être précise comme l'est un degré thermométrique. Elle varie d'un professeur à l'autre et, pour un même professeur, d'un moment à l'autre ; c'est pourquoi — nous nous répétons — on a pu dire qu'une note est une fausse monnaie puisque son étalonnage est impossible, et que les notateurs sont, inconsciemment, des faux-monnayeurs.

Pourtant, ils connaissent bien la subjectivité de leurs notes ; mais c'est une conscience diffuse et fragile qui n'empêche pas la plupart d'entre eux d'affirmer qu'ils « savent » noter, que le 13/20 qu'ils ont attribué à telle copie repère bien sa valeur, à moins que ce ne soit 13 1/2 — ou même 13,25... ou même 13 +. En disant 13 1/2, ils affirment la possibilité de soupeser le devoir avec la précision du 1/40°. Quelle prétention, et quelle illusion !

Le correcteur est dans la situation d'un topographe qui voudrait mesurer une longueur au double-pas, un double-pas que, certes, il aurait étalonné, sur une base de longueur connue... mais

— la longueur à mesurer serait d'une part mal définie dans ses origines et son tracé ; serait-elle rectiligne, circulaire, courbe, capricieusement sinueuse, horizontale, ascendante, descendante ?

— d'autre part l'étalonnage du pas serait incertain, autrement dit il ne serait pas étalonné d'une manière stable ; il dépendrait du moment de la journée, de la fatigue, du repas consommé, du temps orageux, de la montée, de la descente, etc.

Marcel SIRE.

LES TESTS DE CONNAISSANCES

L'usage des tests, en France, est surtout le fait des psychologues. Depuis plusieurs décades, ceux-ci ont apporté à la connaissance de l'enfant, par des méthodes qui leur sont propres, une contribution originale qui s'est souvent révélée d'autant plus efficace que la mise en évidence des aptitudes ou d'autres aspects de la personnalité peut expliquer, nuancer les observations pratiquées quotidiennement par les enseignants et aider à la formulation d'un pronostic.

Si les efforts des pédagogues et des psychologues concourent à un épanouissement plus complet de l'enfant, si leurs recherches ont un but commun, qui est une meilleure compréhension de celui-ci, il est cependant peu contestable que non seulement les uns et les autres utilisent des voies différentes, mais que leurs objectifs immédiats sont habituellement très distincts.

Les tests de connaissances font exception à cet état de choses, précisément parce que ces épreuves ont un caractère mixte, mais l'apport mutuel des deux types de spécialistes n'a pas encore reçu tout le développement ni toute la coordination désirables. L'objet de ces tests est nettement d'ordre pédagogique, mais les techniques qu'ils mettent en œuvre — aussi bien dans leur construction que dans leur utilisation — sont mieux connues des psychologues. L'initiative en ce domaine pouvait donc en principe revenir aussi bien aux uns qu'aux autres. Pour bien des raisons, et en France particulièrement, ce sont les psychologues qui dans l'ensemble ont été les premiers à établir et à utiliser des tests de connaissances.

L'une de ces raisons est certainement liée à la nécessité de faire face aux phénomènes associés à la démocratisation de l'enseignement. Placés par leur fonction en dehors de l'école, aux carrefours des divers enseignements, les psychologues chargés de l'orientation des enfants ont été très vite confrontés aux problèmes de répartition, d'évaluation des chances de réussite, pour lesquels il leur fallait une référence stable. Ils ne sont pas les seuls pour autant à devoir céder à la pression des circonstances, et les enseignants aussi sont contraints de participer à des opérations d'orientation qui dépassent très largement le cadre de leurs classes respectives. Mais alors que ceux-ci s'attachent à un système de « mesure » appuyé sur des traditions, les premiers dénoncent le caractère incertain de cette mesure et lui préfèrent une échelle objective et un classement selon cette échelle.

L'idée de mesure dans le domaine des connaissances peut paraître théoriquement contestable. Selon que l'on veut davantage contrôler les processus dynamiques d'apprentissage qu'estimer la valeur de l'acquis, on peut se demander si l'évaluation du travail d'un élève à un moment donné et dans un contexte donné atteint le but recherché. Ces considérations soulèvent des problèmes plus proprement pédagogiques, et nous nous bornerons à discuter ce qui a trait aux aspects plus formels des problèmes de mesure.

La note donnée en classe est un outil pédagogique. Elle a une dimension temporelle, que n'a pas la note d'examen ; elle condense et matérialise certains aspects des relations maître-élèves et, en tant que telle, il est légitime qu'elle reflète des variations personnelles, qu'elle ait un caractère subjectif. On pourrait également soutenir qu'aucune nécessité autre que le besoin de se conformer à un règlement, ou la commodité de l'habitude, n'impose à cette note une échelle déterminée. C'est d'ailleurs ce qui se passe quelquefois lorsque des élèves connaissent les exigences de leurs professeurs disent que le 8 de ce professeur vaut un 12 de la part d'un autre professeur. D'un point de vue strictement pédagogique, il n'y a là rien de choquant. Ce double caractère d'incertitude, à notre avis, grève plus lourdement la valeur de la note dès que celle-ci joue dans l'orientation des élèves un rôle déterminant.

Les tests de connaissance n'ont pas une place très large dans l'orientation des adolescents français. L'Angleterre, dont l'évolution dans le domaine de l'éducation est sensiblement parallèle à la nôtre, et pays traditionaliste aussi, entreprend maintenant de traiter certaines parties des examens dans un esprit très proche de celui des épreuves préconisées par les psychologues, et avec

l'aide de ceux-ci. Cela ne se fera pas sans réticences ni sans regret. Il est difficile de se défaire d'habitudes de pensée qui se justifiaient peut-être à une époque où l'enseignement s'adressait à une fraction restreinte et socialement définie de la population, où l'entrée à l'université s'accomplissait dans une atmosphère de cooptation. Actuellement, la demande croissante d'éducation et la multiplicité des voies par lesquelles elle est dispensée requièrent un système plus général. Encore une fois, on peut contester la nécessité de classer, de hiérarchiser les élèves, mais ce souhait est, peut-on dire, sur un autre plan. Dans la réalité telle qu'elle se présente aux adolescents au cours de la scolarité secondaire, il semble difficile de s'affranchir de tout principe de différenciation. La Suède, entre tous les pays qui ont tenté de donner corps à une conception vraiment démocratique de l'enseignement, s'est posé le problème sans pouvoir lui donner une solution qui satisfasse pleinement à la fois au modèle théorique et aux exigences qui découlent de l'organisation sociale.

Le point de vue des partisans des tests de connaissances n'est pourtant pas une attitude de compromis ; il participe plutôt à des implications générales d'un nouvel humanisme. C'est le poids des dimensions sociales dans des problèmes qui autrefois apparaissaient surtout dans leurs répercussions personnelles qui les amène à considérer comme importants les aspects métriques des symboles que nous utilisons.

Le plus évident peut-être de ces aspects entraîne l'obligation que l'indice par lequel on caractérise un élève soit **fidèle**. On peut exprimer par des méthodes statistiques, le degré de fidélité ou d'homogénéité d'une notation ou d'une classification. La notion de fidélité a donc un caractère précis et nettement défini. Mais, pour rester au niveau d'une explication simple, nous l'introduirons ici de manière intuitive, en disant qu'elle exprime, par exemple, le degré d'accord entre les deux séries de notes d'un professeur donné à un même ensemble de devoirs à quelques semaines d'intervalle, ou encore la cohérence des appréciations données à un même devoir par des professeurs différents. Il a été de nombreuses fois prouvé que les notes scolaires ont une fidélité faible, ce qui revient à dire que la valeur attribuée à telle composition d'un élève aurait été différente si le correcteur avait noté à un autre moment, ou si le devoir avait été jugé par un autre correcteur. Une autre obligation entraînée par les examens est que les notes attribuées — ou plus exactement les processus de notation employés — soient **valides**. La notion de validité est complexe ; elle peut s'appliquer également à un indice mesurant

le degré d'accord entre un système de notes et un critère extérieur, ou exprimer un niveau de signification de la note par rapport à ce qu'on veut réellement atteindre par la notation. Dans le premier sens, on pourrait examiner par exemple s'il existe une relation nettement positive entre les notes obtenues au baccalauréat, pour un élève donné, et sa réussite ultérieure dans l'enseignement supérieur. Dans le second sens, on pourrait se demander si la note accordée à un élève, sur une copie d'Histoire, par exemple, exprime bien le niveau de l'élève en Histoire, autrement dit dans quelle mesure cette note particulière représente bien le niveau général de l'élève en cette matière.

C'est à partir de ces deux notions de fidélité et de validité que nous tenterons de justifier les tests de connaissances en tant que méthode d'examen (le mot examen étant pris dans le sens le plus large possible), en considérant la nature et les qualités de l'information qu'ils apportent sur l'élève, comparativement à celle que donne, dans les mêmes circonstances, la notation traditionnelle.

Ce sont les efforts faits en vue d'une meilleure fidélité qui ont certainement déterminé les premières formes d'épreuves à correction objective. Le but des novateurs étant de restreindre au maximum la dispersion des notes accordées à un même devoir, ceux-ci se sont appliqués à standardiser les devoirs proposés, à les découper en questions formulées de manière telle que les réponses fournies pouvaient être interprétées sans équivoque, à l'aide d'un barème exactement défini. Cette préoccupation a paru primordiale aux premiers constructeurs de tests, et si leur but était atteint — une notation beaucoup plus fidèle que celle des notes scolaires — ce gain était contrebalancé par une « atomisation » des connaissances. En dépit des améliorations introduites depuis lors, beaucoup de tests de connaissances établissent encore le classement des élèves à partir de leur niveau estimé sur l'acquisition d'éléments ou de mécanismes de base, plus que sur leurs capacités de synthèse d'un sujet donné. On peut alors se demander si dans ces conditions les tests de connaissances sont aussi valides que les notes traditionnelles.

Nous avons évoqué plus haut deux sens assez différents de la notion de validité. Au sens d'accord avec un critère de réussite scolaire d'ordre général, la validité des tests de connaissances n'est pas inférieure à celle des notes scolaires attribuées dans le même temps. Plusieurs études l'ont démontré : par exemple, une enquête menée sur les élèves du Loiret, du CM2 à la fin de la classe de cinquième (1) a mis en évidence une relation

plus étroite entre le classement des élèves à partir des tests subis au CM2 et l'appréciation globale du chef d'établissement à la fin de la cinquième, qu'entre les appréciations des instituteurs de CM2 et le même critère.

Mais, si la valeur prédictive des notes scolaires à moyen ou long terme est relativement basse, celle des tests employés jusqu'à présent n'est pas très élevée non plus. Sans doute, il serait vain d'attendre, par quelque moyen que ce soit, une méthode parfaitement prédictive en ce domaine ; mais la modestie des résultats enregistrés montre qu'il existe certainement une marge de progrès possible. Ces progrès, dussent-ils être entrepris par l'une ou l'autre méthode, ne passent peut-être pas par les mêmes impératifs. On peut penser en effet que c'est la faible fidélité des notes scolaires qui limite leur validité prédictive ; ce n'est pas le cas des tests, et il est possible que ce soit leur nature trop analytique qui aboutit à un résultat analogue. Ceci nous amène à discuter l'autre sens du mot validité.

C'est un reproche très habituel que l'on fait aux tests de connaissances, celui de ne s'attacher qu'aux aspects les moins élaborés de la pensée, de faire surtout appel à la mémoire, voire de n'utiliser dans celle-ci que des processus de reconnaissance. Cela est vrai dans certains cas. Par ailleurs, ce défaut s'assortit généralement d'un avantage, en ce sens que la brièveté des questions et des réponses qu'il autorise est en quelque sorte compensée par le nombre élevé de celles-ci, ce qui permet une investigation des connaissances dans de plus larges champs, différents entre eux mais dépendant d'un même domaine. Par contraste, la composition de type traditionnel apparaît à certains comme beaucoup plus valide, parce que moins structurée, et par la liberté même laissée à l'élève de conduire son raisonnement comme il l'entend, de permettre au correcteur d'apprécier la rectitude et la puissance de la pensée. Mais tous les sujets proposés, ou proposables, sont loin de permettre des jugements comparables entre eux, et surtout, ils n'offrent pas, probablement, les mêmes opportunités à tous les élèves. Dans la perspective continue de la classe, ce n'est pas un inconvénient, parce que l'opinion du professeur se fonde sur une pluralité d'observations. Dans la perspective d'un examen comme le baccalauréat, par exemple, il n'est pas sûr que la note donnée à un élève à partir d'une composition puisse être considérée comme nécessairement représentative de la valeur de cet élève dans la discipline où il a composé.

(1) F. BACHER et M. REUHLIN. Le cycle d'observation. Enquête sur l'ensemble des élèves d'un département - B.I.N.O.P., 1965, XXI, n° 3.

Cette controverse trouve un commun dénominateur dans l'effort réalisé par de nombreux groupes pédagogiques et psychologiques pour clarifier et préciser les objectifs de l'enseignement. Ce que vise tout examen, c'est de vérifier si ces objectifs sont atteints. Diverses taxonomies ont été élaborées ; elles aboutissent toutes à des conclusions pratiques voisines : lorsque l'effort est fait de définir dans ses aspects les plus fins les objectifs pédagogiques, la nature des questions que l'on peut soumettre aux élèves se trouve par le fait même très clarifiée. Le choix du terme « taxonomie » souligne l'optique dans laquelle se font ces recherches et dénote le souci d'une organisation hiérarchique, qui est en elle-même favorable à la formulation explicite du travail demandé.

Le problème de la validité maximum d'une épreuve dans un domaine donné pourrait donc trouver une solution par cette voie, dans une confluence de la recherche pédagogique et de la docimologie. Le test de connaissances deviendrait un travail dans lequel l'élève prouverait qu'il a atteint le but essentiel que ses maîtres s'étaient assignés, et qui serait codifié de telle sorte que l'appréciation donnée en fonction de ce travail suivrait les règles d'une mesure convenable.

Ce projet n'est pas utopique. Un organisme comme The School Mathematics Study Group (U.S.A.) a entrepris la construction de tests de connaissances en mathématiques basés sur une taxonomie qui cautionne la validité de contenu de ceux-ci. Le Projet International pour l'Evaluation du rendement scolaire (I.E.A.) a des préoccupations analogues. En Angleterre, en divers endroits et sous diverses égides, professeurs et psychologues ont entrepris la rédaction d'épreuves de ce type.

Cela ne signifie évidemment pas que ce renouveau dans la manière d'aborder ces problèmes résolve entièrement toutes les difficultés. Des critiques sont possibles, elles sont même souhaitables dans la mesure où elles entraînent un progrès. En dehors des ambiguïtés liées à une terminologie, quelquefois obscure, il est certain que les inférences par lesquelles on passe d'une classification théorique à des questions réelles ne sont pas sans danger. Il faut ajouter aussi que la plus fondamentale des critiques demeure : une mesure basée sur un examen collectif rend-elle la même justice à tous ? N'y a-t-il pas dans les démarches intellectuelles des enfants des différences irréductibles à une conception d'un comportement moyen ? La discussion du postulat implicite dans cette critique dépasse de beaucoup la portée limitée de cet article. Nous nous bornerons à souligner que, si le problème

existe, ses répercussions sur le plan de la mesure risquent d'être beaucoup plus importantes dans les examens traditionnels qu'au travers de tests bien construits. On avance souvent que les examens oraux permettent d'adapter en quelque sorte l'outil de mesure, c'est-à-dire l'interrogation, au type particulier de pensée examiné. Malheureusement, dans la pratique, on a pu constater que « l'examen oral mérite une moindre confiance que l'examen écrit ». (1).

Les tests de connaissances s'orientent donc actuellement vers l'utilisation de questions — il est habituel de parler en ce cas d'item — qui visent des processus hiérarchisés jusqu'aux plus élevés, synthèse, critique en termes d'évidence interne ou externe, invention ou créativité. Leur forme reste classique, les réponses pouvant être libres, ou à choisir parmi plusieurs éventualités. Dans le premier cas, la question doit être formulée de telle sorte que la réponse puisse être donnée de manière claire. Dans le second cas, même si la réponse est sollicitée sous la forme d'un repère simple, (un numéro, une lettre), elle peut recouvrir une élaboration complexe.

Il est facile d'en trouver des exemples dans les examens français de médecine, qui utilisent depuis quelques années les questionnaires à choix multiple. Parmi les plus typiques, citons celui dans lequel deux propositions sont faites, à propos desquelles l'étudiant doit répondre :

A, si les deux propositions sont vraies et qu'il existe une relation de cause à effet dans le sens indiqué,

B, si les deux propositions sont vraies sans qu'il y ait de relation de cause à effet,

C, si la première proposition est vraie mais si la deuxième est fausse,

D, si la première proposition est fausse mais la deuxième est un fait ou un principe accepté,

E, si les deux propositions sont fausses. (2).

L'exemple qui correspond à ce type d'item est le suivant :

Dans un muscle en activité, la différence de teneur en oxygène entre l'artère afférente et la veine efférente est augmentée parce que l'augmentation des besoins énergétiques du muscle entraîne une vaso-dilatation importante. (Réponse B).

(1) H. PIERON, M. REUCLIN et F. BACHER. Une recherche expérimentale de docimologie sur les examens oraux de physique au niveau du baccalauréat de mathématiques. *Blotypologie*, 1962, XXIII, 1-2, 48-61.

(2) d'après Les examens par questions à choix multiple, brochure publiée par l'Institut Pédagogique National, section médicale.

Plus qu'une compréhension d'un concept, on cherche souvent à vérifier si l'élève est capable d'appliquer ce concept. Ceci peut être estimé objectivement. Par exemple on pourra demander si une ou plusieurs propositions représentent une illustration de l'un de ces principes généraux, des deux pris simultanément, ou d'aucun des deux : (1)

A - Toute particule de l'univers exerce une attraction sur une autre particule selon une force proportionnelle au produit des masses divisé par le carré de leur distance.

B - Le mouvement apparent d'un objet relativement proche d'un observateur, par rapport à un objet lointain, varie avec le mouvement propre de l'observateur. (La réponse consiste à placer une croix soit sous A seul, soit sous B seul, soit sous A et B, soit dans la 3^e colonne) :

A	B	Aucun

1 - La distance qui sépare la Terre des plus proches étoiles peut être mesurée.

8 - Le diamètre de l'orbite terrestre est quelquefois utilisé comme « base de départ » dans les calculs astronomiques.

On peut approcher les qualités d'invention, enfin, par des questions du type suivant : on fournit deux données à partir desquelles l'enfant doit trouver « quelque chose » ; puis ce « quelque chose » associé à « autre chose » peut le conduire ensuite à un autre résultat, et ainsi de suite (par exemple, à partir de la longueur et de la largeur d'une brique, que peut-on trouver ? l'enfant peut indiquer ensuite qu'en connaissant la hauteur, on peut trouver le volume ; si de plus on connaît le poids spécifique, on peut trouver le poids, (ou l'inverse), etc.).

Sur le plan de l'utilisation, on fait souvent le reproche aux tests de connaissances de permettre à l'élève astucieux de deviner la réponse exacte par un ensemble d'indications secondaires, souvent sans rapport avec la question elle-même ; nous ne pensons pas que cette « méthode » de travail soit réellement payante si le test est bien fait. On leur fait aussi le grief de favoriser le candidat qui raisonne par élimination des réponses inexactes. Ces

(1) d'après l'exemple donné par P.E. Vernon. XV^e Congrès de Psychologie Appliquée, LJUBJANA, 1964.

défauts nous semblent être des défauts de technique et par conséquent réductibles. Enfin, on peut leur reprocher de faire une part très mince aux capacités d'expression personnelle. Cela est vrai, mais d'une part l'usage des tests n'implique pas que l'information globale fournie par l'examen exclue d'autres méthodes. D'autre part, le reproche inverse pourrait être fait aux examens traditionnels, dans lesquels les capacités d'expression, en raison de la forme des compositions, entrent pour beaucoup dans le jugement de l'élève. Dans la mesure où le baccalauréat est autant un brevet de culture générale qu'un examen de niveau considéré sous l'angle de l'entrée en Faculté, on ne voit pas pourquoi des enfants intelligents mais limités dans leur expression verbale seraient systématiquement défavorisés.

Les tests de connaissances présentent enfin une caractéristique particulière qui les distingue très nettement des notes scolaires. Si les psychologues emploient quelquefois eux aussi le mot « note » à propos d'un résultat à un test, c'est par analogie avec la mesure scolaire, par facilité de langage. Mais à la différence de la notation scolaire, le critère de référence de la valeur attribuée est indépendant de toute échelle absolue de valeurs, dont on a maintenant prouvé l'inconsistance en ce domaine. C'est la réussite de l'ensemble du groupe qui détermine l'échelle à laquelle la valeur de chaque travail individuel sera rapportée. Il existe plusieurs techniques d'étalonnage permettant de calculer cette valeur relative ; elles ne sont pas toutes équivalentes quant à l'information qu'elles apportent mais celle-ci est cependant nettement définie à l'intérieur de chaque système de référence.

Dans les échelles les plus fréquemment employées, le chiffre qui exprime la valeur centrale (5 pour une échelle de 0 à 10 par exemple) correspond à la valeur médiane des notes brutes attribuées à un ensemble défini d'élèves qui ont subi le test dans les mêmes conditions, ensemble pris aussi grand que possible. L'intérêt de ce type de notation est double : il est indépendant de la difficulté de l'épreuve proposée, et c'est le seul qui permette de situer objectivement un élève par rapport à un groupe plus étendu que sa classe, en particulier de comparer les élèves de classes en principe équivalentes relevant d'établissements de types différents. Toute orientation, dès qu'elle est sélective, implique une référence commune, donc extérieure à chaque système particulier de valeurs. Si l'orientation des élèves telle qu'elle est habituellement pratiquée à la fin du premier cycle paraît satisfaisante aux défenseurs du système actuel, c'est qu'elle tient compte essentiellement de vérités de moyennes. Il est vrai qu'en moyenne les

enfants issus de certaines classes sont meilleurs que d'autres. C'est l'opinion générale des professeurs, fondée sur leur expérience professionnelle, qui se traduit très imparfaitement par les notes scolaires, du fait des fluctuations d'échelles, donc de moyennes, entre les classes. C'est aussi le résultat de comparaisons faites au moyen d'une métrique commune, qui confirme globalement cette opinion. Mais l'orientation intéresse des individus, et non des groupes. Le recouvrement des distributions de notes données à partir de critères objectifs (1) sont trop importants pour qu'on puisse nier la nécessité d'unifier l'étalon de jugement, afin que celui-ci soit réellement équitable.

Les structures scolaires, la physionomie sociologique de la classe, la pédagogie subissent actuellement les contrecoups des constantes remises en question d'une évolution générale rapide. Il ne nous semble pas que des recherches visant à contrôler des phénomènes collectifs puissent se passer des méthodes qui jusqu'à présent sont les seules à formuler les règles élémentaires de comparabilité, et sans doute les mieux adaptées à résoudre des difficultés qui sous cet angle sont parmi les plus importantes en psychologie.

Il est de fait que les tests de connaissances sont encore bien souvent un sujet d'opposition entre pédagogues et psychologues. Dans certains cas cette opposition est affaire de principe, l'apport révolutionnaire des tests, au meilleur sens du terme, apparaissant comme négatif à certains esprits nourris de tradition, en ce qu'ils détruisent à leurs yeux une certaine image de l'homme, et de leur mission d'éducateur. Les arguments avancés au long de ces quelques pages seront sans doute de peu de poids en face de ce type d'opposition. Celle qui relève par contre d'une irritation — légitime — en face d'une situation bâtarde n'est peut-être pas aussi définitive. C'est en tout cas notre conviction qu'un réel progrès en termes de mesure ne peut être atteint qu'au travers d'une coopération permanente. C'est déjà ce qui se passe en divers endroits, avec un bénéfice plus grand qu'une simple amélioration de la méthode : celui qu'apporte à chacun des membres de l'équipe une réflexion attentive sur ses propres idées, ses propres connaissances, et ses propres objectifs.

Jacqueline PELNARD-CONSIDERE.

(1) M. REUCLIN et F. BACHER. Enquête sur l'orientation à la fin du premier cycle secondaire - Rapport ronéotypé, I.N.O.P.

LA DOCIMOLOGIE EFFORT D'EXPLICITATION

Depuis les premiers travaux de docimologie réalisés avant 1930 par H. PIERON, D. LAUGIER et J. WEINBERG, une évolution sensible a marqué le développement de cette « étude systématique des examens ». La docimologie a vu son domaine s'élargir : elle ne concerne plus les seuls examens mais, de façon plus générale, toutes les méthodes utilisées pour porter sur les élèves une appréciation. Les recherches entreprises, d'abord exclusivement ou essentiellement critiques, s'efforcent aujourd'hui d'apporter des contributions positives. Cette évolution résulte en partie des circonstances (notamment de l'intérêt accru qui s'est attaché à l'orientation scolaire) mais en partie aussi de la logique propre aux recherches docimologiques elles-mêmes. En effet, si l'on a étudié systématiquement et expérimentalement les méthodes d'examen ou d'appréciation, ce fut d'abord pour **explicit**er le degré d'objectivité des notes ou jugements qu'elles fournissaient. C'est en se prenant à son propre jeu que la docimologie a évolué.

**

Le principe des premières expériences de docimologie (dont la répétition garde pour les enseignants un grand intérêt) est bien connu : deux appréciateurs différents notent indépendamment la même série de copies, et l'on compare leurs appréciations. Pour justifier le principe de telles expériences, on est conduit à distinguer explicitement les rôles pouvant être assignés à de telles appréciations.

Elles peuvent avoir tout d'abord des fins proprement pédagogiques. Chaque enseignant a ressenti la nécessité de faire varier le niveau de ses critères suivant les circonstances de la vie de la classe. Cette variation délibérée des échelles d'appréciation ne relève naturellement que de l'art du pédagogue, et le problème de l'objectivité de telles notations ne se pose pas.

Une appréciation peut constituer aussi un pronostic d'adaptation. Dans ce cas, elle ne tient pas seulement compte du résultat obtenu, mais des conditions dans lesquelles il l'a été (âge, conditions familiales, etc.) ainsi que de la nature et du niveau des exigences à satisfaire dans les études envisagées. Deux résultats identiques, obtenus dans des conditions différentes peuvent légitimement donner lieu à des appréciations-pronostics différentes. Mais la valeur, la « validité » de ces pronostics est ici vérifiable : il suffit de les comparer, le moment venu, aux résultats réellement obtenus.

L'expérience montre que leur valeur moyenne, pour des populations entières d'enfants, reste très faible.

Enfin une appréciation peut avoir pour objet un simple constat des résultats acquis. Dans ce cas, le critère de l'objectivité réside bien dans le degré d'accord entre appréciateurs indépendants, que les expériences docimologiques classiques permettent d'estimer.

Méthodologiquement, ces expériences obligent également à expliciter des notions comme celles de « valeur d'un pronostic » ou de « degré d'accord ». Il ne peut s'agir en effet de fonder ces notions sur une impression globale que chaque participant à ces expériences pourrait avoir éprouvée. Il ne peut suffire de justifier ces impressions par le récit de quelques cas. Alors que l'enseignement lui-même est d'autant plus efficace et vivant qu'il est plus individualisé, une expérimentation ayant pour but de mettre à l'épreuve une technique pédagogique (ici les procédés d'appréciation) n'a de sens que dans la mesure où son résultat décrit et analyse des faits généraux : c'est la population générale des élèves d'un certain niveau que cette technique concerne. Bien entendu, la même technique ne conduira pas au même résultat pour tous les élèves, un grand nombre de circonstances imprévisibles entrant en ligne de compte dans chaque cas particulier. Les jugements que l'on peut porter en la matière sont donc nécessairement de forme statistique. Les statistiques dont il s'agit n'ont rien en commun avec celles qui sont dressées à des fins administratives. La méthode statistique est ici une méthode d'analyse des résultats expérimentaux, qui est utilisée dans tous les domaines

où les phénomènes observés sont affectés par des sources de variation nombreuses et incontrôlables, notamment en biologie et en psychologie. La pédagogie expérimentale fait un large usage des méthodes statistiques dans tous les pays où elle est pratiquée de façon systématique. Son utilisation en docimologie permet de résumer des ensembles d'observations envisagés sous des aspects différents, qu'elle conduit à distinguer explicitement : deux professeurs peuvent noter autour de la même note moyenne, mais utiliser des échelles plus ou moins largement étalées ; ils peuvent utiliser deux moyennes différentes, mais cependant classer les élèves à peu près dans le même ordre, etc. Ces distinctions conduisent à préciser ce que l'on entend exactement par des expressions telles, par exemple, que « degré d'accord ». Elles permettent de distinguer explicitement la valeur que peuvent avoir des notes considérées dans les fonctions différentes : ordonner les élèves d'une classe ou situer chacun d'eux dans l'ensemble des élèves du département ou de l'Académie. L'expérience montre que les notes scolaires traditionnelles n'assument pas ces différentes fonctions avec la même efficacité.

Cet instrument d'analyse des observations qu'est la statistique permet aussi de comparer le degré d'objectivité ou la valeur pronostique de procédés d'appréciation différents. C'est la voie ouverte à une docimologie constructive, qui va rechercher quels sont les procédés les plus objectifs ou les plus valides parmi ceux qui sont employés, proposer et mettre à l'épreuve expérimentale des procédés nouveaux, en précisant explicitement dans quelles circonstances, à quels points de vue et dans quelle mesure ils peuvent apporter une amélioration aux procédures actuelles.

Des recherches entreprises dans cet esprit ont été peu favorables à certaines solutions couramment préconisées sur des bases intuitives. C'est ainsi que le fait de supprimer les notes chiffrées et de les remplacer par des appréciations verbales diminue le degré d'accord entre notateurs indépendants, au lieu de l'améliorer. Les entretiens d'un candidat avec un jury donnent lieu à des différences d'appréciation considérables lorsque les mêmes candidats sont examinés successivement par deux jurys. Les appréciations portées par les professeurs sur leurs élèves à l'issue d'une année de scolarité sont plus difficiles à étudier expérimentalement. On a pu constater cependant que la valeur pronostique des jugements d'instituteurs de Cours Moyen 2^e année, concernant la réussite de leurs élèves en 6^e et 5^e, est très faible, plus faible que celle d'un simple questionnaire de connaissances. Il est très probable que ce genre d'appréciations, à tous les niveaux de la scolarité, est

profondément affecté par les énormes différences qui ont été constatées entre le niveau **moyen** de classes portant la même dénomination, dans des établissements de même type.

En fait, deux catégories de mesures paraissent actuellement possibles pour améliorer, sur des points explicitement délimités, les procédures actuelles d'appréciation.

On peut tout d'abord fournir aux professeurs chargés d'examiner ou d'apprécier, des informations de nature statistique, susceptibles de faciliter leur tâche. Voici quelques exemples :

Si des candidats sont répartis au hasard (ordre alphabétique) entre plusieurs examinateurs les interrogeant sur le même programme, les différences entre les moyennes des notes attribuées par ces différents examinateurs ne peut être que faible, la différence la plus probable étant nulle. On peut donc proposer à chaque examinateur, à l'issue des interrogations, d'ajouter (ou de retrancher) à l'ensemble de ses notes le nombre de points nécessaires pour que sa moyenne individuelle devienne égale à la moyenne générale. Il convient de préciser que cette mesure est seulement la plus probablement équitable, qu'elle constitue seulement le meilleur parti que l'on peut prendre ; et d'autre part qu'elle n'égalise (dans ces limites) que le niveau moyen des appréciations des différents examinateurs, restant sans effet sur la dispersion de leurs notes (qui pourrait être égalisée par un procédé de même nature) et sur leur façon de classer les candidats.

Chaque professeur témoigne d'une relative stabilité quant au niveau moyen de son échelle d'appréciation. Il est sans doute possible d'informer chacun sur la position relative de sa moyenne habituelle, en améliorant la procédure de l'expérience consistant à soumettre un même échantillon de copies à un grand nombre de professeurs et à informer ensuite chacun de la distribution des notes observées pour l'ensemble des participants.

Il est possible également d'informer chaque professeur sur le niveau moyen de la classe qui lui est confiée. Il suffit de soumettre un large groupe de classes à une même épreuve normalisée dans sa présentation et son barème de notation. La publication de l'ensemble des résultats obtenus permet à chacun de situer sa propre classe (l'expérience peut être organisée de façon telle qu'il soit seul à connaître la position de sa classe dans l'ensemble). Cette information laisse au professeur la liberté de classer ses élèves comme il l'entend mais le met en mesure, si ses notes

doivent être comparées à celles qui sont attribuées dans d'autres classes, d'adopter un niveau moyen de notation qui tienne compte de la position de sa classe dans l'ensemble.

On voit que les deux derniers procédés qui viennent d'être suggérés présentent des limites comparables à celles qui ont été explicitées pour le premier : il s'agit toujours d'une approximation (que l'on peut seulement espérer améliorer puisqu'elle est ici définissable et peut être estimée par une procédure expérimentale) et d'une approximation portant seulement sur le choix d'une échelle numérique de notation. Cette seconde limitation est la plus sévère : rien, dans les procédés de cette première catégorie ne permet de rendre plus objectif, c'est-à-dire d'abord plus explicites, les critères qui conduisent un professeur à **classer** ses élèves, ou les copies qui lui sont soumises, dans un ordre plutôt que dans un autre. Ce problème est plus fondamental que le premier, en ce sens qu'il met en cause de contenu même des échelles de valeur utilisées, et plus seulement la forme, numérique ou non, sous laquelle elles s'expriment. Une seconde catégorie de procédés permet de l'aborder, sinon de le résoudre.

On a étudié des épreuves de forme nouvelle dans lesquelles la réponse du sujet à une question ne peut prendre qu'un nombre limité de formes différentes. Souvent, ces différentes réponses possibles sont même proposées au candidat, qui se borne à choisir entre elles. Une convention explicite peut donc ici être établie à l'avance, accordant un certain nombre de points pour chacun des choix possibles. La construction et l'expérimentation préalable d'une telle épreuve est une tâche difficile, exigeant la collaboration entre des spécialistes de la matière sur laquelle porte l'épreuve et un technicien spécialisé dans ce genre de travaux. Les avantages et les limites de ces épreuves doivent être rappelés.

Ces épreuves ne laissent évidemment plus aucune place aux divergences entre utilisateurs. Leur notation peut être confiée à un personnel non qualifié ou, plus avantageusement encore, à une machine. Appliquées à des populations étendues, elles peuvent fournir à chaque enseignant des informations statistiques précises sur la position de ses élèves dans la population à laquelle ils appartiennent. L'utilité proprement pédagogique d'informations de ce type ne doit pas être perdue de vue. Elles peuvent concerner en effet chacun des points du programme et permettre par conséquent à chaque enseignant d'être explicitement informé sur ceux de ces points pour lesquels sa classe présente en moyenne une lacune relative par rapport à son niveau moyen global. De telles

lacunes peuvent fort bien passer inaperçues si l'information du professeur se fonde uniquement sur des comparaisons entre ses propres élèves.

Les limites de ces épreuves normalisées paraissent aussi évidentes, mais elles valent qu'on y réfléchisse. Il n'est pas douteux qu'il est plus facile de construire de telles épreuves pour dresser un bilan de connaissances acquises que pour apprécier la façon dont ces connaissances sont utilisées ou les qualités d'esprit et de sensibilité dont l'enfant témoigne. Pour mesurer équitablement le poids de cette objection, il faut bien entendu tenir compte des progrès qui ont été réalisés dans la construction des épreuves de ce type. Une réponse de forme simple n'implique pas nécessairement que le raisonnement qui conduit à cette réponse soit lui-même simple. Une mise en œuvre complexe de connaissances diverses peut être exigée. D'autres questions peuvent fournir dans leur énoncé, toutes les connaissances nécessaires, et ne porter donc que sur les qualités d'esprit qui permettront de les utiliser (on rejoint là les tests d'intelligence et d'aptitudes des psychologues). Ce type d'objection reste cependant fondé, dans la mesure où l'on peut reprocher aux épreuves normalisées de ne pas refléter fidèlement les objectifs de l'éducation (et, en particulier, de susciter un travail de préparation dont l'esprit soit différent de celui que l'on entend donner à l'enseignement).

La portée de cette objection est alors considérable. Les épreuves dont il s'agit sont en effet, de nouveau, des procédures totalement explicites. Pour montrer qu'elles ne répondent pas aux objectifs de l'éducation, et pour permettre ainsi de les améliorer, il devient inévitable de formuler explicitement ces objectifs. C'est une tâche facile si l'on se borne à utiliser un vocabulaire très général, en évitant de se demander si des éducateurs différents entendent bien de la même façon les mots qu'ils emploient, et s'ils tirent bien les mêmes conséquences concrètes des principes qu'ils énoncent. Mais si, pris au jeu de l'explication, on se pose de telles questions, on est amené à craindre que les difficultés de la notation soient beaucoup plus fondamentales qu'il n'y paraît d'abord, et qu'elles mettent en cause l'idée que chaque éducateur se fait de la tâche qui lui est confiée. Des expériences ont été faites dans ce domaine, dont on voit combien elles élargissent la portée de la docimologie classique. Des professeurs enseignant en seconde ont été priés de définir explicitement l'état des connaissances qu'ils jugeaient souhaitable pour des élèves entrant dans cette classe. Leurs réponses ont été divergentes à un tel point que la définition des objectifs éducatifs de la classe de 3^e

semble à peu près indéterminable. On conçoit que la composition d'une épreuve de connaissances adaptée à ce niveau soit un problème qui ne comporte dans ces conditions aucune solution. Par des méthodes différentes la pédagogie américaine notamment travaille aussi à expliciter une « taxonomie des buts de l'éducation » pour chaque matière et à chaque niveau (BLOOM).

On voit comment la docimologie, en posant la question « Comment apprécier les résultats de l'éducation ? » conduit, par le simple jeu de sa logique propre, à poser une question beaucoup plus fondamentale. « Quels sont les buts de l'éducation ? ». Se trouvera-t-il des éducateurs pour affirmer que tous les buts de l'éducation ne peuvent pas être formulés explicitement, qu'ils ne **doivent** pas l'être ? L'adoption d'un tel postulat constituerait certainement une mauvaise attitude méthodologique. Je me range parmi les éducateurs qui penseront, en outre, qu'ils ont socialement le devoir de chercher à expliciter le rôle qu'ils jouent.

M. REUHLIN.

BACCALAURÉAT ET DOCIMOLOGIE

Les problèmes docimologiques que pose le Baccalauréat sont en principe ceux que pose tout examen. Comment remédier à la subjectivité des jugements portés par les examinateurs ? Comment assurer à tous les candidats des chances égales ? Comment ajuster le verdict des jurys aux exigences définies par le but même de l'examen ?

Cependant, le fait que le Baccalauréat est, comme on dit, un examen de masse donne à ses problèmes certains caractères spéciaux.

D'abord leur gravité évidente. De la solution plus ou moins efficace qu'il est possible de leur apporter dépend chaque année la destinée individuelle de milliers de jeunes gens au terme de leurs études secondaires. En dépend, par voie de conséquence, et tant que le Baccalauréat restera cet instrument de filtrage à l'entrée de l'Enseignement Supérieur, le rendement de cet enseignement.

La prise de conscience des problèmes docimologiques posés à ce niveau est pourtant récente. Il semble que pendant des générations la valeur des résultats obtenus n'ait pas été mise en cause et que les candidats malheureux se soient résignés sans protestation au verdict négatif de leurs juges. Les résultats de l'examen étaient-ils donc plus équitables ? Les examinateurs plus sages ? Les épreuves mieux choisies ? Rien ne permet de le croire, mais, le Baccalauréat, précisément parce qu'il était un examen de masse, bénéficiait de l'heureux effet psychologique de la loi des grands nombres. Il est bien certain, que pour l'ensemble de la France, ou même d'une Académie, les chiffres globaux sont rassurants. Le pourcentage des reçus — en dehors des secousses dues à des changements de régime ou de programme — ne variait, d'une année sur l'autre, que de quelques points sur 100, restant généralement au-dessus de 60 %, ce qui correspondait à peu près à l'idée que l'opinion publique se faisait des chances d'ensemble des candidats.

Mais au moment même où les travaux des « docimologistes » dénonçaient sur le plan scientifique les aléas de nos méthodes traditionnelles d'appréciation, les progrès de l'information découvraient au grand public des cas particuliers, jugés scandaleux. Les Associations de Parents d'Elèves relevaient, de candidat à candidat, de jury à jury, des inégalités de traitement jusque-là méconnues. Enfin l'administration elle-même, justement soucieuse de transformer un simple verdict en instrument possible d'orientation, en décidant

de fournir à chaque candidat, reçu ou refusé, le détail de ses notes, a donné à chacun le moyen de les comparer, non seulement avec celles de ses voisins mais avec celles qu'il croyait avoir méritées. La tentation devenait grande de mettre en accusation ses propres juges.

C'est pourquoi il est devenu nécessaire d'informer le public des problèmes docimologistes en général, de renoncer loyalement à toute prétention d'infaillibilité, mais en même temps de rassurer, dans la mesure du possible, l'opinion publique ébranlée, en la mettant au courant de tous les efforts qui sont faits pour obtenir, malgré tout, d'un examen qui existe depuis 159 ans, le maximum d'efficacité et d'équité.

* *

Le caractère aléatoire des notes d'examen reflète la difficulté, sinon l'impossibilité, de réduire en termes quantitatifs des faits mentaux. Mais comment le public accepterait-il cette évidence, alors que bien des professeurs, eux-mêmes, ne s'en rendent pas compte ? Obligés, à chaque instant de leur enseignement de noter leurs élèves — et donc d'affronter cette difficulté de principe — les plus réfléchis, les plus consciencieux, finissent par la considérer comme résolue. Au reste, tant qu'il s'agit d'enseignement, le problème est relativement simple. Les notes sont un moyen d'encourager, de blâmer, de classer : chaque professeur peut avoir son échelle qui, pour être différente de celle de ses collègues, n'en est pas moins efficace. Mais, le jour de l'examen, il s'agit d'apprécier et de juger par rapport à une norme. Il faudrait donc que la norme fut clairement définie, et l'échelle commune à tous les examinateurs. Et c'est bien là que surgissent les problèmes majeurs.

L'expérience quotidienne des organisateurs du Baccalauréat confirme, hélas, les enquêtes systématiques menées par les docimologistes (1) : si l'on soumet à des professeurs parfaitement qualifiés les mêmes lots de copies de Baccalauréat, les moyennes de leurs notes diffèrent couramment de 2 à 5 points sur 20 et parfois de 5 à 10 points. Il faut donc qu'un élève soit vraiment mauvais pour être sûr qu'il n'atteindra pas la moyenne en telle matière, vraiment bon pour être sûr de la dépasser. Pour les candidats qui ne sont ni franchement bons, ni franchement mauvais, le hasard intervient fatalement dans une large mesure, qu'il faut s'efforcer de réduire.

Que peut-on faire pour harmoniser les échelles de notes des professeurs ?

Comme il s'agit d'un examen de masse, la première façon d'harmoniser les notes d'un groupe de correcteurs dans une matière donnée c'est d'établir, entre eux, une moyenne, et de s'y référer. Il faut partir de la considération, démontrée par la statistique mathématique, que si l'on groupe, au hasard, 150 candidats au Baccalauréat ayant choisi même série et mêmes options, la valeur moyenne des épreuves fournies par chacun de ces groupes de candidats peut être considérée comme équivalente. Il en résulte que si les moyennes des notes attribuées dans une même discipline, par deux correcteurs affectés chacun à un de ces groupes, sont largement divergentes, cette divergence ne traduit pas une inégalité véritable entre les deux groupes de candidats intéressés, mais seulement le fait que les deux correcteurs n'utilisent pas la même échelle de notes ou, si l'on veut, que l'un est plus sévère que l'autre. Les jurys du Baccalauréat rassemblant précisément, depuis quelques années, de 150 à 200 candidats, on peut, après chaque session, faire la moyenne des notes données dans telle discipline par l'ensemble des correcteurs, puis établir la différence entre la moyenne des notes données par chaque correcteur et la moyenne générale.

(1) Le petit livre courageux et irréfutable de M. Henri PIERON « Examens et docimologie » (Presses Universitaires, 1963) en constitue le meilleur exposé, qu'il ne paraît pas utile de reprendre ici.

Quand cet écart est faible, le professeur est défini comme « modéré », quand cet écart est notable au-dessous de la moyenne, le professeur est défini comme « sévère », quand cet écart est notable au-dessus de la moyenne, le professeur est défini comme « indulgent ». Une fois établie, cette classification est précieuse ; elle ne rend pas compte de toutes les variétés d'examineurs — car il faudrait définir aussi les professeurs « timides » qui ne notent presque jamais moins de 8 ou plus de 12 sur 20, et les professeurs « audacieux » qui étalent leurs notes de 1 à 19 sur 20 — mais elle n'en est pas moins le point de départ d'opérations efficaces.

D'abord une information : le professeur est avisé, avant la session suivante, de la moyenne générale des notes données dans sa discipline ainsi que de la sienne propre ; il est donc amené à prendre conscience de sa sévérité ou de son indulgence relatives. Il peut ainsi, et il le fait souvent, en tenir compte de lui-même et aux prochaines sessions rectifier son échelle de notes.

Dans certaines Académies où l'emploi des machines électroniques et l'importance relativement faible des effectifs permettent des calculs très rapides, le professeur peut, au cours même de la session, être informé des moyennes provisoires obtenues et en tenir compte lorsqu'il établit ses notes définitives.

Mais ce genre d'information peut être utilisé d'une façon, si l'on peut dire, plus automatique, lors de la composition des jurys : le Baccalauréat est un examen complexe auquel concourent, dans chaque jury, un minimum de 6 professeurs. Ce pluralisme peut être un inconvénient si on abandonne au hasard sa composition, puisqu'un candidat risque de « tomber » sur un ensemble de professeurs sévères ou un ensemble de professeurs indulgents ; mais une fois les professeurs définis et classés comme on vient de l'exposer, leur grand nombre permet aux machines électroniques de les affecter dans les jurys d'après leur classification, donnant la priorité aux « modérés », compensant la « sévérité » des uns par l'« indulgence » des autres, sans négliger le coefficient affecté à chaque épreuve. Ainsi sont constitués, au moins dans une très importante Académie, des jurys « pondérés » et le candidat qui se présente devant l'un d'eux est assuré, en principe, de n'avoir ni plus ni moins de chances d'ensemble que ses camarades des autres jurys.

Ces précautions sont-elles suffisantes ? Elles réduisent certes, très sensiblement les inégalités à redouter mais ne parviennent pas à les effacer tout à fait. Si les machines ne peuvent faire plus, il reste encore aux hommes le moyen d'intervenir. Il se trouve, en effet, que depuis plusieurs années, les jurys sont présidés par des membres de l'Enseignement Supérieur qui sont responsables chacun de plusieurs jurys, groupés dans un même Centre d'examen. On peut déplorer que ces Présidents ne puissent ainsi, suivre dans le détail les délibérations de chacun de leur jury mais il faut, en revanche, se féliciter qu'ayant à contrôler plusieurs jurys à la fois, ils en puissent comparer les résultats : circulant de l'un à l'autre au moment des délibérations qui suivent l'écrit et, pendant l'oral, calculant et notant des moyennes provisoires, conseillant ici plus de fermeté, là, plus de largeur de vue, bien souvent leur action ramène à une homogénéité acceptable des résultats d'abord divergents. Et, lorsque cette homogénéité est obtenue dans un groupe de jurys, elle définit une moyenne pratique à laquelle il est bien rare que ne correspondent pas les moyennes obtenues de la même sorte dans d'autres Centres, par d'autres groupes de jurys.

Il est donc possible, au Baccalauréat, d'égaliser dans une large mesure les chances des candidats en partant des copies rédigées et des Interrogations subies, mais ces copies elles-mêmes, ces interrogations, correspondent-elles toujours au niveau réel des candidats ? Traduisent-elles automatiquement leur valeur ? On sait malheureusement que non. Traumatisés par l'examen, ou déconcertés par un sujet qu'ils n'avaient pas prévu, il arrive trop souvent que

jeunes gens et jeunes filles ne donnent pas, au jour des épreuves, leur vraie mesure. L'accident est classique, et il en faut tenir compte. C'est pourquoi aucun candidat ne peut être refusé sans que le jury ait examiné son « dossier scolaire ». Ce document, si imparfait qu'il soit encore, fournit pourtant l'appréciation de ses professeurs, évidemment subjective, ses notes de composition, évidemment discutables — nous n'y reviendrons pas — au même titre que des notes d'examen, mais aussi son classement dans chaque matière par rapport à l'ensemble de ses camarades. C'est l'étude de ce classement qui permet bien souvent de déceler l'« accident » et amène les correcteurs intéressés à rendre des points à un candidat malchanceux, et le jury tout entier à lui en prêter pour l'oral lorsque les règlements les y autorisent. Cette interprétation du livret scolaire, parce qu'elle fait intervenir une motivation qui n'est pas le résultat d'une simple opération arithmétique inquiète souvent les candidats et leur famille, prompts à dénoncer comme arbitraire une appréciation non chiffrée. Sur ce point ils ont bien tort ; d'abord parce que cette décision n'intervient pratiquement qu'au bénéfice des candidats, ensuite parce que l'expérience des membres des jurys leur fait discerner, presque à coup sûr, l'accident véritable, enfin, parce que là encore, l'action d'un Président commun à plusieurs jurys lui permet d'harmoniser le comportement de ceux-ci.

Il semble donc que, si le caractère aléatoire de chacune des notes données dans chaque matière reste un fait indiscutable, l'organisation même du Baccalauréat permet de réduire, dans une très large mesure, les inconvénients du phénomène.

**

Le problème qui reste à résoudre est cependant, nous devons le dire, le plus difficile à affronter : c'est celui du niveau même de l'examen. A partir de quel niveau réel d'aptitudes et de connaissances un candidat mérite-t-il d'être reçu ? En dessous de quel niveau doit-il être ajourné ? Ou, en termes plus concrets, à quoi doit correspondre la moyenne 10/20, traditionnellement fixée comme critère d'échec ou de succès ? Tout dépend de la réponse qui sera faite à la question suivante et désormais banale : le candidat doit-il faire la preuve qu'il a achevé « normalement » le cours de ses études secondaires ? Doit-il faire la preuve qu'il est apte à suivre avec profit un Enseignement Supérieur déterminé ?

Contentons-nous, pour l'instant, de décrire un état de fait : le Baccalauréat que passent nos enfants reste, malgré la spécialisation relative de ses séries, un examen de culture générale où la « force » en français peut compenser la « faiblesse » en philosophie ou en mathématiques, où une bonne note en dessin (matière facultative) peut racheter à l'oral une défaillance en latin, où des points négatifs en éducation physique peuvent compromettre le résultat final. Dans ces conditions, les professeurs du second degré qui corrigent toutes les épreuves et composent l'ensemble du jury, se réfèrent plus ou moins consciemment à la notion d'« élève moyen », étant défini comme élève moyen celui qui, dans leur classe ou dans l'ensemble des classes dont ils ont l'habitude, tiendrait le milieu d'un classement par valeur, serait par exemple 15^e sur 30. C'est pourquoi beaucoup de correcteurs au Baccalauréat, et nul ne saurait les en blâmer, commencent par classer leurs copies corrigées, mais non notées, par ordre de valeur ; attribuent la moyenne 10 à celui qu'ils ont classé, disons 50^e sur 100, et répartissent ensuite leurs notes au-dessus et au-dessous de cette moyenne. Si cette façon de procéder était générale et systématique, la moitié des candidats obtiendrait automatiquement 10/20, ce qui donnerait 50 % de reçus et 50 % de refusés. De bons observateurs soutiennent qu'en gros c'est bien ce qui se passe, que cette notion d'« élève moyen » reste en fait le seul critère effectif, même

s'il demeure inconscient, mais qu'il est légèrement modifié par la bienveillance naturelle des examinateurs et la crainte d' « erreurs judiciaires », ce qui explique le pourcentage traditionnel des reçus, plus ou moins égal à 60 %.

Mais, dira-t-on, les Présidents de jury qui appartiennent, eux, à l'Enseignement Supérieur, n'ont-ils pas une autre optique ? Sachant qu'il s'agit de leurs futurs étudiants, ne font-ils pas prévaloir d'autres exigences ? N'ont-ils pas le souci de dégager des aptitudes particulières plutôt qu'un nouveau moyen de connaissances ? Il est certain que le législateur, en faisant présider des jurys composés de professeurs du second degré par un représentant de l'Enseignement Supérieur a voulu, précisément, maintenir, au Baccalauréat, ce souci de sélection. Il faut cependant reconnaître que dans la pratique ce sont les professeurs du second degré qui donnent d'abord toutes les notes et que, si le Président, comme nous l'avons vu, peut intervenir utilement pour les harmoniser, il lui est difficile d'en modifier l'esprit. L'on rencontre bien rarement — si l'on en a jamais rencontré — un Président de jury « scientifique », par exemple, qui intervienne pour faire baisser la note et provoquer l'échec d'un candidat qui ne lui aurait pas paru apte à entrer par la suite dans la Faculté de Sciences que lui-même représente.

En fait, si l'Enseignement Supérieur conserve une certaine action sur le niveau du Baccalauréat et, pourrait-on dire, son « réglage » au point de vue de l'admission en Faculté, c'est par d'autres procédés, par d'autres précautions.

D'abord par le choix des sujets qui peuvent être plus ou moins faciles. Sans doute, dans les matières littéraires, le correcteur, par la considération inconsciente de l'élève moyen, proportionne-t-il presque automatiquement sa note à la difficulté du sujet (par exemple à la longueur de la version latine et au nombre de ses « pièges »). Mais le correcteur des matières scientifiques peut être, lui, tenu par un barème obligatoire qui, figurant sur le sujet, l'empêche d'opérer cette compensation. C'est pourquoi, soit dit en passant, les moyennes des notes des matières littéraires sont, d'une session sur l'autre, plus constantes que les moyennes des matières scientifiques. Il est certain que pour ces dernières, les Commissions de Choix des sujets où l'Enseignement Supérieur est fortement représenté, restent, dans une assez large mesure, maîtresses d'ajuster selon leurs vues le niveau du Baccalauréat. Leur action, en ce domaine, est d'ailleurs délicate : étant donné le grand nombre des candidats dont les notes restent groupées à peu de distance de la moyenne, un relèvement de la « barre » même très léger modifie dans une mesure que les candidats malheureux ont pu parfois juger catastrophique le pourcentage des échecs et des succès. Un coup de pouce en sens inverse peut faire accuser l'Université de « brader » son Baccalauréat.

La difficulté du problème n'est pas méconnue ; c'est pourquoi dans toutes les Académies se réunissent avant la correction des copies, des « Commissions d'entente » où, associés autour de la même table, professeurs de l'Enseignement Supérieur et professeurs de l'Enseignement Secondaire s'efforcent, pour les matières scientifiques, d'affiner les barèmes de notation obligatoires et, pour les matières littéraires, de déterminer les qualités qui doivent hausser un candidat au-dessus de la moyenne, les fautes qui doivent lui interdire d'y atteindre.



S'il est permis de tirer quelques conclusions d'ensemble des considérations qui précèdent, l'on pourrait dire, que si les docimologistes n'ont que trop raison de souligner le caractère subjectif des notes données au Baccalauréat comme à tout autre examen, la bonne volonté des examinateurs et l'ingéniosité des organisateurs permettent cependant, grâce à l'ampleur même des effectifs mis en cause et grâce au nombre des épreuves, de réduire dans une très large mesure des inconvénients qui tiennent à la nature même des choses.

En ce qui concerne, par contre, les critères de niveau, critères dont dépendent en dernière analyse, l'échec et le succès de chacun, il paraît bien difficile de trouver, dans la situation ambiguë où nous restons placés une solution vraiment satisfaisante.

Comment sortir de l'ambiguïté ? Faut-il définir désormais le Baccalauréat comme la consécration d'études secondaires normalement menées à leur terme ? En ce cas, rien n'empêcherait de fixer a priori un pourcentage de succès tel que la possibilité d'un échec ne soit plus qu'une menace pour les paresseux. Mais il faudrait évidemment juxtaposer à ce Baccalauréat de l'Enseignement Secondaire un examen d'aptitude à l'Enseignement Supérieur, passé soit immédiatement après le Baccalauréat, soit au terme d'une année d'études spécialisées, et ne portant que sur les matières où le candidat envisage de poursuivre ses études.

La vraie solution ne serait-elle pas de démythifier enfin le Baccalauréat, de lui faire perdre ce caractère de « rite de passage » analogue aux épreuves douloureuses, de caractère magique qui, dans les sociétés primitives, marquent le passage de l'adolescence à l'âge adulte ? Replacé dans la perspective d'une orientation continue et progressive, le baccalauréat que l'on passerait à 18 ans ne serait plus que l'un des « bilans » périodiques qui permettraient à chacun de mesurer ses aptitudes et ses chances, à la société de choisir selon ses besoins. En retirant au Baccalauréat son caractère redoutable de « verdict » pour en faire un instrument de mesure et d'orientation, on résoudrait du même coup le plus grave des problèmes docimologiques que pose son organisation actuelle.

Jean PASQUIER.

COMMENT AMÉLIORER LA NOTATION

Les études conjointes montrent avec pertinence l'impossibilité de se fier aux notes comme moyen de classement et de sélection.

Une analyse qualitative de la note, comme produit d'un jugement en situation pédagogique permettra peut-être d'esquisser quelque élément de solution.

La note est la concrétisation d'une situation complexe ; le professeur qui note une copie d'élève exprime, en ce geste simple en apparence, une multitude de relations et de régulations complexes. Il s'agit tout d'abord, sémantiquement, d'exprimer une comparaison entre un contenu intellectuel idéal et ce que le devoir noté en recèle : la note exprime la distance entre le corrigé type, tel que l'a conçu le notateur, et l'essai imparfait de l'élève.

Cette distance est donc fonction à la fois d'une norme, posée a priori — le contenu d'un programme — et le niveau d'une production individuelle — celle de l'élève noté.

Il va de soi que cette vue objective des choses est simplifiée. Le programme tout d'abord ne saurait être défini de façon impersonnelle. Aussi précis soit-il, il en est toujours une interprétation par le professeur qui l'assimile et l'enseigne avec sa nature propre, et sa culture : ici apparaissent déjà des facteurs de personnalité qui ne relèvent pas uniquement du contenu intellectuel. Car le programme peut être vu par le professeur en situation de communication ou ne pas l'être. L'exigence plus ou moins grande d'un professeur auprès de ses élèves peut résulter simple-

ment d'un niveau de culture plus ou moins élevé. Mais il peut aussi résulter d'une estimation plus ou moins fine de ce qui, dans le programme abstraitement défini, peut être assimilé par un élève de tel âge donné. Le summum de l'art sera précisément l'attitude du professeur qui, en pleine possession de sa matière, en sait mesurer l'insertion psychologique chez les élèves qui lui sont confiés. Or, cette capacité d'adaptation relève plus de la personnalité et de l'affectivité, que du savoir.

La polarisation sur la matière, l'insensibilité à l'élève comme personne en devenir résulte souvent d'un certain autisme, d'une sécheresse affective manifestée comme justice et sévérité. Inversement un certain laxisme, une générosité habituelle dans la notation peut aussi bien traduire une méconnaissance de la matière enseignée qu'une sensibilité affective du maître aux enfants qui lui sont confiés.

Et certes, ces attitudes fondamentales tissent la trame de fond de variations plus ténues selon les jours et les heures. Le professeur est un homme, comme tous les hommes. Sa « sévérité » ou sa « bienveillance » naturelles peuvent subir les variations de l'humeur momentanée : les notes en sont le reflet.

L'estimation du « devoir » introduit également la plupart du temps des paramètres affectifs nombreux et plus fins.

La présentation de la matière, la correction du style, la lisibilité d'une écriture, l'allure d'un raisonnement, ont souvent une part prépondérante dans un jugement global qui relève de l'estimation plus que de la mesure. Il est certes possible d'esquisser une atomisation de la matière et une explicitation des critères. Dans la pratique habituelle, cet effort d'objectivisation reste toujours relativement inefficace.

Les comparaisons et régulations que nous venons de décrire ne dépassent pas le dialogue élève-maître. Mais une autre dimension intervient presque toujours avec la classe, comme phénomène collectif. Le maître notant ne compare pas seulement un programme et un essai approché mais également les essais entre eux de disciples confrontés collectivement à un même programme. Ce facteur permet d'ailleurs la relativisation évoquée plus haut, dans la mesure du moins où l'attitude fondamentale du maître se prête à cette régulation. Un blocage excessif sur la matière intellectuelle et une conscience trop aiguë de l'idéal souhaitable peut en effet conduire au mépris de l'élève en collectivité et à ces

compositions dont le premier affleure la moyenne sans y parvenir. Il y a des « bonnes » et des « mauvaises » classes et la notation ne peut que refléter ces inévitables comparaisons.

Le maître qui note peut enfin tenir compte de tiers plus ou moins lointains : les parents d'élèves auxquels la note est finalement destinée par l'intermédiaire du carnet de correspondance ou du bulletin trimestriel. Il est juste de dire que ce facteur tient peu de place dans la régulation décrite. On peut s'en réjouir, ou le déplorer. Le professeur notant pense toujours à l'élève, mais rarement aux parents sauf quand une aide indiscreète a par trop perturbé les relations de ce binôme privilégié.

Ces attitudes affectives du notateur sont inconscientes. Elles interviennent spontanément, à l'insu du notateur et des élèves. La note et surtout la notation dans son ensemble les révèlent à l'observateur. Il en va autrement pour l'élève qui prépare, attend, et reçoit la note dans un contexte affectif, conscient et plus ou moins dramatisé selon le caractère, l'éducation, le milieu familial ou le milieu scolaire. Si le maître prétend noter objectivement et n'affectiviser sa notation qu'indirectement et inconsciemment, on peut dire que l'élève reçoit toujours la note dans un contexte affectif et qu'en bonne pédagogie traditionnelle, cette affectivité est nécessaire. La note est en effet dans nos établissements scolaires la récompense de l'effort, la punition de la paresse, la sanction de l'incapacité. Il est inutile, pensons-nous, d'insister longuement sur le caractère pernicieux de ces habitudes inhérentes à un système d'éducation fondé exclusivement sur l'émulation. L'objet d'étude et l'étude elle-même deviennent seconds. La note qui devrait être un signe devient l'objet même de tout effort ou des ruses qui permettront de s'en dispenser. Dès que la notation cesse, le travail cesse. Les mauvaises notes accumulées engendrent « les cancre » ceux qui ont perdu tout espoir, pour lesquels le travail scolaire a perdu tout son sens puisqu'il entraîne infailliblement le déshonneur ou, dans le meilleur des cas, ceux qui se font un honneur de toujours avoir de mauvaises notes.

La note, dans le système scolaire actuel, n'est donc jamais un simple constat. Pour l'élève, comme pour le maître, elle est chargée d'affectivité et résume en un instant une situation complexe où s'équilibrent les attitudes et les intentions, où se matérialise aussi une certaine conception philosophique de l'instruction et de l'éducation.

C'est pourquoi les efforts pour changer le système de notation trouvent leur origine dans deux préoccupations étrangères à

ce système traditionnel ; l'une qui découle des préoccupations de planificateurs, soucieux de mieux adapter l'école à sa fonction sociale d'orientation, l'autre qui résulte naturellement d'une philosophie de l'éducation axée sur l'épanouissement de la personnalité et l'exercice fonctionnel des aptitudes individuelles. Les deux courants peuvent sembler se conjuguer parfois. Ils s'opposent en réalité profondément.

Dans la première perspective, l'objectif est de parvenir à désobjectiver la notation de manière à permettre un classement aussi rigoureux que possible des individus confrontés à une tâche déterminée. L'idéal devient ici la composition ou l'examen « aseptisé ». Deux opérations sont jugées indispensables à cet effet. Tout d'abord une analyse aussi poussée que possible de la matière à enseigner de telle sorte qu'aucune place, idéalement, ne soit laissée à l'interprétation du maître ou de l'élève. La matière, objet de notation, sera donc atomisée autant qu'il se peut. La question elle-même devra être posée et la forme de la réponse précisée de telle façon qu'aucune ambiguïté manifeste ne puisse subsister. L'épreuve devra donc se présenter comme une série de questions successives portant sur un point précis de la matière explorée. Les réponses seront à choisir parmi plusieurs possibles offertes à l'attention de l'élève. Ce choix fixé rend impossible les imprécisions et les interprétations, du moins pour le correcteur. La manière dont le choix a été fait par l'élève reste évidemment obscure mais ce point est considéré comme sans intérêt, seule pouvant et devant compter la manifestation objective, en acte, de l'activité intellectuelle suscitée.

Ce système doit avoir une seconde vertu : désobjectivant le jugement porté sur le contenu, il permet également l'objectivité dans la comparaison et le classement. Les questions à choix multiples permettent une quantification rigoureuse, partant de la fusion, pour comparaison dans une même population, d'élèves issus de classes les plus diverses. Les scores possibles sont ainsi ordonnés et exploités statistiquement. Les machines peuvent intervenir pour le classement. Les niveaux exigibles ne sont plus seulement définis par un programme théoriquement et abstraitement défini, mais par l'assimilation effective de ce programme par les populations scolaires réelles confrontées à ce programme.

Tel est le système déjà largement introduit à l'étranger et qui, peu à peu, chemine dans l'enseignement supérieur français. Les résistances, il est vrai, sont vives et les reproches faits méritent qu'on s'y arrête.

Il serait aberrant en effet d'infléchir les fins et la pratique de l'enseignement pour les besoins de la classification et de la quantification. Or tel est bien le cas, de l'avis de bon nombre d'enseignants français, notamment parmi ceux qui reprochent au système en vigueur de ne pas faire une place suffisante à l'éducation. A leurs yeux, l'introduction des questions à choix multiple ne ferait qu'accuser encore, en les rationalisant, les défauts majeurs du système actuel. La quantification objective conduit infailliblement à l'atomisation des réponses. Elle privilégie du même coup le savoir parcellaire au détriment de l'activité intellectuelle d'organisation et de synthèse. Or l'enseignement traditionnel, polarisé par la matière, n'a déjà que trop tendance à exagérer l'importance du savoir et à stériliser faute d'emploi l'activité propre de l'élève. Les QCM sont dans la ligne de l'encyclopédisme dévorant dont souffrent déjà à l'excès nos enseignants de tous niveaux. Et certes, il est très difficile de porter un jugement objectif sur les aptitudes à l'organisation, sur la créativité, sur le degré de disponibilité et sur la structuration mentale fondamentale. Ce n'est certes pas une raison pour s'en désintéresser et faire dépérir ces qualités faute de les valoriser par la notation et l'examen.

Il semble bien d'ailleurs que tout le problème soit là, dans cette fâcheuse tendance à suspendre toute l'activité pédagogique à l'obtention de la note, à l'examen, et au concours. C'est la raison pour laquelle le désir de dépasser le système de notation actuel peut s'orienter non plus vers la recherche d'une plus grande objectivité mais vers la dédramatisation de la notation et son adaptation à une pédagogie d'éveil et de formation de la personnalité.

On s'efforcera, dans cette perspective, d'éliminer les incidences du classement sur le développement des attitudes d'émulation. L'important, dans la note, ne sera plus sa place relative dans l'ensemble des notes de la classe, de l'établissement, ou des candidats, mais la signification que chacun lui donne dans un destin personnel. Il s'agit ici de permettre à chacun de se situer d'abord par rapport à la norme objective impersonnelle, ensuite par rapport à lui-même dans une perspective temporelle. Il convient également de dépasser les aléas d'une notation instantanée et parcellaire pour réintégrer des facteurs permanents d'aptitude et de disponibilité générale. Une notation doit donc être à la fois moins ponctuelle et plus fine. Elle doit résumer un jugement porté sur une disponibilité durable mais elle doit aussi essayer de diversifier l'éventail de ces disponibilités.

Ainsi se précise la nécessité d'un dossier scolaire qui puisse

situer la notation dans un contexte général (milieu familial, santé, équilibre affectif), qui puisse résumer un devenir intellectuel par rappel des années antérieures, qui puisse diversifier les appréciations en dépassant le verdict brutal et aveugle de la « composition », en permettant une quantification des aptitudes diverses mises en œuvre par l'exercice de chaque matière (par ex. : correction du style, de l'orthographe, imagination, sensibilité littéraire, sens et goût de l'observation...), qui puisse enfin permettre de situer l'élève parmi ses condisciples par cette façon complexe et nuancée.

Ces nécessités multiples justifient également pour une part l'abandon de la notation traditionnelle sur 20 pour l'adoption plus souple et plus parlante du système des « catégories ». La notation en 4, 5 ou 7 catégories présente en effet plusieurs avantages. Elle permet au notateur une prise de conscience plus facile de son équation personnelle par la possibilité offerte de grouper les notes attribuées dans le cadre d'une classe (30 à 40 élèves) en histogramme significatif (la dispersion sur 20 et l'usage des demi points rend souvent impossible cette répartition). Elle oblige également le notateur à prendre du recul, à ne pas se perdre dans le détail d'une notation, l'autorisant ainsi à essayer la quantification du qualitatif c'est-à-dire celle des activités fondamentales d'organisation, de raisonnement, de créativité. Elle permet une représentation graphique plus parlante du devenir intellectuel de chacun, représentation indispensable pour l'auto évaluation ou pour l'orientation.

Est-ce à dire que le dossier scolaire et la notation en catégorie doivent exclure radicalement le système des QCM ? Nous ne le pensons pas. L'un et l'autre procédé ont leurs vertus distinctes et complémentaires dans la mesure où ils cherchent à dépasser la complexité anarchique et aveugle d'une notation composite en rationalisant l'un ou l'autre des aspects de cette notation. Le besoin de mesurer des connaissances conduit aux questions à choix multiple. Le besoin d'estimer des aptitudes générales conduit à la notation en catégorie et au dossier scolaire. L'important est de ne pas exclure l'un au profit de l'autre et d'équilibrer avec sagesse l'importance respective de ces deux facteurs selon le niveau de l'enseignement et les finalités propres à ce niveau. Un enseignement technologique ou paratechnologique devra privilégier les QCM. Un enseignement de culture générale, la notation en catégories. Le dossier scolaire devrait être le lieu de rencontre d'apports ainsi diversifiés.

Louis LEGRAND.

Bibliographie

- BACHER (F.) et REUCHLIN (M.). — Le cycle d'observation. Enquête sur l'ensemble des élèves d'un département (Loiret 1960-62). *B.I.N.O.P.*, 1965, 3.
- BONNARDEL (R.). — Etude des « aptitudes » scolaires dans l'enseignement secondaire au niveau de la seconde et de la troisième. *J. de Psychologie*, 1953, 1, 62-86.
- CHATEAU (J.). — Sur quelques notes d'examen. *Enfance*, 1957, 1, 67-69.
- CHAUFFARD (C.). — Contribution aux études docimologiques. *B.I.N.O.P.*, 1948, mai-juin 83-84.
- DAGUE (M.). — Le régime des examens en France, problèmes docimologiques. *Bull. de psycho.*, 1951-52, 7, 405-407.
- LAUGIER, PIERON (H.), Mme PIERON, TOULOUSE, WEINBERG. — Etudes docimologiques sur le perfectionnement des examens et concours. Paris, C.N.A.M., 1934, 89 p.
- GILBERT (L.). — La mesure des acquisitions scolaires. *Cahiers de Pédagogie et d'O. P.*, 1953, 1, 20-33.
- LARCEBEAU (S.). — Eléments de construction d'une fiche scolaire. *B.I.N.O.P.*, 1955, 2, 133-143.
- LARCHEVEQUE (P.). — Sur la mesure du rendement scolaire. *Pédagogie-Orientation*, février et avril 1950, 28-32.
- LAUGIER et WEINBERG. — Le facteur subjectif dans les notes d'examen. *Année psychol.*, 1930, 229-241.
- MAUCORPS (P. H.). — Le groupe de psychométrie pédagogique. Pour l'ère nouvelle, Bourrellet, Paris, 1946, 5, 93-95.
- PIERON (H.). — La docimologie dans les examens et les concours. *Lycées et collèges*, 1956, 4, 110-121.
- PIERON (H.). — La docimologie ou le problème des examens. *Courrier rationaliste*, 1962, 1.
- PIERON (H.). — Données docimologiques sur la double correction et la typologie des correcteurs. *Biotypologie*, 1957, 1-2, 128-145.

- PIERON (H.). — Education nouvelle et docimologie. *Univers. Nouv.*, 1931, 52, 69-73.
- PIERON (H.). — Examens et Docimologie. P.U.F., Paris, 1963, 191 p.
- PIERON et LAUGIER. — Valeur sélective du certificat d'études. Comparaison de cet examen avec une épreuve par tests. Conférence internationale de psychotechnique, IV, 1927, Paris, Comptes-rendus, voir p. 507.
- PIERON (H.), PICHOT, FAVERGE, STÖTZEL. — Méthodologie psychotechnique, P.U.F., 1952, voir PIERON, p. 123.
- PIERON (H.) et REUCHLIN (M.). — Le problème de la signification et de la valeur des examens oraux. *Biotypologie*, 1962, 1-2, 1-16.
- PIERON (H.). — O. P. et docimologie. *B.I.N.O.P.*, 1929, 6, p. 161.
- PIERON (H.). — La technique des examens et la nécessité d'une « docimologie ». *Prophyl. Ment.* IX, 39, juillet-décembre 1934, 138-141.
- PIOBETTA (J.B.). — Examens et concours. P.U.F., Paris, 1943, 113 p.
- RAPPORT fait en conclusion des travaux de la commission de contrôle créée en vertu de la résolution adoptée par le Sénat le 21-4-1966 et chargée d'examiner les problèmes d'orientation et de sélection dans le service public de l'enseignement. Paris, J. O., 1966, 2 vol.
- REUCHLIN (M.). — Une enquête sur l'état des connaissances dans les classes des cours moyens 2^e année. *Enfance*, 1958, 2, 161-163.
- REUCHLIN (M.). — Etudes docimologiques. De l'enseignement primaire à l'enseignement secondaire. *B.I.N.O.P.*, 1958, n° spéc.
- REUCHLIN (M.). — Hommage à H. PIERON. Examens et concours. *Cahiers rationalistes*, 1965, 227, 73-120.
- REUCHLIN (M.). — L'orientation pendant la période scolaire. Idées et problèmes. L'éducation en Europe, série II, Strasbourg, Conseil de l'Europe, 1964, 376 p.
- SYMPOSIUM au XIII^e congrès de l'A.I.P.A., Rome, 1958. Compte rendu p. 9, section III : Recherches docimologiques du Service de Recherches de l'I.N.O.P. *Travail humain*, 1959, 1-2, 1-158.
- SYMPOSIUM au XV^e congrès de l'A.I.P.A., Ljubljana, 1964 : Les problèmes méthodologiques de l'évaluation des connaissances scolaires. *Travail humain*, 1965, 3-4, 193-390.
- VALIN (E.). — Examen des examens. *B.I.N.O.P.*, 1960, 5, 315-324.

THÈMES D'EXPÉRIENCES

I. — EXPERIENCES A REALISER DANS UN ETABLISSEMENT

1. — Un professeur corrige un paquet de copies de ses propres élèves ; il les note ; il se garde d'annoter les copies et d'y porter la note ; il repère l'ordre dans lequel il les a corrigées.

— Il corrige les mêmes copies à peu d'intervalle, après avoir brouillé l'ordre initial.

— Il compare les deux séries de notes. Les deux séries sont-elles comparables ? Y a-t-il eu des « effets de contraste » ?

2. — Un professeur corrige des copies, de préférence d'élèves qui lui sont inconnus, sans les annoter ; les notes sont mises sur « feuille à part ». Il répète 2, 3, 4 fois l'opération, à intervalles d'un mois. Les séries de notes sont-elles comparables ?

3. — Plusieurs professeurs corrigent, sans les annoter, les mêmes copies. Ils confrontent leurs notes et tirent les conclusions qu'impose cette confrontation.

CAS PARTICULIER : Celui de la double correction.

Deux professeurs corrigent les mêmes copies ; comparent les notes mises ; discutent et se demandent quelle est la valeur logique (et morale) de la moyenne des deux notes. Ne pourrait-on pas trouver la « note définitive » sans prendre la moyenne des 2 notes ?

4. — Dans une classe, on fera subir aux élèves des tests de connaissances étalonnés (en mathématiques, par exemple) et on comparera les résultats de ces tests à ceux des résultats sco-

lares traditionnels. On discutera de l'importance de la convergence (ou non) des résultats donnés par les divers moyens d'appréciation des travaux scolaires.

II. — EXPERIENCES A REALISER DANS UN SECTEUR

1. Faire corriger 3 copies sur un même sujet, a priori bonne, moyenne, faible, par tous les professeurs (de la discipline et de la classe) du secteur.

Rassembler les notes pour chacune des trois copies.

Etablir les histogrammes.

Les diffuser.

Expérience à répéter à plusieurs reprises pour que les professeurs arrivent à corriger, si possible, ce qu'on pourrait appeler leur « équation personnelle ».

2. Pratiquer dans toutes les divisions d'une même classe d'un secteur des épreuves de connaissances standardisées et étalonnées (1) afin de comparer la valeur statistique de ces divisions.

A partir de ces épreuves, dresser l'histogramme de la classe, pour le secteur, dans la discipline considérée.

III. — SUGGESTIONS DIVERSES.

1. Organiser des conférences pour les jeunes enseignants afin de les initier à la science de l'examen. Leur donner des connaissances élémentaires de statistique ; leur faire construire et interpréter des histogrammes.

2. Réunir les professeurs dans des commissions de travail pour qu'ils comparent leurs systèmes de notation et qu'ils envisagent les moyens de les équilibrer.

3. Au cours de ces réunions, se demander si la nature même des exercices scolaires traditionnels n'est pas à mettre en cause, autant, peut-être, que la note même qu'on leur attribue.

Tenter de définir des exercices ayant des dominantes ; par exemple, le sens de l'observation que peuvent révéler des exercices de sciences, de travaux manuels, de géographie, de français, etc.

4. Rompre avec certaines traditions, à titre expérimental :

— abandonner la notation de 0 à 20,

— abandonner les compositions traditionnelles,

— abandonner les moyennes générales,

— abandonner le classement général.

Pratiquer les interrogations écrites et orales et le classement en catégories.

(1) Ces tests donnent donc chaque élément d'un ensemble scolaire en valeur absolue.



VIE DE SÈVRES

● MISSION A L'ETRANGER :

QUEBEC :

M. AUBA, Directeur du Centre International d'Etudes Pédagogiques s'est rendu en mission au Québec, en compagnie de M. Louis LEGRAND, Chef du département de la Recherche Pédagogique à l'Institut Pédagogique National. Cette mission avait pour but d'étudier les mesures à prendre en vue d'appliquer l'article 3 du procès-verbal des décisions arrêtées le 14 septembre 1967 par les autorités des deux pays, article qui prévoit la création d'un « Centre franco-québécois de développement pédagogique ».

Le Centre aura pour tâche :

a) de confronter les expériences pédagogiques des deux parties, en matière de programmes, de méthodes, de manuels, de matériel didactique, de formation des maîtres, etc.

b) d'organiser des expériences conjointes dans le domaine pédagogique.

c) d'étudier l'harmonisation des scolarités à tous les niveaux.

d) de coordonner l'organisation des stages de formation et de perfectionnement des maîtres, et des colloques sur les problèmes pédagogiques.

Les deux gouvernements participeront au fonctionnement du Centre par le détachement de spécialistes, la fourniture de matériel pédagogique et le financement de ces opérations. La France

est disposée à recevoir les pédagogues québécois au Centre International d'Etudes Pédagogiques de Sèvres, où ils trouveront un milieu de recherche pédagogique susceptible de les intéresser et d'où ils pourront rayonner dans les divers établissements qu'ils souhaitent voir.

Les thèmes susceptibles de faire l'objet d'une recherche commune sont extrêmement nombreux et intéressent tous les ordres d'enseignement. Pour le présent, l'enseignement du français, des mathématiques et des sciences, la solution des problèmes concernant les manuels scientifiques et techniques, la confrontation des deux réformes française et québécoise ont été reconnus comme principaux objectifs du futur Centre franco-québécois de développement pédagogique.

● NOUVELLES DU CENTRE :

M. QUIGNARD, jusqu'ici Directeur des Etudes du Centre International d'Etudes Pédagogiques, vient d'être nommé proviseur du Lycée-Pilote de Sèvres, en remplacement de Mme FORASTE, récemment promue Inspectrice Générale. Nous sommes particulièrement heureux que M. le Proviseur QUIGNARD reste lié à notre maison et symbolise, par sa personnalité même, les liens étroits qui unissent le Centre et le Lycée-Pilote. Dans une fort humoristique allocution, M. AUBA, en félicitant M. QUIGNARD au nom de tout le personnel du Centre, lui a affirmé qu'il nous mettait dans la situation de Grandgousier pleurant d'un œil sa femme Gargamelle et riant de l'autre à la vue de son fils Gargantua ! Il traduisait ainsi de très juste et plaisante façon la joie et l'émotion de tous.

C'est M. JANICOT qui remplace M. QUIGNARD comme Directeur des Etudes Pédagogiques du Centre. Jusqu'ici professeur au Lycée-Pilote de Toulouse-Bellevue, M. JANICOT nous était bien connu par ses activités pédagogiques. Nous lui adressons tous nos vœux de bienvenue dans une Maison qu'il connaissait depuis 15 ans déjà grâce aux nombreux stages auxquels il avait participé.

● STAGES ETRANGERS :

Comme chaque année le Centre International d'Etudes Pédagogiques de Sèvres a accueilli ses amis étrangers : 47 Brésiliens en janvier, 13 Hongrois en février, 20 Tchécoslovaques en mars, tous professeurs de français dans leur pays.

Venus en rangs compacts essentiellement féminins et jeunes les Brésiliens ont dès leur arrivée rempli le Centre de mouvement, de chants et de danses, d'accolades désordonnées et chaleureuses, tout cela tempéré par une certaine mélancolie de ne pas participer directement aux préparatifs du Carnaval de Rio ou à la finale du championnat de football. Ce dynamisme si sympathique, si rayonnant et si chaud, nous l'avons retrouvé tout au long du stage. Ils ont participé aux nombreuses conférences, visites et expériences qui leur ont été offertes ou proposées avec une bonne humeur et un sérieux à la fois dignes de tout éloge : exposés sur les problèmes de l'enseignement du français, visites d'établissements scolaires ou d'orientation, conférences de littérature, d'art ou de civilisation, débats sur des problèmes linguistiques, étude — faite par eux-mêmes — des problèmes politiques français analysés à travers les journaux et les réactions des Parisiens par eux interviewés dans la rue, ou le métro ou la cour du Lycée. Ils ont visité nos musées de peinture, admiré nos palais royaux et impériaux, parcouru les rues du Marais, hanté les studios de l'O.R.T.F. et même escaladé, par une nuit de verglas, les pentes ce soir-là hostiles de Montmartre. Nous sommes certains que cette année les élèves brésiliens auront beaucoup à écrire, et analyser et traduire : « C'était un soir de verglas, à Montmartre, près du Lapin Agile, et le sol... le sol se dérobaît sous nos pieds. C'était très merveilleux et très angoissant aussi, et comment !... »

Et puis ils se sont battus dans le parc de Saint-Cloud, à coup de boules de neige et pour se convaincre que Paris n'était pas toute la France, ils ont accepté de passer une semaine en province, les uns à Biarritz, les autres à Toulouse.

Voyages semble-t-il réussis, car en écoutant leurs volubiles et enthousiastes échanges d'impressions et de souvenirs dans le train Bordeaux-Paris, on finissait par se demander qui avait vu la Cité de Carcassonne et qui la Chambre d'Amour, qui avait dégusté un cassoulet et qui du jambon de Bayonne, qui avait dansé le fandango et qui « la trompuso ». Bien sûr, ce fut aussi un voyage

d'études, et de visites, et de contacts personnels. Au Lycée de Biarritz on vit beaucoup de classes, au Lycée de Toulouse-Bellevue, on vécut de 8 à 17 heures la journée d'un élève français, réfectoire compris. A Biarritz on vit beaucoup de professeurs et de parents d'élèves, à Toulouse on fut beaucoup invité dans des familles non-enseignantes.

Et puis vint le dernier soir, avec un grand repas, abondant et chaleureux, et une distribution des prix généreuse et spirituelle où chacun des responsables du Centre reçut un cadeau somptueux et un compliment bien tourné. Et puis nous dansâmes tous ce que certains Français appelèrent avec une discutable nuance d'auto-satisfaction et une correction grammaticale douteuse : « la » samba.

Tout cela fut très vrai, et très gai, et un peu mélancolique et c'est avec une petite, toute discrète, pointe de jalousie que nos amis brésiliens nous demandaient :

« Et les Hongrois, quand est-ce qu'ils arrivent ? »

Et bien, ma foi, ils arrivèrent le lendemain même, sans accouades ni danses, avec discrétion, réserve et souriante timidité.

Et avec eux la vie de stage recommença, toujours la même avec sa succession d'exposés, de visites, de débats et de conférences, jamais la même avec des goûts nouveaux, des opinions divergentes et des personnes différentes. Graves et consciencieux, nos amis hongrois écoutaient beaucoup, notaient énormément, vivaient intensément tout ce qui leur était proposé.

Graduellement s'accrut leur activité pédagogique et ils purent réaliser eux-mêmes un remarquable montage audio-visuel en circuit fermé de télévision avec ébauche fort réussie de commentaire et de questions programmées : le Métro de Paris. Mais malgré la joie calme et évidente qu'ils éprouvaient à goûter le programme offert, il manquait quelque chose à leur bonheur : la mer... A une exception près, aucun d'entre eux n'avait vu la mer. Et c'est sous le soleil, au bord des champs de Normandie enneigés, qu'ils la découvrirent. Au bruit du ressac, au cri des mouettes et dans l'odeur de l'eau ils purent s'endormir et se réveiller dans un hôtel sur la plage, à Etretat.

Les yeux encore éblouis, le sourire encore un peu rêveur, bouleversés par cette rencontre avec les vagues toujours recommencées, il leur fallut cependant nous dire adieu, après un très émouvant et très digne festival de danses et de chants.

Et comme la mer, la vie de Sèvres recommença au soir même de leur départ avec le débarquement des professeurs tchécoslovaques.

Mais ils viennent à peine d'arriver et vous trouverez, comme on dit, la suite dans le prochain numéro.

A. J.

DECORATIONS :

Mme l'Inspectrice Générale HATINGUAIS a reçu récemment du gouvernement portugais la médaille de l'Instruction. Le diplôme de cette décoration a été confié à l'Instituto de Odivelas, et le professeur Carlota Almeida de CARVALHO, au nom de Mme la Directrice et des professeurs de cette Institution, en a offert les insignes à Mme l'Inspectrice Générale HATINGUAIS. A cette occasion, M. Jean BASDEVANT, Directeur Général des Relations Culturelles au Ministère des Affaires étrangères, écrit à Mme l'Inspectrice Générale HATINGUAIS : « Je viens d'apprendre avec grand plaisir que le gouvernement portugais vient de vous décerner la médaille de l'Instruction publique portugaise. Je m'en réjouis et vous félicite très sincèrement de cette distinction amplement méritée. Je connais en effet vos efforts en vue d'établir des relations aussi étroites que possible avec nos partenaires portugais sur le plan pédagogique et de faire connaître au Portugal les activités du Centre International d'Etudes Pédagogiques de Sèvres. »

Trois membres du personnel du Centre International d'Etudes Pédagogiques ont été décorés des palmes académiques : **M. CAYOL** (qui nous a quittés depuis pour prendre la direction du Centre Culturel français d'Amsterdam), **Mme STOURDZE**, conseiller pédagogique pour l'enseignement du français langue étrangère. **Mme PECOIL**, chef du service des stages français.

Nous adressons à ces fidèles collaborateurs de notre maison, toutes nos amicales félicitations.

Le 24 janvier dernier, le pianiste André KRUST a donné au Centre, salle Jean ZAY, un récital d'une haute tenue. Scarlatti, Schumann, Schubert, Debussy, se succédèrent au programme, pour le plus grand plaisir de nos invités français et étrangers.

LA CHORALE DES METIERS DE LA MUSIQUE

(SECTIONS TECHNIQUES ARTISTIQUES DU LYCEE DE SEVRES).

Depuis plusieurs années déjà, la chorale des Métiers de la Musique, sous la direction de Michel FLEURANT, donne une série de concerts annuels auxquels de grands artistes prêtent leur concours. Le 1^{er} février dernier en la salle des Fêtes de Sèvres, la chorale a interprété une cantate de Bach, sous la direction de Bernard WAHL à la tête de son orchestre de chambre de Versailles.

Au début de juillet prochain, la chorale participera, à Saint-Lizier, dans l'Ariège, à un concert organisé par Madame de SAINT-BLANQUAT, directrice honoraire du Lycée de Sèvres, au profit de la réfection des orgues anciennes de la ville. Les élèves de la classe d'orchestre du Conservatoire de Toulouse prêteront leur concours à cette manifestation.

J. L.

La Documentation Française a récemment inauguré ses nouveaux locaux de l'Hôtel de Mailly, 31, quai Voltaire, par une très belle exposition composant un panorama d'ensemble de toutes ses productions. Aimablement invité par M. le Directeur de la Documentation Française, le personnel du service « Recherches et Information », a été reçu et guidé à travers cette exposition

par Mme ENJALBERT. Cette visite marqua le renouveau des relations amicales entre les deux maisons. Plusieurs groupes de nos stagiaires étrangers ont été déjà reçus quai Voltaire et ont largement bénéficié de toutes les ressources documentaires de cet établissement enrichi et rénové. Une partie de l'exposition sera prêtée au Centre à l'occasion des stages « d'Initiation économique et sociale » et d' « enseignement de l'histoire ». Un film décrivant les travaux de la Documentation française sera également présenté à cette occasion.

CONDITIONS D'ADHÉSION

Envoyer le montant de l'adhésion (membres adhérents : 10 F, membres bienfaiteurs : 20 F) au compte chèques postaux : Paris 6959-99 «Amis de Sèvres» 1, avenue Léon-Journault - 92-SÈVRES.

Pour l'étranger, s'adresser à nos correspondants Hachette à l'étranger :

AFRIQUE DU SUD : CONSTANTIA CONTINENTAL BOOKSELLERS. P.O.B. 7753 Johannesburg. — **ALLEMAGNE FEDERALE** : W.E. SAARBACH G.m.b.H. Follerstrasse 2, D. 5. Koeln. — **ARGENTINE** : LIBRERIA HACHETTE. Rivadavia 739-45, Buenos Aires. **AUTRICHE** : MORAWA, Wollzeile 11, 1010 Wien. — **BELGIQUE** : AGENCE ET MESSAGERIES DE LA PRESSE. 14-22 rue du Persil, Bruxelles. — **BRESIL** : LIVRARIA HACHETTE S.A. DO BRASIL. 299 Avenida Erasmo Braga-4° andar, Caixa Postal 1969, Rio de Janeiro. — **CANADA** : LIBRAIRIE HACHETTE. 554 est, rue Sainte Catherine, Montréal 24 P.Q. — **CHILI** : LIBRAIRIE FRANÇAISE. Huerfanos 1076 Casilla 43 D, Santiago. — **COLOMBIE** : LIBRERIA CENTRAL. Apartado aereo 3484, Bogota. — **CONGO** : SOCIETE CONGOLAISE HACHETTE. B.P. 2510, Brazzaville. — **CHINE** : WAIWEN SHUDIAN. P.O. Box 88, Pékin. — **CUBA** : CUBARTIMPEX. Simon Bolivar n° 1, Palacio Aldama Bldg aptdo 6540, La Havane. — **ESPAGNE** : SOCIEDAD GENERAL ESPAGNOLA DE LIBRERIA. Evaristo San Miguel 9, Madrid 8. — **ETATS-UNIS** : EUROPEAN PUBLISHERS REPRESENTATIVES INC. 132 West 43rd Street, New York 36 N.Y. — **ETHIOPIE** : INTERNATIONAL PRESS AGENCY : G.P. Giannopoulos, Hailé Sélassié Star Sq. P.O.B. 120, Addis Abeba. — **FINLANDE** : RAUTATIEKIRJAKAUPPA Oy. Kampinkatu 2, Helsinki. — **GRANDE-BRETAGNE** : THE CONTINENTAL PUBLISHERS ET DISTRIBUTORS LTD. 25 29 Worship Street, London E.C. 2. — **GRECE** : LIBRAIRIE KAUFFMANN. 28, rue du Stade, Athènes. — **HONGRIE** : KULTURA, P.O.B. 149, Budapest 62. — **IRAK** : IRAQ STORE Cie. P.O. Box 26, Baghdad. — **IRAN** : I.A.D.A. 151 Khiaban Soraya, Téhéran. — **ISRAEL** : LIBRAIRIE FRANÇAISE ALCHEH. 62 Nahlat Benyamin B.P. 1550, Tel Aviv. — **ITALIE** : MESSAGERIES ITALIENNES. Via Giulio Carcano 32, Milan. — **LIBAN** : LIBRAIRIE A. NAUFAL. Rue de l'Emir Béchir, Beyrouth. — **LUXEMBOURG** : MESSAGERIES PAUL KRAUS. 5, rue Hollerich, Luxembourg. — **MEXIQUE** : LIBRAIRIE FRANÇAISE. Paséo de la Réforma 12, Mexico. — **NORVEGE** : NARVESENS KIOSKKOMPANI. Stortingsgata 2, Oslo. — **PAYS-BAS** : VAN DITMAR'S. Schiestraat 32-36, Rotterdam. — **PEROU** : PLAISIR DE FRANCE. Av. Nicolas de Pierola 958, Lima. — **PORTUGAL** : LIVRARIA BERTRAND. 72-75 rua Garrett, Lisbonne. — **R.A.U.** : LIBRAIRIE HACHETTE. 45 bis, rue Champollion, Le Caire. — **SUEDE** : WENNERGREN WILLIAMS AB. Drottningatan 71 D, Stockholm. — **SUISSE** : NAVILLE et Cie S.A. 5-7 rue Lévrier, Genève. — **ROUMANIE** : CARTIMEX. 14-18, rue A.-Briand, P.O.B. 134-L35, Bucarest. — **TCHÉCOSLOVAQUIE** : ARTIA. V° Smeckach 30, P.O.B. 790, Prague 2. — **TURQUIE** : LIBRAIRIE HACHETTE. 469 Istiklal Caddesi, Beyoglu, Istanbul. — **URUGUAY** : MONTEVERDE et Cie S.A. 25 de Mayo 577, Montevideo. — **VENEZUELA** : LIBROS Y ARTE. A.L. Avila, La Florida, Caracas. — **YUGOSLAVIE** : JUGOSLOVENSKA KNIGA. Terazije 27, Belgrade.



CENTRE INTERNATIONAL D'ETUDES PEDAGOGIQUES 1, avenue Léon Journault, Sèvres

Le Bulletin n° 2 des "Amis de Sèvres" (année 1968) a publié une étude docimologique sur la correction par cinquante-quatre professeurs de "Sciences" d'un problème de géométrie exécuté le 2 mai 1967 par un élève de 3e.

Nous remercions notre collègue, Monsieur Jean AUBA, Directeur du Centre International d'Etudes Pédagogiques de Sèvres de bien vouloir nous permettre d'apporter sur ce travail l'opinion de l'Inspection Générale de Mathématiques.

1. Les correcteurs. - L'article analysé n'a pas spécifié le niveau mathématique des correcteurs.

L'Inspection Générale de Mathématiques est bien placée pour savoir que dans le premier cycle des établissements du second degré exercent actuellement

- I. { - quelques rares Agrégés de mathématiques
- des certifiés de mathématiques

Mais aussi

- II. { - des certifiés de sciences physiques
- des certifiés de sciences naturelles
- des professeurs de C.E.G.
- des instituteurs
- des maîtres auxiliaires, licenciés ou non ; parmi les licenciés se trouvent des licenciés de sciences physiques ou de sciences naturelles.
- des contractuels de tout âge et de toute formation (l'effectif de cette catégorie diminue nettement depuis deux ans)

Certes nous savons le dévouement dont font preuve un grand nombre des maîtres de la seconde catégorie* (agrégés et certifiés de mathématiques, que pour simplifier, nous appellerons "mathématiciens professionnels", présentent des garanties de compétence. Or depuis une dizaine d'années, grosso modo, 20% seulement des heures d'enseignement des mathématiques dans le premier cycle (lycées et C.E.S.) sont confiées à des mathématiciens professionnels. Cette situation grave va commencer à s'améliorer dès cette année, et cette amélioration s'accroîtra l'an prochain du fait de l'arrivée d'importantes promotions de jeunes certifiés ; on peut prévoir qu'en 1968-69 le nombre de mathématiciens professionnels exerçant dans le premier cycle augmentera de près de 10%.

Lorsqu'il est écrit (page 11, loc. cit.) "un correcteur n'est-il pas un faux monnayeur ? ", la réponse est affirmative si ce correcteur n'a pas la compétence requise ; c'est pourquoi nous trouvons regrettable que l'enquête publiée n'ait pas spécifié la qualification des correcteurs. La dispersion (surprenante)

.../...

* néanmoins seuls ceux de la première catégorie

des notes n'est-elle pas due en partie à la dispersion de la qualification des correcteurs ?

2. Le texte du problème.- Ne portons aucun jugement sur l'intérêt ou l'opportunité du sujet, mais, puisqu'il s'agit de docimologie, disons qu'il existe deux façons de présenter un problème de géométrie :

- ou bien on introduit dans l'énoncé de chaque question, au fur et à mesure, les éléments (points, droites, cercles, ...) qui entrent en jeu

- ou bien, dans un chapeau initial, on donne les définitions de tous les éléments qui seront utilisés jusqu'à la fin du problème.

La première façon aide davantage l'élève ; la seconde l'oblige à construire dès le début une figure très complète qui peut s'avérer inutilement compliquée pour résoudre les questions initiales ; elle permet par contre de voir si l'élève sait reconnaître des situations simples dans une réalité complexe.

C'est cette seconde façon qui avait été choisie dans le problème étudié : la première question ne fait intervenir que le demi-cercle de diamètre OB et la demi-droite IX ; la seconde question fait intervenir le demi-cercle de diamètre AB et le point Q ; la troisième question seulement fait intervenir le point P.

Nous ignorons d'ailleurs si la figure reproduite page 8 (loc. cit.) a été fournie aux élèves ou si elle a été réalisée par l'auteur de la copie.

Le texte renferme en outre un piège à la troisième question, car il affirme que les triangles OBM et BIP sont égaux ; or l'usage, veut, dans l'étude de l'égalité et de la similitude des triangles, qu'on énumère les sommets dans les ordres correspondants, il aurait donc fallu dire "les triangles OBM et PIB".

En résumé, les élèves n'étaient pas spécialement aidés par le texte qui leur était proposé.

3. La solution du problème.- Il est normal que, dans une enquête docimologique, on ne fournisse pas une solution impeccable, mais pourquoi avoir donné une solution d'aussi peu de valeur ?

Dans la solution de la question 1 l'élève parle de l'hypoténuse d'un triangle avant d'avoir justifié que celui-ci était rectangle : il y a là une faute grave qui doit être pénalisée. L'élève commet la même faute dans la question 2, mais aggravée de l'affirmation, sans preuve, que le triangle Q.M.I. est isocèle, la solution de la question 2 est donc sans valeur.

.../...

Dans la solution de la question 5, les deux faits fondamentaux utilisés :

- P.Q. parallèle à A.B.
- égalité des triangles Q.O.I., P.I.B. sont affirmés et non démontrés ; en outre la phrase finale utilise des propriétés surabondantes pour justifier l'existence du parallélogramme O.I.P.Q. ; la solution de la question 5 est donc sans valeur.

En résumé seules les solutions des questions 3 et 4 ont une valeur normale, malgré quelques maladroites.

4. Le choix d'un barème. - Deux faits sont bien connus des professeurs de mathématiques :

a) Il est possible, pour noter, la solution d'un problème de mathématiques, d'adopter un barème de correction éliminant presque totalement l'influence de la personnalité du correcteur -et c'était le cas pour le problème étudié.

b) Au contraire, l'absence d'un barème de correction entraîne, de la part de divers correcteurs, une grande diversité de notes, mieux encore, le même correcteur, jugeant la même copie à des moments différents risque, en l'absence d'un barème, de mettre des notes différentes.

La lettre d'envoi aux correcteurs qui figure page 8 (loc. cit.) ne faisant pas mention d'un barème, on peut en déduire que l'enquête a laissé chacun des cinquante-quatre correcteurs se fixer lui-même un barème : la diversité des notes obtenues n'est donc pas surprenante.

Dans le problème étudié au cours de cette enquête, on pouvait attribuer 4 points à chacune des 5 questions -total 20 points), ce qui donne le barème.

$$4+4+4+4+4$$

Mais on pouvait adopter bien d'autres barèmes suivant le but du problème.

Un professeur voulant en fin d'année, se rendre compte des progrès de ses élèves pourra répartir les points (20) suivant le barème :

$$2+3+4+6+5$$

Si le même professeur veut au contraire rehausser le moral de ses élèves après une période difficile, il pourra utiliser le barème :

$$3+4+5+5+3$$

.../...

Si l'on veut s'attacher aux raisonnements géométriques et si l'on estime que les relations métriques dans les triangles rectangles sont peu intéressantes, on utilisera le barème :

$$3+4+5+3+5$$

Si l'on veut au contraire contrôler l'acquisition des relations métriques, on utilisera un barème tel que :

$$2+3+3+9+3$$

5. Observation finale.— Suivant les barèmes qui ont été proposés au paragraphe 4, le même correcteur au même moment, pourra donner à la solution étudiée l'une des notes :

10, 11, 12, 9, 13

Les bonnes notes (15, 16, 17) mises par certains correcteurs (page 11, loc. cit.) sont le fait de maîtres, au tempérament optimiste qui n'ont pas vu les fautes commises par l'élève soit par suite d'une lecture rapide, soit parce qu'ils ne sont pas mathématiciens professionnels.

Les mauvaises notes (4, 6) mises par d'autres correcteurs sont le fait de maîtres au tempérament pessimiste qui ont cru trouver partout des fautes, soit par suite d'une lecture rapide, soit parce qu'il ne sont pas des mathématiciens professionnels.

En conclusion, l'expérience décrite dans le bulletin n°2 des "Amis de Sèvres" ne peut, en ce qui concerne les mathématiques, être retenue comme ayant une valeur quelconque en matière de docimologie.

G. CAGNAC