



RE 8

**INSTITUT DE RECHERCHE SUR L'ENSEIGNEMENT  
DES MATHÉMATIQUES**

Groupe de Recherche d'Epernay

# **ÉVALUATION**

**DOCIMOLOGIE - ORIENTATION - TAXINOMIE**

**UNIVERSITÉ  
de  
REIMS**

INSTITUT DE RECHERCHE SUR L'ENSEIGNEMENT DES MATHÉMATIQUES

---

(Groupe de Recherche d'EPERNAY)

EVALUATION

Docimologie. Orientation. Taxinomie

(Essai de synthèse)

Mme ARSENE Michèle  
GIANORDOLI Monique  
LEFORT Annick  
MANDRILLE Anne  
PATIN Régine  
M. ARMENGAUD Gérard  
BACQUIAS Michel  
CAREL Daniel  
CORTINA Raymond  
DE FINANCE Hubert  
DEMALANDER Jean  
JOBARD Michel  
LAFON Serge  
MACIEJEWSKI Christian  
MANCIER Gilles  
MAYDAT Jean  
NIMIER Jacques  
SCHACHERER Germain  
THOUMY Henri

ANNEE SCOLAIRE 1976-1977

---



"NUL DOUTE : L'ERREUR EST LA REGLE ; LA VERITE EST L'ACCIDENT  
DE L'ERREUR".

GEORGES DUHAMEL



Nous remercions tous nos collègues qui ont contribué à la réalisation de cet ouvrage, en répondant aimablement à un questionnaire-sondage. La fréquence de leurs réponses a montré combien ils s'intéressent à l'évaluation. Nous espérons ne pas les décevoir en leur présentant, dans ce qui suit, quelques aspects de la docimologie, de l'orientation et de la taxinomie.



## INTRODUCTION

---

Le travail dont nous vous proposons les résultats s'insère dans une problématique qu'on pourrait appeler "Recherche-Action". Il n'illustre pas une recherche pure mais un essai de communication. En effet, le but poursuivi par le groupe n'est pas d'apporter à la science des résultats nouveaux mais de se familiariser avec des méthodes de travail et de s'en servir afin de provoquer une prise de conscience de certains problèmes, de leur complexité et de leur proximité. Nous n'avons pas la prétention d'apporter la "bonne parole" aux lecteurs mais de simplement leur livrer quelques réflexions.

C'est dans cet esprit que nous avons poursuivi un thème qui nous est cher et que nous avons commencé à exploiter en 1975, en effectuant une recherche sur l'influence de la formulation dans l'acquisition d'un concept mathématique. Il s'agissait d'observer les variations de comportement des élèves devant différents types de questions posées par le professeur. Cette année nous suivons notre trajectoire en observant la variation de réussite mesurée par les notes, ce qui nous conduit à la docimologie et en particulier, à la façon, sous jacente, dont la personnalité du professeur est impliquée. Mais il va de soi que le problème de la notation débouche naturellement sur celui de l'orientation et qu'il est difficile d'en évoquer tous les aspects sans évoquer la taxinomie éducative.

Le lecteur peu initié à ces problèmes risque, à la lecture de certaines expériences et résultats, d'être choqué et douter du sérieux des travaux réalisés depuis déjà près de cinquante ans. Qu'il nous soit permis d'insister sur le fait fondamental que tous les chercheurs, psychologues, médecins et professeurs qui ont participé aux expériences que nous relatons l'ont fait avec la plus grande honnêteté et en utilisant au mieux leurs compétences. Nous insistons là-dessus car nous touchons au plus près la personnalité du pédagogue. Pour avoir expérimenté dans ce domaine et effectué des conférences nous savons que certains mots, certaines réalités blessent. Alors les boucliers se lèvent ; on se protège à leur ombre.

Afin de mesurer le degré de connaissance des méthodes d'évaluation et de leurs critiques nous avons demandé la collaboration des collègues de plusieurs établissements de la Marne, en leur proposant de répondre à quelques questions. Voici une synthèse des réflexions, toujours pertinentes, recueillies parmi les 197 réponses à la question : "Remettez-vous les notes en cause ? Expliquez votre position".

Soixante pour cent des professeurs interrogés ne remettent pas en cause les notes. Certains ne précisent pas leur position comme si cela allait de soi ; d'autres pensent simplement qu'on ne peut rien "proposer de mieux" ; la plupart enfin, sans remettre les notes en cause, accompagnent cette position de restrictions plus ou moins importantes.

#### INTERET DES NOTES

Pour l'élève

"Pour qu'un élève sache s'il a acquis une notion"

"L'absence de notes est toujours ressentie comme une situation frustrante"

Le professeur a donc l'impression de "donner" quelque chose.

Pour les parents

"Les notes sont souvent pour les parents la seule façon d'apprécier le travail ou le niveau de l'enfant".

Pour le professeur

"Elle aide le professeur à juger de façon pratique ce qui est à revoir dans ce qui a été expliqué".

"Elle aide le professeur dans son jugement quand les élèves sont trop nombreux"

"C'est le seul moyen assez précis de renseignements transmissibles sur la valeur globale de l'élève"

"C'est un repère malgré leurs imperfections".

#### NOTE = STIMULATION

Ce sujet fait l'objet d'un paragraphe spécial car il y a deux tendances opposées :

"Les élèves travaillent souvent pour la note ce qui est un stimulant"

"Soyons réalistes : l'attrait de la carotte fait courir le monde ! Notes et classement apportent de l'émulation, de l'ardeur et non pas des classes de vieux, de nouilles, d'aigris et de surcroît de bons à rien ; à tout niveler on arrive à la grisaille anonyme, à un troupeau bêlant de ratés incapables du moindre effort personnel, individuel ; la vie est un combat et pas de la rigolade".

Pour d'autres, au contraire, la note entraîne souvent une motivation négative.

"La note détourne l'attention de l'élève vers l'esprit de compétition, de rivalité et l'amène à une fausse idée de lui-même"

"Elles créent dans la classe une compétition souvent pénible"

"Les élèves sont trop motivés par la notation et non par le travail lui-même. C'est grave"

"Les élèves travaillent en fonction d'un résultat immédiat et par exemple en langues vivantes, ils oublient souvent que le but final est tout de même l'acquisition d'un instrument de communication, d'un bagage culturel, voire d'un atout sur le marché du travail"

Et ces remarques qui évoqueront des exemples vécus :

"Il suffit de penser au comportement négatif des élèves après l'arrêt des notes"

"L'enfant qui a des notes moyennes est peu enclin à faire des efforts".

#### INSUFFISANCE DES NOTES...

La majorité des professeurs, qu'ils remettent ou non les notes en cause, sont conscients de leurs imperfections en tant que moyen d'évaluation :

"En aucun cas elles ne suffisent à elles-mêmes"

"Les notes sont des supports, des points de repère tout au plus"

"La note correspond à un travail précis et à rien d'autre"

"La note n'est qu'un mauvais reflet de ce que les élèves peuvent faire"

"L'enfant n'est pas un numéro ; on oublie bien vite sa personnalité au profit de sa note"

"Elles ne peuvent traduire le comportement général de l'élève dans le groupe, ce qui est aussi important qu'un très bon niveau"

"Dans les examens ou concours elles peuvent avoir une influence disproportionnée sur l'avenir de l'élève"

"La note est souvent représentative d'un savoir ou d'un savoir-faire. Elle ne reflète donc pas la personnalité d'un élève ou de ses facultés latentes. Les données affectives et sociales n'apparaissent pas. Or l'éducation véritable doit tenir compte de ces données"

"Elles charrient avec elles trop de notions de sélection, de classement et de rejet"

... OU LIMITES DE LEUR UTILISATION

Certains professeurs pensent que ce ne sont pas les notes qu'il faut remettre en cause mais leur "mauvaise" utilisation, le culte de la note ou "notolâtrie"

Comme le lecteur aura pu s'en rendre compte, les avis divergent et peuvent être radicalement opposés. La passion de certaines prises de position montre que les problèmes évoqués vont au-delà de la technique pédagogique et que les solutions "idéales" demeurent pour l'instant utopiques. Mais ces divergences ne sont pas le seul fait des maîtres ; les contrôles, examens et concours caractérisent notre société. Chacun prend parti. Ainsi P. Valéry s'insurge contre la "diplômanie" en temps qu'ennemie de la culture et contre le contrôle en général. "Si je me fonde sur la seule expérience et si je regarde les effets du contrôle en général, je constate que le contrôle, en toute matière, aboutit à vicier l'action, à la pervertir... Je vous l'ai déjà dit : dès qu'une action est soumise à un contrôle, le but profond de celui qui agit n'est plus l'action même, mais il conçoit d'abord la prévision du contrôle, la mise en échec des moyens de contrôle. Le contrôle des études n'est qu'un cas particulier et une démonstration éclatante de cette observation très générale. "Par contre J.F. Revel estime que" le droit au succès dans les examens (...) est, de plus en plus, placé sur le même plan que le droit à la sécurité sociale, aux vacances, à la retraite (...). Un candidat au brevet de pilote à réaction ne songe pas à faire valoir qu'il a 12 enfants à nourrir, une vieille mère sensible, pour obtenir son brevet, non plus qu'à faire demander par les journaux qu'on lui décerne son brevet, même s'il n'a pas pu piloter le jour de l'examen, vu qu'il est extrêmement impressionnable et que le jour de l'épreuve il perd régulièrement ses moyens. Mais un candidat au baccalauréat (et toute l'opinion française derrière lui) estime que la dernière des choses qui doit entrer en ligne de compte dans cet examen est ce en vue de quoi il est fait, à savoir : constater dans les moins mauvaises conditions possible de secret et d'impartialité, si M. X... sait ou non dire bonjour en anglais.

## APERCU HISTORIQUE DE LA DOCIMOLOGIE

C'est à H. Piéron que l'on doit le terme de "docimologie", qu'il définit comme "étude systématique des examens", au cours de ses premières recherches entreprises en 1920. Ses travaux furent alors concrétisés par une enquête qu'il réalisa en Juin 1922 avec sa femme, H. Laugier et D. Weinberg, auprès d'écoliers se présentant au certificat d'études primaires.

La première enquête internationale fut financée par la Carnégie Corporation. Elle portait sur "les conceptions, les méthodes, la technique et la portée pédagogique et sociale des examens et concours". En 1931 des commissions nationales furent constituées, mais seules les sections anglaises et françaises publièrent à partir de 1936 des études importantes. Burt, Spearman, Thomson, Hartog et Rhodes participèrent aux travaux de la commission anglaise. Le professeur Laugier conduisit des expériences sur le baccalauréat pour la commission française. La dernière conférence internationale de la Carnégie Corporation eut lieu en 1938. Kandel y fit un exposé remarquable que les examens et leurs substituts aux U.S.A. La commission française tenta de se maintenir pendant quelques années avec Desclos et D. Weinberg.

Ces travaux, mal connus, eurent cependant une certaine influence. Le ministre de l'éducation nationale J. Zay entreprit en 1937 d'améliorer la concertation entre les correcteurs du baccalauréat et, sur l'initiative de l'inspecteur général Monod, introduisit la notion d'orientation dans l'enseignement français et fixa, en 1938, par un décret loi, le statut de l'orientation professionnelle.

En 1950, une enquête sur le niveau intellectuel des enfants d'âge scolaire est réalisée par l'Institut National d'Etudes Démographiques dirigé par A. Sauvy ; le Groupe Français d'Education Nouvelle, présidé par H. Piéron, se penche sur les problèmes posés par l'examen d'entrée en 6ème ; et le Conseil Supérieur de la Recherche Scientifique et du Progrès Technique commande une série de travaux sur la correction des épreuves écrites dans les examens, travaux dirigés par H. Piéron et M. Reuchlin.

En 1955, le professeur H. Piéron entreprend sa fameuse enquête sur le baccalauréat dans l'Académie de Paris et publie ses résultats dans l'ouvrage "Examens et docimologie".

En 1967, la conférence des ministres européens de l'éducation et, en 1968, le XVI<sup>e</sup> colloque de l'Association Internationale de Pédagogie Française remettent en question toutes les formes d'examens existantes.

Les recherches en se multipliant ont influencé les différentes réformes de l'orientation scolaire et professionnelle et les conditions d'examen ou d'admissibilité. A l'origine, ces recherches étaient essentiellement critiques et se bornaient à mettre en évidence les défauts des procédures existantes. Elles devinrent, par la suite, des recherches positives en proposant de nouvelles méthodes d'évaluation, plus sûres, en élargissant le domaine de leurs investigations. La tendance actuelle est d'intégrer des problèmes d'évaluation dans l'ensemble des problèmes éducatifs, et de ne pas séparer l'évaluation de l'éducation elle-même.

### IMAGE SOCIALE DES EXAMENS

Si on recherche un rapport entre l'examen et la structure sociale de l'époque où il est apparu, il faut donner une définition précise de l'examen. Peut-on par exemple considérer les rituels d'initiation des sociétés "primitives" comme une sorte d'examen lié à un contexte social ? D'autre part il est difficile de préciser l'époque exacte de son apparition sous une forme élaborée, ou du moins à l'image de ce que nous connaissons aujourd'hui, c'est à dire des examens et des concours organisés sur le plan national et voulus accessibles à tous. Il nous semble cependant intéressant de mettre en parallèle deux époques historiquement très éloignées qui ont vu apparaître ce genre d'épreuves.

"Si le procédé des examens et la sélection garantit une culture, le jeune chinois possédait certainement la plus haute culture. Car aucun élève ne put être examiné et contrôlé avec tous les raffinements (isolement des candidats, des examinateurs) que l'esprit a jamais inventés...  
... Un système fort complexe d'examens et de grades conduisait l'étudiant privilégié des examens élémentaires passés au chef-lieu de district, à ceux du comté, puis à ceux de la province, puis à celui de Pékin, enfin à l'examen final qui conférait l'admission à l'Académie Impériale. Car le chemin est long à travers tous les degrés du mandarinat" (R. Gal : Histoire de l'éducation. P.U.F.) En ce qui concerne la structure sociale de cette époque (VI<sup>e</sup> siècle avant J.C.) R. Gal précise qu'une "crise sociale, qui amène la destruction du pouvoir des seigneurs et une certaine émancipation du peuple, favorise l'effort d'éducation remarquable entrepris dans cet empire replié sur lui-même et satisfait". Lao Tseu et Confucius fixent la morale de cette époque, morale soucieuse de l'éducation du peuple.

Si on rapproche cette époque de celle qui, dès 1880, vit en Occident le développement de systèmes complexes d'examens et de concours, on peut proposer les hypothèses suivantes. Lorsque l'équilibre social permet, en principe, l'accès aux fonctions sociales supérieures, se pose le problème de la sélection donc de la mise au point d'un système de sélection. Par contre dans un système social où les différentes classes de la société sont imperméables (castes de l'Inde par exemple) un tel système est inutile.

Une organisation des examens est donc au départ animée d'un souci "démocratique". J. Delannoy remarque dans le numéro des Cahiers Pédagogiques consacrés à l'examen des examens qu'en 1882 les idées qui sous tendent la prise en main de l'instruction publique par l'état " étaient progressistes : elles marquaient en effet la volonté d'en finir avec le régime de la faveur,

des bénéfiques et de la vénalité des charges".

Mais, de par sa fonction même, ce système se trouve formateur de la sélection sociale. Il s'oppose donc à la perméabilité des différentes classes. En ce sens il dénature la fonction éducative. Dès 1896 E. Combes remarquait : "les élèves ne faisaient pas d'études pour se former, mais pour subir les épreuves de l'examen". Il réclamait déjà la suppression du baccalauréat. P. Valéry va plus loin : "Je n'hésite pas à le déclarer, le diplôme est l'ennemi mortel de la culture".

On peut aller, et on est allé, plus loin : l'examen n'est-il pas un instrument de la classe privilégiée pour contrôler l'accès à son statut et n'admettre que des individus ayant assimilé les valeurs de cette classe ?

"Loin d'essayer, comme elle le prétend parfois, de combler les différences culturelles, l'école va au contraire s'appuyer sur elles, les transformer par des notes, le classement, les examens en inégalités naturelles. On peut ainsi comprendre le maintien des examens, de la notation et de tous les procédés de classement malgré toutes les critiques. Ils ont pour rôle de légitimer par le mérite les inégalités sociales". (Avanzini La pédagogie au XX<sup>e</sup> siècle).

Ainsi l'examen est-il au coeur de l'ambiguïté de l'acte éducatif. Il apparaît "comme une inévitable nécessité dans un grand état fortement centralisé qui dispose d'un nombre infini d'emplois et dans une démocratie pleine de mouvement et d'activité" (F. Pécaut Dictionnaire de Pédagogie 1882). Mais est-il, pour autant, le moyen, le moins imparfait possible, de contrôler démocratiquement certains savoirs, savoir faire, aptitudes et par là de permettre un juste accès à un rôle social fondé sur le mérite ? Ou bien est-il, au contraire, une barrière, artificielle et arbitraire, élevée pour défendre les classes privilégiées. Là encore on ne peut répondre sans résoudre d'abord la question fondamentale de toute démarche éducative. Quels sont les objectifs de l'éducation ?

DOCIMOLOGIE



### EXPERIENCES ET ANALYSE CRITIQUE

Second ou avant dernier ?

Admis à l'agrégation ou éliminé ?

En 1930 le professeur Laugier sème un malaise pernicieux dans les milieux universitaires, en effectuant une expérience de multicorrection de copies d'agrégation d'histoire puisées dans les archives. 166 copies ont été corrigées par 2 professeurs travaillant séparément, sans connaître leurs appréciations respectives. Les deux avaient une longue expérience et corrigeaient méticuleusement. Les résultats furent surprenants. La moyenne des notes du premier correcteur dépassait de près de deux points celle du second. Le candidat classé avant dernier par l'un était classé second par l'autre. Les écarts de notes allaient jusqu'à 9 points. Le premier correcteur donnait un 5 à 21 copies cotées entre 2 et 14 par le second ; le second donnait un 7 à 20 copies cotées entre 2 et 11,5 par le premier. La moitié des candidats reçus par un correcteur était refusée par l'autre.

Cette expérience caractéristique de docimologie a amené des chercheurs de plus en plus nombreux à s'interroger sur les sources d'erreurs des procédures d'évaluation traditionnelles.

### DIFFERENTS TYPES D'ERREURS

F. Bacher (1969) distingue trois sources d'erreurs.

"La première source d'erreurs est due à l'évaluateur lui-même qui note autour d'une moyenne plus ou moins élevée et qui disperse plus ou moins ses notes autour de cette moyenne. De plus les évaluateurs ne classent pas dans le même ordre une même série de travaux ou de réponses.

La seconde source d'erreurs tient, dans les examens traditionnels, au choix du sujet même de l'examen. Il est en effet peu satisfaisant de généraliser à l'ensemble des candidats une constatation "ponctuelle", fondée sur l'un seulement des innombrables sujets qui auraient pu lui être proposé.

La troisième source d'erreurs vient des élèves. D'un jour à l'autre, d'un moment à l'autre, se produisent des fluctuations aléatoires de la capacité qu'il s'agit d'évaluer. De plus une certaine forme d'évaluation peut défavoriser de façon systématique un type d'élèves (la question s'est posée notamment à propos des épreuves orales)" (Reuchlin, 1.67 p. 215) (1)

#### PREMIERE SOURCE D'ERREURS : L'EVALUATEUR

##### Les évaluateurs ne sont pas d'accord entre eux.

En 1932 eut lieu une des premières enquêtes docimologiques. La Commission Carnégie effectua une expérience de multicorrection en prélevant, au hasard, cent copies dans les archives du baccalauréat à Paris. Ces copies furent distribuées à 6 groupes de 5 examinateurs. Les disciplines concernées étaient : le français, la philosophie, le latin, les mathématiques et la physique. On demanda aux examinateurs de noter les copies et de fournir un rapport sur :

- les qualités exigées
- le classement des dites copies
- la méthode de notation utilisée.

Les résultats montrèrent une forte dispersion des notes attribuées à chaque copie par les correcteurs. Aucune copie ne reçut deux fois la même note. L'écart maximum des notes dépassa les prévisions. Une copie de français est notée 3 et 16 ; en philosophie et en latin l'écart maximum est de 12 points. Les mathématiques et la physique, réputées pour des sciences exactes, ne sont pas épargnées : l'écart maximum est respectivement de 9 et 8 points. D'autre part, le classement des copies variait fortement d'un correcteur à l'autre.

En 1975 l'Institut de Recherche sur l'Enseignement des Mathématiques de Grenoble entreprend une expérience analogue de multicorrection. Un échantillon de 6 copies photocopées de mathématiques (niveau B.E.P.C.) est soumis à 64 correcteurs, avec un barème, très précis, sur 40 points. Les résultats confirment ceux de l'enquête précédente, effectuée 43 ans plus tôt. La dispersion des notes atteint près de 20 points.

(1) - Soucieux de ne pas alourdir l'exposé par des références répétées, nous renvoyons le lecteur à la page 88 où il trouvera une bibliographie sommaire des principaux ouvrages cités.

Le Centre International d'Etudes Pédagogiques de Sèvres, l'Association des Professeurs de Français, le Centre Pédagogique Régional de Toulouse, l'I.R.E.M. de Toulouse, pour ne citer que les plus marquants ont mené des expériences du même type et sont arrivés aux mêmes conclusions.

De plus, Laugier et Weinberg ont montré que la double correction est illusoire. En effet il faudrait, pour obtenir une note "exacte" (1) :

- 127 correcteurs en philosophie
- 78 " " en composition française
- 28 " " en anglais
- 19 " " en version latine
- 16 " " en physique
- 13 " " en mathématiques (2)

Mais les divergences entre correcteurs sont-elles seules en cause ?

L'évaluateur n'est pas d'accord avec lui-même

"Pas d'accord entre eux, les examinateurs ne se montrent pas davantage fidèles à eux-mêmes lorsqu'il s'agit de juger une seconde fois un devoir après un intervalle plus ou moins prolongé. Les enquêteurs anglais en avaient été frappés. Laugier et Weinberg en France, avaient fait la même constatation. A leur demande un professeur de physiologie de la Faculté des Sciences accepta 37 copies -dactylographiées et anonymes - qu'il avait corrigées trois ans et demi auparavant.

(1) - Une note "exacte" étant une moyenne de notes telle que l'adjonction d'une autre note ne modifie pas sensiblement cette moyenne.

(2) - Ne faut-il pas une dizaine de juges pour noter une épreuve de gymnastique ou de patinage artistique et bien plus pour désigner un prix Nobel ?

Dans 7 cas seulement, il remit la même note au même devoir. Dans les 30 autres cas, il y eut des divergences comprises entre 1 et 10 points. L'admissibilité, avec ses nouvelles notations, aurait été modifiée ; la moitié des précédents admissibles aurait été refusée et la moitié des refusés déclarée admissible. Le degré d'accord de ce professeur avec lui-même ne fut pas plus élevé qu'avec deux de ses collègues chargés de la même tâche : les coefficients de corrélation atteignant respectivement 0,58, 0,59 et 0,56. On poursuivi plus loin l'expérience, et elle aboutit à un fait plus troublant encore. On demanda à une bachelière, Paulette, intelligente mais ignorant tout de la question traitée, de noter à son tour ces compositions de physiologie, après les avoir lues une fois pour se faire une idée du sujet. Ses notes eurent une corrélation de 0,51 avec celles attribuées par les professeurs compétents. La bachelière ne se trouvait pas plus en désaccord avec les spécialistes que ceux-ci entre eux". (in Scienc et Vie, 1968, n° 610).

"Prenons le cas d'un correcteur en face d'une copie anonyme d'examen. Il se forme dans la conscience du correcteur l'image d'un couple affectif ; il est là devant l'élève qu'il ne connaît pas mais qu'il essaie de voir à travers l'image d'un cancre ou d'un élève intelligent, ou d'un imbécile, ou d'un original... C'est en fonction de l'harmonie ou du désaccord de ce couple affectif que le correcteur va le plus souvent noter.

En corrigeant, il se forme d'une façon plus ou moins distincte une image d'élève sur laquelle il applique progressivement une étiquette-type en fonction de sa répugnance ou de son admiration pour lesquelles peu de chose suffit, une jolie phrase, une de ses idées favorites qu'il retrouve exprimée, ou un cliché qui lui déplaît.

On constate des différences très grandes... à l'occasion de la même copie corrigée une deuxième fois à quelques jours d'intervalle, la première note ayant été effacée. Dans ce dernier cas, selon l'humeur du moment, la lecture de la veille, les soucis du jour, l'amabilité, la gentillesse ou l'insolence dont ont pu faire preuve les élèves dans la classe du jour précédent, les réactions sentimentales du correcteur ne sont pas toujours les mêmes et un nouveau couple affectif apparaît dans sa conscience, un couple dont les deux partenaires se sont trouvés changés." (M. Marchand, Le couple de l'éducateur et de l'élève dans leurs relations concrètes, 3-7).

### L'évaluateur à l'oral

Les réactions affectives du couple examinateur-examiné s'amplifient à l'oral.

En 1962 une expérience fut entreprise en classe de mathématiques élémentaires. "Trois professeurs interrogent chacun une série de 40 élèves de ces classes pendant qu'on enregistre au magnétophone ces interrogations. Mais attention :

- un enregistrement de l'interrogation de 20 élèves est soumis à 16 professeurs.
- un autre enregistrement de 20 élèves a été répété par un parleur unique, remplaçant aussi bien l'élève que l'examineur, et l'enregistrement de ces interrogations répétées a été également soumis à ces mêmes 16 professeurs pour notation.
- enfin 20 élèves avaient rédigé par écrit leur réponse à des questions dont le texte leur avait été fourni, et les 16 professeurs furent appelés à noter leurs copies".

On constata de fortes divergences : "l'accord fut le plus grand en présence des textes écrits avec le coefficient (de corrélation) moyen de 0,813... Dans l'écoute directe des interrogations, l'accord est moindre (0,727) alors que la base objective est bien exactement la même... La présentation répétée où l'on n'entend pas la voix des élèves interrogés a suscité le plus de désaccords." (PIERON 1-63, p. 148 à 154)

D'autre part, au "Colloque sur l'évaluation" tenu à Strasbourg en Mars 1976, Bonniol a présenté les résultats d'une recherche menée par le "groupe de recherche en docimologie d'Aix" sur les évaluations des examinateurs au baccalauréat. Utilisant deux variables expérimentales, elles-mêmes dichotomiques :

- "connaissance ou non connaissance du livret scolaire des candidats par les examinateurs lors des épreuves orales
- connaissance ou non connaissance par les mêmes examinateurs de la note qu'ils avaient attribuée aux candidats lors de l'écrit".

Il fut constaté que les corrélations ne sont pas meilleures lorsque l'examineur connaît le livret scolaire, "ce qui suggérerait que le livret scolaire n'est guère pris en compte en réalité," et "qu'elles sont meilleures lorsque l'examineur connaît la note qu'il a lui-même attribuée à l'écrit, ce qui suggère une tendance à ne pas se déjuger lors de l'oral quand il possède l'information sur sa première évaluation".

Le désaccord s'amplifie quand on passe de la note-diagnostic à la note-pronostic. Les interviews destinées à prévoir le devenir d'un candidat présentent peu de fiabilité. "Hartog et Rhodes citent un exemple frappant de l'absence de fidélité de telles épreuves: lors d'un examen pour le Civil Service, on décida de noter par ce procédé la vivacité d'esprit, l'intelligence et l'apparence générale des 16 candidats : ceux-ci furent soumis à deux groupes indépendants de quatre jurés expérimentés qui s'étaient préalablement mis d'accord sur la procédure et l'interprétation. Le coefficient de corrélation entre les moyennes des jugements fut seulement de 0,41 ; le premier classé par le second groupe était onzième pour le premier et le premier du premier groupe treizième pour le second". (Hotyat, in Revue Française de Pédagogie, Janvier 1968).

"Les rapports nationaux signalent que les examens oraux sont très répandus tant à l'admission qu'au terme des études secondaires. En outre, beaucoup de professeurs ont l'intime conviction de l'exactitude et de la haute signification des jugements qu'ils portent sur les candidats. En réalité, les épreuves orales sont les moins dignes de confiance dans les examens importants, surtout lorsque ceux-ci ont pour objet l'admission des élèves.

1°) Elles accordent rarement des chances égales à tous car chacun est en général soumis à des questions différentes, soit par tirage au sort, soit par l'inspiration du moment chez l'examineur. Même si l'interrogateur donne des coups de sonde dans d'autres domaines à l'occasion des réponses du candidat, les situations ne peuvent être comparées et les jugements sont arbitraires, spécialement aux abords de la note critique.

2°) L'examen oral impressionne les candidats surtout les élèves jeunes. Il agit différemment selon les formes de personnalité : certains sont stimulés et portés à briller ; d'autres, à l'opposé, sont inhibés et s'expriment moins aisément que dans les circonstances normales ; quant aux timides, il leur arrive de réagir par des poussées d'apparence agressive de nature à indisposer la plupart des interrogateurs.

3°) Ceux-ci ont d'ailleurs un comportement habituel très varié : il suffit de recourir à nos souvenirs pour retrouver tout un échantillon d'examineurs d'oral : les ouverts et les distants, d'une froideur glaciale ; les posés et les rapides qui mitraillent le candidat de rafales de questions ; les taciturnes et les loquaces allant jusqu'à amorcer les réponses ; les calmes et les agressifs manifestant leur mauvaise humeur aux faiblesses de l'étudiant.

On comprend que les possibilités multiples de relations qui peuvent s'établir entre l'élève et l'examineur suscitent au départ tel ou tel climat de nature à modifier sérieusement le rendement du premier et la note du second.

En outre l'interrogateur, engagé dans l'entretien avec le candidat doit répartir son attention entre la matière et les réponses ; aussi, au moment d'estimer la note, est-il rare qu'il ait gardé une idée précise et complète de l'exposé pour un jugement sûr". (In Hotyat, 1. 41 p. 61).

### Les préjugés de l'évaluateur

Il est certain que la personnalité de l'évaluateur intervient, dans une large mesure, dans la fluctuation des notes.

L'I.R.E.M. de Grenoble, au cours d'une expérience de multi-correction, a mis en évidence dans quelle proportion âge, sexe, grade et profil influent sur la dispersion et la moyenne des notes. L'échantillon des correcteurs était composé de 66 professeurs dont :

- 34 hommes
- 32 femmes
- 2 agrégés
- 31 certifiés
- 23 P. E. G. C.
- 10 maîtres auxiliaires
- 16 professeurs nés avant 1935
- 27 professeurs nés entre 1935 et 1945
- 23 professeurs nés depuis 1945

Pour une copie de mathématiques, de niveau B.E.P.C., "...mal écrite et raturée... avec de nombreux résultats exacts et correctement prouvés", on obtint les résultats suivants :

- peu de différences significatives en ce qui concerne le sexe,
- les notes des P.E.G.C. sont inférieures à celles des certifiés,
- la dispersion des certifiés est la plus faible,
- les professeurs du second cycle sont plus "généreux" que ceux du premier cycle,
- la dispersion des notes du professeur né entre 1900 et 1935 est de loin la plus importante : 4,718 pour 3,982 (1935-1945) et 2,818 (1945...)

Cette expérience de L'I.R.E.M. de Grenoble n'est pas unique. En effet, le Monde de l'Education (N° 18, juin 1976) dans un article intitulé "comment notent les correcteurs au Baccalauréat" a dégagé, à la suite d'interviews, non directives, de correcteurs dans différentes disciplines et Académies, une "classification de l'espèce". On y distingue 3 catégories principales

- les JUGES
- les TRADITIONALISTES
- les LIBERAUX

et deux catégories marginales :

- les débutants
- les incertains.

On y met aussi en relief l'importance du facteur "discipline" ; les plus sujettes à caution étant, comme les expériences précédemment relatées l'ont déjà montré, les disciplines littéraires.

L'influence des préjugés du maître - "effet de halo" - a été observée par de nombreux chercheurs. Les américains, dont nous citerons les principaux travaux, se sont particulièrement distingués dans ce domaine.

"R. Weiss a sélectionné deux rédactions faites par des élèves de 4° primaire ; les travaux ont été dactylographiés sans modification aucune puis soumis pour correction à deux groupes de 46 instituteurs enseignant aussi en 4° primaire. Le commentaire suivant accompagnait la distribution des travaux au premier groupe : "Voici deux compositions écrites par des élèves de 4° primaire ; le travail n° 1 est l'oeuvre d'un élève moyen qui aime lire les bandes dessinées, son père et sa mère sont employés ; le travail n° 2 a été fait par un enfant doué, son père est rédacteur d'un quotidien connu". Pour le second groupe de maîtres, les rôles ont été inversés. La correction devait se faire suivant une échelle à 5 degrés : T.B., B., M., S., I. L'orthographe, le style, le fond puis l'ensemble devaient être corrigés séparément.

Pour les 4 aspects considérés, les notes attribuées au travail pour lequel on a créé un préjugé favorable ont été significativement supérieures aux autres. Pour l'orthographe qui semblait le plus devoir échapper à l'effet de halo, on observe qu'au travail de l'élève présenté comme doué, 16 % des correcteurs accordent la note T. B. et aucun la note I. ; si le même élève est présenté comme moyen, les correcteurs n'accordent aucun T. B. mais 11 % notent insuffisant". (in Landsheere, 1-45).

Plus significatives encore sont les expériences de R. Rosenthal (1-74). "Les expérimentateurs devaient étudier le comportement de rats et les noter, au cours d'un programme comprenant des techniques classiques (apprentissage du labyrinthe et boîte de Skinner). Il fut dit à la moitié des expérimentateurs que les rats qu'ils devaient étudier avaient été spécialement choisis pour se comporter brillamment, à l'autre que les rats avaient été spécialement choisis pour être stupides ; bien entendu, les rats avaient été choisis au hasard. Dans les deux études de ce type réalisées par Rosenthal les expérimentateurs croyant leurs sujets brillants obtinrent un bien meilleur apprentissage de leurs rats que ceux qui pensaient que leurs sujets avaient été choisis pour être stupides".

Dans un travail récent, Rosenthal est allé encore plus loin en reproduisant une situation réelle de l'enseignement. "Dans 18 classes d'une école primaire américaine, 20 % des élèves choisis rigoureusement au hasard, furent signalés à leurs professeurs comme ayant eu des résultats particulièrement brillants à un test non verbal d'intelligence générale, permettant prétendument de prédire leur épanouissement intellectuel. Les enfants ne savaient rien, seuls les professeurs étaient au courant. Au bout d'une année, ces enfants avaient réellement un gain de quotient intellectuel franchement supérieur, en moyenne, à ceux du groupe de contrôle". (in Atome n° 242 article de H. Pequinot, 2-108).

Citons les travaux de C. Chase qui ont mis en évidence l'influence de la qualité de l'écriture sur les notes attribuées aux rédactions (The impact of some obvious variables, in journal of Educational Measurement, 1968) et ceux de Wilson qui ont montré que les enfants des zones géographiquement défavorisées sont cotés au plus bas, par le fait même qu'ils appartiennent à ces zones.

On peut rapprocher l'effet de halo des résultats décrits par C. Baudelot et R. Establet dans "L'école capitaliste en France" (5-1 p. 81). "Les probabilités de scolarisation pour un enfant de la bourgeoisie dans le réseau Secondaire-Supérieur et le réseau Primaire-Professionnel sont radicalement différentes : 50 % contre 14 %. Aussi différentes sont, pour un enfant de la classe ouvrière, les probabilités d'être scolarisés dans le réseau Primaire-Professionnel et le réseau Secondaire-Supérieur : 54 % contre 14 %... C'est-à-dire que les enfants de la bourgeoisie ont autant de chances d'être scolarisés dans le Secondaire-Supérieur que ceux de la classe ouvrière dans le Primaire-Professionnel."

## DEUXIEME SOURCE D'ERREURS : LE SUJET

Nous venons de voir que l'évaluation est fortement teintée par la personnalité de l'évaluateur. Mais cette source de fluctuation est loin d'être l'unique cause d'erreurs. Le choix même du sujet d'un contrôle ou d'un examen, les conditions de passation, interviennent également dans une large mesure. Des éléments, choisis pour éviter toute variation intempestive, tels que barèmes et coefficients de pondération sont loin de remplir les conditions de fiabilité souhaitées.

### Le barème

La plus grande partie des épreuves d'un examen fait intervenir un barème plus ou moins précis selon la matière. Ce barème a pour but d'uniformiser la codification des appréciations. Mais ce barème que les correcteurs sont "priés de respecter scrupuleusement" est-il un instrument fiable ? Plusieurs équipes de chercheurs se sont penchés sur ce problème.

L'I. R. E. M. de Rennes a fait corriger 22 copies de mathématiques (c'est l'épreuve de mathématiques, où le barème a une place fondamentale, qui a le plus souvent été testée) du B.E.P.C. par 10 professeurs. Cinq d'entre eux l'ont fait avec barème, les autres sans barème. L'analyse des résultats a mis en évidence trois points :

- les utilisateurs du barème ont corrigé plus sévèrement
- l'écart entre les notes extrêmes est moindre quand on tient compte du barème
- le barème ne supprime pas la dispersion des notes.

Devant la déception de voir le barème se limiter à un rôle vaguement modérateur, on a été amené à l'affiner, dans l'espoir d'obtenir une meilleure précision dans l'évaluation. Le Groupe de Recherche de Montauban dirigé par Cransac et Dauvisis s'est penché sur ce problème en 1975. Il leur a fallu plus de 15 heures de travail pour se mettre d'accord et élaborer un barème très précis (sur 120 points). Malgré cet effort

"pour une copie dont la note varie de 4 à 13 sur 20, chaque correcteur concerné a justifié sa notation et il a été impossible de donner raison à l'un ou à l'autre. Certains ont donné tour à tour raison à l'un puis à l'autre". D'autre part, le barème devient illusoire quand dans "une réponse peu claire, l'un voit un bon raisonnement, l'autre un raisonnement faux ; ce qui est inquiétant". Nouvelle déception. L'I.R.E.M. de Strasbourg reprend cette expérience. On raffine le barème. On demande aux correcteurs "un travail d'appréciation de l'écart des copies en les confrontant deux à deux afin d'améliorer leur évaluation." Le groupe de recherche espère ainsi "faire des observations des copies deux à deux". Résultats : la fiabilité de la correction n'est pas améliorée. Sans se décourager on fait une "tentative de codage des comportements de réponse" avec "détermination d'un barème associé à ce codage". Là encore c'est l'échec (1). Mais ces travaux remarquables amènent les chercheurs vers une autre direction : l'analyse plus soigneuse des sujets proposés.

#### Le choix du sujet

La présentation et la formulation du sujet d'une épreuve peuvent-être source d'erreurs et la note d'un candidat ne pas refléter sa valeur propre. En effet, peut-on évaluer sérieusement les qualités d'un élève qui n'a pas su faire un problème dont les questions sont dépendantes et la première piégée ? Peut-on évaluer sérieusement les qualités d'un élève qui disserte "hors sujet" quand, au cours d'une expérience de multicorrection, une copie de français a été annotée "sujet bien traité" par un correcteur et "sujet non traité" par un autre correcteur ? La formulation d'une question portant sur un concept mathématique donné détermine souvent l'exactitude de la réponse, tout comme, selon Piaget, la façon de poser une question orale à un candidat influence la réponse de ce candidat.

(1) - On retrouve dans ces études sur le barème l'évaluateur : "le barème imposé est appliqué en fonction de la personnalité des correcteurs. Certains l'appliquent à la lettre en faisant fi de leurs sentiments et de leur expérience pédagogique, d'autres ne peuvent ou ne veulent le faire et modulent le dit barème." (Ronceray, De l'influence du barème, I.R.E.M. Rennes)

Le Groupe de Recherche d'Epernay de l'I.R.E.M. de Reims a effectué une enquête sur ce thème en 1975. Il a mis en évidence la variabilité de l'exactitude des réponses selon les différents items d'un concept donné. Citons un exemple frappant : le concept de l'équation. A la question : "On achète deux pains, on donne dix francs, la marchande rend une pièce de cinq francs et une de un franc ; quel est le prix d'un pain ?, on obtient 93,44 % de réponses exactes dans une population d'élèves de troisième. A la question : "Résoudre dans  $R:2x+6=10$ ", le pourcentage de réponses exactes tombe à 81,49 %. Enfin si la question prend forme suivante : "Résoudre dans  $R : 6000x+2000=10000$ ." on obtient plus que 60,34 % de bonnes réponses. Il suffit donc de multiplier des données numériques 2, 6, 10 par 1000 pour que un quart des élèves ayant bien répondu à la question précédente se trompe à la dernière. Les résultats de cette enquête incitent "à la prudence dans le choix des critères qui permettent de déterminer le degré d'acquisition et de compréhension d'un concept mathématique. On peut également se demander si la compréhension d'un cours ne dépend pas étroitement de sa formulation, surtout en ces temps d'inflation et desophtisation du langage mathématique. Enfin on peut s'interroger sur le désir inconscient d'échec ou de succès du professeur vis-à-vis de ces élèves;" (in, Influence de formulation dans l'acquisition d'un concept mathématique, I.R.E.M. de Reims, 1976).

On retrouve la formulation source d'erreurs au Colloque sur l'Education à Strasbourg en mars 1976. La "lisibilité" a dominé les travaux sur l'analyse des sujets d'exams, particulièrement des sujets du baccalauréat. Les sujets proposés aux commissions de choix sont en effet souvent loin de représenter toutes les garanties. Ainsi "à Dijon, sur 30 propositions de sujets du Bac, 20 sont immédiatement rejetées comme non conformes" ; "un sujet de bac de Poitiers reproduit la première question du concours d'entrée à l'Ecole Normale Supérieure de deux ans avant..."

Ces travaux sont à rapprocher du rapport de la Régionale Parisienne de l'Association des Professeurs de Mathématiques, paru dans le n° 302 de la revue de l'A. P. M., sur le problème de terminale C (Paris, Juin 1975). Il s'agissait de souligner le caractère démesuré d'un texte de cinq pages, autant que ceux des concours d'entrée aux Grandes Ecoles ; de plus certaines questions étaient "piégées" et la partie du programme traitée fort réduite. Mais pour éviter des erreurs d'évaluation trop injustes "des efforts ont été faits dans le choix du barème pour améliorer, statistiquement, le rendement de l'épreuve." (sic)

#### COEFFICIENTS DE PONDERATION

Dans presque tous les examens et concours interviennent des disciplines affectées d'un coefficient de pondération. Or, selon la discipline, l'échelle de notation est soit ouverte, soit rétrécie. La probabilité d'obtenir un 2 ou un 18 en mathématique est supérieure à celle d'obtenir les mêmes notes en français. Un simple calcul montre alors que le coefficient de la discipline ne joue plus son rôle :

- français, coefficient 4 : note obtenue 8, soit 32 points,
  - mathématique, coefficient 2 : note obtenue 14, soit 28 points,
- ce qui donne un total de 60 points sur 120.

"On peut voir des élèves, conscients du phénomène, négliger une matière dotée d'un coefficient élevé. Témoin ces souvenirs de R. Ikor, rapportés dans "Les cas de conscience d'un professeur" : "J'avais accepté de faire une classe impossible, tout au moins pour un professeur de français : la préparation à H.E.C. Or il se trouve que le français, malgré un fort coefficient au concours n'y sert presque de rien, parce que les examinateurs resserrent à l'extrême leurs notes autour de la moyenne. Les résultats se jouent donc sur les matières scientifiques. Très vite ils me firent savoir qu'ils se souciaient du français comme d'une guigne... Je sais bien qu'ils n'avaient pas tort. Quel intérêt pour eux de s'échiner en français pour passer au plus de 9 à 10, alors qu'en mathématiques des 16, des 18 étaient possibles ?" (in Science et Vie n° 610, 1968)

D'autre part, l'étude des coefficients réels, appliqués à la correction des examens, publiée par l'Institut National d'Etude du Travail et d'Orientation Professionnelle en 1958, montre que l'importance relative donnée à chaque matière par ces coefficients est faussée par le fait que l'échelle des notes utilisée ne couvre pas la marge entière de 0 à 20. Ainsi, dans l'épreuve de dictée, les notes s'échelonnent de 0 à 20. Par contre, l'appréciation de la présentation de cette même dictée est cotée entre 8 et 16. Si l'on affecte dictée et présentation du même coefficient théorique, on voit que la dictée a deux fois plus d'importance que la présentation. Le coefficient "réel" de la dictée est ainsi double de celui de la présentation, la différence des moyennes provenant de la différence de dispersion des notes. Voici, toujours selon l'I.N.O.P., quelques exemples de coefficients théoriques et réels de diverses épreuves :

	<u>Coefficient théorique</u>	<u>coefficient réel</u>
Dictée	3	4 à 5
Questions	4	2,8 à 4
Compte-rendu	3	1,9 à 2,8
Calcul	6	5,6 à 7
Présentation	1	0,4 à 0,9

TROISIEME SOURCE D'ERREUR : L'EVALUE

"Une cause non négligeable de fluctuation réside dans deux facteurs difficiles à maîtriser et à estimer : la variabilité des candidats et des conditions extérieures lors du déroulement des opérations. La réaction affective de chaque élève à la situation diffère selon sa stabilité émotionnelle du moment, son état de santé, la pression familiale ; en outre, du fait du temps limité, un incident mineur, même le bris d'une pointe de crayon, constitue un handicap inaperçu par le professeur. Dans des épreuves comme la dictée, la place occupée dans la salle peut fausser les résultats. De tels facteurs d'irrégularité sont spécialement actifs au cours des périodes de rapide évolution physique ou intellectuelle ; en particulier, plusieurs auteurs les ont mis en évidence au moment de la puberté". (Hotyat, in Revue Française de Pédagogie, Janvier 1968)

"Ce qui frappe aussitôt, c'est l'extrême variété et complexité des réactions individuelles des élèves confrontés aux épreuves et aux estimations qu'en font les maîtres. Il apparaît alors à l'évidence que l'examen, voire la simple composition de contrôle, sont d'abord des "événements" à caractère historique, qui s'inscrivent dans le destin personnel de l'élève, dont il est difficile de les dissocier. Qu'on pense à tel enfant de sixième pour qui chaque épreuve scolaire marque l'heure d'une nouvelle et solennelle remise en question de son statut à l'égard du professeur, et plus encore de ses parents, tandis que tel autre, au même âge et dans la même classe, vit les examens comme une ennuyeuse corvée, dont il faut se tirer économiquement pour conserver le minimum de quiétude indispensable à la survie en milieu d'école. Qu'on pense à la panique de quelques uns devant le baccalauréat, à l'excitation conquérante de quelques autres, à la conduite soumise et humblement séductrice d'un certain nombre, au fatalisme de ceux-ci, à l'ingéniosité défensive, parfois malhonnête, de ceux-là, et ainsi de suite. Il est clair que ces conduites pèsent lourdement sur le résultat final de l'examen écrit ou oral, même si l'on admet, par impossible, que le maître demeure, pour sa part, à l'abri de toute variation ou inconsistance dans l'appréciation qu'il porte. Si l'on considère alors que de tels effets ne sont intelligibles qu'en référence à la personnalité de l'élève, et que cette personnalité ne se comprend à son tour qu'en fonction d'un contexte socio-familial et d'un devenir d'ensemble qui lui est rigoureusement propre, on admettra qu'il y a de quoi inquiéter un esprit

soucieux de "justice" et rendre quelque peu suspectes les tentatives de réduction statistique concernant la notation, encore que les règlements rendent obligatoires classements et notations. A la limite en effet, il n'y a plus, en cette voie que des cas rigoureusement individuels, et donc incomparables". (Guillaumin, Docimologie et éducation, Les sciences de l'éducation, avril-septembre 1969).

"Lors d'un sondage belge sur le travail après la classe, la question suivante avait été posée à des élèves de 6°, 4°, 3°, 1° d'enseignement moyen et de dernière année d'école normale (2031 sujets). "Lors d'un examen ou d'une interrogation écrite, devenez-vous nerveux au point de ne pas obtenir des résultats correspondant à vos moyens et à vos efforts ?" Trois degrés étaient prévus pour les réponses : rarement, souvent, presque toujours. Plus de la moitié des interrogés signalaient ressentir souvent ou presque toujours un tel handicap. Contrairement à ce que l'on avait pu attendre, les fréquences étaient à peu près égales chez les garçons (54,2 %) et chez les filles (54,6 %) apparemment plus émotives. En revanche, les élèves étant répartis en trois groupes selon leurs résultats dans leur classe, les fréquences étaient respectivement les suivantes : 38,2 % dans le tiers supérieur, 59,3 % dans le deuxième tiers et 67,5 % dans le tiers inférieur.

Cependant, des difficultés ont rarement un caractère pathologique : en Grande Bretagne, un relevé des cas soumis aux Child Guidance Clinics, entrepris dans une aire géographique courant 212000 écoliers de 5 à 16 ans, révéla que sur 5 705 sujets ayant eu besoin de consultations des services d'hygiène mentale, le psychologue ou le psychiatre avaient signalé les examens comme facteurs de trouble dans 41 cas seulement". (Hotyat, Revue Française de Pédagogie, Janvier 1968).

DOCIMOLOGIE CONSTRUCTIVE

I INTRODUCTION

Depuis les premiers travaux ci-dessus rappelés de Pieron, Laugier et Weinberg dans les années vingt, de nombreuses critiques ont été formulées à l'égard des différentes sortes d'examens, qu'ils soient externes comme le B.E.P.C., le Baccalauréat, les examens de médecine, les concours des grandes écoles, ou bien internes à l'établissement scolaire - compositions, interrogations de contrôle, devoirs sur table.

En France, le débat est resté souvent confus, plus philosophique que visant des applications concrètes ; les considérations pertinentes sur la docimologie "n'ont guère mordu sur le public, et pas même sur celui des premiers intéressés, les pédagogues" (Guillaumin 5-14 p. 130). Pourtant, ailleurs et particulièrement dans les pays anglo-saxons, au Canada, en Belgique, l'évaluation des résultats de l'enseignement fait partie intégrante de toute tentative de réforme.

Comme le fait remarquer Reuchlin (1-67 p.234). "Il paraît difficile d'échapper à la nécessité pédagogique et sociale d'évaluer les résultats de l'éducation".

Nous n'entrerons pas dans la querelle qui consiste à opposer les systèmes fondés sur les notes scolaires ou sur des appréciations plus ou moins détaillées, Remondino (cite par BACHER in 2.8 p. 85) a d'ailleurs montré que dans des conditions bien définies une correspondance étroite s'établit entre les 2 systèmes de notation.

Les difficultés que l'on rencontre à l'occasion de cette évaluation ont été longuement rappelées précédemment et tous les auteurs les décrivent avec plus ou moins de détails (cf Landsherre 2-74 p. 72-73).

Il demeure que l'évaluation est nécessaire et une association de parents d'élèves n'a pas manqué d'en faire le thème d'étude de l'une de ses commissions dans le cadre d'un projet d'école (La Famille et l'Ecole n° 166 - octobre 1975 F C P E Paris) ; non seulement son importance est affirmée pour les examens externes, et l'orientation, mais elle est au moins implicite lors des actions pédagogiques visant au rattrapage, au soutien, à la réduction des retards scolaires.

Il faut bien voir que l'objet des examens et des contrôles des connaissances n'est pas univoque ; il est rare qu'une épreuve ne vise qu'un seul but que ce soit de manière explicite ou sous entendue.

Toute épreuve est tout d'abord un bilan, un inventaire des connaissances acquises par les sujets qui la subissent, à un moment donné qui marque une étape dans la progression d'un enseignement.

On diagnostique ainsi les faiblesses, les insuffisances, les manques non pas seulement au plan des connaissances mais aussi le plus souvent dans la manière de les utiliser, et l'on en infère très vite des manques ou des insuffisances au plan mental, psychologique. Le diagnostic effectué fonde des jugements qui vont au delà même des faits observés.

Bilan et diagnostic ont une valeur au moment où ils sont formulés ; mais les épreuves qui les permettent, remplissent aussi le plus souvent, une autre fonction qui est celle là prédictive. Il s'agit alors du déroulement futur d'une scolarité, ou des choix qu'il convient de faire dans un système scolaire donné. Cette fonction supplémentaire, c'est le pronostic d'une réussite future, à court terme ou à moyen terme, ou tout au moins l'ouverture vers de nouvelles études.

La nécessité apparaît donc d'une docimologie qui soit constructive, qui apporte plus de rigueur dans l'évaluation, qui se fonde "sur une démarche explicite permettant l'étude objective de faits contrôlables" (Reuchlin 1-67 p. 234).

Elle implique d'ailleurs de rechercher "une formulation explicite aux objectifs de l'éducation" (ibid.) ce qui sera envisagé dans un autre chapitre.

En l'état actuel des études cette face de la docimologie n'est pas à l'abri de la critique mais elle témoigne d'un effort de clarification et montre des directions.

Nous envisageons donc successivement dans ce chapitre :

- l'amélioration des examens externes de type traditionnel et quelques remèdes aux distorsions qui les affectent.
- les épreuves normalisées de connaissances dont la méthodologie s'inspire largement de celle que les tests psychologiques ont permis de développer d'une manière rigoureuse.
- les épreuves où l'expression libre est recherchée qu'il s'agisse de dessin ou de rédaction dans la langue maternelle.
- enfin l'utilisation qui peut être faite des notes et appréciations recueillies dans le domaine scolaire et l'harmonisation des jugements, indispensables lorsqu'ils émanent des divers enseignants d'un même établissement ou d'établissements différents (procédure dite de "modération").

Suivant en cela F. Bacher (2-8 p. 75), nous nous garderons cependant d'oublier que l'évaluation par des notes de la production des élèves peut être :

- un moyen pédagogique individualisé et dans ce cas la notation devra tenir compte non seulement du niveau de connaissance des capacités, des savoir-faire acquis par l'enfant ou l'adolescent, mais aussi des efforts, des progrès, en un mot de l'effet pédagogique recherché ; nous avons là une évaluation personnalisée.

- un moyen de comparer les élèves entre eux, d'établir une hiérarchie de niveau comme dans un classement en cours d'année, un examen de fin de scolarité, un examen d'entrée dans un cycle d'étude, un concours etc... Il est alors indispensable, si l'on ne veut courir des risques qui peuvent être graves, que ces notes soient comparables d'un élève à l'autre.

- enfin le moyen de vérifier si ces élèves atteignent les normes définies (par exemple par des objectifs de connaissances, de capacités à acquérir). Il convient dans ce cas également que les notes d'élèves différents soient établies sur des critères qui les rendent comparables ; en particulier des élèves ayant reçu la même note doivent se situer identiquement par rapport aux normes déterminées non par rapport aux élèves mais directement en terme de contenu et d'objectifs de l'enseignement.

Distinguer ces cas n'est pas toujours facile dans la pratique, mais cette distinction est indispensable si l'on veut dans chaque situation examiner, et tenter de résoudre les problèmes qui se posent.

AMELIORATION DES EXAMENS EXTERNES DE TYPE TRADITIONNEL

"Les examens externes semblent devoir subsister encore longtemps et il n'est pas inutile de chercher à atténuer leurs imperfections".

(1-6 p. 38)

Les responsables des principaux examens à l'échelon national ont procédé à différentes améliorations ou aménagements dans l'espoir d'atténuer les injustices ou erreurs normalement liées à ce type d'épreuve.

Les propositions ne manquent pas à ce sujet : les chercheurs ne sont depuis longtemps penchés sur ces problèmes et ont donné de nombreuses solutions dont quelques unes ont été mises en application.

Fraisse nous propose par exemple l'utilisation des statistiques : "Pour l'instant, le plus utile serait d'initier - ou plutôt de convertir - les examinateurs à quelques notions élémentaires de statistiques". (1-32 p. 91)

Bacher lui fait écho : "Une simple correction statistique permettrait d'égaliser les échelles de notes des correcteurs et d'éliminer l'effet des différences de sévérité". (1-6 p. 38)

Fraisse répond par un exemple : "L'office du baccalauréat a tenté il y a quelques années de mettre en garde ses administrés. Les correcteurs recevaient communication de la moyenne générale des notes ainsi que de l'écart étalon pour toute l'épreuve. Ils étaient invités à comparer à leur propre moyenne et à compenser leur indulgence ou leur sévérité relative au moment de la délibération" (1-32 p. 91).

Fraisse encore montre un procédé voisin utilisé dans certaines facultés : "Le président demande à tous ses collaborateurs d'observer un pourcentage fixé à l'avance, ou du moins, de s'aligner sur la moyenne du groupe. Ce procédé, si empirique qu'il soit, est le seul qui assure aux candidats une certaine égalité des chances" (1-32 p. 91).

Nous laissons bien sûr à Fraisse la responsabilité de ses affirmations.

Bacher, de son côté, propose plus simplement une information des correcteurs : "Un moyen moins global consiste à informer tous les correcteurs, dans une discipline, de la façon dont se répartissent leurs notes par rapport aux notes données par l'ensemble de leurs collègues. Chaque correcteur peut alors, le cas échéant, modifier ses notes pour les rendre plus comparables à l'ensemble de celles qui ont été données.

Cette transformation n'implique pas une nouvelle correction des copies, dont le classement est inchangé". (1-6 p. 39)

D'autres, comme Pannetier, envisagent concrètement la situation et donnent une manière d'opérer pour normaliser les notes, méthode qui peut sembler brutale mais qui peut résumer assez bien le problème du réajustement des notes :

"Le mode opératoire est le suivant :

- 1) noter les copies sans autre souci que l'objectivité, donc sans tenir compte notamment de la difficulté du sujet.
- 2) tracer l'histogramme des notes attribuées.
- 3) analyser l'histogramme. Généralement, plusieurs groupes apparaissent dans la distribution. C'est sur l'une des séparations que le correcteur placera la frontière entre "suffisants et insuffisants" dont les nombres peuvent être très différents (si aucun groupement ne se révèle dans la distribution, toute la responsabilité du choix de la frontière incombe au correcteur).
- 4) quelle que soit la note correspondant à cette frontière, la remplacer en principe par la note 10.
- 5) réajuster toutes les autres notes par un glissement de points en plus ou en moins". (4-5 p. 23)

Les conseils restent bien souvent lettres mortes et Fraise poursuit d'un ton amer :

"Cette pratique a disparu... Les professeurs sont simplement invités à calculer leurs moyennes, recommandations, toute pléthorique car elle n'est suivie d'aucun effet". (1-32 )

Bacher nous donne une explication de cet abandon :

"Cette correction des notes n'est habituellement pas admise, car chaque correcteur tient à rester maître de ses appréciations". (1-6 p. 38)

Les erreurs de correction ou de notation, n'arrivent-elles qu'aux autres ?

Aussi est-ce indirectement que l'on peut espérer réduire ces différences :

"Un premier moyen consiste à équilibrer les jurys ; lorsqu'un des examinateurs est connu pour sa sévérité, on peut affecter au même jury un examinateur connu pour son indulgence". (1-6 p. 39)

Le moyen semble difficile à mettre en pratique dans la mesure où les examinateurs ne sont pas toujours connus à l'avance et dans la mesure où le caractère humain n'est pas une constante sur laquelle on peut placer une étiquette définitive. Il faut donc se tourner vers d'autres pratiques, comme par exemple l'établissement d'un barème :

"On sait qu'il s'agit par exemple de convenir d'un barème analytique appliqué par tous les notateurs, ou de noter en commun quelques devoirs afin d'harmoniser les points de vue, ou de procéder à une double correction. De façon générale, l'introduction d'un barème atténué de façon considérable les divergences entre estimateurs".

(1-67 p. 220) (Sur ce point, l'accord n'est pas général : on se reportera au chapitre sur le barème).

Bacher ne peut qu'approuver :

"L'établissement d'un barème, la discussion préalable par plusieurs correcteurs de la note à attribuer à quelques copies, permettent aussi de limiter les effets d'une inégale sévérité ; mais leur intérêt est surtout de faciliter l'adoption, par l'ensemble des correcteurs, d'échelles de valeurs communes" (1-6 p. 39)

Il est donc bien difficile de corriger les défauts de nos examens actuels. Il faut se méfier des rectifications trop hâtives ou trop sommaires.

Husson devient virulent lorsqu'on lui parle de moyenne des notes :

"La moyenne arithmétique n'est qu'un indice grossier et même dans bien des cas fallacieux. On oublie, quand on se fie à elle qu'il ne s'agit pas dans un examen de parvenir à un pourcentage global satisfaisant de reçus, mais de prendre des décisions individuelles aussi équitables que possible. Deux correcteurs peuvent parvenir à la même moyenne en distribuant très différemment les bonnes et les mauvaises notes par un procédé purement mathématique est non seulement illusoire mais dangereuse : elle substitue à des notes, subjectives sans doute, mais dont chacune a une signification, une manipulation de nombres opérant à l'aveugle". (5-14 p. 104)

Alors, si les examens posent tant de problèmes au niveau de la correction et de la notation, ne pourrait-on pas modifier leur nature même, de façon à augmenter considérablement l'objectivité de cette correction ? Nous allons tenter de répondre à cette question.

LES EPREUVES NORMALISEES DE CONNAISSANCE

"De nouveaux modes d'évaluation des élèves entrent dans nos moeurs pédagogiques. Les résultats de l'observation continue prennent progressivement le pas sur l'examen de fin d'année. L'innovation est louable. Elle ne traduit nullement la volonté de supprimer la mesure à l'école, mais, au contraire, de la rendre moralement plus juste et scientifiquement plus exacte". (Landsheere - 1-45 p. 7)

Tous les chercheurs en docimologie se sont intéressés à la méthode des tests :

F. Bacher : "On peut largement s'inspirer de la méthode des tests pour rechercher des modes d'évaluation des connaissances ne dépendant pas du jugement du correcteur". (1-6 p. 41)

Et A. Bonboir : "Le test est un outil dont la valeur est relative à l'usage que l'on veut en faire, admettant qu'il convient de l'employer conformément aux règles imposées par sa nature même" (1-11 p. 133)

L'élaboration des épreuves exige le plus grand soin et nous allons suivre pas à pas le plan de travail préconisé par Lansheere dans le cas d'épreuves normalisées :

"Les grandes phases de la construction d'un examen sont :

- définir l'objet et les objectifs ;
- rédiger les questions ;
- standardiser la présentation, l'administration et la correction ;
- étalonner ;
- éprouver la fidélité et établir la validité de l'examen (1-45 p. 59).

Nous allons donc envisager dans le détail la construction d'épreuves de connaissance en s'appuyant sur les recherches effectuées par quelques uns de nos éminents docimologues.

ELABORATION D'UNE EPREUVE NORMALISEE

1 - D'abord définir l'objet et les objectifs.

"La construction d'un test commence par la définition des caractères que l'on veut mesurer... Les épreuves de connaissances, elles, portent d'abord sur le contenu d'un programme" (1-6 p. 42)

"Dans tous les manuels récents, on recommande à celui qui veut construire une épreuve de connaissance d'établir un plan conçu selon deux dimensions : contenu du programme, objectifs" (1-6 p. 44)

"Tout dépend de l'usage envisagé. De cet usage dépend le choix des items suivant leur difficulté et leur pouvoir discriminatif" (1-11 p.141)

"Qu'est-ce qu'un item ? Ce mot, d'origine informatique, désigne un élément de réponse que l'on peut juger par tout ou rien, c'est-à-dire de façon binaire : réponse acceptée ou refusée" (FRAPPE 5-18 p. 93)

2 - Ensuite rédiger les questions.

"Deux méthodes ont été proposées pour formuler les questions : les questions à choix multiple (Q.C.M.) et les questions à réponse ouverte et courte (Q.R.O.C.)" (FRAPPE 5-18 p. 106)

En ce qui concerne les Q.R.O.C., nous renvoyons le lecteur au paragraphe intitulé "Epreuves à expression libre".

"Il convient de rassembler un nombre élevé de questions couvrant tout le champ sur lequel l'évaluation doit porter ; d'imaginer dans certains domaines des formes de questions exigeant de l'élève un travail d'élaboration et de mise en oeuvre de ses connaissances, et non un simple effort de mémoire" (1-67 p. 221)

"Le travail du constructeur de test consiste en premier lieu en une analyse du travail demandé à l'élève" en s'inspirant par exemple "d'exercices trouvés dans les manuels scolaires ou proposés à l'occasion d'examens. Cette analyse doit conduire à la formulation d'une série de questions plus limitées et plus précises, telles que la réponse à fournir puisse être choisie parmi une série de réponses proposées (Q.C.M.) ou soit suffisamment délimité pour qu'on sache sans ambiguïté si la réponse fournie par un élève est correcte ou non (Q.R.O.C.)" (1-6 p. 42-43)

"Savoir poser des questions est probablement la capacité la plus nécessaire au professeur. Mais c'est un art difficile, les erreurs dans le choix du niveau de langage et les obscurités fréquemment rencontrées dans le libellé des exercices d'application et les questions d'examen en témoignent". (1-45 p. 83-84)

"Autant que possible, il faut essayer les questions avant de les utiliser pour un examen. On découvre ainsi les imprécisions, les défauts de rédaction, les erreurs matérielles et, aussi, le niveau de difficulté". (1-45 p. 85)

"De nombreuses objections ont été faites à l'emploi des questions à choix multiple. Certaines semblent peu fondées : ce n'est pas parce qu'un mode de réponse est simple que l'opération permettant de trouver la bonne réponse l'est aussi... Peu à peu, des formes de réponse plus élaborées ont été proposées : la bonne réponse peut correspondre à une combinaison de choix, ou à un classement..." (1-6 p. 47)

"Une autre façon d'atténuer les effets du hasard, et en même temps d'accroître l'information apportée par les réponses, qu'elles soient bonnes ou mauvaises, consiste à construire l'épreuve de façon telle que toutes les éventualités de réponse répondent à certains critères... la série des réponses comprend par exemple une réponse juste, une réponse approximative, une réponse représentant une information inexacte, une réponse correspondant à une lourde ignorance, et enfin la réponse "je ne sais pas"..." (1-6 p. 47)

"Un autre type d'épreuve dans lequel la bonne réponse est déterminée bien que les diverses éventualités ne soient pas proposées est celui dans lequel on demande à l'élève de détecter et de corriger des erreurs". (1-6 p. 48)

"Lorsque la question posée délimite suffisamment les possibilités de réponse pour qu'il n'y ait pas de doute sur l'exactitude ou l'inexactitude de celle proposée par l'élève, on peut enfin utiliser des réponses libres". (1-6 p. 48)

### 3 - Correction et notation.

"Un problème est encore mal résolu pour nous : c'est celui de la pondération des items. Parmi les notions que nous essayons de contrôler, faut-il établir une hiérarchie en attribuant un coefficient plus élevé pour certaines d'entre elles. Deux possibilités peuvent être prises en considération. Ou bien, on introduira une plus grande proportion d'items portant sur les notations qui paraissent plus importantes, ou bien on n'attribuera pas le même nombre de points à chaque item. Ces deux solutions présentent des inconvénients. Dans le premier cas, les tests seront forcément plus longs pour la même quantité d'information... Dans le second cas, le coefficient qu'il faut attribuer à un item n'est pas facile à déterminer". (5-14 p. 45)

Les réponses à cette question sont nombreuses et variées. Nous nous contentons de présenter un exemple particulièrement significatif. Il s'agit des propositions du Dr FRAPPE qui s'est intéressé au concours d'Internat de Médecine ; propositions qui pourraient s'appliquer à n'importe quel autre examen.

"Nous pouvons, en corrigeant les épreuves par items, tenter de noter les candidats. Une copie est supposée d'autant meilleure qu'elle contient plus d'items réussis. Le nombre d'items réussis est le "score brut" du candidat. Une simple règle de trois permet de ramener ce score brut sur 20 et d'en faire une note". (FRAPPE 5-18 p. 93)

Cette méthode très simple peut se révéler insuffisante :

"La notation par le score brut n'est pas satisfaisante car elle attribue la même valeur à tous les items, ce qui supprime toute notion de pondération de ces derniers". (FRAPPE 5-18 p. 94)

Le Dr FRAPPE définit alors une méthode originale de pondération des items suivant leur importance :

"La correction par items renseigne le jury sur la proportion de candidats qui ont réussi chacun des items. On peut ainsi définir un "coefficient de réussite" pour chaque item. Ce coefficient de réussite est le pourcentage de candidats ayant réussi chacun des items. Si aucun des candidats n'a réussi l'item, ce coefficient est nul. Il est égal à 100 % si tous les candidats l'ont exprimé... Pourquoi ne pas attribuer tout simplement son coefficient de réussite à l'item considéré ? De la sorte un candidat qui rate ce que la majorité de ses concurrents réussissent est lourdement sanctionné... A l'inverse, comment accorder la moindre importance à un item qui n'a été réussi que par très peu de candidat ? Dans ces conditions, il suffit, pour chaque candidat, de déterminer les items qu'il a réussis et de faire la somme de leurs coefficients de réussite. Barrès et Riss ont nommé cette somme "Score Pondéré individuel". (FRAPPE 5-18 p. 95)

Cette méthode est particulièrement applicable dans le cas de concours car "le coefficient de réussite est une variable continue pouvant être calculée avec la précision que l'on désire". (FRAPPE 5-18 p. 96)

Les arguments défavorables à cette méthode ne manqueront pas : par exemple

"En notant les candidats d'après leur réussite, on risque de n'attribuer aucune importance à un élément jugé capital par le jury. Cela se produit en effet lorsque peu de candidats ont réussi cet élément "important"... Cependant, si cet élément a été réussi par très peu de candidats, on peut penser qu'il n'a pas été enseigné dans les différentes Facultés avec l'insistance que seuls les membres du jury de ce concours auraient souhaitée. Il est alors permis de douter de son "importance". (FRAPPE 5-18 p. 101)

"De tout cela il est possible de conclure que le bon candidat doit, avec cette notation, d'abord éviter de manquer ne serait-ce qu'un seul item "facile", c'est-à-dire à coefficient de réussite élevé. Ce n'est qu'après avoir rempli cette condition qu'il se distinguera de ses concurrents en réussissant en plus quelques items moins faciles". (FRAPPE 5-18 p. 102)

La démarche de correction diffère sensiblement de la méthode habituelle :

"Remarquons qu'il n'est pas possible d'attribuer une note à un candidat avant d'avoir corrigé l'ensemble des copies de l'épreuve" (5-18 p. 103)

Il faut en effet déterminer d'abord le coefficient de réussite de chaque item.

"Il est souhaitable, par ailleurs, que la correction s'effectue item par item, et non candidat par candidat. La correction d'un item donné pour l'ensemble des candidats semble en effet devoir être plus stable dans le temps que la correction de tous les items pour un candidat" (5-18 p. 105)

Lorsque les problèmes de notation seront résolus, il faudra encore suivre les règles de Landsheere.

#### 4 - Eprouver la fidélité et la validité de l'épreuve.

"Le contrôle de la fidélité relève de la même préoccupation : idéalement, un même examen (écrit ou oral) passé plusieurs fois, sans que l'élève ait le temps d'apprendre des choses nouvelles, devrait toujours conduire au même résultat". (1-45 p. 169)

"L'étude de la fidélité de l'épreuve pose aussi quelques problèmes. On ne peut l'estimer par ce test dans la mesure où des modifications systématiques au cours du temps sont attendues sous l'effet de l'enseignement... La meilleure méthode pour étudier la fidélité d'une épreuve de connaissances semble être de construire et d'appliquer deux épreuves parallèles en terme de contenus et d'objectifs". (1-6 p. 51)

"L'examen doit bien échantillonner l'ensemble. En principe, plus on ajoute de questions, plus la fidélité augmente" (1-45 p. 170)

"La fidélité d'un examen ne garantit en rien sa validité. Un mètre mal construit, qui mesurerait cinq centimètres de trop, indiquerait la même longueur à chaque mesure, mais conduirait néanmoins à une conclusion fautive. En bref, valider un examen, c'est prouver qu'il mesure effectivement ce pour quoi il est proposé". (1-45 p. 172)

Il est encore de nombreuses qualités métriques qu'il faut mesurer pour augmenter les chances d'objectivité d'un examen :

- la facilité des questions, par le pourcentage de réussite ;
- l'efficacité des questions et leur pouvoir discriminatif ;
- la sensibilité : épreuves globales portant sur un programme complet ou seulement sur une partie du programme.

Pour plus de détails, nous renvoyons le lecteur aux ouvrages mentionnés en bibliographie.

### LES AVANTAGES

L'unanimité se fait autour de la correction des épreuves normalisées :

"La notation est entièrement automatique, ce qui exclut toute possibilité de divergence entre estimateurs". (1-67 p. 221)

"Les questions à choix multiples, courantes dans les tests cognitifs, ont été d'autant plus souvent utilisées pour les épreuves de connaissances que le développement de celles-ci a coïncidé avec celui des machines à corriger les tests, puis des ordinateurs. Cette forme de réponse se prête, en effet, particulièrement bien aux modes de correction automatiques". (1-6 p. 47)

"La correction rigoureusement standardisée prend peu de temps et il est possible d'envisager une forme compatible avec un contrôle mécanique". (1-11 p. 137)

Le mot objectivité se retrouve lui aussi dans de nombreuses citations au chapitre des avantages :

"Le test est objectif, sensible aux différences individuelles et entre groupes (classes). Les repères sont bien établis et ainsi, les comparaisons rigoureuses sont possibles". (1-11 p. 138)

On parle aussi beaucoup de gain de temps :

"Dans un temps relativement court, les tests permettent d'examiner toute la classe relativement à une matière étendue (celle enseignée au cours de l'année, du trimestre, du semestre)". (1-11 p. 137)

"Ces épreuves de connaissance permettent un diagnostic rapide, sur un champ étendu, des lacunes observables dans une classe, par rapport à un large ensemble de classes de ce niveau". (1-67 p. 221)

"Elles constituent en outre un mode d'évaluation utilisable dans beaucoup d'autres circonstances : détermination du niveau des élèves pour l'individualisation de l'enseignement, analyse de leur progression, contrôle d'expériences pédagogiques, étude des effets d'un changement de structure ou de programme, etc..." (1-11 p. 40)

"Par rapport aux épreuves d'examen classiques, l'épreuve normalisée de connaissances, composée d'une série relativement étendue d'exercices, présente l'avantage de mieux couvrir l'ensemble du programme". (1-11 p. 42)

"Un autre apport d'intérêt général réside dans le fait que la valeur des épreuves normalisées de connaissances est considérée comme un objet relevant de méthodes scientifiques". (1-11 p. 56)

C'est aussi l'opinion de Reuchlin :

"Les épreuves normalisées ont cependant le grand avantage d'ouvrir une voie dans laquelle il est possible de progresser par des procédures expérimentales. Elles ont un autre avantage : elles dissocient le problème de l'objectivité de l'estimation au sens de l'accord entre estimateurs (problème qu'elles résolvent complètement) du problème de l'adéquation de l'estimation à ses objectifs, c'est-à-dire aux objectifs de l'éducation". (1-67 p. 223)

LES LIMITES

Après une longue liste d'avantages, il faut présenter aussi quelques inconvénients de ces tests :

L'objection la plus fréquente :

"Ces épreuves détourneraient les élèves de l'entraînement à la rédaction. C'est incontestable, et il convient en général de ne pas employer seul ce type d'épreuves". (1-67 p. 222)

Nous renvoyons une nouvelle fois le lecteur au chapitre suivant.

"On a fait remarquer que certaines bonnes réponses pouvaient être obtenues par hasard... On peut en atténuer la portée en augmentant le nombre des réponses proposées... ou en pénalisant les mauvaises réponses". (1-67 p. 222)

"La forme des épreuves peut aussi faciliter le copiage... Le problème est surtout d'acceptation par les étudiants de l'examen et de sa forme... Le bachotage constitue une autre difficulté très sérieuse". (1-6 p. 54-55)

Il existe en effet en Angleterre et aux Etats-Unis des recueils de questions ou même des écoles spécialisées qui pratiquent un bachotage efficace "sans pour autant donner aux élèves qui le subissent une supériorité durable".

"Il arrive assez souvent que les épreuves de connaissances, lorsqu'elles sont utilisées, le soient en fin d'année et présentent le même caractère ponctuel dans le temps que n'importe quel examen... Chacun sait que ses propres performances varient d'une occasion à l'autre, sous l'effet de dispositions temporaires ("forme", fatigue, difficultés personnelles, etc...)... Couvrir l'ensemble du programme constitue un progrès par rapport aux examens dans lesquels un seul sujet est proposé, mais ne supprime pas la nécessité de procéder à des examens répétés". (1-6 p. 53)

D'autres inconvénients proviennent de la forme même de l'examen :

"Une familiarisation antérieure avec ce type d'épreuve améliorerait la performance et il est nécessaire de la donner à tous les candidats pour les mettre dans des conditions comparables au moment de l'examen... Il est de toutes façons probable qu'il existe des différences individuelles dans la rapidité d'adaptation... A des niveaux élevés, l'épreuve normalisée de connaissances exige, par la complexité de la formulation des questions, un haut niveau de compréhension de lecture, tandis que les examens traditionnels privilégient la facilité d'expression". (1-6 p. 54 et 56)

Enfin il existe des inconvénients d'ordre purement technique :

"Les tentatives d'utilisation d'ordinateurs pour la correction d'épreuves à réponse libre en sont encore à leurs débuts. Les difficultés viennent de ce que, même dans une situation très structurée pour laquelle la bonne réponse est bien déterminée, de multiples formulations de cette réponse sont souvent possibles. Par exemple, une même formule mathématique peut être donnée sous des formes très diverses". (1-6 p. 49)

"La complexité des consignes données à l'élève, et parfois la multiplicité de ces consignes dans une même épreuve, introduisent en outre une difficulté qui n'a rien à voir avec l'objet de l'évaluation". (1-6 p. 54)

"La préparation des tests est longue et coûteuse ; celle du programme pour le contrôle par machine - très coûteuse - requiert une mise au point laborieuse et minutieuse ; souvent tous les documents doivent être revus - et parfois transposés - si on veut que la machine les accepte". (1-11 p. 137)

D'autre part, il est bien évident que ces épreuves ne peuvent servir qu'une fois.

Landsheere parle aussi d'objectivité trompeuse :

"La subjectivité est loin d'avoir été complètement éliminée : elle subsiste, au moins en partie, dans la rédaction des questions et dans la décision concernant la réponse à considérer comme correcte". (1-45 p. 99)

"La traduction des objectifs en comportements observables ainsi que la définition des critères d'appréciation ne sont pas uniques et peuvent susciter des désaccords. Rendre ces choix explicites, mettre en évidence les points de désaccord, en faciliter la discussion constitue peut-être l'apport le plus important du développement des épreuves normalisées". (1-6 p. 56)

"Plus fondamentalement, on a reproché aux épreuves normalisées de n'atteindre qu'un savoir fragmentaire, et non les capacités d'élaboration, de raisonnement, d'expression ! Il reste vrai que les élaborations mentales les plus complexes et surtout l'élégance de la pensée et de l'expression se laissent mal appréhender par de telles épreuves". (1-67 p. 222)

"Actuellement les épreuves normalisées ne permettent pas de contrôler tous les objectifs d'un enseignement et leur emploi exclusif pourrait même conduire les enseignants et les élèves à négliger certains objectifs importants". (1-6 p. 58)

#### CONCLUSION

"Les tests portant sur les acquis dérivant de l'acte pédagogique lui-même se présentent comme le moyen radical proposé par une docimologie positive soucieuse de rendre objectif le classement qui résultera de l'appréciation du travail des "écoliers" et de fonder le conseil ou l'action qui orienteront les sujets dans une voie où leurs chances de succès sont optimales. Ainsi, leur fonction est triple : établir un constat ; analyser de façon détaillée et complète l'état constaté ou faire un diagnostic inspirant l'action ; établir un pronostic, soit sous la forme d'une prédiction statistique, uniquement, soit dans l'exercice d'une prévision de tous les instants". (1-11 p. 132)

"Quoi qu'il en soit, il reste qu'actuellement - à tort ou à raison - les tests contrôlant l'acquis pédagogique, les tests de maturité spécifique exceptés, se bornent à l'examen du domaine cognitif et, à l'intérieur de celui-ci, l'accent est surtout mis sur les connaissances". (1-11 p. 137)

### Utilisation des épreuves à expression libre

Toutes les formes de la pensée intervenant dans un domaine d'étude peuvent être valablement testées par les méthodes étudiées dans le sous-chapitre précédent. Néanmoins, les Français sont en général partisans d'épreuves à réponses rédigées, se fondant sur cette remarque que les épreuves standardisées du type "test" se concilient mal avec l'expression de la pensée" (F. Hotyat 1-42 p. 5).

L'usage généralisé des épreuves standardisées dans les pays de langue anglaise a pu faire craindre que les enseignants en viennent à négliger la pratique de la rédaction étant donné l'interaction de la pédagogie et des examens.

Un article récent du monde (26 II 1977 p. 11) fait état de l'opinion d'universitaires américains qui enregistrent un déclin du niveau de rédaction chez les étudiants et de la création d'enseignements nouveaux dans plusieurs universités visant à améliorer la rédaction et la composition en langue anglaise ; il serait sans doute osé de mettre en cause l'utilisation généralisée des tests de connaissances, mais il demeure que l'évaluation de ce genre d'épreuve où l'expression est libre, se doit d'être faite à l'aide de critères objectifs communicables.

Par ailleurs, de nombreuses productions scolaires sont à évaluer en tant que telles, globalement. "Certains produits, ou certaines conduites résultant de l'apprentissage systématique sont apparemment inaccessibles au processus d'analyse que requiert l'élaboration des tests" (A. Bonboir 1-11 p. 143).

Il s'agit là non seulement des rédactions en langue maternelle, mais également de l'écriture, du dessin, de l'expression orale en langue étrangère, des travaux manuels, qu'ils soient artistiques ou artisanaux.

A) Nous examinerons d'abord ce que Anne Bondoir appelle les échelles de production scolaire.

Ce sont des spécimens de production classés en une série graduée qui va, en plusieurs échelons, d'une réussite de qualité médiocre jusqu'au résultat de la meilleure qualité.

Il faut donc obtenir une hiérarchie de valeur entre les documents soumis à l'examen des juges, c'est-à-dire transformer en grandeurs numériques des valeurs subjectives et établir une série de degrés dans des phénomènes dont l'intensité ne se laisse que difficilement apprécier. L'établissement d'une telle échelle suppose l'utilisation de statistiques portant sur les classements effectués par plusieurs juges, ou sur la comparaison deux à deux des différentes productions par les mêmes juges.

On remarquera que dans cette situation l'appréciation est globale et que c'est l'avis subjectif de la majorité qui sert de référence. Ceci se justifie en l'absence de tout autre moyen analytique permettant de fonder un jugement.

B) A la frontière entre les questionnaires de connaissances dits "à choix multiples" et les textes du type essai ou rédaction se situent des questionnaires qui demandent une réponse rédigée, suffisamment déterminée par la nature de la question, et d'un développement relativement court, limité.

Cette méthode permet d'effectuer rapidement une double correction et permet également l'étude statistique des divergences qui peuvent apparaître entre les notateurs (et qui résulte quelquefois d'une rédaction ambiguë de la question).

Cette méthode, entre autres, est couramment utilisée à la Faculté de Médecine du Kremlin Bicêtre et à celle de Tours.

Le relevé des notes peut facilement être fait sur fiche mécanographique et les résultats statistiques sont alors fournis par l'ordinateur.

Le Docteur Mergot de l'Université de Montréal (in 1-42 p. 75-109) donne un exemple précis d'utilisation de la méthode du questionnaire à réponse ouverte courte (Q.R.O.C.) qu'il a notablement enrichie de telle sorte que l'épreuve ne privilégie pas le contrôle des connaissances, mais au contraire, oblige le sujet à créer quelque chose de neuf, à manifester que "les études accomplies au cours de l'année ont augmenté ses "pouvoirs" et activé sa créativité" (1-42 p. 76).

A cet effet, chaque question est conçue de telle sorte que "chacune d'elle exige la solution d'un petit problème (pas forcément mathématique) indépendant des autres questions, dont la solution peut être construite en partant d'une bonne connaissance du cours, mais cette solution ne doit pas être présente dans les cours" (ibid p. 77).

L'intérêt de cette méthode réside également dans la facilité d'appréciation des résultats et dans l'étude ultérieure de la répartition des notes aux différentes questions, qui permet d'éliminer des questions trop faciles ou trop difficiles dont le pouvoir de discrimination est faible.

C) Le véritable problème posé cependant au début de ce sous-chapitre est l'appréciation d'une épreuve "relativement libre (...) qui fournit une certaine information au sujet de la structure, de la dynamique et du fonctionnement des expériences mentales du sujet" (Encyclopedia of Educational Research cité in 1-11 p. 147).

Dans cette situation, on ne dispose pas de normes, ni de réponses types permettant la comparaison ; par ailleurs, la diversité des textes s'oppose à ce qu'on puisse étudier la totalité des réponses d'une population donnée "tous les élèves des classes de seconde" par exemple).

Les nombreuses tentatives faites pour améliorer la notation des épreuves à expression libre ont mis en évidence quatre sources de variation à prendre en considération (2-97).

- la personnalité des correcteurs (tendance à varier dans le temps, à donner des notes trop sévères, trop indulgentes, trop localisées dans l'échelle des notes).

- la méthode de correction (impression générale, ou méthode analytique sur critères définis).

- les sujets de devoirs proposés aux élèves (degré de difficulté ou de familiarité).

- l'élève lui-même dont la "forme" est variable suivant les moments, les conditions de l'examen, son état psychologique ou physiologique.

Les tentatives faites pour améliorer la notation des épreuves à expression libre s'inscrivent dans deux grandes tendances. L'une consiste à se rapprocher de la méthode précédente, en restreignant la liberté de l'élève et celle du correcteur, afin de diminuer les causes de désaccord.

A l'opposé, l'autre délibérément joue la carte de la liberté totale d'expression et de correction ; on s'efforce alors de stabiliser l'appréciation par la multiplication du nombre des correcteurs et aussi en accroissant le nombre de sujets traités par l'élève.

Dans la première situation, il ne s'agit nullement de normaliser la situation et les critères de jugement comme dans un test, mais de diminuer l'importance des désaccords entre notateurs.

S'il est vrai que l'imposition d'un sujet de devoir commun à tous les élèves facilite la comparaison, il faut noter que chaque sujet favorise certains élèves et en défavorise d'autres ; cet inconvénient peut-être écarté si l'on demande à tous de traiter plusieurs sujets

de façon relativement brève ; si l'on corrige d'abord les travaux des élèves portant sur un même sujet, et si l'on recommence avec l'ensemble des copies pour le sujet suivant, on élimine l'effet de halo dans les appréciations portant sur un même élève. De même une première lecture, rapide de l'ensemble des productions permet d'ajuster son jugement à tout le groupe.

Avec des élèves avancés dans leur scolarité cette technique permet de plus en plus difficilement d'apprécier les qualités d'exposition, d'organisation de la pensée, voire d'argumentation dans une question complexe.

Il conviendra alors de préciser lors de la correction les aspects à prendre en considération, en utilisant un barème, et éventuellement une adaptation statistique de l'échelle des notes, plus ou moins sévères, plus ou moins dispersées de chaque notateur.

Toute correction analytique suppose que l'on distingue entre les qualités de forme (orthographe, correction grammaticale, élégance de l'expression) et celles qui sont liées au contenu (qualité de l'organisation des idées, ainsi que leur richesse, leur originalité, leur pertinence).

Par une analyse factorielle portant sur une enquête réalisée auprès de professeurs du domaine littéraire et sur la correction de plusieurs centaines de copies, Remondino (2-115 p. 27) a mis en évidence les dimensions suivantes qui interviennent dans la notation.

- Présentation graphique (lisibilité, esthétique, présentation)
- Usage de la langue (exactitude de l'orthographe, de la morphologie, de la syntaxe, propriété du langage, style)
- Contenu et organisation de l'exposé (structure, richesse, pertinence des idées, précision de l'information, degré d'exhaustivité)
- Aspects personnels du fond (originalité, maturité, imagination).

On remarquera qu'un petit nombre seulement de dimensions sont notées indépendamment les unes des autres et que de ce fait des distinctions plus fines sont assez illusoire.

Une autre recherche expérimentale de FRENCH citée par F. BACHER (1-6 p. 62) montre que l'accord entre les correcteurs sur l'importance relative des différents aspects dégagés par l'analyse est loin d'être réalisé, même si un facteur général assez important traduit l'existence de valeurs communes.

Cette recherche met en évidence l'importance accordée aux points suivants :

- les idées ;
- la forme et l'organisation ;
- le caractère personnel (flavor) ;
- la correction grammaticale (méchanics) ;
- le choix des mots, l'expression (wording).

A l'opposé, en Grande Bretagne on a tenté de donner la plus grande liberté possible à l'élève et au correcteur, en prenant en considération de nombreux travaux de rédaction et en combinant plusieurs appréciations portées sur chacun d'eux.

Si laisser le choix aux élèves entre plusieurs sujets rend plus difficile la comparaison entre élèves différents, elle donne par contre à chacun la possibilité de montrer ses capacités d'organisation de la pensée et d'expression de celle-ci sur un sujet qui lui plaît. On pourra d'ailleurs augmenter la fidélité de la méthode en faisant faire plusieurs compositions à chaque sujet portant chacune sur un genre différent (narrations, exposés d'une question, analyse critique, commentaire composé, résumé de texte, etc...).

La correction du même devoir par plusieurs notateurs agissant indépendamment l'un de l'autre améliore grandement la fidélité des évaluations même si la note stabilisée, on l'a vu, n'est atteinte que par un grand nombre de corrections (Pieron).

De plus, nous avons vu que chaque correcteur n'accorde pas la même importance aux dimensions différentes qui interviennent dans son jugement (voir ci-dessus) ; faire appel à plusieurs correcteurs revient donc à pondérer (d'une manière non explicite d'ailleurs) ces différentes dimensions ; on peut alors laisser aux correcteurs une entière liberté de jugement.

Dans une revue critique des recherches anglaises sur la notation des devoirs de rédaction d'élèves de 10-12 ans, A. Nguyen Xuan (2-97 p; 254-266) montre que les facteurs principaux qui jouent dans la variation des notes de composition de ces élèves sont l'expérience du correcteur et la méthode de correction.

Sans insister sur l'intérêt que présente une préparation méthodique à la correction des devoirs incluse dans la formation des enseignants, cet auteur remarque que la méthode analytique se révèle plus logique, mais lourde et que les critères sont parfois assez arbitraires faute d'une étude expérimentale suffisante.

Par contre la méthode "d'impression générale contrôlée" plus rapide, moins astreignante qui utilise la somme des notes données par quatre correcteurs a chaque devoir permet d'atteindre une bonne validité (méthode de Devon) ; Wiseman (3-29) remarque par ailleurs que l'expérience des correcteurs s'acquiert assez rapidement si chacun prend soin de lire préalablement à la correction une quantité assez importante de ses copies en tenant compte des critères habituels pour un devoir de rédaction.

Nous concluerons sur ce point en citant la remarque de F. Bacher (1-6 p. 75) "les épreuves à expression libre utilisées seules ne permettent pas une aussi bonne prédiction de la réussite scolaire ultérieure que les épreuves normalisées", mais leur adjonction à d'autres épreuves plus standardisées accroît la valeur de la prédiction car la composition dans la langue maternelle atteint des dimensions différentes de la personnalité de l'élève que les épreuves normalisées ne touchent pas.

Il convient cependant de garder en mémoire que les méthodes plus ou moins élaborées qui viennent d'être évoquées, si elles contribuent notablement à améliorer la fidélité et la validité des épreuves de composition, se heurtent à des difficultés de coût (en nombre de correcteurs et en temps) qui les rendent souvent difficilement utilisables.

### L'utilisation des notes et appréciations scolaires

Un grand nombre d'appréciations sont normalement portées par un professeur sur ses élèves à l'occasion des différents exercices que comporte son enseignement. Il semble donc à priori satisfaisant pour apprécier le niveau des élèves d'utiliser cet ensemble de renseignements plutôt que de fonder une opinion sur une épreuve d'examen unique nécessairement affectée par des erreurs (anxiété - base de jugement étroite...)

La prise en considération d'un dossier scolaire par un jury traduit cette préoccupation.

Mais "la difficulté vient de ce que les notes ou les appréciations données par des professeurs différents ne sont pas directement comparables" Bacher (1-6 p. 68).

Nous ne reprendrons pas le détail de ces sources d'erreurs étudiées dans un précédent chapitre ; rappelons simplement :

- les situations diverses dans lesquelles sont portées une note ou une appréciation ;
- l'effet de halo qui résulte de la personne de l'élève ;
- la différence qui existe dans les critères de référence qu'utilise implicitement tout enseignant ;
- l'absence de concertation sur le niveau exigible des élèves ;
- les interactions élèves-professeurs ;
- l'objectif pédagogique qui est différent du contrôle de niveau.

Cependant l'obstacle majeur "tient au fait qu'un professeur peut bien comparer entre eux des élèves qu'il connaît ; il ne peut les situer par rapport à des élèves qu'il ne connaît pas" F. Bacher (1-6 p. 69).

Or, le niveau moyen des élèves varie considérablement d'une classe à l'autre. Dans des épreuves de connaissances la différence entre classes extrêmes a été de 37 points sur 80 en français, de 31 points sur 44 en mathématiques dans l'enquête Nationale de Reuchlin Bacher (3-11 P.19).

Aux U.S.A. Flanagan (2-50) constate à l'issue d'une enquête étendue que nombre d'élèves classés dans le quart inférieur pourraient se retrouver dans le quart supérieur s'ils changeaient d'école.

Pour utiliser ces notes et appréciations il faut faire la part de ce qui tient à l'élève et ce qui tient tant au professeur qu'un niveau de l'établissement fréquenté.

On appelle "modération" l'ensemble des méthodes qui ont pour objet de tempérer la notation en tenant compte des différences de niveau entre classe et les différences de sévérité entre professeurs.

"L'idée générale est de parvenir à connaître le niveau relatif des classes sur une échelle commune" (Bacher 1-6 p. 70).

C'est seulement lorsque cette échelle commune sera établie que l'on pourra "calibrer" les notes attribuées par les différents professeurs à leurs élèves.

Les recherches les plus avancées dans ce domaine sont utilisées en Suède, Suisse, Grande-Bretagne. On en trouvera un exposé détaillé dans le livre de Landsheere (1-45 p. 179-210).

Nous n'exposerons donc ici que les grandes lignes des différentes méthodes :

A) "Pour situer les classes par rapport les unes aux autres, le plus simple est de connaître les résultats de l'ensemble des élèves à une épreuve commune" (Bacher, ibidene).

On peut utiliser à cet effet soit un test d'intelligence, soit une épreuve normalisée de connaissances qui présentent l'avantage d'une notation tout à fait objective.

Il est également possible d'utiliser une épreuve scolaire, à condition que la correction soit assurée par une équipe de notateurs extérieurs aux établissements et que toutes précautions soient prises pour une objectivité suffisante de la notation.

Connaissant le résultat de ses élèves sur l'échelle commune des notes, chaque professeur utilisera la même zone de l'échelle pour ses appréciations ultérieures selon différentes procédures dont le détail figure dans la littérature spécialisée.

Les notes de l'épreuve commune servent seulement à situer la marge des notes appropriée que chaque professeur utilisera dans sa classe ensuite selon ses propres critères de jugement.

Les notes de l'épreuve commune ne serviront pas à juger individuellement les élèves.

On peut objecter à cette méthode qu'en toute logique il serait nécessaire d'utiliser une épreuve commune qui recouvre les mêmes dimensions d'appréciation que les notes du professeur. Le calibrage réalisé n'est que partiellement justifié, mais de nature en tout état de cause à réduire l'erreur due aux différences de niveau entre classe.

On a fait également remarquer que se contenter d'une seule épreuve commune est insuffisant, mais en utiliser plusieurs rend le processus plus lourd et contraignant.

B) D'autres méthodes n'utilisent pas d'épreuves communes : en Grande-Bretagne au cours de certains examens internes aux établissements, des épreuves communes à plusieurs établissements d'une région sont données à l'ensemble des élèves.

Ces épreuves sont corrigées par leurs professeurs, mais un échantillon des devoirs produits par différentes écoles est constitué et corrigé à nouveau par une équipe de correcteurs extérieurs spécialement préparés ; c'est à partir de l'ensemble de ces épreuves que l'on peut établir l'échelle des niveaux des différentes classes ou différentes écoles.

C) Avantages et limites des méthodes de modération.

L'intérêt principal des méthodes qui viennent d'être évoquées est de permettre d'utiliser les notes et appréciations des professeurs tout au long de l'année en accroissant de façon importante la validité des épreuves scolaires.

Elles perturbent peu l'enseignement et renforcent notablement le poids des jugements "calibrés" des professeurs en particulier dans les situations d'orientation.

Mais les méthodes de modération ne corrigent que les différences de niveau, et éventuellement la différence de dispersion des échelles de note de chaque professeur, elles ne peuvent atténuer les désaccords existants quant aux objectifs de l'enseignement.

Egalement on peut regretter que cette méthode n'ait aucun effet sur l'interaction élève-professeur et les variations qu'elle peut produire dans le degré de sévérité ou d'indulgence de ce dernier. Le remède est ici à rechercher, en particulier, dans une formation adaptée des enseignants.

## PERSPECTIVES

Au terme de ce chapitre ne paraîtra-t-il pas hors de propos de se poser la question : est-il bien nécessaire d'évaluer les résultats scolaires ?

L'expérience de la Suède rapporté par Lehman (2-80 page 34) incite au moins à la réflexion.

"Pas de sélection (...) pas de redoublement, pas de classes de niveau ni d'évaluation individuelle jusqu'à la 8ème année incluse" (15 ans).

Conséquence : "les enfants ne vivent pas leur scolarité dans un chantage permanent à l'échec".

Comment ne pas être séduit par cette éducation dans l'épanouissement ?

Et pourtant nous avons vu antérieurement et J. Beauté le rappelle (1-8 p. 14) ; il y a dans l'évaluation des résultats d'un élève une possibilité d'action pédagogique qu'on ne saurait négliger indépendamment de tout examen. D'ailleurs, et Landsheere s'en félicite (1-45 p. 237), il restera toujours des traits humains dont la mesure rigoureuse restera hors du possible.

Il n'en demeure pas moins que l'exigence d'évaluations comparées correspond à un caractère fondamental de notre civilisation. Si notre culture continue à s'intellectualiser, à croire en complexité, comment pourrait-elle y renoncer ?

Glaeser remarque justement (1-35 p. 6) "Dans une société évoluée, il est des responsabilités techniques qu'on ne peut confier à n'importe qui sans se préoccuper de sa compétence dûment contrôlée".

Cette nécessité sociale apparaît clairement à la lecture de Reuchlin : "l'enseignement en l'an 2000" (PUF 1973).

Une réflexion sur les objectifs de l'éducation conduit directement à se donner des moyens de vérifier si les objectifs définis en termes de savoirs et de pouvoirs sont atteints ou non, et s'ils ne le sont pas de mesurer l'écart afin de se donner les moyens de le réduire. Les "compteurs" qui permettront cela devraient être construits en partant des principes développés ci-dessus dans ce chapitre. Ce type d'évaluation ne peut d'ailleurs se fonder sur de simples comparaisons entre élèves mais sur des épreuves correspondant à des exigences posées a priori. BACHER (2-11 p. 48) et BONORA (5-3 p. 155) précisent que : "une solution réelle du problème ne peut être trouvée en dehors (...) de la détermination pour chaque enseignement d'une plate forme d'objectifs explicites sur laquelle l'ensemble des professeurs soient d'accord".

Pour cela "une explicitation est une publication des objectifs à chaque niveau permettraient sans doute des ajustements de proche en proche et réduiraient considérablement les désaccords".

Les pays anglo-saxons ont été amenés avant tout programme d'évaluation à établir des listes d'objectifs éducatifs.

Même si pour les planificateurs et économistes de l'éducation "les objectifs à atteindre s'expriment uniquement dans la plupart des travaux en fonction de la durée d'éducation offerte à un effectif déterminé d'élèves. La nature des activités éducatives occupant cette durée et les résultats éventuels de ces activités ne sont pas pris en considération.

On peut douter que cette façon de définir les objectifs de l'éducation conserve son intérêt dans l'avenir" (Reuchlin ouvrage cité p. 98).

Condition nécessaire de tout effort d'évaluation sérieux la définition d'objectifs devrait s'accompagner d'une formation aux tâches d'examineur. "On ne naît pas examinateur, mais on le devient" paraphrase Mialaret (1-42 p. 165), de même "un bon professeur n'est pas nécessairement un bon examinateur" ce qui revient à dire que l'action pédagogique est autre que l'évaluation objective du résultat d'une formation.

Une pédagogie par objectifs qui doit conduire chaque élève vers la maîtrise d'un programme devra tenir compte des rythmes d'acquisition qui ne sont pas identiques et conduire à une certaine individualisation de l'enseignement rendant ainsi plus indispensables des moyens de contrôle permettant une analyse plus fine.

Les méthodes d'enseignement programmé (assisté par ordinateurs dans les pays industriellement avancés) intègrent dans le processus d'apprentissage celui de contrôle à chaque étape.

On trouvera dans BACHER (1-6 p. 76-79) quelques exemples qui montrent entre autre que dans ces expériences "les contacts avec les enseignants constituent une part essentielle de la méthode ; déchargés de la transmission d'information ils peuvent centrer leurs efforts sur la discussion, la mise en perspective, la synthèse, le conseil personnel".

S'il fallait conclure au moins provisoirement, nous pourrions dire que quelles que soient les réformes pédagogiques ou institutionnelles, les exigences d'une civilisation plus complexe, l'abaissement des frontières ou simplement l'action pédagogique rendront toujours nécessaire l'évaluation et le contrôle du profit qu'un élève a tiré de l'enseignement qui lui est donné ; la docimologie devra donc pouvoir répondre d'une façon très simple aux nouveaux besoins qui apparaissent et qui exigent la conception de moyens adaptés.

Loin d'être une technique presque achevée, ses moyens sont encore imparfaits et de grands progrès seront encore nécessaires.

ORIENTATION



ORIENTATION

"Le problème de l'orientation scolaire et professionnelle est le problème central de l'enseignement. C'est de sa solution que dépendent l'épanouissement et le bien de l'individu, puisqu'orienter, c'est s'efforcer de savoir de quelle manière on développera au maximum les forces latentes de chaque personnalité en formation, chercher dans quel sens chaque être humain réalisera sa plénitude. C'est elle aussi qui assurera la bonne mise en place de l'individu dans la société, et par conséquent la bonne marche de la machine sociale".

(R. GAL - l'Orientation - Nouvelle Encyclopédie Pédagogique - Page 2)

HISTORIQUE

"On s'est passé d'orientation tant que le monde et la vie ont connu une certaine stabilité où le destin de l'homme était déterminé si l'on peut dire, du dehors, par sa naissance, son origine sociale, familiale. Il naissait fils d'artisan, de paysan, de commerçant, de noble et de ce simple fait son destin était tout tracé : il n'avait plus qu'à suivre la volonté sociale comme une sorte de prédestination incontestable. En vertu de ce sort que la naissance lui avait réservé et que la société ne lui contestait pas, il était élevé, instruit, éduqué. Le seul problème qui se posait à lui était un problème d'adaptation de l'individu, adaptation forcée à cette destinée imposée du dehors". (id. p. 9)

Le développement des techniques, l'augmentation des connaissances, l'apparition de nouveaux métiers, nécessitent une nouvelle formation.

"L'appareil scolaire et universitaire qui organise cette formation ne peut plus avoir pour seul objectif la diffusion d'une culture désintéressée, comme ce pouvait être le cas à une époque où les études prolongées n'étaient permises qu'à une minorité privilégiée sans souci professionnel. L'immense majorité des élèves et des étudiants viennent aujourd'hui chercher à l'école et à l'université les moyens d'accès à une activité professionnelle aussi hautement qualifiée que possible. Si l'université se refusait à établir des évaluations individuelles relatives à cette formation, les utilisateurs eux-mêmes ne manqueraient pas de le faire.

La science et la technique modernes mettent à la disposition des individus une puissance considérable... L'éducateur peut-il se décharger du rôle d'expert que la société lui demande d'assumer ? Qui pourrait jouer ce rôle de façon mieux informée et plus objective ?" (1-67 p. 210).

Il se crée alors un certain nombre d'examens de sélection, ou d'orientation.

"A côté des examens tendant, en somme, à établir des classements hiérarchiques, il en est qui visent à définir, plus ou moins complètement un état individuel, une structure de personnalité, pour en tirer des conséquences éventuelles. Au cours de ceux-ci, on peut être appelé à déterminer un certain niveau de connaissances ou de capacité intellectuelles. C'est le cas de l'examen d'orientation, où le risque d'erreur est grave". ('1-63 p. 55)

Jadis, "les professeurs jugeaient souverainement du passage en classe supérieure, ceux qui étaient éliminés n'avaient qu'à s'en prendre à eux-mêmes. N'étaient-ils pas venus volontairement dans les lycées pour obtenir le bac ?... Il s'agissait pour l'élève de s'adapter à l'enseignement qu'il avait souhaité". (3-22 p. )

Actuellement, tandis qu'on supprime des examens, (entrée en sixième, B.E.P.C. ...), on développe à leur place l'orientation.

A l'origine, elle apparaît à deux sorties du système éducatif : par l'orientation professionnelle, mise en place de 1922 à 1963, qui guide les élèves vers un métier, à la fin du primaire : et par le BUS qui, à partir de 1932 les guide après le bac.

En 1963, les C.E.S. apparaissent, avec dans un seul établissement la juxtaposition de trois filières jusqu'alors séparées, et l'orientation entre à l'école : on doit orienter les élèves à travers ces trois filières.

"Auparavant, la sélection principale s'opérait en dehors de l'école : les jeunes s'engageaient dans l'enseignement complémentaire, secondaire, l'apprentissage en technique pour des raisons familiales, de voisinage, de relation, de hasard. Une fois entrés, les élèves pouvaient accepter la règle du jeu de leur institution scolaire, tout comme dans les grandes écoles, une fois admis, chacun est assuré d'en sortir avec au moins le titre d'ancien élève. Désormais, c'est à l'intérieur de l'école que joue la compétition pour la situation sociale de sortie, puisque celle-ci est conditionnée par la filière d'enseignement où l'élève est inscrit dès la sixième. Bien plus, la compétition est devenue la structure principale de la vie scolaire, élèves et parents jugent l'enseignement donné en fonction des possibilités qu'il ouvrira pour accéder au niveau le plus noble". (rapport de la commission JOXE)

### L'ORIENTATION ACTUELLE

Tout au long de la scolarité, le profil de l'élève est consigné dans un dossier consulté à chaque palier d'orientation : entrée en 6°, fin de 5°, fin de 3°...

A chacun de ces paliers, le conseiller d'orientation peut soumettre les élèves à des tests d'aptitude, dont les résultats sont communiqués aux familles et au conseil de classe. Celui-ci fait la synthèse des vœux des familles, des motivations des élèves, de l'avis du conseiller d'orientation et des résultats scolaires, et décide de l'orientation.

Telle qu'elle est actuellement pratiquée, l'orientation supprime la rigueur de l'examen et évite le bachotage. Elle est répartie sur toute la durée des études, et permet ainsi d'apprécier l'évolution et les progrès de l'élève.

"Or, c'est cette capacité de progrès, de perfectionnement qu'il faut saisir, beaucoup plus que les connaissances acquises ou non à un certain moment de la scolarité qui jugent non pas seulement l'enfant mais aussi l'effet du milieu familial et social dans lequel il a vécu ses premières années ou encore le résultat de son éducation antérieure". (R. GAL - L'orientation scolaire et la recherche des aptitudes - Recherche Pédagogique p. 10)

Mais cette forme d'orientation repose pour une grande part sur les notes et appréciations mises par les professeurs au cours de la scolarité. Or on a vu combien notes et appréciations sont sujettes à caution.

"L'appréciation des résultats scolaires, telle qu'elle est pratiquée, est insuffisante pour le but qu'elle veut se fixer : l'orientation. Car, en fait, elle répond mal au contrôle des connaissances, et ne tient pas compte de facteurs importants, le caractère, le fait social, l'aspect santé. Dans l'être jeune, rien n'est encore fixé, c'est souvent la contradiction : c'est le problème de la maturation". (compte rendu de l'inspecteur général SIRE à l'IPN le 13.2.69)

"On retrouve deux autres dangers : celui de ficher, d'étiqueter à tout jamais les élèves dans des catégories bien connues des parents (élève faible en mathématique, élève dysorthographique, élève moyen, élève qui peut mieux faire"... ) et celui de laisser croire que la note est le reflet des véritables aptitudes des individus". (2-133 p. 46)

"L'école ne met pas en relief des valeurs ou des préférences positives, mais procède essentiellement par la constatation négative d'insuffisances. Elle n'oriente pas par la connaissance du profil mental et affectif d'ensemble ; elle élimine d'après le niveau "moyen" défini par une échelle de notation illusoire. Le résultat est qu'une telle constatation est sans signification pour les choix nécessaires et n'impose que des abandons. Chaque bifurcation en cours d'études, au lieu d'apparaître comme un aiguillage horizontal entre des voies différentes, mais pas nécessairement inégales, est ressentie comme un échec, une décantation, une baisse verticale de niveau". (L. CROS l'Education Nationale n° 863 - 864 - 865 p. 9)

Est-il simple pour l'élève et sa famille de choisir une orientation ?

"G3, T2, F10 ... Pour les non initiés que sont la quasi-totalité des parents et des élèves, l'orientation a l'allure d'une bataille navale - ou d'un catalogue de self-service particulièrement bien fourni - Etes-vous un goinfre prêt à absorber des mathématiques à haute dose, plus du latin et du grec ? nous avons cela. Votre estomac est-il si menu qu'il ne peut absorber qu'une seule langue légèrement "renforcée" ? C'est possible. Avez-vous une vocation de "tulliste en voilettes et dentelles dites de Lyon" ? Il y a précisément un CAP qui y prépare. On trouve tout dans la boutique de l'orienteur...

Par le jeu des options et des sections, il n'existe pas moins de cinq cent quatre vingt possibilités d'orientation dans l'enseignement secondaire près de trente baccalauréats de l'enseignement général et technique, quelque trois cents CAP, sans compter les brevets de techniciens, les diplômes régionaux, les sections en voie de disparition et d'autres qui ne mènent à aucun diplôme. On comprend que l'orientation soit avant tout un épouvantable casse-tête pour dénicher "la" section correspondant précisément au cas - toujours unique - que pose son enfant, pour éviter les chausse-trapes des sections impasses ou dépotoirs, et pour décrocher le gros lot du bac le plus prestigieux". (C. ARDITTI - Le Monde de l'Education - février 1976 - p. 6)

"L'orientation pose aux professeurs, tout au long du premier cycle et en particulier en fin de troisième, de difficiles problèmes. Doivent-ils, comme certains le pensent, devant ces difficultés, abandonner au conseiller d'orientation scolaire et professionnelle cette tâche ? Nous pensons, au contraire, que l'orientation doit être un travail d'équipe". (1-16 p. 17)

Or, si les professeurs n'ont jamais appris à noter ils ont encore moins été formés pour orienter. Beaucoup même n'ont pas conscience de préparer l'orientation : parmi les professeurs qui ont répondu à notre sondage, 40 % estiment qu'ils ne mettent "JAMAIS" de notes pour préparer une orientation ;

16 % pensent le faire rarement ;

alors que 11 % le font "souvent" et 11 % "toujours".

L'enseignant ressent donc un certain désarroi.

Conscient ou non du fait qu'il est juge et parti, tout à la fois "alphabétiseur et sélectionneur, l'ambiguïté fondamentale d'un système compétitif, l'amène à sélectionner sur ce qu'il doit transmettre". (2-86 LURCAT L. l'échec et le désintérêt scolaire - orientations Fr (1975) 15, n° 53,51-71)

"La réforme de l'enseignement, pour imparfaite qu'elle soit, a eu le mérite de mettre en pleine lumière les problèmes de l'orientation jusqu'ici résolus par le déterminisme sociologique. Mais en affirmant la nécessité de substituer une orientation positive et consciente aux processus mécaniques d'élimination par les notes et les examens, le législateur n'a point facilité la tâche des maîtres confrontés à des problèmes qu'ils ne sont pas armés pour résoudre. L'orientation scolaire exige en effet des professeurs, devenus orienteurs, deux compétences nouvelles : d'une part, une connaissance précise de ce qui n'est pas leur univers et leur spécialité, en particuliers celle des structures scolaires extérieures à leurs sections et à leur établissement, et celle des réalités économiques qui les enveloppent et les justifient ; d'autre part, une attention, au sein même de son enseignement, à ce qui déborde de sa discipline, à ce qui peut dans les réussites et dans les échecs de son élève annoncer quelque aptitude positive à autre chose qu'à l'exercice de la discipline qu'il enseigne". (R. GAL- L'orientation scolaire et la recherche des aptitudes - recherche pédagogique 1969 p. 5)

Enfin, l'observation du professeur est limitée à un petit nombre d'élèves, il ne peut pas comparer avec beaucoup d'autres classes. Faute de définitions claires des objectifs à atteindre à chaque niveau, et des connaissances et aptitudes nécessaires à un élève pour accéder à telle ou telle section, chacun oriente selon ses propres critères.

D'après une enquête d'opinion auprès de professeurs de mathématiques enseignant dans les classes de 3ème ou de 2ème, sur 260 professeurs consultés, à la question : "effectuer et simplifier"

$$(x^2 - 4y^2) \left( \frac{x+2y}{x-2y} - \frac{x-2y}{x+2y} \right)$$

24 pensent que TOUS les élèves qui entrent en 2ème devraient donner la réponse correcte

pour 31 professeurs, 90 % des élèves doivent bien répondre

pour 71 , 80 %

pour 43 , 70 %

pour 40 , 60 %

pour 43 , 50 %

pour 29 , 40 %

pour 15 , 30 %

pour 6 , 20 %

pour 1 , 10 %

En fait, 28 % DES ELEVES DE 3ème donnent une réponse correcte.

Si on compare les exigences médianes des professeurs concernant l'entrée en 2ème aux connaissances effectivement acquises par les élèves en fin de 3ème, on est évidemment frappé par le décalage entre les unes et les autres. Mais on peut se demander si la grande dispersion des réponses des professeurs permet de dire qu'il existe une norme.

Il est possible que des professeurs différents, placés dans les mêmes conditions conçoivent de façons diverses ce que doit être le niveau d'une classe de 2ème. Il faut admettre alors qu'il n'y a pas, dans le corps professoral une doctrine commune quand au rôle que cette classe joue.

Il est possible aussi que l'entrée en 2ème ne soit pas considérée par les professeurs en dehors des conditions propres à la région ou à l'établissement qu'ils connaissent et que la grande diversité de ces conditions ne permette pas d'établir une norme à l'échelon national. (3-12 L'orientation à la fin du 1er cycle REUHLIN-BACHER Puf 1969).

Le conseiller d'orientation, étant "extérieur" au corps professoral est bien accueilli par les élèves ; sa fonction est indispensable parce qu'il est pour eux, la distance prise à l'égard du système scolaire... Il pourra souligner des corrélations, des harmonies, des désaccords entre pédagogie et psychologie. Néanmoins la fonction de la psychologie scolaire est d'informer, de donner des points de repère, mais elle n'est pas d'établir des pronostics infaillibles... Les critiques contre l'utilisation des tests en psychologie scolaire (Monde de l'éducation 3-76 BRESSON ; le quotient intellectuel) peuvent être irritantes par leur aspect caricatural, mais elles ne sont pas sans fondement". (PARTOES M. L'avis d'un conseiller d'orientation Pédagogie p. 41-46)

Un tel aveu est grave. Peut-on encore faire confiance à la psychologie scolaire pour résoudre les cas difficiles d'inadaptation scolaire ? Le principal rôle de l'orienteur scolaire sera un rôle d'informateur car il possède une documentation importante sur les différentes filières et leurs débouchés.

Cependant ses informations sont livresques, fournies par l'ONISEP pour l'essentiel. Il existe en outre un décalage entre l'information qui guide à l'entrée dans une filière et les débouchés réels à la sortie.

"Les filières de l'enseignement secondaire et supérieur continuent de reposer sur la préparation à des emplois en perte de vitesse, alors que les branches en extension (dans le secteur tertiaire en particulier) ne correspondent pas à des formations existantes.

L'enseignement technique s'efforce de donner aux jeunes des formations leur permettant d'acquérir un diplôme et un métier, mais on constate que des filières demeurent longtemps après que leurs débouchés directs aient disparu, et que les responsables de l'économie sont incapables de fournir des indications précises et fiables sur l'évolution à moyen terme du marché de l'emploi et des qualifications". (C. ARDITTI - Le Monde de l'éducation - février 1976 p. 9)

De plus, chaque conseiller d'orientation se voit confier un trop grand nombre d'élèves : 1300 à 1500 en moyenne chaque année. Il lui est impossible de suivre individuellement chacun d'eux, et finalement il n'est plus qu'un "technicien du tri et du renseignement". (La plupart des conseillers estiment que, pour remplir efficacement leur rôle, il faudrait que le nombre d'élèves dont ils ont à s'occuper ne dépasse pas 600).

De nouvelles contraintes apparaissent pour l'orientation, avec la carte scolaire : les élèves sont dirigés en priorité vers des établissements de leur zone géographique. Et ces contraintes n'agissent pas de la même façon pour tous, suivant le milieu social.

"La hiérarchie entre les sections se double bien souvent d'une autre hiérarchie - non officielle, mais réelle - entre les établissements. En effet, l'égalité théoriquement assurée par la carte scolaire est encore loin d'être réalisée. Certains établissements sont connus pour ouvrir plus aisément les portes des meilleures classes préparatoires aux grandes écoles. Les sections B n'ont pas le même prestige selon qu'elles sont dans un lycée classique ou un lycée technique. La différence entre les établissements est surtout frappante dans le premier cycle.

Les élèves de premier cycle de lycée ont de fortes chances d'entrer en seconde dans le même établissement. Ils ont l'avantage d'être "sur place", les professeurs sont souvent certifiés ou agrégés ; certains donnent des cours dans les deux cycles de l'établissement et sont ainsi plus à même de faciliter le passage de l'un à l'autre. "Autrefois, lorsque j'enseignais uniquement dans un CES, je me conformais strictement au programme, dit un professeur. Aujourd'hui, j'enseigne aussi dans le second cycle. Je tiens davantage compte de ce que mes élèves auront à apprendre en seconde et en première. Ils arrivent ainsi mieux préparés au lycée".

Inversement, les élèves ayant accompli leurs études secondaires dans un CES auront plus de difficultés à entrer en seconde au lycée. Encore la situation est-elle différente selon qu'il s'agit d'un CES nouvellement créé, où les enseignants peuvent être en grande partie certifiés, ou d'un CEG, où les instituteurs et les professeurs de collège sont en majorité. Les parents avisés savent qu'il faut tenter par tous les moyens (interventions personnelles, choix d'une première langue rare) de ne pas mettre leur enfant au CES de leur secteur, lorsque celui-ci est moins bien coté... Or, la répartition entre les différents types d'établissements de premier cycle est très inégale : les beaux quartiers parisiens n'ont pratiquement que des lycées, alors que les zones rurales ne connaissent que des CEG... Dans la "couronne ouvrière" de la périphérie nord de Paris (Nanterre, Colombes, Gennevilliers, Villeneuve, Clichy, Levallois, Puteaux) il y a, pour quatre cent mille habitants, mille six cents places dans deux lycées d'enseignement général, mille six cent cinquante dans trois lycées techniques et cinq mille dans dix huit établissements techniques courts. A Neuilly (soixante dix mille habitants), où la population est surtout "bourgeoise", on trouve trois mille trois cents places dans deux lycées d'enseignement général, et deux cents en CET. Il n'y a pas de lycée technique...

On peut déplorer également "le nombre trop faible de places dans les CET", ce qui empêche un nombre important d'élèves d'accéder à la section choisie, même s'ils ont obtenu une décision favorable du conseil de classe.

L'orientation définitive échappe, là encore, au corps professoral.

Enfin, il faut noter le rôle considérable de "détournement" que joue l'enseignement privé, qui permet aux parents fortunés, mieux avisés ou plus ambitieux pour leur progéniture, de tourner les rigueurs d'une orientation "négative" vers une classe de transition ou vers l'enseignement technique.

Pour beaucoup de parents - même attachés à la laïcité - l'enseignement privé constitue ainsi une "soupape" nécessaire et compense l'anonymat et l'arbitraire des décisions d'orientation de l'administration publique".

(C. ARDITTI - Le Monde de l'éducation février 1976 p. 8-9)

Une autre inégalité sociale vient du fait que certaines sections ne sont pas créées dans toutes les villes, dans l'enseignement public, alors qu'elles y existent dans le privé.

Une hiérarchie s'établit entre les matières et les différentes sections.

Jadis "le bon élève" était un littéraire Latin-Grec ; maintenant l'élite se regroupe autour des mathématiques : "les filières scientifiques ont au cours de ce siècle, remplacé les filières comme filières "nobles" de l'enseignement du second degré... ; la primauté des filières scientifiques est évidemment liée au développement des techniques et au dépérissement certain de l'humanisme du début de ce siècle. "L'honnête homme" de 1980 n'est plus celui qui lit PLATON, VIRGILE ou RACINE ; c'est celui qui sait utiliser un ordinateur ou interpréter les photographies d'une chambre à bulles, ou comprendre les mécanismes de la biochimie moléculaire. Les carrières scientifiques, médicales ou techniques sont particulièrement recherchées. L'orientation des élèves en fin de 3ème prend - de ce fait - une importance essentielle". (Conférence faite le 15.2.1977 à la M.J.C. de Vitry-Le-François par Monsieur M. DAVID - Directeur de l'I.R.E.M. de Reims).

Quelques très bons élèves qui, par goût, préféreraient des sections techniques ou littéraires vont cependant s'orienter vers la section C. Ce phénomène s'amplifie de lui-même, car les sections très demandées peuvent sélectionner plus sévèrement que les autres. Elles acquièrent ainsi un prestige de plus en plus grand, au détriment des autres sections qui ne reçoivent plus que les élèves "qui ne peuvent aller ailleurs". On est ainsi amené à penser qu'il faudrait revaloriser certaines filières pour atteindre un équilibre où chaque section aurait à peu près le même nombre de "doués".

Les conditions d'âge sont une contrainte de plus : par exemple, on n'admet en première année de préparation au CAP que les élèves ayant entre 14 et 16 ans à leur sortie de cinquième.

"Pour être parfaitement objective, l'orientation suppose que les élèves soient capables, dès la classe de troisième, de faire des choix clairs entre les avenir professionnels qui leur soient suffisamment perceptibles. Or, le plus souvent il n'en est rien. Les professions citées correspondent généralement à un rêve, une mode ou une influence passagère. Pilote, photographe, berger (surtout au printemps), jardinière d'enfants, éducateur, animateur, psychologue... tels sont les souhaits qui reviennent le plus souvent parmi les six mille appels que reçoivent chaque année les conseillères d'Inter-service parents sur les problèmes d'orientation. Le plus souvent, les questions sur les carrières se font à partir de l'orientation "subie" par les élèves : qu'est-ce que je peux faire avec un bac B ?..., "avec les langues vivantes ?...". La représentation des métiers est très floue. Les enfants - surtout dans les villes - savent peu de chose des activités professionnelles pratiquées par la plupart des adultes, à commencer par celles de leurs parents. L'information sur les métiers "accroche" peu l'intérêt des enfants. Les "foires au métier" ou "carrefours d'informations" organisés dans les établissements - qui font souvent la part trop belle aux carrières socialement ou psychologiquement prestigieuses - attirent peu de monde.

D'après une enquête de l'office national d'information sur les enseignements et les professions (ONISEP), 50 % des élèves de classe terminale n'ont aucune idée de ce qu'ils veulent faire. Parmi ceux qui en ont une, la moitié envisage une profession médicale ou para-médicale.

Ainsi l'orientation est-elle plus "subie" que prise en charge de façon fructueuse par les élèves et les parents. Mais peut-il en être autrement ? Si perfectionnée soit-elle, l'orientation est limitée par les pressions conjuguées des traditions académiques, des préjugés et des clivages sociaux, et du développement économique. Les premières sont internes au système. Elles expliquent, par exemple que les mathématiques se soient peu à peu substituées au latin pour maintenir une élite sélectionnée sur des critères intellectuels, ou dans le fait que les enseignants, dans leur comportement et leur pratique professionnelle, tendent à maintenir un système fondé sur la spécialisation et la hiérarchie des disciplines. Les secondes sont extérieures. Elles sont liées à l'importance de l'éventail des salaires, au maintien et au discrédit des tâches manuelles, à la survivance de professions jouissant de privilèges sociaux et financiers considérables".

(C. ARDITTI Le Monde de l'éducation - Février 1976 p. 9)

## CONCLUSION

Comme on l'a vu ; l'orientation actuelle a pour avantage de supprimer le caractère abrupt de l'examen, mais reste encore très imparfaite.

"En vue d'une orientation future, la note, la moyenne et le classement permettent de formuler un pronostic sur l'avenir scolaire immédiat ou plus lointain de l'élève car c'est bien par référence à ses notes que sera finalement prononcé son passage dans la classe supérieure ou dans le cycle suivant. C'est indéniable, mais nombreux sont ceux qui regrettent que l'orientation ne se fasse que par référence aux notes obtenues précédemment. On a souvent, pour stigmatiser cette procédure d'orientation, qualifié ce système "d'orientation par l'échec" parce qu'il consiste à éliminer successivement des voies nobles du cycle long, puis du cycle long lui-même, et enfin du second cycle court les élèves "présumés inaptes" à y poursuivre des études avec quelques chances de succès ; comme si un enfant qui avait eu à 11 ou 12 ans des difficultés en Mathématiques ou en Français, était condamné, devenu adolescent de 16 ou 18 ans, à demeurer ignare en ces matières...

Qu'est-il le plus important en vue de l'orientation ?

De juger et classer les élèves en excellents, bons, moyens, passables et médiocres ou de chercher à mieux cerner leur propre personnalité, leurs qualités humaines, les points forts sur lesquels on peut appuyer un pronostic et les faiblesses qu'il faudra chercher à compenser. Les comparaisons de notes, de classements ne donne pas ces importantes différences de comportement, éléments pourtant déterminants au moment où il faudra décider pour eux d'une orientation professionnelle...

Et pourtant la notation mériterait d'être étudiée avec sérieux car d'elle dépend l'avenir scolaire et professionnel de nombreux élèves. La note est généralement le seul lien qui unit la famille à l'école, les parents aux professeurs et maîtres de leurs enfants ; elle peut en outre aider les élèves à progresser plus ou moins aisément dans leurs études". (Les parents et la notation novembre 1972)

Certains pensent que l'orientation pourrait consister en un simple avis qui laisserait les familles libres de leur choix, mais "à leur risque", un peu comme le chirurgien qui donne son avis à son patient, lui indique les risques et les avantages d'une opération, en laissant le malade libre de son choix. "Pourquoi ne pas laisser un enfant "essayer" une option qui, à priori, n'apparaît pas faite pour lui ? Si, en cas de conflit persistant, entre le conseil de classe et la famille, l'école offrait la possibilité d'une tentative, et que cette tentative aboutisse à l'échec, la décision définitive serait prise, à posteriori. L'orientation serait alors mieux acceptée par l'élève et la famille". (avis d'une conseillère d'orientation pédagogie p. 41-46)

D'autres reprochent, comme on l'a vu, à l'orientation de gêner le professeur dans sa pédagogie, obligeant l'élève à se sentir en "examen permanent". L'orientation trop précoce dans l'année scolaire (pour des raisons administratives) perturbe le 3ème trimestre et démobilise trop tôt les élèves (tel enfant de 5ème sachant dès le 10 mai qu'il est admis en 4ème aura tendance à relâcher son effort).

D'autres encore regrettent que l'orientation actuelle ait déçu leurs espérances. Malgré tous les efforts faits pour que tous les élèves, quelle que soit leur situation sociale, puissent accéder à la filière qui coïncide le mieux avec leurs vœux et leurs aptitudes, on constate

une faible mobilité sociale :

		Catégorie socio-professionnelle du fils					
Catégorie socio-prof. du père							
	Agricult.	Patron (ind. et commerce)	cadre supérieur	cadre moyen	employé	ouvrier	
Agricult.	38,8 %	6,5 %	1,9 %	3,8 %	4,9 %	34,9 %	
Patron (ind. et com.)	1,9 %	28,3 %	10,7 %	12,5 %	8,6 %	32,6 %	
Cadre sup.	1,9 %	6,7 %	41,7 %	24,4 %	6,7 %	13,7 %	
Cadre moyen	0,5 %	5,1 %	20,9 %	31,7 %	11,3 %	25,1 %	
Employé	0,9 %	7,7 %	10,1 %	18,4 %	15,1 %	41,9 %	
Ouvrier	1,1 %	6,0 %	3,3 %	10,6 %	9,3 %	63,9 %	
% dans la population	11,2 %	9,8 %	7,0 %	11,1 %	8,3 %	44,9 %	

"Malgré le "boom" de l'éducation, la mobilité de la société française demeure faible. Il y a bien des gens qui "montent" et d'autres qui "descendent", mais l'amplitude de ces mouvements est limitée.

Une enquête le montre clairement (I.N.S.E.E. 1970).

"TEL PERE, TEL FILS" telle est la devise de la société française. Quelle que soit la catégorie sociale considérée, c'est dans ses propres rangs qu'elle recrute en priorité. Principalement si on considère les catégories extrêmes de l'échelle sociale : 63,9 % des fils d'ouvriers deviennent ouvriers à leur tour ; quant aux fils d'agriculteurs, 38,8 % font comme leur père et 34,9 % deviennent ouvriers.

Dans les catégories intermédiaires, les choses sont plus fluides : 20,9 % des fils de cadres moyens grimpent dans la catégorie supérieure tandis que 25 % "descendent" chez les ouvriers (31,7 % gardent le statut de leur père). L'instabilité est encore plus grande chez les employés : 28 % "montent" vers les cadres supérieurs ou moyens, tandis que 41,9 % "descendent" chez les ouvriers (15,1 % seulement se reproduisent).

Les patrons de l'industrie se prolétarisent fortement puisque :

30,6 % de leurs fils deviennent ouvriers

10,7 % seulement accèdent à la catégorie de cadres supérieurs

28,3 % font comme leur "père".

(I.N.S.E.E. - enquête de 1970 in "économie et statistique" n° 41 1/73).

Terminons avec Jean GUEHENNO (sur le chemin des hommes)

"Un mot fit fortune il y a quelques années, parmi les pédagogues : orientation, et c'est même devenu une nouvelle spécialité d'être orienteur..."

L'orientation a pour objet de déterminer les dons particuliers de chaque enfant et de le mettre sur son vrai chemin. Quel merveilleux talent qui permet de reconnaître dès la onzième année d'un enfant s'il sera membre de l'Institut ou devra faire fortune dans le petit commerce ! Alain peut-être le dernier de nos philosophes, expliquait que, supposé que ces nouveaux orienteurs aient en effet les moyens (mais il n'en était pas très sûr) de discerner avec précision les tendances et les aptitudes des individus, il ne faudrait utiliser un tel pouvoir que pour cultiver en eux celles qu'ils n'ont pas".

TAXINOMIE



## INTRODUCTION

C'est après la 2ème guerre mondiale que des enseignants américains cherchèrent à étudier de façon plus systématique les objectifs de l'éducation, dans la mesure où une tentative de classification de ces objectifs semblait nécessaire. Il n'était plus suffisant de se contenter d'une critique de l'évaluation, encore fallait-il savoir (entre autres) ce qui était évalué et dans quels buts.

En 1948 à Boston (USA) un groupe d'examineurs de l'enseignement supérieur se réunit autour de Bloom dans l'intention modeste d'élaborer un document de travail simple permettant de classer les questions d'examen. A la suite de travaux plus importants fut publié en 1956 le premier ouvrage sur la question, intitulé "Taxinomie des objectifs pédagogiques".

A l'origine la taxinomie est la "science des lois de classification des formes vivantes" (Robert), par classification en général, puis le résultat même. L'extension de l'Informatique, l'apparition des ordinateurs, la scolarisation de nouvelles couches de jeunes ont-elles contribué au développement de l'étude de ces questions. Le problème reste ouvert.

L'ouvrage de Bloom obtint tout de suite un grand succès et contribua au développement de nombreuses études, suivies par la publication de nouvelles taxinomies, dont la liste n'est sûrement pas close aujourd'hui et dont Gilbert et Viviane De Landsheere ont fait une étude détaillée dans leur ouvrage "Définir les buts de l'Education" qui a fortement inspiré ce chapitre.

## NOTIONS SUR QUELQUES TAXINOMIES

Dans son livre Bloom tente d'élaborer une taxinomie des objectifs pédagogiques en vue d'une classification des buts du système éducatif, afin d'évaluer les problèmes de l'éducation, de préparer les programmes scolaires, de préparer les exercices d'apprentissage.

L'enseignant se trouve souvent confronté aux questions fondamentales suivantes :

- 1) Quels buts ou objectifs pédagogiques l'école (ou le cours) doit-elle se proposer d'atteindre ?
- 2) Quel apprentissage permettra d'atteindre les buts proposés ?
- 3) Comment prévoir les expériences d'apprentissage dans une succession logique et continue pour aider l'élève à les saisir autrement que comme des expériences isolées ?
- 4) Comment mesurer l'efficacité des expériences par le recours aux examens et autres procédés systématiques de contrôle ?

La taxinomie de Bloom apporte quelques lumières sur ces questions. En voici un résumé succinct.

Pour des raisons purement didactiques, Bloom distingue et sépare trois domaines (affectif, cognitif et psychomoteur). Dans le domaine cognitif (ensemble des activités relatives au rappel des connaissances, à l'exercice de la pensée, à la résolution des problèmes, à la création), le premier exploré, il considère 6 catégories plus ou moins nettement hiérarchisées.

1) Acquisition des connaissances

(par exemple : connaissance de la terminologie, connaissance des principes et des lois).

2) Compréhension

(dans un sens assez restreint, par exemple : capacité de transformer du matériel mathématique verbal en énoncés symboliques et vice-versa).

3) Application

4) Analyse

(objectifs qui requièrent une application non routinière de concepts).

5) Synthèse

(production d'une oeuvre personnelle).

6) Evaluation

(formulation de jugements sur la valeur du matériel et des méthodes utilisées dans un but précis).

La taxinomie de Bloom a fait prendre une conscience nouvelle et très vive de l'énorme place que l'enseignement continu à réserver à la connaissance (de mémoire) et de la très faible stimulation des processus cognitifs supérieurs.

Bonora (5. 3 p. 177) remarque : "La taxinomie de Bloom présente l'avantage important d'être un instrument assez clair et exhaustif entre les mains des divers spécialistes susceptibles de l'utiliser pour communiquer entre eux, pour construire un programme scolaire, une séquence d'enseignement ou un instrument d'évaluation".

Pourtant certaines insuffisances ont été relevées. Il semble difficile de trouver des exercices spécifiques à chaque catégorie, même dans un domaine de stricte connaissance, de fait les catégories présentent un certain flou et ne s'excluent pas toujours mutuellement, ce qui permet à Hutin (H p. 36) d'indiquer : "Il est en effet bien difficile de distinguer, dans certains cas, ce qui est connaissance de ce qui est compréhension. Peut-on dire d'autre part qu'il y a compréhension sans qu'il y ait analyse ? Selon le plan d'ensemble de Bloom, le savoir apparaît un peu comme les éléments d'une machine. Or, les différents aspects de la connaissance ne se juxtaposent pas comme les pièces d'un mécano. Il existe entre eux un système très complexe de relations et d'interpénétrations.

La théorie de Bloom est peut-être utile sur un plan général, pour l'élaboration des programmes, mais il nous est apparu qu'elle était difficilement utilisable pour l'élaboration d'un test."

En outre certaines catégories peuvent être mises en cause : "Des réserves sont parfois exprimées à l'égard de l'inclusion, dans la taxinomie, d'objectifs correspondant à des processus supérieurs, tels que "l'évaluation critique", ces processus ne pouvant donner lieu à des apprentissages scolaires. Cet argument repose sur les deux hypothèses suivantes :

- on ne peut distinguer ces processus supérieurs par rapport à l'intelligence générale ;
- l'intelligence étant une caractéristique attachée à la personnalité de l'individu, innée et constante, ne peut être améliorée par aucun traitement pédagogique.

Cette deuxième assertion est très discutable, dans la mesure où les courants actuels de la psychologie mettent souvent en évidence le rôle essentiel du milieu dans le développement de l'intelligence, et on peut donc espérer que l'école - qui est un élément de ce milieu - contribue à l'élévation du niveau intellectuel des enfants qui lui sont confiés". (Bonora 5. 3 p. 178)

Signalons encore une critique adressée à la taxinomie de Bloom : celle-ci accorderait une place beaucoup trop large à la connaissance (de mémoire) par rapport aux processus mentaux supérieurs.

Il faut noter que Bloom avait par avance répondu à certaines de ces critiques, très conscient qu'il était des limites de son entreprise et regrettant parfois l'utilisation de son instrument pour un usage auquel il n'était pas destiné.

A la suite de ces réactions, quelques auteurs s'engagent dans la voie qui venait d'être tracée, certains pour améliorer l'instrument de Bloom, d'autres pour faire oeuvre plus originale dans le cadre d'une conception plus théorique, ou au contraire plus pragmatique.

En particulier la taxinomie de De Block (A, 1975) dont la dominante se situe dans le domaine cognitif, conçue dans un réel souci pédagogique et inspirée par celle de Guilfort (E. 1967), déborde ce domaine dans une tentative d'intégration des champs de recherche. Il est assez difficile de résumer en quelques lignes son travail, qui semble se situer dans la ligne de celui de Bloom ; retenons seulement que les processus sont étudiés suivant trois directions simultanées, ce qui rend la taxinomie de De Block extrêmement subtile. Déjà le modèle de Guilfort se déployait dans trois directions (les opérations, les contenus, les produits ou résultats).

Sans se vouloir exhaustif, on peut encore citer deux taxinomies dont l'une (Krathwohl 2. 71, 1964) intéresse plus particulièrement le domaine affectif et l'autre (Harrow G, 1972) le domaine psychomoteur.

En collaboration avec Bloom et d'autres chercheurs, Krathwohl fait paraître en 1964 un ouvrage dont le titre "Taxinomie des objectifs pédagogiques II Domaine affectif" indique bien qu'il se situe dans la ligne des travaux de Bloom. Voici les catégories retenues :

1. Réception (présence)

Sensibiliser l'élève à l'existence de certains phénomènes et certains stimuli, c'est-à-dire l'inciter à les recevoir ou à y faire attention.

2. Réponse

Réponses qui suivent la simple attention prêtée aux phénomènes. On souhaite qu'un élève soit suffisamment engagé dans un sujet, un phénomène ou une activité pour chercher à le découvrir et avoir plaisir à l'approfondir.

3. Valorisation

Comportement qui est assez solide et stable pour prendre les caractéristiques d'une croyance ou d'une attitude. L'élève manifeste ce comportement avec suffisamment de cohérence, dans les circonstances appropriées, pour qu'on estime qu'il détient une valeur. Intériorisation d'un ensemble de valeurs spécifiques idéales : Le comportement est motivé, non par le désir de plaire ou d'obéir, mais par l'engagement individuel à la valeur fondamentale déterminant le comportement.

4. Organisation

Organiser les valeurs en système, déterminer les interrelations qui existent entre elles, établir celles qui sont dominantes et plus profondes.

5. Caractérisation par une valeur ou un système de valeurs

Les valeurs ont une place dans la hiérarchie des valeurs de l'individu. Elles sont organisées en une sorte de système intrinsèquement cohérent. Elles ont réglé le comportement de l'individu assez longtemps pour que celui-ci s'y soit adapté.

Chacune de ces catégories comporte un certain nombre de sous-catégories. Le principal reproche fait à la taxinomie de Krathwohl est que le caractère de grande abstraction et grande généralité qu'elle présente rend malaisée son utilisation pratique.

"Une taxinomie du domaine psychomoteur" titre de l'ouvrage de A. J. Horrow, fut publié en 1972. G. et V. de Landsheere en donnent la présentation synthétique suivante (1.46 p. 181)

Niveaux

1. Mouvements réflexes	Base de tous les mouvements	Non appris
2. Mouvements naturels ou fondamentaux	Combinaison de mouvements réflexes	Les combinaisons existantes seront cependant utilisées dans les mouvements volontaires.

N. B. 1 et 2 ne font pas l'objet d'objectifs pour l'éducation, du moins dans les cas normaux.

---

3. Aptitudes perceptives	C'est à ce niveau que commence normalement l'enseignement	Se développent par maturation et apprentissage En effet les expériences d'apprentissage
4. Aptitudes physiques		- aiguisent les perceptions - développent les aptitudes physiques

---

5. Habiletés motrices	Il existe à ce niveau un continuum d'habileté	Dépendent : - du contrôle des mouvements fondamentaux, - de l'efficacité de la perception, - du niveau de développement des aptitudes physiques
-----------------------	---	--

---

6. Communication non verbale	Il existe à ce niveau un continuum d'habileté	Quand le sujet dispose d'un répertoire d'habiletés motrices il est prêt pour la création de mouvements esthétiques Le 6. 2. représente le sommet de la hiérarchie : expression par la danse, le mime etc...
------------------------------	---	--

---

Ceci n'est bien sur qu'un résumé, les sous-catégories foisonnent dans le modèle de Harrow. Sa classification ne permet pas toujours une distinction nette entre les divers comportements, certains relevant de plusieurs sous-catégories. De même qu'une des grandes qualités

de la taxinomie de Krathwohl est de mettre l'accent sur la dépendance des domaines affectifs et cognitifs, un des avantages de celle de Harrow est qu'elle souligne la dépendance cognitif-psychomoteur au niveau des aptitudes perceptives.

Il existe bien sûr d'autres taxinomies, et il n'est pas sûr du tout que les exemples cités plus haut soient les plus intéressants. Le lecteur courageux, qui sera parvenu jusqu'ici sans manifester un trop grand accablement pourra consulter l'ouvrage de Gilbert et Viviane de Landsheere (1. 46) pour plus de précision.

### Intérêt et limites

On peut constater actuellement que les processus cognitifs ont été étudiés de façon relativement approfondie, dans la mesure où les apprentissages cognitifs présentaient pour l'école un intérêt primordial. L'évaluation de ces apprentissages offrait d'ailleurs moins de difficultés que la "mesure" des émotions par exemple. Et d'ailleurs ne faut-il pas noter chez de nombreux enseignants, peu ou mal formés à l'approche de ces problèmes (problèmes dont la connaissance reste encore faible), un certain désarroi devant les manifestations affectives des élèves, voire un refus de les prendre en compte ? Landsheere (1.46 p. 194) remarque que "dans le passé, l'affectivité a parfois été cultivée de façon intense, mais unilatéralement, dans des secteurs servant des endoctrinements".

Quant au domaine des activités psychomotrices, activités qui se manifestent particulièrement à travers certaines disciplines (E.P.S., musique, arts plastiques, travaux manuels éducatifs...), chacun sait quelle place et quelle importance les programmes scolaires leur accordent ! Il faut souligner ici l'importance des processus psychomoteurs dans tous les aspects de l'éducation, ils affectent aussi bien les apprentissages que la formulation par l'élève des résultats qu'il obtient : comment nier que l'adresse et l'organisation d'un élève aient, au même titre que ses connaissances et ses aptitudes, une grande importance dans la réalisation d'un travail écrit (en temps limité, ou même illimité), d'un schéma ou d'une manipulation ?

Les premières taxinomies ont distingué nettement trois domaines mais la plupart des auteurs (sinon la totalité de ceux-ci) agissant ainsi pour des motifs d'ordre pédagogique ou didactique, sont bien conscients que cette distinction reste artificielle, qu'il est impossible de découper la personnalité d'un individu en plusieurs secteurs parfaitement isolés et que dans tout comportement résident des éléments cognitifs, affectifs et psychomoteurs.

Si certains auteurs continuent à explorer le domaine cognitif, si certains encore tentent de faire progresser la classification des processus affectifs et psychomoteurs, il en est d'autres qui essaient d'intégrer ces trois domaines, et c'est probablement l'une des principales directions de recherche que l'on puisse entrevoir. Il est clair que ces derniers auront fort à faire...

Comment dire aujourd'hui si l'apport des taxinomies déjà élaborées se révélera fondamental dans le cadre d'une synthèse de ces travaux (peut-être d'abord à l'intérieur de chaque domaine, considéré isolément), ou si ces modèles seront rapidement dépassés et subsisteront comme les premiers balbutiements d'une science qui évoluera très vite?

Si ses limites doivent être précisées, les mérites de la taxinomie doivent être reconnus. Landsheere souligne (1.46) : "La façon d'utiliser les taxinomies n'est pas toujours bien comprise. Souvent on a voulu en faire un instrument destiné à formuler des objectifs, alors qu'il est d'abord destiné à analyser des objectifs existants... La taxinomie est d'abord une invitation à préciser quels comportements seront recherchés, installés, encouragés, renforcés de préférence à d'autres. C'est le pas décisif vers l'opérationnalisation.

Mais en pointant dans une taxinomie les comportements susceptibles de traduire l'objectif poursuivi, on peut ainsi apercevoir des comportements pour lesquels on ne dispose pas d'objectifs. A ce moment, la taxinomie devient une source d'inspiration d'objectifs nouveaux. C'est son second rôle". (p. 65)

Et plus loin (P; 195) " Dans l'état actuel des choses, le mouvement taxinomique en éducation est de toute façon bénéfique. Car, même si les instruments sont imparfaits, voire erronés, ils suscitent néanmoins une réflexion, une mise en question, un dépassement dont l'éducation ne peut que profiter.

A mi-chemin entre les grandes options idéologiques et les micro-objectifs, les taxinomies semblent jeter des ponts entre la philosophie et la technique de l'éducation. Ce n'est pas un de leurs moindres mérites".

Mais tous ces mérites ne doivent pas masquer les limites de l'entreprise ; il faut encore rappeler que la taxinomie ne saurait être considérée comme un instrument pour la formulation de nouveaux objectifs.

Certaines taxinomies (par exemple celle de Bloom et De Block) ont été caractérisées et critiquées comme excessivement "mentalistes" dans la mesure où les catégories retenues désignent des "facultés", "aptitudes" ou "capacités" (compréhension, analyse, synthèse... intégrer, comprendre, savoir...). Sans revenir sur les controverses qui peuvent exister et se développer à propos de cette terminologie, constatons que ces dénominations ne peuvent avoir de signification opérationnelle directe (en terme de comportements) et qu'elles laissent une part importante à la subjectivité, même si les auteurs se sont efforcés par des exemples concrets de préciser leur terminologie. N'existe-t-il pas, d'ailleurs, un danger que ces catégories soient confondues avec les objectifs eux-mêmes ? Ce risque rejoint une préoccupation exprimée par Krygowska (4. 3 p. 68/69) :

"Il est possible de contrôler et d'évaluer à l'aide de moyens connus les connaissances, les habiletés algorithmes, le savoir de résoudre les problèmes typiques selon des schémas précédemment assimilés, bien que même dans ce domaine il y ait encore beaucoup de difficultés. Les progrès concernant "le développement", "l'initiation", "la familiarisation" etc... au contraire ne conviennent pas aux moyens primitifs de l'évaluation objective, dont on dispose actuellement. C'est pourquoi la tentative extrême à exprimer toutes les acquisitions par des mesures objectives, dans certaines conditions, peut avoir une influence nocive sur l'enseignement. Elle renverse l'ordre du procéder didactique : "se poser les objectifs et chercher les moyens adéquats du contrôle et de l'évaluation de leur réalisation" ; ce renversement s'exprime donc dans la tendance de poser des objectifs de l'éducation mathématique en fonction des outils du contrôle et de l'évaluation, dont on dispose".

En opposition avec ces taxinomies mentalistes, certaines taxinomies plus pragmatiques, classant uniquement des comportements, ont pu être considérées comme des instruments relativement frustes, catalogues de comportements avant d'être des moyens de classification et des incitations à la réflexion.

F. Bacher (1-6 p. 45) remarque : "D'autre part, pour contrôler qu'un objectif a été atteint, il est nécessaire de le traduire en comportements observables ; le plus souvent, cette opération conduit à prendre en considération des objectifs beaucoup plus limités que ceux qu'ont d'abord été formulés ; et bien qu'il existe une filiation directe entre un objectif et les comportements qui permettent de dire qu'il a été atteint, définir ces comportements ne suffit pas à définir complètement l'objectif lui-même, qui constitue un concept plus abstrait".

On peut encore signaler un écueil, qui risque de fausser l'utilisation des taxinomies : il serait faux de croire que celles-ci peuvent être employées dans l'absolu, sans que soient pris en compte le passé de l'individu concerné et le contexte psycho-sociologique dans lequel il s'inscrit ; tel élève, résolvant une question, mettra en jeu des processus certainement plus complexes que tel autre, plus vieux ou mieux familiarisé avec le type de problème posé.

Un élève de 2ème C résolvant l'équation " $3x - 6 = 0$ " ne progressera pas suivant les mêmes processus mentaux qu'un élève de 5ème, à supposer que ce dernier vienne à bout du problème...

Ici interviennent aussi la formulation des questions posées et la forme du langage utilisé : un élève ne saura pas répondre "correctement" à une question, soit parce que la formulation en est ambiguë ou hermétique, soit parce que lui-même est limité par ses propres insuffisances en vocabulaire ou en syntaxe (exemple vécu : la confusion faite par un élève de 1ère entre les termes "simultanément" et "successivement").

En tout état de cause il est parfaitement illusoire de penser que la taxinomie a permis de résoudre scientifiquement, techniquement et définitivement le problème de la formulation des objectifs pédagogiques. Mais elle a créé les conditions techniques d'un approfondissement de la réflexion sur ces derniers, de même que sur leur opérationnalisation.

## REFLEXIONS SUR LES OBJECTIFS

Il y a un siècle les buts de l'enseignement, alors que l'enseignement long était destiné aux enfants de sexe masculin issues des couches sociales favorisées, pouvaient paraître relativement clairs, de même que les objectifs généraux de l'éducation où règnaient les catégories de la morale. Ainsi, encore en 1897, les objectifs généraux de l'enseignement primaire en Belgique.

### Objectifs cognitifs :

- 1) Eveiller constamment l'esprit d'observation, de réflexion et d'invention.
- 2) Habituer les enfants à exprimer simplement, mais correctement, leurs propres observations, leurs propres jugements.

### Objectifs affectifs :

- 1) Inculquer les principes de la morale.
- 2) Inspirer le sentiment du devoir, l'amour de la patrie, le respect des institutions nationales, l'attachement aux libertés constitutionnelles.
- 3) Habituer à observer les règles de la bienséance.

Depuis lors le développement scientifique et l'extension de la scolarité ont contribué à une réflexion approfondie sur ces problèmes, mais il faut pourtant avouer que si tous les enseignants peuvent s'accorder sur quelques formulations très générales (en 1947 aux U.S.A. : "Comprendre les idées des autres et exprimer les siennes de façon pertinente et claire"), qui ne serait embarrassé pour répondre à la question suivante :

"Quel objectif précis poursuivez-vous lorsque vous proposez telle activité à vos élèves ?"

Le développement de la taxinomie, qui a permis de constater quelle (faible) place occupe réellement la stimulation des processus cognitifs les plus complexes, et de s'interroger en conséquence sur la distorsion entre les intentions avouées et les comportements réels qui se manifestent au sein du système scolaire, débouche naturellement sur un approfondissement des objectifs pédagogiques, de sorte que ceux-ci soient formulés en termes opérationnels.

Citons à cet égard J.C. Flanagan (2-51 p. 763) :

"Pour déterminer ou évaluer l'efficacité d'un programme d'éducation, il est nécessaire de mesurer, ou du moins de classer les résultats..."

Certains de ces résultats... peuvent être mesurés facilement et directement. Ce sont ceux qui se proposent un objectif d'information comme la connaissance ou la signification des mots, des règles de grammaire, l'acquisition de techniques mathématiques, l'aptitude à appliquer les principes scientifiques à des problèmes nouveaux.

D'autres effets de l'éducation sont plus difficiles à déterminer ou à évaluer, c'est le cas du développement du jugement, de la sensibilité...

La première étape de cette évaluation consiste toujours à définir les objectifs du programme d'éducation en termes opérationnels. A moins de savoir quels changements le programme d'éducation cherche à provoquer dans le comportement des élèves, il est impossible de dire s'il est efficace ou non.

Un objectif éducatif qui se proposerait de changer l'élève d'une façon qu'on ne peut observer ni directement, ni indirectement est extérieur au domaine de la connaissance... et sans importance pratique".

Revenons avec Landsheere (1-46 p. 201-202) dans le vif du sujet.

"Voyons dans le concret une filière complète d'objectifs :

Fin de l'éducation : Un homme parfaitement développé dans une société en constant progrès.

Un but : Protéger l'environnement (valeur sous-jacente : un environnement non pollué est nécessaire à la survie de l'humanité. L'environnement est donc précieux).

Un objectif intermédiaire : Connaître les principaux agents de pollution.

Un objectif traduit en comportements observables : L'élève recueillera dans son environnement proche 5 preuves concrètes de pollution.

Micro-objectif : L'élève saura lire le taux de C.O. contenu dans l'atmosphère sur le cadran d'un analyseur portatif.

La formulation complète d'un objectif opérationnel comprend cinq indications précises :

- 1) Qui produira le comportement souhaité ?
- 2) Quel comportement observable démontrera que l'objectif est atteint ?
- 3) Quel sera le produit de ce comportement (performance) ?
- 4) Dans quelles conditions le comportement doit avoir lieu ?
- 5) Quels critères serviront à déterminer si le produit est satisfaisant ?

Exemple :

- 1) l'élève
- 2) saura construire
- 3) un poste de radio à transistors
- 4) en choisissant lui-même les pièces au magasin, en se référant au schéma adopté
- 5) l'appareil devra capter correctement des émissions d'au moins cinq émetteurs différents sur ondes moyennes et de cinq émetteurs sur ondes longues".

L'idée générale est donc qu'un objectif est rendu opérationnel lorsqu'ont été définis des méthodes, des étapes et des moyens qui permettent de s'assurer que cet objectif a été atteint. Il va sans dire qu'à tous les niveaux surgissent des questions, qui souvent opposent les théoriciens des différentes écoles. Quelques-unes de ces questions vont être ici effleurées.

Un premier problème se pose immédiatement : tout apprentissage est-il passible d'analyse exhaustive préliminaire et de mesure rigoureuse ? Si la réponse semble négative (on ne peut prédire quel est le produit attendu des processus créatifs), cela ne signifie nullement qu'on ne peut pas éduquer à la créativité.

Un autre problème réside dans la mesurabilité des résultats d'un processus éducatif : tous les résultats sont-ils mesurables ? et d'ailleurs est-il possible de distinguer clairement les résultats intentionnels de ceux qui ne le sont pas ?

Un troisième problème est encore plus fondamental : tous les objectifs peuvent-ils être explicités ? En fait, quelle que soit la réponse à cette interrogation, rien ne s'oppose à leur recherche, et à l'opérationnalisation des objectifs qui auront été reconnus, précisés et retenus.

Soulignons enfin qu'on ne peut éluder la question si débattue de la méthode et du contenu. Si chacun s'accorde à dire que l'éducation ne doit pas procéder du dressage (mais dans les faits n'est-ce pas encore trop souvent le cas ? on pense immédiatement à l'importance extrême que tendent à prendre les savoir-faire au détriment du savoir et de la connaissance véritable), peut-on dire qu'il est possible d'apprendre sans apprendre quelque chose ? Une contribution importante est apportée par F. Bacher (1-6 p. 45-46) :

"On a parfois l'impression, lorsqu'on lit des discussions récentes sur les buts de l'enseignement, que le corps de connaissances qui constitue la base d'une discipline tend à être négligé par certains qui ne voient plus, dans l'éducation, qu'une occasion de développement de capacités cognitives s'appliquant indifféremment à n'importe quel contenu. Ces capacités mettraient

celui qui les possède en mesure d'acquérir n'importe quelles connaissances au moment où il en aurait besoin.

Il y a tout de même lieu de s'inquiéter des conséquences de l'adoption d'une position aussi extrême, qui néglige au moins trois faits. Le premier est que les connaissances, dans une discipline, constituent un ensemble structuré qui ne peut-être maîtrisé, puis enrichi et mis à jour, qu'à la suite d'une intégration progressive et souvent lente de connaissances par l'individu. Le second fait est que l'exercice et le développement de capacités élevées, dans un domaine, nécessite l'acquisition de mécanismes et d'informations suffisamment disponibles pour permettre des rapprochements, des comparaisons et une restructuration rapide d'éléments nombreux éventuellement complexes ; l'acquisition de connaissances constitue un préalable nécessaire. Le troisième fait tient à ce que l'utilisation des mêmes outils intellectuels se spécifie en fonction de la discipline considérée, de sorte qu'une formation de niveau élevé ne peut être seulement générale".

Nous faisons, grâce au lecteur épuisé de la taxinomie élaborée en 1970 par D'HAINAUT (B) afin de réaliser une synthèse générale de la méthodologie utilisée dans la définition des objectifs...

Ainsi depuis quelques années, les buts et les méthodes de l'éducation ont pu être l'objet de recherches et d'études qui ont permis de préciser un certain nombre de problèmes, en particulier d'étudier le problème de la formulation des objectifs en des termes plus rigoureux que les principes abstraits et généraux qui dominaient la littérature pédagogique.

#### CONCLUSION (en forme d'ouverture)

Si l'étude des objectifs pédagogiques a pu ainsi se développer dans de nouvelles directions, il n'en reste pas moins qu'il est impensable d'isoler les objectifs pédagogiques des objectifs généraux de l'éducation, problème à caractère social qui ne concerne plus seulement les techniciens de l'éducation, mais intéresse et engage l'organisation sociale dans son ensemble.

Comment penser en effet que les objectifs généraux définis pour l'enseignement primaire belge à la fin du 19ème siècle, et qui ont été énoncés plus haut (voir p. 79 du présent texte), requièrent les mêmes objectifs et les mêmes moyens pédagogiques que les objectifs généraux pour l'enseignement primaire définis dans le même pays en 1936, puis en 1973 ?

Objectifs affectifs (1936) : 1) Exercer une action éducative, stimuler des intérêts, révéler des valeurs, provoquer la libération spirituelle et l'élévation de l'âme.

2) Préparer nos enfants au gouvernement de leur pensée comme à celui de leur conduite.

Objectifs affectifs (1973) : Dix objectifs classés dont :

3) Pouvoir accueillir et orienter le changement.

10) Susciter la créativité, la pensée divergente, la liberté, la disponibilité, la responsabilité, la faculté de s'assurer, l'engagement.

On pourra trouver une exposition plus complète dans l'ouvrage de Gilbert et Viviane De Landsheere (1-46 p. 50-53), lesquels soulignent (p. 52) :

"Alors que 1897 a pour évident souci le respect de l'ordre établi, l'acquisition des vertus bourgeoises, le conformisme ou, si l'on préfère, la convergence, 1973 marque le triomphe de l'homme qui reste certes solidaire du groupe social, mais se voit accorder le droit de s'épanouir, de s'exprimer, de contester".

Aussi c'est en dernière analyse un problème de société que la définition des buts de l'éducation, et les psychologues et chercheurs se s'y sont pas trompés. Les quelques citations qui suivent permettront d'en rendre compte. Nous les citons en guise de conclusion à ce chapitre, au titre d'éléments de réflexion.

"Ou l'éducation trouve-t-elle sa finalité ?

Au-delà de toutes les nuances, deux courants de pensée se dessinent nettement.

Pour les uns, les buts de l'éducation découlent (déductivement) d'une conception arbitraire de l'homme, de ses caractéristiques immuables, de sa "vraie" nature, c'est-à-dire de sa nature essentielle.

Pour les autres, le but de l'éducation est la conquête quotidienne du milieu mis en service de l'individu et de la collectivité. Dans cette perspective, les buts ne se définissent pas déductivement à partir d'un principe ou d'une vérité abstraite, mais naissent des nécessités de l'action, de la contingence.

Certains vont même jusqu'à considérer que l'identité entre les fins et l'action est si étroite qu'une formulation explicite des fins devient pratiquement impossible ?...".

"Dans un domaine qui concerne pratiquement tous les groupes sociaux (enseignants, élèves, mais aussi parents et employeurs) on conçoit que le choix des buts pourrait résulter d'une concertation de tous les intéressés.

Du point de vue de la mesure, qui est ici le nôtre, il semble assez évident que cette réflexion est la condition nécessaire à la construction d'instruments adoptés à leur objet : le contrôle des effets attendus de l'action éducative. Cette formulation suggère d'ailleurs l'existence possible d'effets non prévus, favorables ou défavorables, qu'il serait important d'apprécier aussi".

F. Bacher (1-6 p. 45)

"L'éducation est un phénomène social. Le détenteur du pouvoir, dans une société donnée, n'accepte pas des dispositions éducatives détruisant l'ordre qui lui donne le pouvoir.

J. Goodlad a donc raison de distinguer trois niveaux de décision en matière d'objectifs :

- les décisions sociales, prises par le pouvoir politique,
- les décisions institutionnelles, prises par les autorités pédagogiques, dans la ligne des décisions sociales,
- les décisions d'enseignement, principalement prises par les maîtres".

G. et V. De Landsheere (1-46 p. 29)

"Les finalités de la pédagogie sont vues comme contradictoires ; en effet, d'une part, on les définit par rapport à une fonction d'adaptation à la société, dans ce sens le rôle du pédagogue consiste à transmettre des systèmes de valeurs, des normes et un savoir déterminé en partie par celles-ci.

D'autre part, on ressent que ces finalités ne correspondent qu'à un versant de la pédagogie ; beaucoup se refusent à ce rôle d'adaptateur, étant eux-mêmes en conflit avec certaines des valeurs et normes sociales. Ces pédagogues tendent à définir leur rôle comme "formation à la liberté" ; ils estiment que la société est en mutation, et que, notamment, le principe d'autorité est une des valeurs contestées.

Ainsi la pédagogie aurait pour finalité de faciliter un changement de la société en développant chez les enseignés des aptitudes à l'autonomie..."

J. Filloux - A. Lévy (C p. 214)

"Même s'il se trouve à la base de toute décision éducative l'objectif n'est pas le premier déterminant. Il résulte lui-même d'une option plus fondamentale encore : les valeurs de l'individu et les idéologies de la société.

Aussi longtemps que ces valeurs ne sont pas reconnues, explicitées, défendues, mises en question aussi, l'ambiguïté (volontaire ou non) grève le processus éducatif. De là il faut toujours partir et là il faut toujours revenir.

Aussi quel que soit le niveau auquel il travaille à un moment donné (objectif intermédiaire, objectif opérationnel...) l'éducateur et, le plus rapidement possible, le "learner" (l'élève) devraient à chaque moment de décision, reposer la question de compatibilité et de cohérence avec les fins, les buts et les valeurs qui les sous-tendent.

Sans ce retour continuels aux origines, la voie s'ouvre aux incohérences, aux aberrations, aux trahisons aussi".

G. et V. De Landsheere (1-46 p. 259)

"Plus les objectifs seront formulés de façon vague, au départ, plus l'action de l'école risque évidemment de s'éloigner des intentions de ses organisations, à moins que l'imprécision de la formulation ne soit un artifice permettant de faire passer pour noble et humanitaire une politique scolaire autoritaire, servant les intérêts d'une minorité, et réalisé systématiquement grâce à des cadres pédagogiques seuls chargés de définir les modalités éducatives, de les imposer et de les faire respecter".

G. et V. De Landsheere (1-46 p. 8)

"Les conditions de la vie politique, économique et sociale dans notre pays amènent aujourd'hui à faire le point des changements nécessaires dans la conception du système éducatif, pour préparer les voies de la France de demain".

R. Haby (F p. 1)

"La nouveauté réside dans un ensemble de travaux visant à formuler les buts de l'éducation en des termes suffisamment précis, comportant une traduction opérationnelle. Des recherches expérimentales ne se proposent nullement de choisir des buts à l'éducation. Ce choix ne relève pas de la science, mais de la politique : seuls les citoyens sont habilités à assigner ses buts à un système éducatif faisant partie intégrante de la société et ne pouvant fonctionner que grâce aux moyens que cette société décide de lui consacrer...".

M. Reuchlin (1-67 p. 228)

Notre ambition, en rédigeant ce chapitre, était d'apporter quelques éléments d'information sur les travaux en cours. Il est fort possible que ce (trop) long texte ait dépassé le cadre initialement prévu et que certains aspects techniques lui aient conféré un caractère un peu indigeste.

Cependant, nous espérons (très immodestement) qu'il a été en mesure d'inspirer un lecteur des réflexions profitables...

### CONCLUSION

Comme on a pu le constater plusieurs composantes peuvent déterminer la relativité de l'évaluation : difficulté de préciser ce qui est noté ou jugé, la personnalité de l'élève, celle du professeur ; difficulté de trouver un terrain d'entente pour définir des critères unanimement reconnus et acceptés. On pourrait à la suite de cette lecture être tenté par des solutions expéditives comme la suppression pure et simple des notes et par la même des examens, des concours, de la sélection et d'une façon générale de tout système contraignant. Cette solution n'est pas nouvelle ; certains l'ont déjà envisagée et la défendent encore avec beaucoup de vigueur et de sincérité. Nous ne partageons pas cette vue parce que la suppression des notes ne résoud pas tous les problèmes. En particulier celui de la "fonction de défense" contre une proximité "professeur-élève" et l'on sait qu'il n'est ni opportun ni souhaitable de supprimer brutalement une défense. On aurait plutôt intérêt à approfondir la connaissance que nous avons de la fonction "évaluation", à prendre conscience que nous, professeurs, évaluons sous la pression de déterminants que nous ne contrôlons pas toujours. Alors, peut-être, une certaine prise de distance vis à vis de l'évaluation permettra un assouplissement de cette fonction et d'entreprendre, dans chaque discipline, un travail de grille de critères définis par les enseignants.

BIBLIOGRAPHIE

- 1.6 BACHER (F)  
La docimologie  
in REUCHLIN (M) Traité de psychologie appliquée Volume VI  
Paris 1975
- 1.8 BEAUTE (J.)  
Français et docimologie  
Documents et recherche. Lettres n° 12 (pp. 13-15) Hatier 1974
- 1.9 BLOOM (B.S.)  
Taxonomy of educational objectives  
David Mac Kay, New-York 1956  
cité dans Cahiers Pédagogiques n° 83 "notes de contrôles" p. 12  
(Paris SEVPEN 1969)
- 1.11 BONBOIR (A.)  
La docimologie  
Paris PUF 1972 (194 p) Collection SVP
- 1.16 BRASSART (R.)  
Classement ou niveau ?  
L'éducation nationale n° 825 (pp 16-17) 4 mai 1967
- 1.32 FRAISSE (S.)  
Les examens : autocritique d'un examinateur  
Cahiers pédagogiques n° 53 mars 1965 (pp 91-93)
- 1.35 GLAESER (G.)  
Problèmes moraux et politiques de l'évaluation pédagogique.  
Bulletin inter IREM n° 12 pp 5-12 édité par l'IREM de  
Lyon (juillet 1976)
- 1.41 HOTYAT (F.)  
Les examens et les moyens d'évaluation dans l'enseignement.  
Documents pédagogiques internationaux de l'institut de  
l'UNESCO pour l'éducation. Bourrelier 1962
- 1.42 HOTYAT, BONBOIR, BACHER etc  
Docimologie et éducation. Actes du Colloque de LYON 1968  
Les sciences de l'éducation pour l'ère nouvelle.  
Didier, Paris, Montréal, Bruxelles. 2-3 revue internationale  
avril-sept. 1969
- 1.45 LANDHEERE (G. de)  
Evaluation continue et examen. Précis de docimologie  
Bruxelle-Labor, Paris - F. Nathan 1972 (238 p)
- 1.46 LANDSHEERE (V. et G.)  
Définir les objectifs de l'éducation  
PUF 1976 (294 p) Collection Pédagogie d'aujourd'hui

- 1.63 PIERON (H)  
Examens et docimologie  
PUF (180 p) 1963
- 1.67 REUCHLIN (M.)  
Problème d'évaluation  
in Debesse et Mialaret, traité de sciences pédagogiques volume IV  
Paris 1973
- 1.74 ROSENTHAL  
Pygmalion  
Casterman 1971
- 2.8 BACHER (F.)  
La normalisation de la notation  
BINOP 1969, 25, 75-90
- 2.11 BACHER  
Congrès de Berlin sur les possibilités et les limites de  
l'application des tests dans les écoles.  
BINOP, 24, 42-49, 1965
- 2.50 FLANAGAN (J.)  
Project Talent. Some early findings from a survey  
N.E.A. Journal January 1964
- 2.51 FLANAGAN (J.C.)  
L'évaluation de l'efficacité des programmes d'éducation.  
Bull. Psychol. 20, 763-769, 1967
- 2.71 KRATHWOHL (D.R.) BLOOM (B.S.) MASIA (B.B.)  
Taxonomy of educational objectives. The classification of  
educational goals. Handbook II affective domain.  
New-York, D. Mac Kay and Co, 196 p, 1964
- 2.74 LANDSHEERE (G. de)  
Introduction à la recherche en éducation  
Paris, A. Colin Bourrelier (3<sup>e</sup> éd.) 1970
- 2.80 LEHMAN (E.)  
L'enseignement à l'étranger : la Suède  
Bulletin inter-IREM n° 13, oct. 1976, p. 34, édité par l'IREM de Lyon
- 2.86 LURCAT (L.)  
L'échec et le désintérêt scolaire.  
Orientation, Fr. (1975) 15, n° 53, 51-71
- 2.97 NGUYEN HUAN (A.)  
Recherches anglaises sur la notation des devoirs de rédaction  
BINOP, 19, 254-266, 1963
- 2.108 PEQUIGNOT (H.)  
L'équation personnelle du juge  
Semaine des hopitaux, 1966, n° 14 (mars 20) Suppl 11, pp 4-11
- 2.115 REMONDINO (C.)  
Etude factorielle sur la notation de compositions scolaires  
portant sur la langue maternelle.  
Le travail humain, 22, 27-40, 1959

- 2.133 Les parents et la notation  
l'école des parents, 1972, n° 9 pp 42-50
- 3.1 AVANZINI (G.)  
La pédagogie au 20ème siècle.  
Privat. 1975 (400 p)
- 3.7 MARCHAND (M.)  
Le couple de l'éducateur et de l'élève dans leurs relations  
concrètes.  
Université d'Alger ; Fac. de lettres  
Impr. L. Fouque Oran (283p) 1954
- 3.11 REUHLIN (M.) BACHER (F.)  
L'appréciation des élèves par leurs professeurs (en classe de  
3ème) Revue française ped. n° 2, 19-25, 1968
- 3.12 REUHLIN (M.) BACHER (F.)  
L'orientation à la fin du premier cycle secondaire  
Paris PUF 1968
- 3.22 VANDERMEERSCH  
L'orientation scolaire : "De l'examen à l'orientation"  
Pédagogie Fr. 1975, n° 4, 9-51.
- 3.29 WISEMAN (S.)  
The marking of English composition in grammar school selection  
Brit, J. educ. Psychol. 19, 200-209, 1949
- 4.3 KRYGOWSKA  
De 10 à 16 ans Congrès international de Karlsruhe, Août 1976  
Bull. n° 7 de l'IREM et n° 22 de la régionale A.P.M.E.P.  
1976, pp 61-75
- 4.5 PANNETIER  
Notes, histogrammes et normalisation  
Education Nationale n° 825, pp 20-23, mai 1967
- 5.1 BAUDELLOT ESTABLET  
l'école capitaliste en France  
Ed. Maspéro "cahiers libres" 1971
- 5.3 BONORA (D)  
Les buts de l'éducation  
in traité de Psychologie appliquée, sous la direction de  
REUHLIN (M.) pp 139-191 PUF 1973
- 5.14 Docimologie et éducation  
Actes du colloque de Lyon 1968  
Les sciences de l'éducation. Didier, 2-3 revue internationale  
N° spécial (avril-sept. 1969)
- 5.18 FRAPPE (F.)  
Le jeu de l'Internat et du hasard (Analyse docimologique du  
concours d'internat des hôpitaux de Tours en 1975)  
Thèse Université de Tours - Faculté Tours n°t Tours 1976-82

- A DE BLOCK (A.)  
Taxonomie van Leerdoelen  
Anvers, Amsterdam, Standaard Uitgeverij 1975
- B D'HAINAUT (L.)  
Un modèle pour la détermination et la sélection des objectifs  
pédagogiques du domaine cognitif.  
in Enseignement programmé 1970, 11
- C FILLOUX (J.) LEVY (A.) etc...  
(Association d'étude pour l'exansion de la recherche scientifique)  
Reflexions sur quelques expériences de stages de perfectionnement  
d'enseignements, de formateurs et d'éducateurs.  
in : Les enseignants du second degré. Leur situation dans  
l'établissement scolaire. Quatre enquêtes préparatoires au  
colloque d'Amiens 1968.
- D GUILFORD (J.P.)  
A system of psychomotor abilities  
in American Journal of Psychology, 71, 1958
- E GUILFORD (J.P.)  
The Nature of Human Intelligence  
New-York, Mac Grow Hill, 1967
- F HABY (R)  
Pour une modernisation du système éducatif  
La Documentation Française Les cahiers français  
n° spécial hors série février 1975
- G HARROW (A.J.)  
A Taxonomy of the Psychomotor Domain  
New-York, D? Mac Kay, 1972
- H HUTIN (R)  
in Docimologie et éducation (actes du colloque de Lyon 1968)  
Les sciences de l'Education (2-3 avril-septembre 1969)

## TABLE DES MATIERES

	Page
<u>INTRODUCTION</u>	1
Sondage	2
Historique	5
Image sociale des examens	7
<u>DOCIMOLOGIE</u>	
I. Expériences et analyse critique	9
1) Première source d'erreurs : l'évaluateur	10
Les évaluateurs ne sont pas d'accord entre eux	10
L'évaluateur n'est pas d'accord avec lui-même	11
L'évaluateur à l'oral	13
Les préjugés de l'évaluateur	16
2) Deuxième source d'erreurs : le sujet	19
Le barème	19
Le choix du sujet	20
Le coefficient de pondération	22
3) Troisième source d'erreurs : l'évalué	24
II Docimologie constructive	
Introduction	26
Amélioration des examens traditionnels	29
Les épreuves normalisées de connaissances	32
Utilisation des épreuves à expression libre	43
Utilisation des notes et appréciations scolaires	49
Perspectives	52
<u>ORIENTATION</u>	
Historique	55
Orientation actuelle	57
Conclusion	65
<u>TAXINOMIE</u>	
Introduction	69
Notions sur quelques taxinomies	69
Intérêt et limites	75
Réflexions sur les objectifs	79
Conclusion (en forme d'ouverture)	82
CONCLUSION	87
BIBLIOGRAPHIE	88