

EVALUATION CONTINUE ET EXAMENS.

Précis de docimologie.

DU MEME AUTEUR

Introduction à la recherche en éducation, Paris, A. Colin-Bourrellet ; Liège, Thone, 1970, 3^e éd. revue et augmentée.

Einführung in die pädagogische Forschung, Weinheim, Julius Beltz Verlag, 1969.

Les tests de connaissances, Bruxelles, Editest, 1965.

Rendement de l'enseignement des mathématiques dans douze pays (en collaboration avec T.N. POSTLETHWAITE), Paris, Institut Pédagogique National, 1969.

Comment les maîtres enseignent. Analyse des interactions verbales en classe, (avec la collaboration de E. BAYER), Bruxelles, Ministère de l'Education Nationale, Administration des Etudes, 1969.

H. BENJAMIN, *La pédagogie paléolithique ou Préhistoire de la Contestation*. Préface et adaptation française, Collection « Education 2000 », Paris, F. Nathan ; Bruxelles, Labor, 1970.

Gilbert
Professeur

EVALUATION ET

PRECIS D

Editions LABOR
Rue Royale, 342
1030 BRUXELLES

EUR

Paris, A. Colin-Bourrellet ; Liège,

Weinheim, Julius Beltz Verlag,

t, 1965.

ques dans douze pays (en col-
Paris, Institut Pédagogique Na-

interactions verbales en classe,
xelles, Ministère de l'Education

Préhistoire de la Contestation.
n « Education 2000 », Paris, F.

Gilbert DE LANDSHEERE
Professeur à l'Université de Liège

EVALUATION CONTINUE ET EXAMENS

PRECIS DE DOCIMOLOGIE

Editions LABOR
Rue Royale, 342
1030 BRUXELLES

Fernand NATHAN
Editeur
PARIS

INTRODUCTION.

De nouveaux modes d'évaluation des élèves entrent dans nos mœurs 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.

Ce changement déborde largement le domaine de la notation et coïncide avec des transformations profondes de l'éducation. L'événement n'est pas fortuit. La civilisation contemporaine, l'économie de notre société réclament un homme pourvu de qualités et de connaissances nouvelles : en pareil cas, les innovations pédagogiques s'imposent irrésistiblement.

La contestation des examens traditionnels a toutefois créé un malentendu grave. Il ressemble de façon frappante à celui qui a surgi, entre 1920 et 1940 surtout, lorsque l'adoption des idées, mal digérées, de la pédagogie fonctionnelle des Claparède, Dewey, et autres « progressistes » conduisit au culte de l'improvisation, au mépris de la discipline rigoureuse, au centrage sur des intérêts n'existant que dans l'esprit des théoriciens...

Avec l'avènement de J. Dewey, l'école aurait pu prendre comme devise : « L'effort est mort, vive l'effort ! ». De même, serions-nous tenté de proposer aujourd'hui : « Les examens meurent, les examens sont morts, vivent les examens ! ».

Assurément, les vieux examens qui empoisonnaient l'atmosphère et la matière de cycles entiers d'études et décidaient de la carrière scolaire ou professionnelle en quelques heures, — voire en quelques minutes ! — doivent disparaître. Assurément aussi, la majeure partie des travaux docimologiques publiés jusqu'ici n'ont été que mises au point de systèmes

perfectionnés pour continuer à mal faire les choses. Car ils ne s'appuyaient pas sur une remise en cause de tout l'enseignement.

Mais si, par contre, on ménage, dans chaque activité d'enseignement de base, la place qui revient à l'évaluation, au *feedback* comme disent les psychologues d'aujourd'hui, les examens existent alors, de façon *quasi* permanente dans la vie scolaire. Tantôt ils consistent en brèves évaluations, tantôt ils prennent la forme d'épreuves plus longues, couvrant parfois des quantités considérables de matières.

L'essentiel est qu'ils ne s'insèrent pas, en corps étrangers, dans le processus d'éducation, mais qu'ils en fassent partie intégrante. Pour cette même raison d'ailleurs, ils ne se cantonnent donc pas étroitement au domaine de la connaissance, mais visent à saisir l'ensemble de la personnalité.

Toute réforme profonde s'accompagne, presque nécessairement, d'une période d'adaptation, de tâtonnements, d'errements. Ainsi s'explique le désarroi de beaucoup de maîtres qui essaient actuellement le système d'évaluation continue.

Ainsi s'explique aussi qu'un tabou frappe soudain le mot examen et, qu'au moment où nous écrivons ces lignes, on délivre çà et là des certificats de fin d'études sur la base d'évaluations ponctuelles, qualitativement et quantitativement insuffisantes, mal intégrées de surcroît, et sur la foi d'impressions personnelles à côté desquelles les vieux examens napoléoniens semblent des monuments d'objectivité ! Mais il ne s'agit, nous en sommes convaincu, que d'une crise d'adaptation.

Il n'est donc nullement paradoxal qu'après avoir recommencé — brièvement — le procès des examens et nous être déclaré ardent partisan de l'observation continue et du bilan de fin de cycle, nous consacrons le présent ouvrage... à l'organisation des examens. Mais le lecteur sait maintenant que, si le mot est resté le même, son sens a profondément changé.

*
**

Beaucoup d'enseignants se méfient encore des tests et des autres instruments qui ambitionnent de quantifier l'humain. Non sans raison d'ailleurs, car ces outils ont fait, dans les dernières décennies, leur maladie de jeunesse. Et comme leurs utilisateurs n'ont pas toujours su compenser l'imperfection des techniques par une grande modération de jugement et une intelligence profonde des situations d'ensemble, bien des erreurs ont été commises.

Mais si la prudence et le sens clinique ne p... grès considérables ont maintenant mesurer bien satisfaisante.

Aujourd'hui, d'aucuns non plus parce que tout qu'ils n'en comprennent ce qu'on ignore, est un ré... mathématique des métho... rimentale, souvent indige... tion.

Le temps de l'oppos... ciens de l'éducation est r... leurs chiffres ; les seco... leurs procédures ; tous... efforts pour le plus gran... nauté.

Que ceux qui, en... ques chiffres se rassuren... seront un bagage suffisant

Comme d'habitude, sont plus difficiles à ex...

Certes, tous les sec... ne seront pas livrés, ma... rencontrer ont une valeu... en outre, un indéniable p... ultérieurement la lecture... cialisés.

aire les choses. Car ils ne cause de tout l'enseigne-

ans chaque activité d'en- revient à l'évaluation, au es d'aujourd'hui, les exa- permanente dans la vie ves évaluations, tantôt ils gues, couvrant parfois des

pas, en corps étrangers, qu'ils en fassent partie ailleurs, ils ne se canton- ine de la connaissance, a personnalité.

agne, presque nécessai- de tâtonnements, d'erre- beaucoup de maîtres qui valuation continue.

u frappe soudain le mot ivons ces lignes, on déli- des sur la base d'évalua- antitativement insuffisan- foi d'impressions person- mens napoléoniens sem- ais il ne s'agit, nous en d'adaptation.

l qu'après avoir recom- s examens et nous être n continue et du bilan de sent ouvrage... à l'orga- sait maintenant que, si a profondément changé.

ent encore des tests et nt de quantifier l'humain. outils ont fait, dans les eunesse. Et comme leurs enser l'imperfection des de jugement et une intel- emble, bien des erreurs

Mais si la prudence reste nécessaire, si l'esprit critique et le sens clinique ne perdront jamais leurs droits, des progrès considérables ont néanmoins été accomplis : on peut maintenant mesurer bien des comportements humains de façon satisfaisante.

Aujourd'hui, d'aucuns pourraient rejeter la mesure objective, non plus parce que toute validité lui fait défaut, mais parce qu'ils n'en comprennent pas l'économie. Dénier toute valeur à ce qu'on ignore, est un réflexe de défense bien connu. L'aspect mathématique des méthodes à adopter et de la littérature expérimentale, souvent indigeste, vient encore compliquer la situation.

Le temps de l'opposition entre psychométriciens et praticiens de l'éducation est révolu. Les premiers doivent humaniser leurs chiffres ; les seconds, introduire plus de rigueur dans leurs procédures ; tous deux doivent unir et harmoniser leurs efforts pour le plus grand profit de l'étudiant et de la communauté.

*

**

Que ceux qui, en feuilletant ce livre, ont aperçu quelques chiffres se rassurent : les quatre opérations arithmétiques seront un bagage suffisant !

Comme d'habitude, les procédés dont il va être question sont plus difficiles à expliquer qu'à appliquer.

Certes, tous les secrets de la mesure et de la statistique ne seront pas livrés, mais les notions simples que nous allons rencontrer ont une valeur pratique éprouvée. Elles détiennent, en outre, un indéniable pouvoir de démythification et faciliteront ultérieurement la lecture de travaux docimologiques plus spécialisés.

PREMIERE PARTIE

DEFINITIONS

I. DOCIMOLOGIE, DOCIMASTIQUE ET DOXOLOGIE.

La *docimologie* est une science qui a pour objet l'étude systématique des examens, en particulier des systèmes de notation, et du comportement des examinateurs et des examinés.

La *docimastique* est la technique des examens.

Au début, la docimologie a revêtu un caractère négatif en critiquant les modes de notation et en montrant expérimentalement le manque de fidélité et de validité des examens.

Par la suite, elle est entrée dans une phase constructive en essayant de proposer des méthodes et des techniques de mesure plus objectives ou, au moins, plus rigoureuses, et en mettant au point les moyens de rendre les notes comparables, de façon à assurer plus de justice scolaire.

La *doxologie* est l'étude systématique du rôle que l'évaluation joue dans l'éducation scolaire.

J. Guillaumin¹ lui assigne, notamment, les objectifs suivants :

- Etude des effets inhibiteurs ou stimulants des différentes formes d'examens ;
- Etude des réactions émotionnelles des élèves et, de là, des réactions intellectuelles aux jugements du maître ;
- Etude des effets de l'opinion du maître concernant les élèves sur son enseignement et sur l'apprentissage scolaire ;
- Etude des processus mis en jeu et des effets obtenus par l'automatisation, par l'internotation, par la notation d'équipe, par l'absence de notation.

¹ J. GUILLAUMIN, L'aspect interpersonnel de la notation : de la docimologie à la doxologie pédagogique, in *Bulletin de la Société A. Binet et T. Simon*, 86, 1968, 250-275.

II. EXAMENS ET CONCOURS. OBSERVATION ET EVALUATION CONTINUES.

L'examen et le concours sont difficilement dissociables de l'idée d'épreuve, devenue d'ailleurs leur synonyme après avoir désigné plus généralement la souffrance, le malheur, le danger qui révèlent le courage et la résistance.

Dans l'examen, l'admission est déterminée par une note que le candidat doit atteindre ou dépasser, tandis que dans le concours le nombre de places proposées est fixé d'avance. La présence de la *menace*, du *danger*, celui du refus, est, dans les deux cas, indéniable. Pour peu que la procédure soit entachée d'imperfections graves, nous ne sommes pas très éloignés de l'ordalie.

Dans la notion de concours et d'examen, nous percevons une *charge agressive* totalement absente des concepts d'observation et d'évaluation continues. La sérénité, la bienveillance, l'indulgence aussi, imprègnent le maître qui suit, avec une sympathie d'où toute sévérité n'est pas exclue, le long cheminement de ses élèves vers l'équilibre du moment et la pleine accession à l'état adulte, dans l'avenir.

Les examens marquent les fins d'étapes ; les concours ouvrent les portes aux élus : ce sont des points dans la durée, des événements dans le processus de l'éducation.

L'apprentissage qui précède ces événements consiste fondamentalement en une succession continue de comportements et de *feedbacks*, c'est-à-dire d'informations éclairant sur leur validité, leur pertinence. Il ne nous appartient pas d'engager ici une longue discussion des modalités et des effets du renforcement des conduites. L'important, c'est que, sans lui, l'apprentissage ne semble pas pouvoir se produire.

L'évaluation, au sens restreint que nous lui donnons dans le présent ouvrage, *mérite donc une place importante dans l'enseignement, dont elle fait partie intégrante. Elle a toujours, directement ou indirectement, rapport avec le progrès, en extension ou en qualité, de l'apprentissage.*

L'évaluation joue trois rôles :

1° Un rôle pronostique : l'élève est-il pourvu des qualités intellectuelles et caractérielles, et des connaissances nécessaires pour aborder une matière nouvelle ou un cycle d'études supérieures ? Est-il là où il doit se trouver ? Répondre à ces questions équivaut à prédire le succès dans l'étape qui va commencer.

2° Un rôle de jaugeage :

- a) Contrôle des acquis
- b) Evaluation du progrès même ;
- c) Situation de l'élève
 - dans sa classe ou
 - dans l'ensemble école ?
 - dans des ensembles vince, pays ?

Il ne s'agit pas nécessairement d'un concours, mais d'une position relative.

3° Un rôle diagnostique :

Pourquoi un apprentissage ? Quelles matières ou techniques, finalement, quels sont les instruments nécessaires ?

Les instruments nécessaires au concours ne diffèrent pas de ceux que nous utiliserons pour les examens. Aussi, pour les examens, nous utiliserons les mêmes instruments que pour les concours : examen proprement dit, procédure pédagogique adaptée, de connaissances pour obtenir un diplôme.

Examens internes et externes

Au sens strict, l'examen est truit et noté par le maître. Les élèves qui ont reçu cet examen sont ceux de la classe ou de l'école.

Au sens plus large, on entend par examens internes des examens indépendamment de la classe, de la section ou de la section.

Par examens externes, on entend les examens et notées par des jurys locaux, régionaux ou nationaux. Les examens nationaux sont ceux du baccalauréat, les examens cantonaux, les examens cantonaux, Angleterre, jusqu'à ces derniers, à l'entrée dans l'enseignement.

CONCOURS. ÉVALUATION CONTINUES.

facilement dissociables de leur synonyme après avoir vaincu, le malheur, le danger...

déterminée par une note à dépasser, tandis que dans l'évaluation continue, celle-ci est fixée d'avance. Le succès, celui du refus, est, dans l'évaluation continue, la procédure soit entaillée, soit pas très éloignée

l'examen, nous percevons l'absence des concepts d'observation, de sérénité, la bienveillance, le maître qui suit, avec une attention exclue, le long chemin du moment et la pleine

d'étapes ; les concours comptent des points dans la durée, de l'éducation.

l'événement consiste fondamentalement de comportements d'observations éclairant sur leur appartenance pas d'engager des effets et des effets du renforcement que, sans lui, l'apprentissage conduit.

nous lui donnons dans une place importante dans l'évaluation. Elle a toujours, avec le progrès, en exten-

pourvu des qualités intellectuelles et des connaissances nécessaires à l'entrée dans un cycle d'études. Répondre à ce succès dans l'étape qui

2° Un rôle de jaugeage :

- a) Contrôle des acquisitions ;
- b) Évaluation du progrès, cas où l'on compare l'élève à lui-même ;
- c) Situation de l'élève à un moment donné :
 - dans sa classe ou son groupe de travail ?
 - dans l'ensemble des classes parallèles d'une même école ?
 - dans des ensembles plus vastes : ville, canton, province, pays ?

Il ne s'agit pas nécessairement de procéder à un examen ou à un concours, mais de faire le point, de déterminer la position relative.

3° Un rôle diagnostique :

Pourquoi un apprentissage parfait ne s'est-il pas produit ? Quelles matières ou techniques l'étudiant domine-t-il insuffisamment, quels sont les processus mentaux en cause ?

Les instruments nécessaires à l'évaluation, à l'examen ou au concours ne diffèrent pas toujours, mais bien la façon de les utiliser. Aussi, pour éviter des précisions verbales fastidieuses emploierons-nous, dans les pages qui suivent, le mot examen dans deux sens différents, que le contexte éclairera toujours : examen proprement dit et, plus généralement, toute procédure pédagogique ayant une mesure d'apprentissages ou de connaissances pour objet.

Examens internes et examens externes.

Au sens strict, l'*examen interne* dans une branche est construit et noté par le maître qui l'a enseignée, et subi par les élèves qui ont reçu cet enseignement, dans le cadre de la classe ou de l'école.

Au sens plus large, on qualifie d'internes les examens organisés indépendamment dans chaque école, qu'il existe ou non une coordination ou une unification par branche et par niveaux et sections.

Par *examens externes*, on désigne les épreuves organisées et notées par des jurys indépendants des écoles, à l'échelon local, régional ou national. Les plus connues de ces épreuves sont celles du baccalauréat français ; citons encore, en Belgique, les examens cantonaux, en fin d'études primaires, et, en Angleterre, jusqu'à ces derniers temps, l'*Eleven + Examination* à l'entrée dans l'enseignement secondaire.

III. NOTES ET SCORES.

La distinction entre notes et scores rendrait, pensons-nous, un grand service en éducation.

Dans une dictée, l'élève peut commettre un certain nombre de fautes dont il ne nous appartient guère d'évaluer l'existence : elles sont ou ne sont pas. Toutefois, le nombre de fautes relevées n'a, en soi, aucune signification éducative : faire cinq fautes dans la dictée de Mérimée témoigne d'une extraordinaire connaissance des arcanes de l'orthographe ; en d'autres conditions, le même résultat annonce une faiblesse grave. Une information relativiste est donc aussi nécessaire.

Par *score*, nous désignerons les résultats objectifs obtenus à un test ou à toute autre forme d'évaluation par *compte ou décompte de points, selon des règles fixes* : nombre de fautes en dictées, résultats à un test standardisé.

Par *note*, nous entendrons une appréciation synthétique traduisant l'évaluation d'une performance dans le domaine de l'éducation.

La note peut être objective ou subjective, mais elle est toujours relative. Attribuer la note A à un élève dont la performance se situe à tel niveau dans un étalonnage national relève de la première catégorie ; marquer sa composition d'un *bien* relève de la seconde.

SCORES.

ores rendrait, pensons-nous,

omettre un certain nombre
guère d'évaluer l'existence :
s, le nombre de fautes rele-
ation éducative : faire cinq
émoigne d'une extraordinaire
ographe ; en d'autres condi-
e faiblesse grave. Une infor-
cessaire.

es résultats objectifs obtenus
d'évaluation par *compte ou*
les fixes : nombre de fautes
ardisé.

appréciation synthétique tra-
nce dans le domaine de l'édu-

subjective, mais elle est tou-
un élève dont la performance
onnage national relève de la
omposition d'un *bien* relève

DEUXIEME PARTIE

L'ACCUSATION ET LA DEFENSE

CHAPITRE 1.

CRITIQUE DES EXAMENS.

Une critique détaillée des examens traditionnels a été faite par H. Piéron dans son excellent ouvrage *Examens et docimologie*¹. Tout enseignant doit l'avoir lu.

Le présent chapitre contient deux types de remarques. Les unes sont de brefs rappels d'imperfections bien connues, inlassablement mises en lumière par les docimologistes ; les autres concernent des phénomènes moins bien étudiés : stéréotypie, effet de halo, effet oedipien de la prédiction,... Tantôt, seul l'examen est en cause ; tantôt, c'est toute la pédagogie dans laquelle il s'insère.

1. Corps étrangers dans l'éducation, au service d'une pédagogie dépassée.

Qu'il s'agisse d'interrogations périodiques ou d'examens trimestriels ou annuels, l'évaluation se réduit souvent à un contrôle de rétention de connaissances laissant inexplorés, non seulement les aspects les plus importants de l'intelligence, mais aussi à peu près tous les traits de personnalité qu'une éducation bien comprise doit cultiver.

Nous nous trouvons, en fait, devant les séquelles d'un système pédagogique où, comme T. Brameld aime à le dire, les leçons servent de bandes transporteuses de connaissances et de valeurs, sélectionnées en fonction d'un rôle prédéterminé à jouer dans une société non démocratique.

Dans ce cadre, l'examen constitue une sorte de contrôle de fabrication, de vérification de la conformité au moule, au gabarit, bref aux spécifications imposées par l'autorité.

L'épanouissement de la personne humaine n'occupe guère de place dans ces préoccupations, car l'éducation n'est pas conçue pour elle, mais pour le service d'un régime.

¹ Paris, P.U.F., 1963.

Les conséquences malheureuses de cette situation ont été analysées mille fois. Au lieu de servir l'élève, de l'informer fonctionnellement de la valeur de ses comportements en cours d'apprentissage, de faire naître une adaptation meilleure, d'apporter donc une évaluation acceptée au point que l'étudiant y participe sincèrement et spontanément, l'examen est aussi mal accueilli qu'une déclaration à l'impôt qui, elle non plus, n'est certainement pas génératrice de civisme.

En outre, le rejet profond de la procédure combinant son effet avec la pauvreté intellectuelle des questions posées conduit à un résultat aisément imaginable : dans les quinze mois qui suivent un examen, un oubli de 80 % des matières factuelles n'est pas exceptionnel.

Ainsi, si l'on excepte la fonction sociale jouée par l'examen, le fiasco est complet : on n'a ni éduqué, ni instruit.

2. Anxiété et stress.

Dans une civilisation où la réussite scolaire conditionne la réussite matérielle et sociale, au point que M. Young a cru devoir dénoncer les dangers de la « méritocratie », l'examen qui décide du passage de classe ou de l'obtention du diplôme est redouté par l'enfant et sa famille. Même en cours d'année, les épreuves de contrôle de connaissances et d'aptitudes sont abordées avec tension et appréhension, ce qui n'est certes pas la condition idéale, et dénature profondément le rôle éducatif de la mesure des apprentissages.

Circonstance aggravante, la tradition veut que, chez nous, les « grands » examens portant sur toutes les branches se déroulent consécutivement en quelques jours. Lors d'une session universitaire, j'eus, une année, le triste privilège d'être interrogé sur quatorze cours en deux jours de suite... Dans l'enseignement secondaire, le même phénomène de concentration se produit souvent aussi en cours d'année, lorsque l'approche de la remise aux parents du bulletin périodique suscite une floraison d'interrogations.

Dans les deux cas, l'accumulation des épreuves et leur préparation, souvent contrariées par l'apport de matières nouvelles jusqu'au dernier moment (des professeurs ne respectent pas toujours les périodes de révision ou ne les exploitent pas de la façon la plus heureuse), fatiguent considérablement le corps et l'esprit.

D'aucuns prétendent que ces conditions difficiles revêtent elles-mêmes une valeur éducative : la vie moderne ne nous épargne guère le stress, et il est bon d'y être préparé. D'accord, à condition que les choses se passent dans la clarté !

Avant d'organiser un
ment son but : s'agit-il
sances ? la résistance a
d'user des connaissances
objectifs différents et ap

3. Inégalité - injustice.

Dans notre système
cun les questions d'exa
cipe est excellent et dar
où, traditionnellement, t
des commissions extéri
chaque enseignant la
les méthodes anglo-sax
les nôtres, car des mes
la comparabilité des ex
le pays. Nous consacro
procédés utilisés pour a

La liberté, à peu pr
à des situations injustes

Les écoles de notr
dont les populations p
fort différentes. Tel pet
daire ne compte que q
par leurs origines soc
établissement situé da
« fortes » peuplées d'une
lectuellement, les adole
une école voisine, qui a

Avec le temps, et l
sévérité ou de « généro
Marion Coulon observai
même type sont succes
la sévérité des examens
cienneté de l'établisseme
raliser, mais qui oserait
de tout fondement ? B
ambitionnent pour leur
se soucier de la valeur
il est plus aisé de chan
ou de modifier l'orientat

Quoi qu'il en soit,
et les examens au nive
avec la conséquence d
tion faible serait parmi
groupe fort.

es de cette situation ont été servir l'élève, de l'informer es comportements en cours une adaptation meilleure, eptée au point que l'étudiant nément, l'examen est aussi l'impôt qui, elle non plus, de civisme.

la procédure combinant son des questions posées con- able : dans les quinze mois 80 % des matières factuelles

h sociale jouée par l'examen, uqué, ni instruit.

essite scolaire conditionne la point que M. Young a cru la « méritocratie », l'examen ou de l'obtention du diplôme lle. Même en cours d'année, issances et d'aptitudes sont sion, ce qui n'est certes pas rofondément le rôle éducatif

adition veut que, chez nous, sur toutes les branches se lques jours. Lors d'une ses- ée, le triste privilège d'être deux jours de suite... Dans me phénomène de concen- cours d'année, lorsque l'ap- u bulletin périodique suscite

lation des épreuves et leur ar l'apport de matières nou- es professeurs ne respectent ion ou ne les exploitent pas ntiguent considérablement le

conditions difficiles revêtent e : la vie moderne ne nous on d'y être préparé. D'accord, assent dans la clarté !

Avant d'organiser un examen, il importe de définir claire- ment son but : s'agit-il de contrôler l'acquisition des connais- sances ? la résistance au *stress* ? ou la capacité de restituer et d'user des connaissances en situation de *stress* ? Ces trois objectifs diffèrent et appellent des épreuves différentes.

3. Inégalité - injustice.

Dans notre système scolaire, les professeurs rédigent cha- cun les questions d'examen destinées à leurs élèves. Le prin- cipe est excellent et dans des pays comme la Grande-Bretagne où, traditionnellement, tous les examens-clés étaient confiés à des commissions extérieures à l'école, on tend à accorder à chaque enseignant la responsabilité du contrôle. Toutefois, les méthodes anglo-saxonnes présentent plus de rigueur que les nôtres, car des mesures précises sont prises pour assurer la comparabilité des examens dans la région, voire dans tout le pays. Nous consacrons d'ailleurs une partie de ce livre aux procédés utilisés pour atteindre ce but.

La liberté, à peu près totale, laissée à nos maîtres conduit à des situations injustes.

Les écoles de notre pays constituent des mondes isolés dont les populations présentent des caractéristiques parfois fort différentes. Tel petit établissement d'enseignement secon- daire ne compte que quelques élèves, en majorité handicapés par leurs origines socio-économiques. A l'opposé, tel autre établissement situé dans une grande ville voit ses sections « fortes » peuplées d'une majorité favorisée socialement et intel- lectuellement, les adolescents en difficulté étant orientés vers une école voisine, qui a la réputation d'être plus compréhensive.

Avec le temps, et la concurrence aidant, des traditions de sévérité ou de « générosité » se créent dans les établissements. Marion Coulon observait récemment que, quand des écoles de même type sont successivement créées dans une même région, la sévérité des examens est inversement proportionnelle à l'an- cienneté de l'établissement. Nous ne pouvons certes pas géné- raliser, mais qui oserait affirmer que la proposition est dénuée de tout fondement ? Beaucoup de parents l'ont compris, qui ambitionnent pour leur enfant un diplôme déterminé, sans trop se soucier de la valeur intrinsèque des études : en cas d'échec, il est plus aisé de changer d'école que d'installer un traitement ou de modifier l'orientation.

Quoi qu'il en soit, les maîtres adaptent leur enseignement et les examens au niveau de leur classe (ce qui est louable), avec la conséquence qu'un élève bien noté dans une popula- tion faible serait parmi les très faibles ou échouerait dans un groupe fort.

L'effet cumulatif de tous les ajustements circonstanciels aboutit parfois à des situations scandaleuses. Il y a quelques années, nous avons pu consulter des statistiques établissant que, pour un même type et un même niveau d'enseignement secondaire, les chances de réussite d'un élève variaient de un à huit selon l'école fréquentée. Le cas est probablement extrême, encore manquons-nous d'études précises permettant de savoir ce qui se passe réellement.

Il faut y insister, l'adaptation du maître au niveau de ses élèves n'est pas, en soi, critiquable ; elle est, au contraire, caractéristique du bon éducateur.

L'injustice apparaît lorsque des notes purement relatives sont utilisées en situation de concurrence extérieure, ou sont l'unique critère pour l'obtention d'un diplôme dont peut dépendre l'avenir des élèves.

De toute façon, on constate des variations considérables dans la *quantité des matières couvertes* et dans la *qualité des réponses exigées*. Ici, la restitution de mémoire d'une partie d'un court syllabus suffit, alors que là, les questions exigent analyse, synthèse et jugement personnel à propos d'une matière abondante.

Tel maître ne pose qu'une ou deux questions, plus choisies pour la facilité de correction qu'elles offrent que pour leur importance réelle, tel autre s'efforce de parcourir tout le cours, au risque de transformer l'examen en un marathon. En sport, on sait que cette longue course doit être réservée à des individus exceptionnels ; pourquoi, dans le domaine de l'instruction, les trente élèves d'une classe seraient-ils tous des coureurs de fond ?

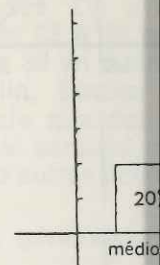
Avant de poursuivre, trois résultats de recherches nous permettront d'illustrer les phénomènes dont il vient d'être question.

R. Gjorgjevski¹ a montré, par une expérience simple, comment, par relativité, des professeurs évaluent différemment une même performance.

Cinq correcteurs, professeurs de la même branche, ont noté indépendamment cent épreuves écrites provenant d'une école secondaire. On a ensuite extrait quinze copies qui toutes avaient obtenu la note « bien ». Elles ont été confiées, pour

¹ Voir N. ROT et Z. BUTAS, Les distributions des notes scolaires comparées aux distributions des résultats obtenus aux tests de connaissances, in *Le travail humain*, XXII, 1-2, 1959.

nouvelle correction, à spontanément adopté de la moitié des cas environ répartition nouvelle¹ :



Poursuivant son ex même série de cent co bonnes, et un groupe d

Dans chaque group Deux fois cinq profes nes » copies glissées de 3 sur 5 à 2,40 ; da passée de 3 à 3,87.

On doit, par ailleur 4 860 élèves répartis da instituteurs ont été inv catégories : très bons, fait passer un même t élèves. Le dépouilleme degré de connaissance étaient jugés très bons

Comme le souligne dire (...) que l'instituteu sonne les points du p chacun de ses élèves. C che à chaque faiblesse dère non plus au sein « faible », mais par ra divergences d'appréci

¹ Le nombre de copies est signification statistique.

² Voir *Le Travail humain*,

ajustements circonstanciels
andaleuses. Il y a quelques
des statistiques établissant
ême niveau d'enseignement
e d'un élève variaient de un à
s est probablement extrême,
écises permettant de savoir

du maître au niveau de ses
ble ; elle est, au contraire,

es notes purement relatives
urrence extérieure, ou sont
un diplôme dont peut dépen-

des variations considérables
vertes et dans la *qualité des*
on de mémoire d'une partie
ue là, les questions exigent
ersonnel à propos d'une ma-

deux questions, plus choisies
elles offrent que pour leur
e de parcourir tout le cours,
n en un marathon. En sport,
oit être réservée à des indi-
s le domaine de l'instruction,
raient-ils tous des coureurs

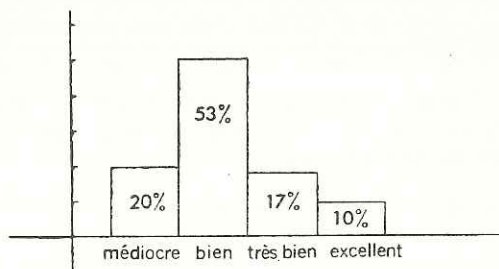
ultats de recherches nous
mènes dont il vient d'être

par une expérience simple,
seurs évaluent différemment

s de la même branche, ont
ves écrites provenant d'une
trait quinze copies qui toutes
Elles ont été confiées, pour

ons des notes scolaires comparées aux
e connaissances, in *Le travail humain*,

nouvelle correction, à quatre autres professeurs. Ceux-ci ont
spontanément adopté des exigences nouvelles : seulement dans
la moitié des cas environ, la note est restée « bien ». Voici la
répartition nouvelle ¹ :



Poursuivant son expérience, Gjorgjevski a extrait, de la
même série de cent copies, un groupe de douze jugées très
bonnes, et un groupe de douze jugées médiocres.

Dans chaque groupe, il a glissé trois copies jugées bonnes.
Deux fois cinq professeurs ont évalué. La moyenne des « bon-
nes » copies glissées dans les « très bonnes » est descendue
de 3 sur 5 à 2,40 ; dans le groupe opposé, la moyenne est
passée de 3 à 3,87.

On doit, par ailleurs, à M. Reuchlin ² une étude portant sur
4 860 élèves répartis dans 397 écoles primaires de France. Les
instituteurs ont été invités à classer leurs élèves en quatre
catégories : très bons, bons, moyens, médiocres. On a ensuite
fait passer un même test de français et de calcul à tous les
élèves. Le dépouillement des résultats a montré que le même
degré de connaissances était le fait d'élèves qui, selon l'école,
étaient jugés très bons, bons, moyens ou médiocres.

Comme le souligne pertinemment M. Reuchlin : « On peut
dire (...) que l'instituteur, certainement, connaît mieux que per-
sonne les points du programme qui sont acquis ou non par
chacun de ses élèves. Ce qu'il ignore, c'est la gravité qui s'atta-
che à chaque faiblesse, à chaque lacune, lorsqu'on la consi-
dère non plus au sein d'une classe qui peut être « forte » ou
« faible », mais par rapport à l'ensemble du pays. De là, les
divergences d'appréciation mises en lumière par l'enquête. »

¹ Le nombre de copies est trop peu élevé pour conduire à des résultats ayant une
signification statistique.

² Voir *Le Travail humain*, XXII, 1-2, 1959, p. 12 sq.

Les graphiques ci-dessous synthétisent les résultats.

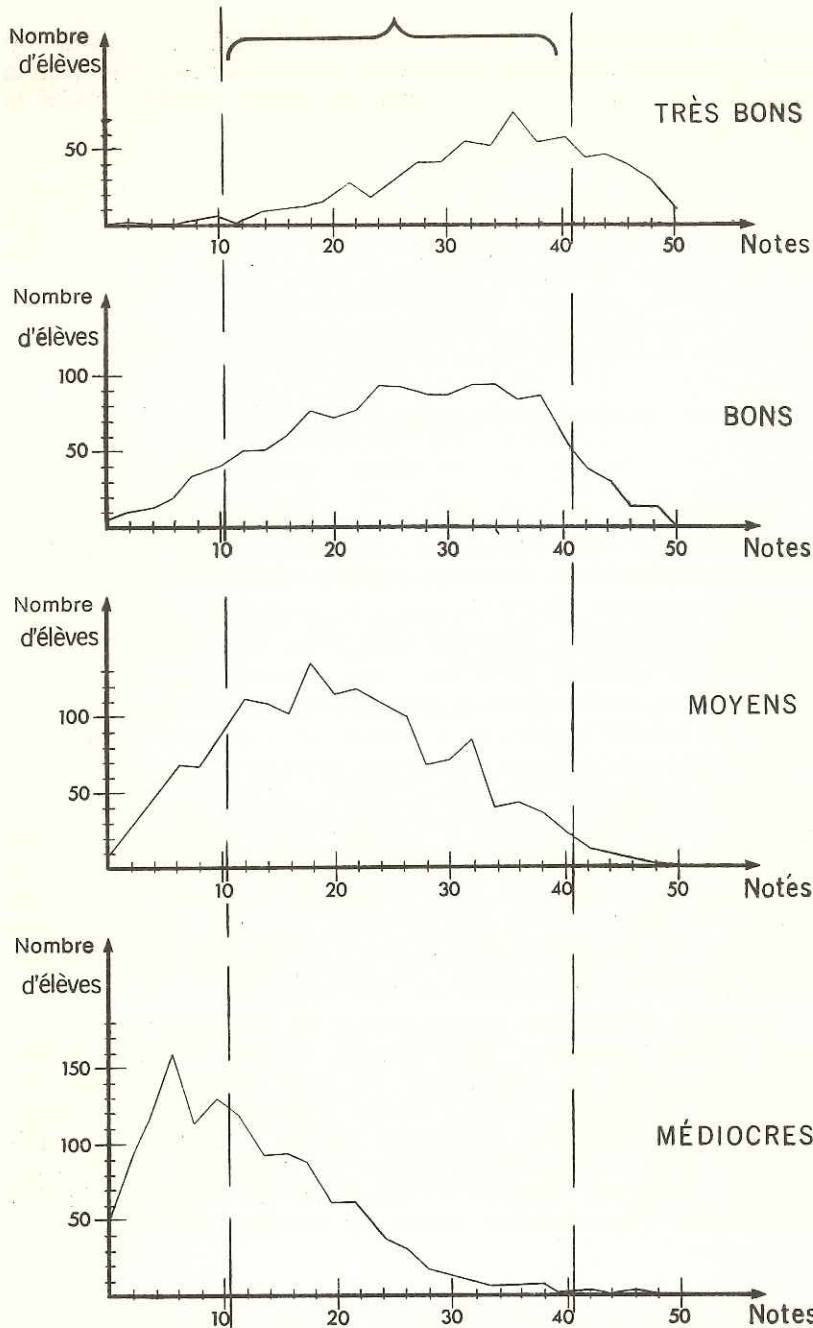


Fig. 1. — Les quatre graphiques permettent de comparer la distribution des notes obtenues, dans l'épreuve de calcul, par 654 élèves jugés « très bons » par leurs maîtres, par 1 303 élèves « bons », 1 551 « moyens » et 1 300 élèves jugés « médiocres ». On voit que les distributions de notes se recouvrent largement : dans la zone de notes qui va de 10 à 40, le même niveau de connaissances peut être qualifié, selon le cas, de très bon, moyen ou médiocre.

Voici encore deux

« Dans une enquête représentatif d'élèves de (daire), on a fait passer un res et une épreuve de con de 406 classes. On a co s'étaient de 23 à 60 en 80) et de 7 à 38 en mathé 44) (Reuchlin, Bacher (1964) signale que des é de leur école seraient cla quentaient d'autres écoles

4. L'échec, générateur d

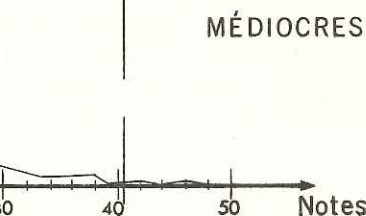
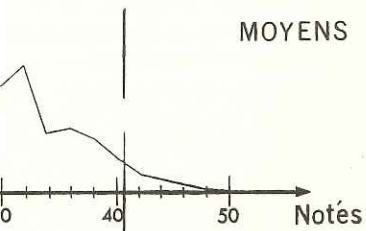
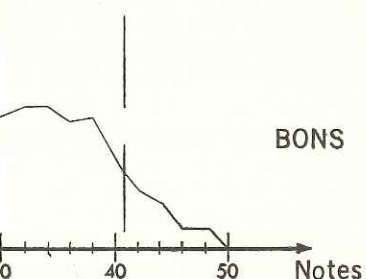
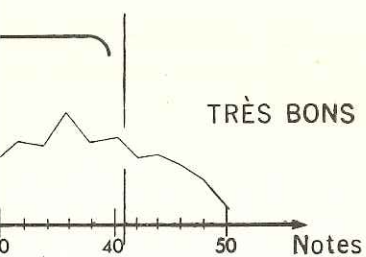
On peut objecter qu chargeront de rétablir la raient pas respectée. Les aussi simples. D'abord pa ragés injustement de l'e exemple, n'auront peut-é chance à l'université. En de gagner sur tous les tab

Les recherches actue gique : le succès engend prépare un échec demain valorisent un capital inte « à bon marché » a par de ne pas pousser les p où il se trouve et de pe sorte de nouveau souffle. daire ne voient-ils pas le ter spectaculairement dè gique de l'adolescence p

R. Rosenthal et L. J étude d'ensemble sur l'a tion, ainsi qualifié parce n'avait pas prédit qu'il éloigné de sa famille. C tué. En d'autres mots, l avait été prédite.

1 F. BACHER, La normalisatio spécial de la revue Les Sciences
 2 *Pygmalion in the Classroom*, ment, New York, Holt, Rinehart and

ynthésisent les résultats.



nt de comparer la distribution des notes
ves jugés « très bons » par leurs maîtres,
1 300 élèves jugés « médiocres ». On voit
largement : dans la zone de notes qui va
ut être qualifié, selon le cas, de très bon,

Voici encore deux exemples donnés par F. Bacher¹ :

« Dans une enquête française portant sur un échantillon représentatif d'élèves de troisième (fin du premier cycle secondaire), on a fait passer une épreuve de connaissances littéraires et une épreuve de connaissances mathématiques aux élèves de 406 classes. On a constaté que les moyennes de classe s'étalaient de 23 à 60 en français (pour une épreuve notée sur 80) et de 7 à 38 en mathématiques (pour une épreuve notée sur 44) (Reuchlin, Bacher 1968). (...) Aux Etats-Unis, Flanagan (1964) signale que des élèves classés dans le quart inférieur de leur école seraient classés dans le quart supérieur s'ils fréquentaient d'autres écoles de la même région. »

4. L'échec, générateur d'échecs.

On peut objecter que les études supérieures ou la vie se chargeront de rétablir la justice que les notes scolaires n'auraient pas respectée. Les choses ne sont malheureusement pas aussi simples. D'abord parce que les élèves écartés ou découragés injustement de l'enseignement secondaire général, par exemple, n'auront peut-être plus l'occasion de défendre leur chance à l'université. Ensuite parce que les favorisés risquent de gagner sur tous les tableaux.

Les recherches actuelles confirment un vieil adage pédagogique : le succès engendre le succès et l'échec d'aujourd'hui prépare un échec demain. Portés par la réussite, certains élèves valorisent un capital intellectuel limité. Le passage de classe « à bon marché » a parfois la double conséquence salutaire de ne pas pousser les parents à retirer leur enfant de l'école où il se trouve et de permettre à l'étudiant de retrouver une sorte de nouveau souffle. Combien de fois les maîtres du secondaire ne voient-ils pas le rendement scolaire d'un élève remonter spectaculairement dès que la crise physique et psychologique de l'adolescence perd de son acuité ?

R. Rosenthal et L. Jacobson² ont récemment publié une étude d'ensemble sur l'aspect de l'*effet œdipien de la prédiction*, ainsi qualifié parce que si, à la naissance d'Œdipe, l'oracle n'avait pas prédit qu'il tuerait son père, il n'aurait pas été éloigné de sa famille. Connaissant son père, il ne l'aurait pas tué. En d'autres mots, la tragédie se produisit *parce qu'elle avait été prédite*.

¹ F. BACHER, La normalisation de la notation, in *Docimologie et Education*, numéro spécial de la revue *Les Sciences de l'Education*, 2-3, 1969, pp. 51-52.

² *Pygmalion in the Classroom, Teacher Expectations and Pupil's Intellectual Development*, New York, Holt, Rinehart and Winston, 1968.

Il semble que, dans une mesure certes difficile à apprécier, un élève se comporte en fonction du jugement que le professeur porte sur lui. Même les animaux en dressage n'échappent pas à cette règle. Rosenthal rappelle l'expérience suivante : on constitue deux groupes de souris blanches génétiquement identiques. Au moment où on les remet aux étudiants chargés de les dresser, une remarque indique que le premier groupe est composé d'animaux particulièrement bien doués, alors que le second est de pauvre qualité. Les résultats du dressage confirmeront ce pronostic... fantaisiste.

L'expérience suivante, conduite par Rosenthal lui-même, s'inscrit dans le même contexte.

On a annoncé à des instituteurs des six années primaires de l'Oak School que d'éminents chercheurs venaient de mettre au point un test d'« épanouissement » (en réalité, un simple test d'intelligence, ¹ peu connu et sans vertu particulière, a été utilisé).

Tous les élèves de l'école y ont été soumis, et on a signalé ceux qui étaient sur le point de s'épanouir intellectuellement, pronostic fantaisiste.

Le quotient intellectuel des élèves ainsi désignés s'est élevé d'une façon significative dans les trois années inférieures ; en outre, des progrès supérieurs à la moyenne ont été observés en lecture et en arithmétique ².

5. Rupture entre enseignement et examen.

L'examen doit être le reflet de l'enseignement donné. Imaginons qu'au lieu de conduire ses élèves à la piscine, un professeur d'éducation physique ait passé son année à disserter sur la natation. Il serait évidemment inadmissible que les notes d'examen soient attribuées d'après les performances natatoires... car seuls réussiraient ceux qui ont appris à nager en dehors du cours.

Pareil agissement paraît impensable. Pourtant, combien de fois n'assiste-t-on pas à une aberration similaire. Le professeur fait un cours où il impose les informations, les opinions et les jugements, puis, à l'examen, il pose des questions dites d'intel-

¹ Mesuré à l'aide du Toga de Flanagan (1960).

² Les chiffres avancés par Rosenthal et Jacobsen sont contestés par R.L. Thorndike. Par contre, la tendance à l'augmentation des QI. et des scores semble reconnue par tous les spécialistes. C'est évidemment ce qui importe. Pour plus de détails, voir *American Educational Research Journal*, 5, 4, 1968, 708.

ligence, entendons qui n'ont précisément pas été quoi les étudiants seraient seuls, à l'examen, la solution, s'ils n'ont pas pu

Ainsi s'expliquent bien. Une enquête suscitée insuffisantes chez un professeur avaient obtenu plus de (d'année) fit rapidement que deux ou trois interrogations questions d'examen, deux, fert, démarches qui, au plaçait, nécessitaient une mènes et un entraînement tion de problèmes originaux

Une discussion amicale révéla qu'il se croyait ob points du programme, qu sité de l'entraînement syst lisé à la solution de probl il ne possédait pas les nécessaires à la conduite

6. Désaccord entre cor

Tous les docimologi divergences entre maître d'élève ou un même ense tant une relativité plus même professeur jugean de temps (fidélité) peuv

Comme tous les d'exemples, nous n'en rappel.

a) La composition franç

Elle est une accusée

- Piéron rapporte qu'une recteurs, tous profess aux résultats suivants

¹ PIERON, o.c., p. 123.

est certes difficile à apprécier, du jugement que le professeur en dressage n'échappent appelle l'expérience suivante : sur des souris blanches génétiquement sélectionnées et remises aux étudiants chargés de dresser, on a remarqué que le premier groupe, généralement bien doués, alors que le second, moins doué. Les résultats du dressage ont été les suivants :

rapporté par Rosenthal lui-même,

sur des six années primaires, les chercheurs venaient de mettre en évidence un « effet de nouveauté » (en réalité, un simple effet de nouveauté sans vertu particulière, a été

qui ont été soumis, et on a signalé que les élèves s'épanouissent intellectuellement,

les élèves ainsi désignés s'estimant dans les trois années inférieures à la moyenne ont obtenu de meilleurs résultats que les autres.

à l'examen.

de l'enseignement donné. Imaginons des élèves à la piscine, un professeur qui a passé son année à dissenter sur l'importance de la natation, il est inadmissible que les notes obtenues par les élèves ne reflètent pas les performances natatoires qui ont appris à nager en

raisonnable. Pourtant, combien de fois avons-nous vu des notes de notation similaires. Le professeur ignore les opinions et les résultats des questions dites d'intel-

les observations sont contestées par R.L. Thorndike. L'absence de reconnaissance des scores semble reconnue par Thorndike. Pour plus de détails, voir *American*

ligence, entendons qui mettent en œuvre des capacités qui n'ont précisément pas été installées pendant les leçons. Pourquoi les étudiants seraient-ils soudain capables de découvrir seuls, à l'examen, la solution d'applications géométriques originales, s'ils n'ont pas pu le faire pendant l'année ?

Ainsi s'expliquent bien des pourcentages d'échecs anormaux. Une enquête suscitée à cause de nombreuses notes insuffisantes chez un professeur de chimie (4 élèves sur 22 avaient obtenu plus de la moitié des points à l'examen de fin d'année) fit rapidement apparaître que le professeur ne faisait que deux ou trois interrogations écrites par an. Sur les trois questions d'examen, deux impliquaient interprétation et transfert, démarches qui, au niveau de difficulté où le maître se plaçait, nécessitaient une compréhension profonde des phénomènes et un entraînement, long et finement contrôlé, à la solution de problèmes originaux.

Une discussion amicale avec le jeune professeur en question révéla qu'il se croyait obligé de traiter en détail de tous les points du programme, qu'il n'avait pas conscience de la nécessité de l'entraînement systématique et au moins semi-individualisé à la solution de problèmes originaux et que, de toute façon, il ne possédait pas les notions psychopédagogiques de base nécessaires à la conduite d'un tel entraînement.

6. Désaccord entre correcteurs.

Tous les docimologistes citent des exemples de grandes divergences entre maîtres appelés à évaluer un même travail d'élève ou un même ensemble de travaux, ce qui permet pourtant une relativité plus sûre. Les fluctuations des notes d'un même professeur jugeant un même travail à quelque intervalle de temps (fidélité) peuvent aussi être importantes.

Comme tous les ouvrages de docimologie fourmillent d'exemples, nous n'en proposons que quelques-uns, pour rappel.

a) La composition française.

Elle est une accusée de prédilection.

- Piéron rapporte qu'une même composition, jugée par 76 correcteurs, tous professeurs de langue maternelle, conduisit aux résultats suivants ¹ :

¹ PIERON, o.c., p. 123.

Note sur 20	Nombre de correcteurs qui l'ont attribuée
de 0 à 1	1
2 - 3	6
4 - 5	20
6 - 7	34
8 - 9	10
10 - 11	3
12 - 13	2

- Le Service de la Recherche pédagogique du C.R.D.P. de Lyon a invité, en 1967¹, 150 professeurs de français de l'Académie de Lyon à corriger trois compositions françaises traitant d'un même sujet.

Max. 20	Moyenne	Marge de variation
Devoir I	10 1/4	4 1/2 - 13 3/4
Devoir II	6,5	2 1/2 - 12 1/2
Devoir III	11,6	5 1/2 - 17 1/2

- L'intervention d'hommes cultivés, étrangers aux routines de l'enseignement améliorerait-elle la situation ? P. Diederich et ses collaborateurs² ont comparé la façon de corriger de 10 professeurs de langue maternelle et de 43 autres notateurs : spécialistes des sciences sociales, des sciences naturelles, écrivains, éditeurs, avocats et hommes d'affaires (cadres supérieurs).

Tous ont été invités, sans autre précision, à classer 300 compositions en 9 tas, par ordre de mérite général. Aucun des

1 C.R.D.P., *Docimologie et Examens*, Lyon, I.P.N., 1969.
 2 P. DIEDERICH et al., *Factors in Judgments of Writing Ability*, Princeton, E.T.S., Res. Bulletin RB-61-15, 1961.

tas ne pouvait compter l'ensemble se sont vu a possibles. Aucune comp sements différents...

b) *Mathématiques.*

Le C.R.D.P. de Lyon de mathématiques exerça les copies de 3 élèves a

Max. 20
Elève I
Elève II
Elève III

Si, pour l'élève I, par extrêmes, il reste autant 8/20, alors qu'on pouvait

c) *Médecine.*

Au niveau universitaire

Un même candidat d septembre 1965, une épre simples et précises à trait

Les réponses (anony ment par cinq correcteurs identiques.

Résultat total sur 100	M
Copie de juin	
Copie de septembre	

1 O.c.
 2 Voir *Le Monde* du 6 septembre
 3 Remarquons que les trois ad

Nombre de correcteurs
qui l'ont attribuée

.	1
.	6
.	20
.	34
.	10
.	3
.	2

gogique du C.R.D.P. de Lyon
rs de français de l'Académie
tions françaises traitant d'un

Marge de variation

$4\frac{1}{2} - 13\frac{3}{4}$
$2\frac{1}{2} - 12\frac{1}{2}$
$5\frac{1}{2} - 17\frac{1}{2}$

étrangers aux routines de
la situation ? P. Diederich et
é la façon de corriger de 10
e et de 43 autres notateurs :
es, des sciences naturelles,
hommes d'affaires (cadres

précision, à classer 300 com-
e mérite général. Aucun des

.N., 1969.
of Writing Ability, Princeton, E.T.S., Res.

tas ne pouvait compter moins de 12 copies (4 %). 94 % de l'ensemble se sont vu attribuer 7, 8 ou 9 des 9 classements possibles. Aucune composition n'a reçu moins de cinq classements différents...

b) *Mathématiques.*

Le C.R.D.P. de Lyon¹ a aussi demandé à 150 professeurs de mathématiques exerçant dans une classe de 3^e de noter les copies de 3 élèves ayant à résoudre le même problème.

Max. 20	Moyenne	Marge de variation
Elève I	5,70	0,5 - 11,5
Elève II	16	11,5 - 20
Elève III	8	3,5 - 11,5

Si, pour l'élève I, par exemple, on élimine les deux notes extrêmes, il reste autant de correcteurs pour noter 2/20 que 8/20, alors qu'on pouvait attendre une notation objective.

c) *Médecine.*

Au niveau universitaire, les résultats ne sont pas plus sûrs.

Un même candidat de 2^e médecine a subi, en juin et en septembre 1965, une épreuve écrite comportant cinq questions simples et précises à traiter en 1 h 30.²

Les réponses (anonymes) ont été corrigées indépendamment par cinq correcteurs possédant des titres et des qualités identiques.

Résultat total sur 100	Marge de variation	Décision ³	
		Admis	Refusé
Copie de juin	47 - 78	3	2
Copie de septembre	32,5 - 73	3	2

1 O.c.

2 Voir *Le Monde* du 6 septembre 1966.

3 Remarquons que les trois admis furent les mêmes dans les deux cas.

Si on considère les notes attribuées à chacune des 10 questions, on constate :

1) Ecart maximum :	12	(de 3 à 15 sur 20)
2) Ecart pour 4 questions :	11	
3) Plus petit écart maximum entre cinq correcteurs :	7,5	

d) Divers.

A. Agazzi¹ rappelle un exemple, plus frappant encore, concernant un ensemble de branches.

Six correcteurs ont, chaque fois, noté les examens du baccalauréat (note d'échec : moins de 10 sur 20) :

	Refusés par les six correcteurs	Admis par les six correcteurs	Admis par les uns et refusés par les autres
Version latine	40 %	10 %	50 %
Composition française	21 %	9 %	70 %
Anglais	37 %	16 %	47 %
Mathématiques	44 %	20 %	36 %
Philosophie	9 %	10 %	81 %
Physique	37 %	13 %	50 %

e) Combien de correcteurs pour stabiliser la note ?

Calculer la moyenne entre deux correcteurs apporterait-il une amélioration considérable ? Non, si aucune mesure stricte n'a été prise pour accorder les notateurs. Par combien de professeurs différents faudrait-il faire évaluer le même travail pour que la note se stabilise ? Les nombres suivants, avancés par Piéron¹ n'ont certes pas de valeur absolue. Ils donnent néanmoins une idée de l'ordre de grandeur des nombres...

Composition française	78
Version latine	19
Anglais	28
Mathématiques	13
Dissertation philosophique	127
Physique	16

¹ A. AGAZZI, *Les aspects pédagogiques des examens*, Strasbourg, Conseil de l'Europe, 1967, p. 119. La recherche est due à H. Laugier et D. Weinberg et portait sur le baccalauréat français.

L'histoire et la géographie ces deux branches aussi fluctuations.

On pourrait multiplier les notateurs. Mais, à quoi bon ? Facile, comme trop de choses, est d'être conscient du problème et de ses remèdes.

Quelle est l'explication ?

Elle réside principalement dans la diversité de vue, des dimensions, de la note, à être jugée, et à l'absence de consensus (en a pas du tout) donnée.

« Dans une dissertation, la diversité des idées, leur originalité, la maîtrise du style, etc. Si l'on considère une dimension de la note, la note est située dans un espace à n dimensions, dans cet espace, se trouvent les élèves : un élève occupant une position peut être déclassé dans une position moins élevée que la sienne ne peut dire d'un élève qui occupe la même dimension et une position s'il est supérieur ou inférieur à sa position inverse. »

En outre, même s'ils sont différents, les notateurs varient en sévérité, ce qui a pour effet de halo, etc.

7. Infidélité d'un même notateur

Un même correcteur peut noter de manière différente A des phénomènes aussi complexes que la physique et mentale et l'évaluation du nombre de facteurs plus que la qualité de la relation de l'évaluation (si l'on veut) suivant peut être sous-estimé ou inconsciemment.

Quatorze historiens ont noté quinze compositions,

¹ F. BACHER, *La normalisation*

tribuées à chacune des 10

12 (de 3 à 15 sur 20)
11
7,5

ple, plus frappant encore,
s.

ois, noté les examens du
de 10 sur 20) :

Admis par les six correcteurs	Admis par les uns et refusés par les autres
10 %	50 %
9 %	70 %
16 %	47 %
20 %	36 %
10 %	81 %
13 %	50 %

utiliser la note ?

x correcteurs apporterait-il
n, si aucune mesure stricte
noteurs. Par combien de pro-
valuer le même travail pour
bres suivants, avancés par
absolue. Ils donnent néan-
deur des nombres...

.	78
.	19
.	28
.	13
e	127
.	16

examens, Strasbourg, Conseil de l'Euro-
gier et D. Weinberg et portait sur le

L'histoire et la géographie ne figurent pas dans la liste ; ces deux branches aussi donnent pourtant lieu à de grandes fluctuations.

On pourrait multiplier les exemples de désaccord entre notateurs. Mais, à quoi bon se complaire dans l'accusation facile, comme trop de docimologistes l'ont fait ? L'important est d'être conscient du danger et, surtout, de chercher les remèdes.

Quelle est l'explication des divergences observées ?

Elle réside principalement dans la multiplicité des points de vue, des dimensions, selon lesquels une même épreuve peut être jugée, et à l'absence de directives précises (parfois, il n'y en a pas du tout) données pour la notation. F. Bacher¹ écrit :

« Dans une dissertation, on peut noter l'organisation des idées, leur originalité, la correction de l'expression, l'élégance du style, etc. Si l'on considère chaque aspect possible comme une dimension de la notation, le devoir d'un élève peut être situé dans un espace à n dimensions ; la difficulté vient de ce que, dans cet espace, seul un ordre partiel s'établit entre les élèves : un élève occupant une position élevée sur toutes les dimensions peut être déclaré supérieur à un élève occupant une position moins élevée sur toutes les dimensions, mais on ne peut dire d'un élève qui a une position élevée sur une première dimension et une position faible sur une deuxième dimension s'il est supérieur ou inférieur à un élève occupant des positions inverses. »

En outre, même s'ils considèrent le même aspect, les notateurs varient en sévérité, en stabilité de jugement, en résistance à l'effet de halo, etc.

7. Infidélité d'un même correcteur.

Un même correcteur est-il au moins égal à lui-même ? Non. A des phénomènes aussi évidents que les variations de santé physique et mentale et l'évolution du savoir s'ajoutent un grand nombre de facteurs plus ou moins bien définis : variation dans la qualité de la relation établie avec l'élève, dans le contexte de l'évaluation (si l'on vient de corriger un travail excellent, le suivant peut être sous-évalué,...), dans l'échelle consciemment ou inconsciemment adoptée, etc.

Quatorze historiens ont été invités à noter une deuxième fois quinze compositions, douze à dix-neuf mois après les avoir

1 F. BACHER, *La normalisation des notes*, o.c., pp. 53-54.

notées une première fois. Toute trace de correction avait été effacée. Les professeurs accordaient non seulement des points, mais indiquaient la réussite globale ou l'échec.

Dans 92 cas sur 210, le verdict a été différent d'une fois à l'autre.¹

Il faut toutefois insister sur le fait que des résultats aussi pauvres sont dus au manque de directives rigoureuses précisant les aspects à considérer par les notateurs.

Un schéma pour continuer la recherche.

Il est utile que, de temps en temps, des recherches soient faites sur la fidélité d'un notateur par rapport à lui-même ou sur l'accord entre notateurs, ne fût-ce que pour ranimer la conscience d'un danger menaçant. C'est pourquoi nous croyons opportun de signaler l'excellent plan de recherche dû à deux des meilleurs docimologistes contemporains : F. Yates et D. Pidgeon.

- Sept groupes de 50 enfants achevant l'école primaire ont été constitués. Ils ont été soumis à deux épreuves de langue maternelle :
 - a) rédaction sur un sujet choisi parmi trois ou quatre ;
 - b) questions de compréhension de textes et questions de grammaire (usage de la langue).
- Sept examinateurs expérimentés — cinq hommes et deux femmes — ont noté les travaux.
- Chaque groupe d'enfants a d'abord subi trois examens différents, à une semaine d'intervalle ; ensuite, la première des trois épreuves a été recommencée deux fois à une semaine d'intervalle.

SCHEMA²

GROUPE	Séances d'examen				
	I	II	III	IV	V
1	Aae	Bac	Dab	Aac	Aae
2	Bbf	Cbd	Ebc	Bbf	Bbf
3	Ccg	Dce	Fcd	Ccg	Ccg
4	Dda	Edf	Gde	Dda	Dda
5	Eeb	Feg	Aef	Eeb	Eeb
6	Ffc	Gfa	Bfg	Ffc	Ffc
7	Ggd	Agb	Cga	Ggd	Ggd

¹ HARTOG et RHODES, *An Examination of Examinations*, London, McMillan, 1935, pp. 81 et p. 15.

² Il est dû à D. FINNEY, cité par Yates et Pidgeon, *Admission to Grammar Schools*, Londres, N.F.E.R., 1957, p. 99.

Sept séries de question

Sept examinateurs : a,

Nous ne reprenons pas simplement voulu montrer avec, notamment, une mes

8. Stéréotypes et effets d

Dans la situation so chacun de ses élèves et en fonction d'un effet sou mant le travail ; là, on fait pour donner un choc que le maître agit délibéreme

Il en va tout autrement et de halo.

a) Stéréotypie.

Par stéréotypie, nous moins accusée, qui s'insta

La stéréotypie résulte Un premier travail médioc sera aussi ; si cela se vér médiocre au troisième t suite. Chez le professeur mation se produit d'autar

On aurait tort de cro quement les évaluations tion, composition d'histoir un jugement de valeur). E tifs » que la dictée ortho témoin.

Un professeur de l des dictées. Bientôt, il habituellement le mieux e détermine la fréquence par le correcteur, on cons élèves sont significativem faibles. Dans le premier contrer d'erreurs ; dans l

¹ Voir à ce propos M. ZILLIG *f. d. ges. Psychologie*, 12, 1934, p. R. Piper, 1967, p. 105.

...ance de correction avait été
...nt non seulement des points,
...ou l'échec.

...ct a été différent d'une fois

...fait que des résultats aussi
...directives rigoureuses précises
...notateurs.

...che.

...emps, des recherches soient
...ar rapport à lui-même ou sur
...e que pour ranimer la con-
...est pourquoi nous croyons
...n de recherche dû à deux des
...orains : F. Yates et D. Pid-

...avant l'école primaire ont été
...à deux épreuves de langue

...parmi trois ou quatre ;
...de textes et questions de
...e).

...s — cinq hommes et deux

...ord subi trois examens diffé-
...e ; ensuite, la première des
...ée deux fois à une semaine

A²

d'examen		
III	IV	V
ab	Aac	Aae
oc	Bbf	Bbf
cd	Ccg	Ccg
de	Dda	Dda
ef	Eeb	Eeb
fg	Ffc	Ffc
ga	Ggd	Ggd

...aminations, London, McMillan, 1935, pp.
...Pidgeon, Admission to Grammar Schools,

Sept séries de questions : A, B, C, D, E, F, G.

Sept examinateurs : a, b, c, d, e, f, g.

Nous ne reprenons pas le détail des résultats, mais avons simplement voulu montrer un exemple de contrôle bien mené, avec, notamment, une mesure de l'apprentissage.

8. Stéréotypes et effets de halo.

Dans la situation scolaire habituelle, le maître connaît chacun de ses élèves et peut donc doser, nuancer ses notes en fonction d'un effet souhaité : ici, on encourage en surestimant le travail ; là, on fait preuve d'une sévérité exceptionnelle pour donner un choc que l'on espère salutaire. Dans ces cas, le maître agit délibérément, en toute conscience.

Il en va tout autrement dans les phénomènes de stéréotypie et de halo.

a) Stéréotypie.

Par *stéréotypie*, nous entendons une immuabilité plus ou moins accusée, qui s'installe dans le jugement porté sur l'élève.

La stéréotypie résulte d'une contamination des résultats. Un premier travail médiocre incline à penser que le second le sera aussi ; si cela se vérifie, la tendance à accorder une note médiocre au troisième travail s'accroît encore, et ainsi de suite. Chez le professeur surchargé de corrections, la déformation se produit d'autant plus facilement.

On aurait tort de croire que la stéréotypie influence uniquement les évaluations à base subjective accusée (dissertation, composition d'histoire et, en général, réponses impliquant un jugement de valeur). Elle atteint des exercices aussi « objectifs » que la dictée orthographique. L'expérience suivante en témoigne.

Un professeur de langue maternelle fait régulièrement des dictées. Bientôt, il connaît les élèves qui réussissent habituellement le mieux et le moins bien cet exercice. Si l'on détermine la fréquence des fautes « oubliées », non perçues par le correcteur, on constate que les oublis en faveur des bons élèves sont significativement plus élevés que pour les élèves faibles. Dans le premier cas, le maître s'attend à ne pas rencontrer d'erreurs ; dans le second, il les guette¹.

¹ Voir à ce propos M. ZILLIG, Beliebte und unbeliebte Volksschülerinnen, in Arch. f. d. ges. Psychologie, 12, 1934, p. 32, cité par E. HOHN, Der schlechte Schüler, Munich, R. Piper, 1967, p. 105.

Si un manque de conscience professionnelle intervient, la situation peut devenir très grave. Un cas de stéréotypie accusée dont un élève était victime pour les versions latines nous avait été signalé (enseignement secondaire). A titre de vérification, nous avons fait faire les devoirs ultérieurs successivement par un autre élève de la classe, par un élève de même niveau pédagogique, classé premier dans une autre école, puis par un licencié en philologie classique : la note n'a pas varié d'un demi-point sur vingt...

Nous avons répété ce genre d'expériences à d'autres occasions. Combien de parents, aussi, parfois très compétents dans le domaine où ils faisaient occasionnellement le travail de leur enfant, ont éprouvé quelque déception ou étonnement quand ils ont eu connaissance de « leur » note.

En déduire que tous les maîtres incriminés de telle façon manquent du sens des responsabilités serait pourtant injuste. C'est plus les méthodes d'évaluation que les hommes qu'il faut mettre en cause.

b) Effet de halo.

L'effet de halo présente un caractère affectif accusé. Souvent, on surestime les réponses d'un élève de belle allure, au regard franc, à la diction agréable. Il ne faut cependant pas généraliser. On connaît des professeurs qui, par anticonformisme, favorisent un certain débraillé ou des originalités qui ne sont pas toujours du meilleur goût.

Soit pour des raisons de lisibilité, soit pour des raisons nettement affectives, l'écriture peut aussi influencer le correcteur. Les spécialistes de la publicité savent depuis longtemps que la présentation du message exerce une influence considérable sur son rendement.

Les expériences suivantes, dues à R. Weiss¹, montrent avec quelle facilité une combinaison de stéréotypie et d'effet de halo peut être artificiellement provoquée.

R. Weiss a sélectionné deux rédactions faites par des élèves de 4^e primaire. Les travaux ont été dactylographiés sans modification aucune, puis soumis pour correction à deux groupes de 46 instituteurs enseignant aussi en 4^e primaire.

Le commentaire suivant accompagnait la distribution des travaux au premier groupe : « Voici deux compositions écrites par des élèves de 4^e primaire. Le travail n° 1 est l'œuvre d'un

¹ R. WEISS, The Reliability of the Number Marking System : An Austrian Study, in J. LAUWERYS et D. SCALON, *Examinations*, Londres, Evans Brothers Ltd, 1969, pp. 101-107.

élève moyen qui aime lire sa mère sont employés. L'élève doué ; son père est rédacteur second groupe de maître

La correction devait être notée en degrés : très bien (1) ; bien (2) ; satisfaisant (3) ; insuffisant (5). L'orthographe et la grammaire devaient être notés séparément.

Pour les quatre aspects de travail pour lequel on a fait une prédiction, on observe que pour l'élève doué, 16 % des correcteurs ont noté la note insuffisante ; si l'élève est moyen, les correcteurs n'ont noté *insuffisant*.

NOTATION DES COMPOSANTES D'UN PREJUGE FAUTIF

Influence positive

Moyenne

Orthographe

2-35

Style

1-8

Fond

1-6

Note globale

2-0

¹ Toutes les différences sont...

professionnelle intervient, la
 Un cas de stéréotypie accu-
 ur les versions latines nous
 ondaire). A titre de vérifi-
 evoirs ultérieurs successive-
 sse, par un élève de même
 r dans une autre école, puis
 ique : la note n'a pas varié

e d'expériences à d'autres
 aussi, parfois très compétents
 occasionnellement le travail
 e déception ou étonnement
 « leur » note.

es incriminés de telle façon
 lités serait pourtant injuste.
 tion que les hommes qu'il

ractère affectif accusé. Sou-
 un élève de belle allure, au
 e. Il ne faut cependant pas
 sseurs qui, par anticonfor-
 aillé ou des originalités qui
 it.

abilité, soit pour des raisons
 t aussi influencer le correc-
 té savent depuis longtemps
 erce une influence considé-

ues à R. Weiss¹, montrent
 n de stéréotypie et d'effet de
 quée.

rédactions faites par des
 ont été dactylographiés sans
 s pour correction à deux
 ant aussi en 4^e primaire.

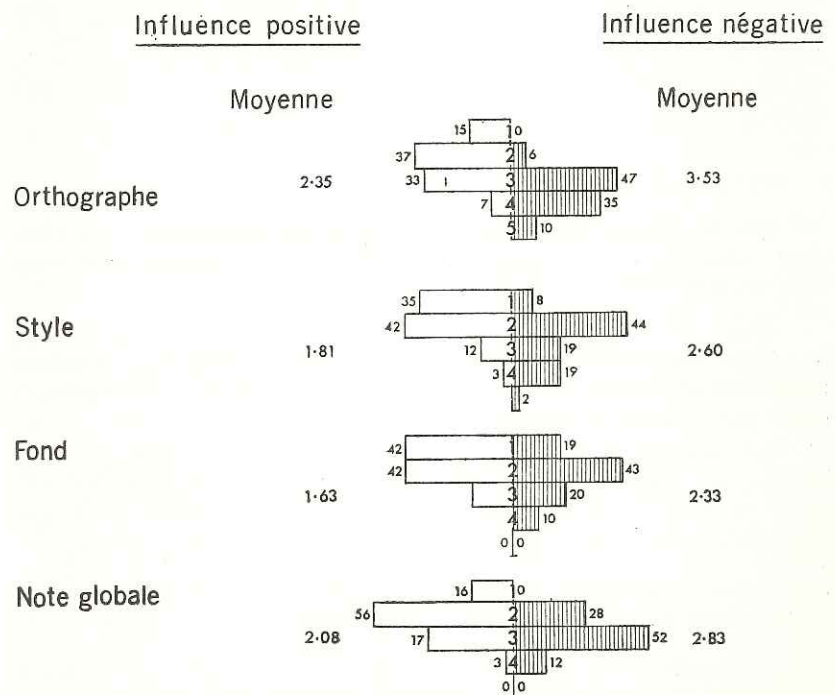
mpagnait la distribution des
 i deux compositions écrites
 travail n° 1 est l'œuvre d'un

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 selon une échelle à cinq degrés : très bien (1) ; bien (2) ; moyen (3) ; suffisant (4) et insuffisant (5). L'orthographe, le style, le fond, puis l'ensemble devaient être notés séparément.

Pour les quatre 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 œdipien de la prédiction, on observe qu'au travail de l'élève présenté comme doué, 16 % des correcteurs accordent la note *très bien* et aucun la note insuffisant ; si le même élève est présenté comme moyen, les correcteurs n'accordent aucun *très bien*, mais 11 % notent *insuffisant*.

NOTATION DES COMPOSITIONS SOUS L'INFLUENCE D'UN PREJUGE FAVORABLE OU DEFAVORABLE¹



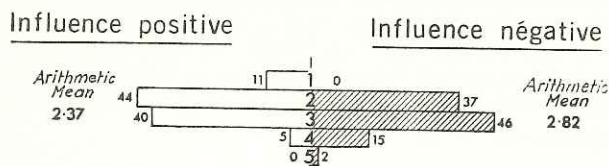
Marking System : An Austrian Study, in
 es, Evans Brothers Ltd, 1969, pp. 101-107.

¹ Toutes les différences sont statistiquement significatives à P.01 sauf une (P.10).

Dans une expérience similaire sur des problèmes d'arithmétique, également en 4^e primaire, R. Weiss observe :

<i>Préjugé favorable</i>	<i>Préjugé défavorable</i>
11 % de très bien	Aucun très bien
5 % de suffisant	15 % de suffisant
0 % d'insuffisant	2 % d'insuffisant

PROBLEME D'ARITHMETIQUE



On voit combien un artifice relativement grossier exerce déjà une influence. Comme le souligne Weiss, la déformation est très probablement bien plus grande encore en situation réelle où le maître connaît l'apparence de l'élève, sa conduite, sa façon de s'exprimer, la profession et le statut social des parents, etc.

9. Manque de validité.

Les examens traditionnels ne sont pas seulement de pauvres moyens d'inventaire ; leur valeur pronostique apparaît tout aussi contestable dans bien des cas.

Nous avons observé, en compagnie du préfet d'un des meilleurs athénées de Belgique, que deux catégories d'élèves réussissaient brillamment dans l'enseignement universitaire : ceux qui, pendant toutes leurs études, semblent s'être joués de la difficulté et sont restés en tête de leur classe, et une partie, non négligeable, de ceux qui n'ont presque jamais enregistré d'échecs, mais se sont néanmoins maintenus juste au-dessus de la note fatidique. Dans les deux cas, les performances dans l'enseignement secondaire sont de bon augure, moins par leur nature apparente que par la facilité d'adaptation, par la plasticité intellectuelle qui y a conduit.

Pronostiquer la réussite ou l'échec en faculté sur la base du pourcentage obtenu au terme de l'enseignement secondaire est donc bien hasardeux.

10. Un instrument d'im

L'examen traditionnel moyen de sélection que l'on trouve dans la plupart des cas, moins en ce qui concerne la qualification so

Ainsi, l'examen dev bilisme social, phénomène J.-C. Passeron à qui nous devons ces observations suivantes¹.

a) Effets irréversibles d

Une fois obtenu, un tel examen assure *grosso modo* son niveau de ré

Or, au moment de la sélection se trouve mesuré sera exigée du travailleur analogique, préparatoire mesurent les examens réels, ce n'est pas ce qui est mesuré, mais plutôt le niveau auquel on doit payer » (p. 7).

Ces observations ne correspondent néanmoins à un exemple, jamais on n'attribue un poste ou au technicien qui en

Le dernier examen d'un individu dans une carrière encore aujourd'hui, fort de sa fixation est en contradiction avec la contemporaine d'éducat

Que l'examen accorde une garantie de compétence par le fait que, dans bien des cas, ils occupent le même poste

« l'une officiellement requises pour l'occupation et le pouvoir), l'autre est la part technique de la tâche du travail » (...). Parfois,

¹ J.-C. PASSERON, Sociolog

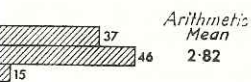
ur des problèmes d'arithmétique
eiss observe :

Préjugé défavorable

Aucun très bien
15 % de suffisant
2 % d'insuffisant

ARITHMETIQUE

Influence négative



relativement grossier exerce
ligne Weiss, la déformation
grande encore en situation
nce de l'élève, sa conduite,
ion et le statut social des

ont pas seulement de pau-
aleur pronostique apparaît
cas.

agnie du préfet d'un des
e deux catégories d'élèves
enseignement universitaire :
des, semblent s'être joués
ête de leur classe, et une
n'ont presque jamais enre-
moins maintenus juste au-
les deux cas, les perfor-
adaire sont de bon augure,
e par la facilité d'adaptation,
conduit.

chec en faculté sur la base
e l'enseignement secondaire

10. Un instrument d'immobilisme social.

L'examen traditionnel se justifie au moins autant comme moyen de sélection que comme outil de promotion de l'éducation individuelle. De plus, cette sélection s'opère, dans beaucoup de cas, moins en fonction de la qualification technique que de la qualification sociale.

Ainsi, l'examen devient un instrument privilégié de l'immobilisme social, phénomène analysé de façon pénétrante par J.-C. Passeron à qui nous empruntons l'essentiel des considérations suivantes¹.

a) Effets irréversibles de la certification scolaire.

Une fois obtenu, un titre scolaire suit l'individu toute sa vie, lui assure *grosso modo* son grade dans la hiérarchie professionnelle, son niveau de rémunération et le pouvoir qu'il détiendra.

Or, au moment de l'obtention d'un diplôme, « la compétence se trouve mesurée non pas au travers de l'activité qui sera exigée du travailleur, mais au travers d'une sorte d'activité analogique, préparatoire, ludique et presque fictive. Ce que mesurent les examens par rapport à la demande professionnelle, ce n'est pas ce que les gens seront censés faire, mais plutôt le niveau auquel il sera socialement exigible de les payer » (p. 7).

Ces observations n'ont certes pas de valeur absolue ; elles correspondent néanmoins à une réalité incontestable. Par exemple, jamais on n'attribue le grade d'ingénieur à un ouvrier ou au technicien qui en a acquis la compétence fonctionnelle.

Le dernier examen qu'il subit avec fruit fixe donc souvent un individu dans une zone socio-professionnelle dont il est, encore aujourd'hui, fort malaisé de sortir. La précocité de cette fixation est en contradiction profonde avec toute la politique contemporaine d'éducation et donc de promotion permanente.

Que l'examen accorde plus une certification sociale qu'une garantie de compétence technique serait d'ailleurs confirmé par le fait que, dans beaucoup d'entreprises, deux personnes occupent le même poste :

« l'une officiellement parce qu'elle a tous les titres scolaires requis pour l'occuper (avec la rémunération, le prestige et le pouvoir), l'autre effectivement, parce qu'elle assume la part technique de la tâche (...), c'est-à-dire qu'elle fait « tout le travail » (...). Parfois, le second conviendrait techniquement

¹ J.-C. PASSERON, Sociologie des examens, in *Education et Gestion*, 1970, 2, 6-16.

au poste, bien mieux que le titulaire. Mais c'est précisément ce qu'empêche l'effet de la certification produit par l'examen qui, une fois pour toutes, et pour toute la vie, tend à définir ce que vaut la prestation professionnelle d'un individu par la valeur de cet individu sur le marché des titres scolaires.»

De nouveau, cette constatation appelle des restrictions. Dans une entreprise, la seule compétence technique suffit rarement aux cadres : la capacité de communiquer, d'exercer des fonctions de *leader* et de relations publiques, d'harmoniser le travail d'administration et de production, etc. sont autant de qualités dont la manifestation échappe souvent à ceux qui ne voient que l'aspect technique des choses. Mais ces qualités ne sont pas toujours réelles, loin s'en faut. Souvent, on a affaire à un formalisme dont nous allons reparler.

b) *Les examens ne sont pas socialement neutres.*

Une enquête de l'INSEE, en France, révèle que si, au cours des dernières années, les chances d'accès à l'enseignement supérieur se sont accrues pour toutes les catégories sociales, « l'augmentation globale du taux de scolarisation s'est ventilée entre les diverses catégories socio-professionnelles dans la même proportion que les inégalités antérieures. Bref, il s'agit d'une simple *translation vers le haut* de la structure des inégalités » (p. 8). Le tableau ci-dessous est révélateur.

Catégorie socio-professionnelle du père	Probabilités d'accès à l'enseignement supérieur	
	1961-62	1965-66
Salariés agricoles	1,1	2,7
Agriculteurs	3,4	8
Ouvriers	1,3	3,4
Employés	9	16,2
Cadres moyens	24,9	35,4
Professions libérales et cadres supérieurs	38	58,7

Comment s'explique ce phénomène en apparence contradictoire avec l'obligation scolaire généralisée, de plus en plus longue, la gratuité des études et la répartition des aptitudes, selon la loi du hasard, dans toutes les couches de la société ?

La différence d'éducation familiale, au cours des premières années, créerait rapidement des différences dans les mécanismes intellectuels, les moyens linguistiques et les attitudes devant le travail : « L'éducation donnée par les familles prédé-

pose d'autant moins à la réappartient à une classe de la culture savante » (p. 1

Les examens « neutres » ont des effets profonds de cette précision. Notre opticien Passeron qui ne reconnaît pas son origine socio-culturelle. Mais, comme lui, nous ne sommes pas socialement neutres.

« Les procédures de sélection préparent en compte les techniques certains aspects n'ont aucune importance très fortement liés aux hiérarchies sociales plutôt que de telle sorte, les examens français sont des épreuves de connaissance des exercices d'usage lettré

La dissertation reste le privilège. Rejoignant la pensée de Max Weber, d'A. Clausewitz, a toujours été un moyen. Il est, en effet, frappant de constater d'importance à une épreuve plus radicalement à toute épreuve. Le formalisme consiste précisément son contenu objectif, mais la manière, la nuance. Cette fonction de fermeture au

La neutralité des examens est tellement mieux garantie, le langage, et bien peu de l'élégance de l'expression technique sur lequel port

Quant à l'examen oral, la belle présentation, du

Notons, pour conclure, que le caractère sélectif paraît caractéristique du XIX^e siècle, en particulier dans une société très hiérarchisée. Il est donc naturel que dans la société dans laquelle nous évoluons s'agit d'un type d'examen conçu

Mais c'est précisément ce qu'on produit par l'examen de la vie, tend à définir ce d'un individu par la valeur des choses scolaires.»

appelle des restrictions. L'absence technique suffit rarement à communiquer, d'exercer des fonctions publiques, d'harmoniser le fonctionnement, etc. sont autant de choses qui ne sont pas souvent à l'origine de succès. Mais ces qualités ne sont pas suffisantes. Souvent, on a affaire à des personnes qui ne savent pas parler.

ment neutres.

ce, révèle que si, au cours de l'accès à l'enseignement supérieur, les catégories sociales, la scolarisation s'est ventilée de manière inégale dans les professions et les métiers antérieures. Bref, il s'agit d'un fait de la structure des inégalités qui est révélateur.

Probabilités d'accès à l'enseignement supérieur
1961-62 1965-66

1,1	2,7
3,4	8
1,3	3,4
9	16,2
24,9	35,4
38	58,7

ène en apparence contra-
néralisée, de plus en plus
répartition des aptitudes,
des couches de la société ?

le, au cours des premières
férences dans les mécanis-
mismes et les attitudes
née par les familles prédis-

pose d'autant moins à la réussite aux examens que ces familles appartiennent à une classe plus éloignée de la culture scolaire, de la culture savante » (p. 11).

Les examens « neutres » se borneraient à enregistrer les effets profonds de cette première éducation. Ce n'est que partiellement exact. Notre opinion est plus nuancée que celle de Passeron qui ne reconnaît pas assez la réalité des handicaps d'origine socio-culturelle auxquels on vient de faire allusion. Mais, comme lui, nous pensons que les examens ne sont pas socialement neutres.

« Les procédures de notation et les types d'épreuves utilisés prennent en compte au moins autant que les aptitudes techniques certains aspects gratuits de la performance, qui n'ont aucune importance technique, mais qui sont en revanche très fortement liés aux habitudes culturelles de telle classe sociale plutôt que de telle autre » (p. 12). Ainsi, constate Passeron, les examens français les plus prestigieux sont *moins des épreuves de connaissances que des épreuves de manières ou des exercices d'usage lettré du langage*.

La dissertation reste, dans cette perspective, le moyen privilégié. Rejoignant la thèse sociologique de Durkheim, de Max Weber, d'A. Clause, Passeron rappelle que le formalisme a toujours été un moyen de défense des classes privilégiées. Il est, en effet, frappant de constater que l'on accorde le plus d'importance à une épreuve, la dissertation, qui échappe le plus radicalement à toute notation objective, technique... Or, le formalisme consiste précisément à définir la culture, non par son contenu objectif, mais par ces impondérables que sont la manière, la nuance. Cette « indéfinition » permet d'exercer une *fonction de fermeture* au profit d'un groupe favorisé.

La neutralité des examens scientifiques n'est d'ailleurs pas tellement mieux garantie, car eux aussi font souvent intervenir le langage, et bien peu d'examineurs restent indifférents à l'élégance de l'expression, pourtant étrangère au problème technique sur lequel porte fondamentalement l'examen.

Quant à l'examen oral, il ne fait que renforcer l'effet de la belle présentation, du langage châtié, de la diction élégante.

Notons, pour conclure, que la floraison des examens à caractère sélectif paraît caractéristique des conditions culturelles du XIX^e siècle, en particulier du développement d'une bureaucratie très hiérarchisée au service de l'économie capitaliste. Il est donc naturel que la nouvelle forme de civilisation vers laquelle nous évoluons s'accompagne d'une mise en question d'un type d'examen conçu pour d'autres conditions.

ciences docimologiques.

accumule les expériences
entre notateurs ou le man-

Mais on l'aggrave artificiel-
ment d'inviter les notateurs à
travailler et l'importance relative
continue à enregistrer pa-
raître l'anarchie des examens.

ent encore aux examens.
sultats sur les Examens
Grande-Bretagne, *déposé*
Nous l'allongeons un peu,
manière exhaustive.

duire les idées des autres
é.

atissages souvent éphémè-
80 % des connaissances,
examen, ont disparu quinze

ux consignes.

parfois le pas sur le con-
sistent les handicaps socio-
le domaine verbal.)

jugement personnel, pour
qui évaluera l'examen.

isés par un esprit de com-
mercenaire.

ce : dans l'examen tradi-
tionnel, reflet des idiosyncra-
sies, jeu des « tuyaux ».

ongue préparation et la
s'accourcissent considéra-

ent le travail en groupe et
la performance indivi-

- La menace d'échec qu'ils font peser, — souvent doublée de la peur d'une sanction familiale, — incite à la fraude. Le succès des fraudeurs constitue un mauvais exemple pour les élèves restés honnêtes.

- Les examens donnent une idée fautive du travail adulte où, plus la question est complexe et difficile, plus on s'entoure de conseils et d'ouvrages de référence.

Pour le professeur :

- Il enseigne en fonction de l'examen, voire des exigences particulières des membres d'un jury extérieur. Nous verrons qu'en bonnes conditions, l'esprit de l'enseignement et de l'examen doit être un. On fait allusion ici à des examens étrangers aux objectifs éducationnels.

- Il est jugulé dans sa méthodologie, si l'esprit de l'examen est étranger aux objectifs pédagogiques qui lui paraissent essentiels.

Comment pratiquer une pédagogie de la découverte, de l'exploration personnelle, — processus lents, mais générateurs d'apprentissages profonds, — si l'examen imposé est de caractère encyclopédique ?

- Esclave du programme, il ne laisse pas les élèves avancer à leur allure propre, et donc digérer la matière.

- Il tend à accorder trop d'importance aux aptitudes et aux connaissances utiles à l'examen.

Par après, on a baptisé *effet de contrecoup* le phénomène par lequel les professeurs modifient la méthode et le contenu de leur enseignement en fonction de l'évolution des examens imposés de l'extérieur.

Des aspects éducatifs importants peuvent ainsi être négligés et l'ont souvent été.

En résumé, les examens traditionnels présentent généralement de graves défauts de construction. Leur validité est contestable. L'évaluation des travaux est, de son côté, grevée de lourdes faiblesses. En outre, les examens peuvent nuire à la santé physique et mentale des élèves. On ajoutera enfin que, dans certains cas, les professeurs font leur cours en fonction d'un examen et non le contraire...

Mais il est temps de donner la parole à la défense.

CHAPITRE 2.

DEFENSE DE LA NOTE SUBJECTIVE ET DE L'EXAMEN.

Que l'on ait abusé des examens au point de vicier l'action éducative de l'école, personne ne le contestera. Que bien des concours furent surtout de sinistres loteries et duperies semble aussi hors de doute. Que des scores se virent attribuer une signification qu'ils n'avaient pas est également patent.

De là à conclure qu'examens et concours doivent être à jamais proscrits pour ne laisser subsister que des séries d'évaluations occasionnelles et les rapports qui les synthétisent, paraît utopique.

Dans la première partie, nous avons vu que les différents modes d'évaluation correspondent à des fonctions particulières. Il suffit d'en reprendre la liste pour constater qu'elles ne peuvent être toutes remplies par une seule procédure.

Par ailleurs, tous les modes d'évaluation impliquent l'établissement d'un score ou d'une note que l'on souhaite parfaitement objectifs ou, mieux, parfaitement contrôlés : entendons par là que nous reconnaissons aux maîtres le droit d'adapter beaucoup de leurs évaluations, autant en fonction de l'élève que de la matière, à condition qu'ils sachent exactement ce qu'ils font.

Avant d'envisager certains avantages souvent reconnus aux examens, une discussion de caractère négatif s'impose.

1. La mesure rigoureuse est peut-être impossible.

Rechercher les voies de la parfaite validité des examens et de la fidélité des évaluations repose sur deux postulats rappelés par J. Guillaumin¹ :

¹ J. GUILLAUMIN, L'aspect interpersonnel de la notation scolaire : de la docimologie à la doxologie, in *Bulletin de la Société A. Binet et T. Simon*, IV, 1968, p. 270.

- 1° « Que les productions soient objectives, quantifiables. »
- 2° « Que les différences de notation soient justifiées par les pratiques pratiquées par les notateurs. »

Si ces deux postulats s'écroulent, Or, de bons esprits, de grands principes sont définis, mais n'a guère pénétré à l'école. Une bonne partie, à une incompréhension de la situation d'enseignement que ou physique des procédés.

Le débat n'est pas nouveau. Les progrès considérables de la psychologie contemporaines sont l'observation, donc à la mesure des sciences humaines ont aujourd'hui dépassé leurs limites et, en pratique, la tendance générale d'utiliser des échelles de notation parfaites. Ils savent aussi que les différences de comportement sont dues à l'observation rigoureuse des performances particulières sur lesquelles on a pu considérer ?

De même que l'on ne peut pas mesurer avec précision les points de vue à considérer, tant les points de vue à considérer sont nombreux, de même on ne peut pas compter de la valeur réelle d'un objet. Même si, en calcul, on est faux, la démarche de la mesure varie considérablement d'un individu à l'autre, ce qui rend difficile à chiffrer objectivement une performance considérée comme œuvre parcellaire¹. Aussi, le droit de la mesure, compte, on le comprend.

Si la réponse à donner est celle de l'histoire, de sciences naturelles.

¹ Signalons néanmoins qu'Ellis F. Sturges, chercheur qui permet d'évaluer automatiquement la corrélation avec les notes traditionnelles, dans ce cas, le chercheur n'a pas seulement des imparfaites des notateurs... E. PAGE et D. PAULUS, *The Analysis of H.E.W.*, Project 6-1318, 1968.

- 1° « Que les productions de l'élève sont par nature mesurables, quantifiables. »
- 2° « Que les différences qu'on peut trouver entre les mesures pratiquées par les notateurs sont susceptibles d'être réduites. »

Si ces deux postulats sont faux, la docimologie classique s'écroule. Or, de bons esprits constatent que, bien que ses grands principes sont définis depuis plusieurs décennies, elle n'a guère pénétré à l'école. Cet échec serait dû, pour une bonne partie, à une incompatibilité entre la nature psychologique de la situation d'enseignement et le caractère mathématique ou physique des procédés de mesure proposés.

Le débat n'est pas nouveau et l'issue est dans le compromis. Les progrès considérables de la psychologie et de la pédagogie contemporaines sont largement dus à l'objectivation de l'observation, donc à la mesure. Toutefois, les chercheurs en sciences humaines ont aujourd'hui une conscience assez claire de leurs limites et, en particulier, de l'impossibilité presque générale d'utiliser des échelles de mesure mathématiquement parfaites. Ils savent aussi que nous ne mesurons le plus souvent que des comportements isolés du tout humain. Comment évaluer rigoureusement des entités hypothétiques comme l'esprit critique, le sens de l'observation, etc., sinon à travers des performances particulières supposées représentatives de l'ensemble considéré ?

De même que l'on n'est jamais parvenu, jusqu'à présent, à mesurer avec précision le rendement de l'enseignement, tant les points de vue à considérer à des moments différents sont nombreux, de même on ne peut rendre analytiquement compte de la valeur réelle d'une performance scolaire complexe. Même si, en calcul, le résultat ne peut être que juste ou faux, la démarche de la pensée et l'effort produit peuvent varier considérablement d'un sujet à l'autre et donc être difficiles à chiffrer objectivement. Quant à la composition française, considérée comme œuvre d'art, elle échappe à l'évaluation parcellaire¹. Aussi, le droit à la subjectivité, à la réaction globale, compte, on le comprend, d'ardents défenseurs.

Si la réponse à donner à des questions de géographie, d'histoire, de sciences naturelles contraint l'élève à s'exprimer,

¹ Signalons néanmoins qu'Ellis Page a réussi à construire un programme d'ordinateur qui permet d'évaluer automatiquement les compositions de langue maternelle. La corrélation avec les notes traditionnelles est élevée. Toutefois, il faudrait savoir si, dans ce cas, le chercheur n'a pas simplement fait adopter par l'ordinateur les démarches imparfaites des notateurs...
E. PAGE et D. PAULUS, *The Analysis of Essays by Computer*, Washington, U.S. Office of H.E.W., Project 6-1318, 1968.

oralement ou par écrit, faut-il toujours faire abstraction de l'élégance du langage, de l'ordonnance de la pensée, de la rigueur du raisonnement ?

Bref, dès que l'on ne se résout plus à ramener la réussite ou l'échec à des critères rigides et souvent schématiques, les productions des étudiants échappent à la quantification automatique, impersonnelle.

Cet argument ne suffit assurément pas pour renoncer à objectiver une partie de la notation des élèves, mais une partie seulement. L'appréciation globale du maître, toute en finesse, sa sensibilité tant à la performance matérielle qu'à l'effort de dépassement et à la faiblesse humaine, doivent garder leur place.

La richesse de l'enseignement réside avant tout dans la qualité de la relation humaine qu'il crée et l'évaluation est un des aspects de cette relation. Si elle s'appauvrit au point de ne laisser subsister qu'une communication impersonnelle, le maître peut être avantageusement remplacé par une machine à enseigner.

2. Les maîtres jugent bien leurs élèves.

Si les maîtres tendent à relativiser leurs jugements par rapport au niveau moyen du groupe, leur classement à l'intérieur de celui-ci n'en possède pas moins une validité élevée.

Il suffirait donc d'appliquer des mesures permettant de rendre ces jugements comparables entre écoles (nous verrons comment au chapitre de la « modération ») pour disposer de points de repère importants.

La sûreté de jugement des maîtres s'explique par plusieurs facteurs :

1. ils fondent leur jugement sur une observation longue et continue ;
2. ils considèrent un beaucoup plus grand nombre de facteurs (notamment de personnalité) que l'examen ;
3. ils peuvent, en particulier, tenir compte de comportements exceptionnels (en telle occasion, tel élève a fait preuve d'une lucidité, d'une originalité peu communes) qui n'apparaîtraient probablement pas à l'examen.

La valeur prédictive des jugements des instituteurs (ajustés pour les rendre comparables) a été bien mise en lumière par une recherche faite par la *National Foundation for Educational Research in England and Wales*¹.

¹ Voir F. YATES and D. PIDGEON, *Admission to Grammar School*, o.c., pp. 57 sq.

En 1951 et en 1952, e
Twickenham ont été exam
dans l'enseignement seco
sultats ont été étudiés deu

Epreuves subies au dé

Tests :

1. Test d'intelligence verb
2. Test standardisé de c
(E1) (Questions à choix
3. Test standardisé de
(Calculs et problèmes)
4. Test d'intelligence non
5. Test d'aptitude spatiale
6. Test d'aptitude spatiale
7. Test de connaissances
plus ouvertes que E1).

Jugement par un instituteur

1. Prédit réussite des ét
échelle à 15 degrés (M)
2. Classe ses élèves selo
mar School. Ces juger
en fonction d'un test d
du système au chapitre

Epreuves utilisées po
ment secondaire : ordre
études, ajusté en fonction
objectifs d'aptitudes et de
tement après l'élaboration

Le tableau ci-dessou
corrélation entre chaque
en cours d'études :

En 1951 et en 1952, environ 1 200 élèves de la région de Twickenham ont été examinés lors de leur sélection à l'entrée dans l'enseignement secondaire (*Grammar Schools*). Leurs résultats ont été étudiés deux ans, puis trois ans après.

Epreuves subies au départ :

Tests :

1. Test d'intelligence verbale (V) ;
2. Test standardisé de connaissances en langue maternelle (E1) (Questions à choix multiple) ;
3. Test standardisé de connaissances en arithmétique (A) (Calculs et problèmes) ;
4. Test d'intelligence non verbale (N/V) ;
5. Test d'aptitude spatiale (Sp. 1) (Espace à deux dimensions) ;
6. Test d'aptitude spatiale (Sp. 2) (Espace à trois dimensions) ;
7. Test de connaissances en langue maternelle (E2) (Questions plus ouvertes que E1).

Jugement par un instituteur de 6^e primaire :

1. Prédit réussite des études à la *Grammar School* sur une échelle à 15 degrés (M) ;
2. Classe ses élèves selon l'ordre de leur réussite à la *Grammar School*. Ces jugements sont ajustés par les chercheurs en fonction d'un test d'intelligence verbale (voir description du système au chapitre de la « modération ») (F).

Epreuves utilisées pour mesurer le succès dans l'enseignement secondaire : ordre de mérite fourni par le préfet des études, ajusté en fonction des résultats à une batterie de tests objectifs d'aptitudes et de connaissances, administrés immédiatement après l'élaboration du classement (S.H.A.).

Le tableau ci-dessous montre, pour le groupe de 1951, la corrélation entre chaque épreuve de prédiction et les résultats en cours d'études :

Prédicteur	Groupe de 1951 : deux ans après	Groupe de 1951 : trois ans après
Jugement de l'instituteur (F)	.821	.748
(M)	.796	.722
Test Verbal (V)	.789	.704
E2	.749	.623
A	.734	.659
E1	.729	.622
N/V	.648	.535
Sp.1	.565	.453
Sp.2	.491	.361

Le jugement ajusté des instituteurs se révèle le meilleur prédicteur. La valeur pronostique élevée du simple test verbal est aussi confirmée.

Il faut y insister : ces observations ne sont valides que dans la situation considérée : la réussite dans les *Grammar Schools*. Nous ne savons pas ce qui se passerait dans d'autres formes d'enseignement. Nous disposons néanmoins ici d'une observation strictement contrôlée dont on n'a nulle raison de croire qu'elle ne puisse s'appliquer à d'autres cas.

3. Validité limitée mais réelle des examens traditionnels.

Quelles qu'en fussent les imperfections, les examens traditionnels n'ont pas empêché notre civilisation d'atteindre un niveau scientifique, jamais égalé dans l'histoire de l'humanité. Le filtrage qu'ils ont opéré s'est donc révélé, au moins partiellement, valide.

La chose est évidente. Avec la restriction faite au chapitre précédent, on peut affirmer que les examens traditionnels ont permis d'identifier les élèves les mieux et les moins bien doués, tri capital dans la forme de civilisation que nous avons connue dans la dernière centaine d'années.

Que la procédure ait été grossière (en ce sens qu'elle a ignoré ou gaspillé des talents, au détriment des couches sociales défavorisées surtout) et injuste envers la population des élèves moyens, mis brutalement dans le même sac, n'enlève rien au fait que, dans une démocratie peu évoluée, les examens ont bien joué leur rôle.

4. S'endurcir pour la vie

Selon les behavioristes elle est effectivement pro- périodiquement, l'étudiant exceptionnel, à bander ses examens, avec toutes ses Car la vie ne lui épargne bon d'y être préparé.

Par ailleurs, la concu fondamentaux de notre existent et l'on peut même en plus là où il y a plus de disponibles, sinon le fav l'inefficacité qui l'accomp

5. Se situer par rapport

Si un classement de un choc, il permet aussi (non dans l'absolu, mais Celui qui a fait de son mi tions ; celui qui ne l'a pa être avant qu'il ne soit faiblesses.

En règle générale, mais une certaine frustra aiguillon.

6. Large synthèse et int

Il est indéniable qu' ensembles de connaissances des synthèses à l'ocasi la première fois, l'éconor parties et, éventuellement disciplines.

7. L'examen externe co

Un examen régional contrôle du travail des pr

S'il ne peut plus être programmes surabondan moins nécessaire que l'é d'études précis dans ses principales.

de 1951 : ans après	Groupe de 1951 : trois ans après
.821	.748
.796	.722
.789	.704
.749	.623
.734	.659
.729	.622
.648	.535
.565	.453
.491	.361

eurs se révèle le meilleur
évée du simple test verbal

s ne sont valides que dans
dans les *Grammar Schools*.
erait dans d'autres formes
anmoins ici d'une observa-
n'a nulle raison de croire
res cas.

Examens traditionnels.

ections, les examens tradi-
civilisation d'atteindre un
ns l'histoire de l'humanité.
révélé, au moins partielle-

restriction faite au chapitre
examens traditionnels ont
x et les moins bien doués,
on que nous avons connue

ère (en ce sens qu'elle a
u détrimment des couches
uste envers la population
dans le même sac, n'enlève
peu évoluée, les examens

4. S'endurcir pour la vie.

Selon les behavioristes, une conduite ne s'apprend que si elle est effectivement produite. Il paraît donc souhaitable que, périodiquement, l'étudiant soit amené à faire un effort exceptionnel, à bander son énergie, à affronter l'ordalie des examens, avec toutes ses imperfections, voire ses injustices. Car la vie ne lui épargnera pas semblables épreuves, et il est bon d'y être préparé.

Par ailleurs, la concurrence, la compétition sont des traits fondamentaux de notre civilisation. L'examen et le concours existent et l'on peut même souhaiter qu'ils s'imposent de plus en plus là où il y a plus de candidats que de places ou d'emplois disponibles, sinon le favoritisme sous toutes ses formes et l'inefficacité qui l'accompagne ont libre jeu.

5. Se situer par rapport aux autres.

Si un classement défavorable à une épreuve peut donner un choc, il permet aussi de se situer par rapport aux autres (non dans l'absolu, mais dans les conditions de l'examen !). Celui qui a fait de son mieux peut ainsi mieux ajuster ses ambitions ; celui qui ne l'a pas fait a l'occasion de découvrir, peut-être avant qu'il ne soit trop tard, les conséquences de ses faiblesses.

En règle générale, le succès profite mieux que l'échec, mais une certaine frustration constitue, pour certains, un utile aiguillon.

6. Large synthèse et intégration des connaissances.

Il est indéniable que les examens portant sur de vastes ensembles de connaissances obligent l'étudiant à construire des synthèses à l'occasion desquelles il perçoit, parfois pour la première fois, l'économie de tout l'édifice, les relations entre parties et, éventuellement, les points communs entre différentes disciplines.

7. L'examen externe contrôle le professeur.

Un examen régional ou national constitue un moyen de contrôle du travail des professeurs.

S'il ne peut plus être question d'asservir les maîtres à des programmes surabondants, fixés en détail, il n'en reste pas moins nécessaire que l'enseignement se déroule selon un plan d'études précis dans ses objectifs et défini dans ses matières principales.

Un temps viendra sans doute où chaque élève travaillera selon un programme individualisé ou semi-individualisé, mais, même dans ce cas, les buts à atteindre devront être définis avant que ne s'engage l'action pédagogique.

Toujours, les parents et les autorités organisatrices auront le droit de vérifier si la mission enseignante a été accomplie.

L'existence d'un contrôle portant sur tous les éléments importants du plan d'études semble aussi une saine sauvegarde contre la laxité, le relâchement. Consacrer le temps nécessaire à l'apprentissage est louable ; musarder est condamnable. Il me souvient d'un professeur de langues qui, en deux ans, n'avait pas dépassé l'introduction du cours... Apparemment, aucun inspecteur ne s'aperçut de la chose.

Enfin, l'examen permet d'établir, au moins dans une certaine mesure, si le professeur ne concentre pas trop ses efforts sur certains élèves, au détriment des autres.

8. L'examen externe, feedback pour le professeur.

Tant pour notre santé mentale que pour la meilleure efficacité de notre action, nous avons besoin d'être informés de la validité de nos comportements.

L'examen bien conçu permet au professeur de juger de la valeur de certains aspects de son enseignement à travers les apprentissages réalisés par des élèves.

Il serait erroné d'évaluer les maîtres uniquement sur les résultats obtenus à court terme. Toutefois, seul ce genre d'évaluation semble actuellement pouvoir être effectué avec quelque rigueur scientifique. Or, tous ceux qui exercent une profession se rattachant aux sciences humaines souffrent, à vrai dire à des degrés fort divers, du manque d'information sûre sur la pertinence de leur action. C'est le souci du magistrat comme de l'assistant social, du professeur comme du prêtre.

Quand un maître peut-il dire avec certitude qu'il a bien rempli sa mission ?

*

**

En résumé, nous ne sommes partisan ni de supprimer complètement les examens, ni de renoncer entièrement à la notation subjective. Il importe d'adopter une façon de faire qui emprunte à chaque procédure ce qu'elle a de meilleur et de plus sûr. Nous allons essayer de voir comment pareil but pourrait être atteint.

chaque élève travaillera
semi-individualisé, mais,
ndre devront être définis
ogique.

rités organisatrices auront
signante a été accomplie.

nt sur tous les éléments
ussi une saine sauvegarde
acrer le temps nécessaire
rder est condamnable. Il
ngues qui, en deux ans,
du cours... Apparemment,
nose.

au moins dans une cer-
centre pas trop ses efforts
autres.

le professeur.

que pour la meilleure effi-
besoin d'être informés de

professeur de juger de la
nseignement à travers les
s.

maîtres uniquement sur les
efois, seul ce genre d'éva-
être effectué avec quelque
ni exercent une profession
s souffrent, à vrai dire à
d'information sûre sur la
ici du magistrat comme de
ne du prêtre.

avec certitude qu'il a bien

ntisan ni de supprimer com-
r entièrement à la notation
çon de faire qui emprunte
e meilleur et de plus sûr.
nt pareil but pourrait être

TROISIEME PARTIE

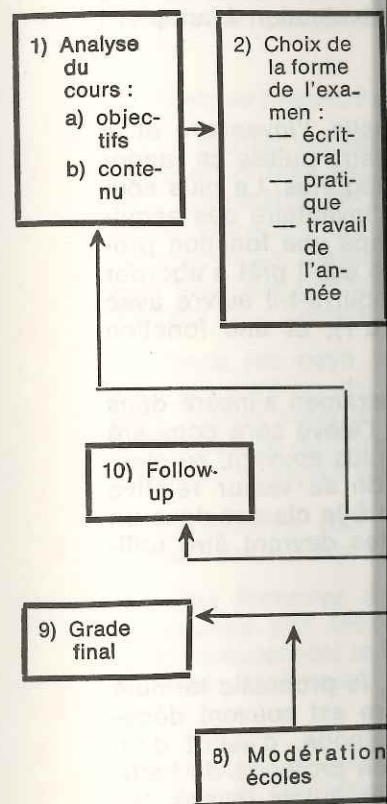
CONSTRUCTION DE L'EXAMEN

Comme les tests s
les maîtres devraient s
n'est, certes, pas possib
gation de routine, les m
tinée à des centaines
base restent les mêmes

Les grandes phase

1. Définir l'objet et les
2. Rédiger les question
3. Standardiser la prés
4. Etalonner ;
5. Eprouver la fidélité
6. Etablir la validité.

Le schéma suivan
parcourir.¹



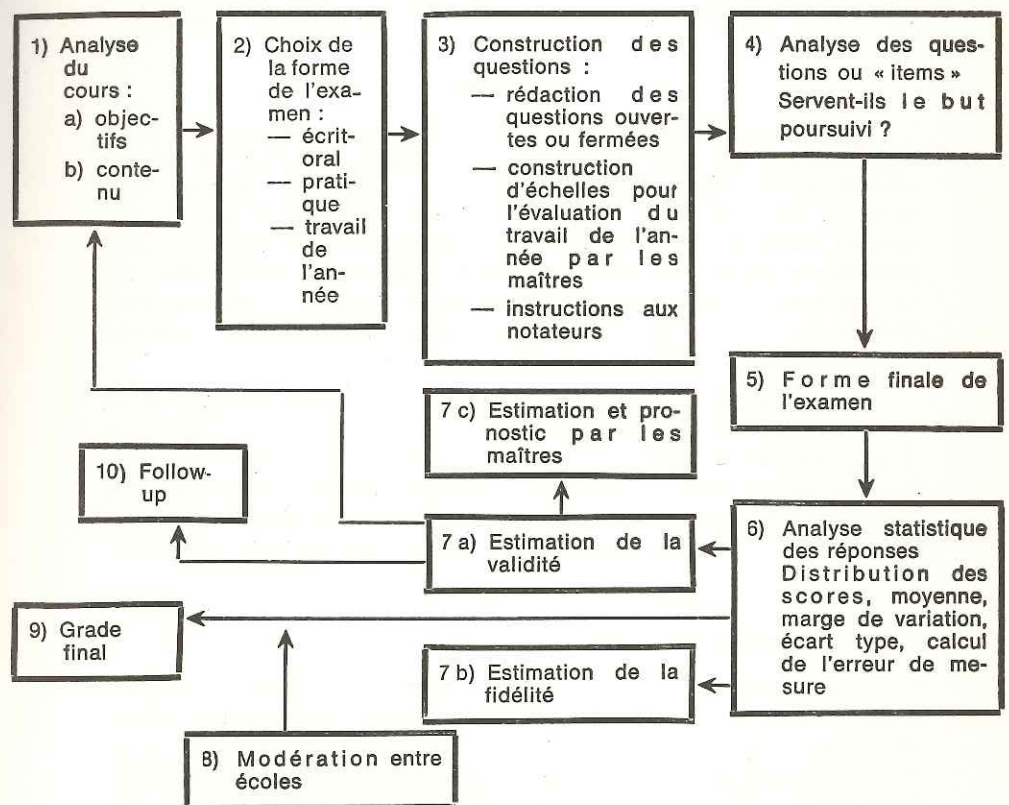
¹ Adapté d'après *Examina*

Comme les tests standardisés, les examens préparés par les maîtres devraient se construire par étapes bien définies. Il n'est, certes, pas possible de déployer, pour une simple interrogation de routine, les mêmes efforts que pour une épreuve destinée à des centaines d'élèves. Pourtant, les exigences de base restent les mêmes.

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

1. Définir l'objet et les objectifs ;
2. Rédiger les questions ;
3. Standardiser la présentation, l'administration et la correction ;
4. Etalonner ;
5. Eprouver la fidélité de l'examen ;
6. Etablir la validité.

Le schéma suivant montre le chemin que nous allons parcourir.¹



¹ Adapté d'après *Examinations Bulletin* n° 3, 1964, p. 33.

CHAPITRE 1.

L'OBJET ET LES OBJECTIFS.

Voir clairement la raison d'être de l'examen que l'on prépare et définir, avec le plus de précision possible, les types d'apprentissages à évaluer est probablement l'étape la plus décisive dans la construction d'une épreuve. Non seulement sa validité en dépend, mais aussi le mode d'évaluation à adopter !

I. L'objet.

Trois objets sont possibles : le *pronostic*, l'*inventaire* et le *diagnostic*. Il est rare qu'une épreuve scolaire puisse se ranger exclusivement dans l'une de ces trois catégories. Le plus souvent, les examens visent à faire le bilan, l'inventaire des acquisitions, mais ils remplissent en même temps une fonction pronostique à court ou à moyen terme (L'élève est-il prêt à aborder un nouveau chapitre du cours ? L'élève pourra-t-il suivre avec succès les cours de la classe supérieure ?), et une fonction diagnostique (Où l'élève achoppe-t-il ?).

Par ailleurs, il importe de savoir si l'examen s'insère dans l'action éducative quotidienne, — cas où l'élève sera comparé à lui-même et au groupe qui l'entoure (le plus souvent, sa classe), — ou s'il sert à informer l'individu de sa valeur relative dans sa communauté ou dans son pays ou à le classer dans un concours, — cas où des normes adéquates devront être utilisées.

A. Le pronostic.

A *long terme* (plus d'un an, au moins), le pronostic formulé à partir des résultats d'un examen scolaire est souvent décevant. Les changements de maître, de méthode, d'esprit d'un cours peuvent provoquer une transformation profonde de l'attitude chez l'élève. De plus, surtout chez les sujets jeunes, les intérêts manifestés connaissent d'importantes fluctuations. Il faut aussi tenir compte de l'évolution psychologique des élèves :

par exemple, les profonds changements qui se produisent bien l'incident de la vie scolaire.

On a maintes fois remarqué, en fait, que les tests d'intelligence sont plus sûrs que les tests de personnalité ; les tests de personnalité qui, surtout après une période de grande stabilité² ; elles sont plus verbal, spatial, numérique.

En outre, si les tests de personnalité comprennent les zones de personnalité les plus manifestées occasionnellement, ils assurent un succès scolaire et académique de sûreté.

A court terme, les tests de personnalité se comportent de la façon suivante :

1) Tests de maturité sociale

Ils portent sur les facteurs de personnalité qui interviennent dans l'apprentissage de la vie sociale, aussi l'utilisation de ce type de tests par le spécialiste. Quelques-uns sont toutefois conçus pour l'initiation à la vie sociale.

Dans les pays où les centres de recherche ont fait leurs efforts dans le domaine de la personnalité, nous avons montré ailleurs que les tests de personnalité pour la plupart des cas.

2) Vérification des conditions nécessaires aux acquisitions

Des épreuves de personnalité construites par les méthodes systématiques la liste des tests de personnalité antérieure⁴.

¹ Voir F. HOTYAT (1962), V. P. VERNON (1957), F. BACHER (1957).

² Certaines modifications peuvent être faites, mais le Quotient intellectuel n'est pas affecté.

³ Voir G. DE LANDSHEERE (1957).

⁴ Chaque notion critique doit être vérifiée.

par exemple, les professeurs d'enseignement secondaire connaissent bien l'incidence de la crise d'adolescence sur le rendement scolaire.

On a maintes fois démontré que, pour le pronostic à long terme, les tests d'intelligence sont au moins d'aussi bons prédicteurs que les tests de connaissances et sont beaucoup plus sûrs que les résultats scolaires¹. Ces tests explorent des aptitudes qui, surtout après l'âge de 8 ou 9 ans, connaissent une grande stabilité²; elles touchent le raisonnement, les facteurs verbal, spatial, numérique, ...

En outre, si les dominantes profondes de la personnalité, y compris les zones d'intérêt (et non des intérêts particuliers, manifestés occasionnellement) sont identifiées, la prédiction du succès scolaire et académique peut atteindre un niveau élevé de sûreté.

A court terme, le pronostic pédagogique s'établit de trois façons :

1) *Tests de maturité spécifique (readiness).*

Ils portent sur les formes de raisonnement, sur les aptitudes qui interviennent dans un apprentissage déterminé, par exemple : apprentissage de la lecture. La construction et souvent aussi l'utilisation de ces instruments complexes sont du ressort du spécialiste. Quelques tests permettant un premier dépistage sont toutefois conçus pour les maîtres, par exemple : le *Test d'Inizan* pour l'apprentissage de la lecture.

Dans les pays de langue française, il serait souhaitable que les centres de recherche augmentent considérablement leurs efforts dans le domaine de la maturité spécifique. Nous avons montré ailleurs combien nous sommes encore mal équipés pour la plupart des branches et à presque tous les niveaux³.

2) *Vérification des connaissances-clés ou notions critiques nécessaires aux acquisitions nouvelles.*

Des épreuves de ce genre peuvent être assez aisément construites par les maîtres, pour autant qu'ils aient dressé systématiquement la liste des erreurs commises lors d'un enseignement antérieur⁴.

1 Voir F. HOTYAT (1962), W. McCLELLAND (1945), E. PEEL et D. RUTTER (1951), P. VERNON (1957), F. BACHER (1965), etc.

2 Certaines modifications peuvent toutefois encore se produire. On sait aujourd'hui que le *Quotient intellectuel* n'est pas fixé une fois pour toutes.

3 Voir G. DE LANDSHEERE, *Les tests de connaissances*, Bruxelles, Editest, 1965.

4 Chaque notion critique doit figurer dans au moins trois exercices, nous y reviendrons.

Ici aussi, des fiches de recherche précisant le prérequis pour aborder les points importants de la matière rendraient aux maîtres d'inestimables services.

3) Essai.

Pour déterminer si un élève est prêt à aborder une matière nouvelle, on lui présente les premiers éléments de cette matière et l'on observe systématiquement les réactions. Une leçon programmée de type mixte (Skinner-Crowder) constitue souvent un heureux alliage d'enseignement et de contrôle analytique permettant un pronostic à court terme.

B. L'inventaire (épreuves de rendement).

Il a pour but premier de mesurer les apprentissages réalisés au cours d'une période plus ou moins longue. Sauf dans les cas de concours ou de *surveys* préparant des réformes, les inventaires remplissent aussi, à l'école, une fonction pronostique et diagnostique.

Il semble, en effet, contraire à toute attitude éducative saine qu'un maître constatant des faiblesses graves n'essaie pas d'en localiser la cause pour y remédier et ne s'inquiète pas des difficultés probables que les faiblesses annoncent pour les apprentissages ultérieurs.

Pour cette raison, notamment, il importe de distinguer soigneusement dans tout inventaire, les notions critiques, indispensables pour certains apprentissages ultérieurs, des notions marginales.

C. Le diagnostic.

Les épreuves de diagnostic ont pour objet de découvrir et d'expliquer les faiblesses et les habitudes défectueuses dans tous les domaines de l'apprentissage scolaire.

Sans nous aventurer dans la psychologie de l'échec (il est, dans certains cas, du ressort du psychanalyste), nous noterons, avec D. Durrell¹, que les causes les plus communes des difficultés scolaires sont :

1. Le manque d'aptitudes pour effectuer une tâche ;
2. La connaissance imparfaite des éléments de base ;
3. Un enseignement mal ajusté au niveau d'aptitudes de l'enfant et à sa vitesse d'apprentissage ;

¹ D. DURRELL, *Analysis of Reading Difficulties*, New York, Harcourt, 1955.

4. L'acquisition de mauvaises habitudes ;
5. L'inaptitude à transférer les techniques apprises ;
6. Le manque de vigueur et de persévérance devant tant d'échecs répétés.

Si des difficultés de la personnalité requièrent une attention particulière, n'oublie pas toutefois que certains des problèmes de l'élève proviennent de l'environnement.

Dans la pratique, le diagnostic devrait être une des premières tâches des maîtres et donc occuper une place importante. C'est qu'une sorte de mythe, parfois fait croire, l'enseignant doit maîtriser les notions-clés au moins, pour assurer à l'élève une performance satisfaisante.

Quelques bons tests de diagnostic que les maîtres ne peuvent pas ne pas utiliser seraient ceux qui ont été construits et évalués par eux-mêmes et qui sont simples et efficaces, spécifiques et sensibles.

On imagine aisément que, pour chacun des élèves, on peut faire un inventaire + ou par - si la réponse est oui ou non, à des exercices portant sur des notions-clés. On colorie en rouge les concepts-clés et en bleu les notions marginales. Il y a quelques années, un inventaire de diagnostic de la maîtrise d'une langue étrangère a été utilisé dans une classe.

Pour que l'épreuve de diagnostic soit efficace, il faut évidemment que l'élève ait une maîtrise précise de la matière et que l'épreuve soit conçue de façon à ne mettre en évidence que les faiblesses.

Comme l'élaboration d'un inventaire de diagnostic est souvent lourde, la collaboration des élèves est souvent souhaitable.

¹ Dans une épreuve de diagnostic, on ne doit pas tester la même notion, à des endroits différents, car l'élève peut éventuellement jouer par une distraction, le hasard, etc.

² Pour un exemple d'application de la méthode analytique d'arithmétique (fin de l'ouvrage), voir l'annexe 1961.

ne précisant le prérequis
la matière rendraient aux

rêt à aborder une matière
éléments de cette matière
réactions. Une leçon pro-
powder) constitue souvent
et de contrôle analytique

ent).

es apprentissages réalisés
longue. Sauf dans les cas
t des réformes, les inven-
une fonction pronostique

toute attitude éducative
faiblesses graves n'essaie
médier et ne s'inquiète pas
faiblesses annoncent pour les

il importe de distinguer
es notions critiques, indis-
es ultérieurs, des notions

pour objet de découvrir
habitudes défectueuses dans
scolaire.

chologie de l'échec (il est,
psychanalyste), nous noterons,
plus communes des diffi-

uer une tâche ;

éléments de base ;

niveau d'aptitudes de l'en-
e ;

, New York, Harcourt, 1955.

4. L'acquisition de mauvaises habitudes qui freinent le progrès ;
5. L'inaptitude à transférer, à découvrir le « système », à généraliser les techniques de solution ;
6. Le manque de vigueur, dans l'attaque des problèmes, résultant d'échecs répétés et du manque d'intérêt.

Si des difficultés scolaires surgissent, l'état physique et la personnalité requièrent aussi la plus grande attention. On n'oubliera pas toutefois que beaucoup de problèmes émotionnels de l'élève proviennent de l'échec au lieu d'en être la cause.

Dans la pratique quotidienne de l'école, le diagnostic fin devrait être une des préoccupations dominantes des professeurs et donc occuper un temps important. Contrairement à ce qu'une sorte de mythe de la courbe de Gauss (voir p. 180) a parfois fait croire, l'enseignement idéal serait celui qui, pour les notions-clés au moins, amènerait tous les élèves d'un groupe à une performance parfaite.

Quelques bons tests diagnostiques ont été publiés et les maîtres ne peuvent les ignorer. Mais, le plus souvent, ils pourraient construire eux-mêmes des instruments analytiques simples et efficaces, spécialement adaptés à leurs besoins.

On imagine aisément l'utilité d'un tableau d'ensemble où, pour chacun des élèves d'une classe, on voit apparaître par + ou par — si la réponse a été ou non correcte pour trois exercices portant sur un des points de la matière¹. Il suffit de colorier en rouge les cases qui contiennent deux ou trois signes moins pour obtenir une première impression de la situation². Il y a quelques années, nous avons dressé et tenu, avec la collaboration des élèves, un tableau de ce genre pour la prononciation d'une langue étrangère. Les résultats furent positifs.

Pour que l'épreuve puisse jouer son rôle diagnostique, il faut évidemment que les exercices portent sur un aspect très précis de la matière et qu'ils soient conçus, si possible, de façon à ne mettre en cause qu'une difficulté à la fois.

Comme l'élaboration de pareils contrôles devient facilement lourde, la collaboration de plusieurs professeurs est souvent souhaitable.

¹ Dans une épreuve diagnostique, on vérifie au moins trois fois la connaissance de la même notion, à des endroits différents du test, afin de réduire sérieusement le rôle éventuellement joué par une distraction momentanée, par une réponse correcte trouvée par hasard, etc.

² Pour un exemple d'application de ce système, voir : BONGRAIN, et al., *Epreuves analytiques d'arithmétique (fin du primaire)*, Morlanwelz, Institut Supérieur de Pédagogie, 1961.

II. Les objectifs.

« Il est curieux de constater que lorsqu'on demande aux personnes ayant charge d'éduquer les enfants de préciser les fins qu'elles poursuivent, on les plonge souvent dans la perplexité. »

P. Osterrieth, *Faire des adultes*, p. 9.

Affirmer que les maîtres doivent non seulement *instruire*, mais aussi *éduquer* est devenu un truisme. Encore importe-t-il de ne pas se contenter de vœux pieux, de préoccupations vagues. Certes, on enseigne ce que l'on est : notre personnalité, notre façon de penser et d'agir influencent directement, et presque malgré nous, nos élèves.

Mais, si nous voulons systématiser notre action, une définition précise des buts poursuivis devient nécessaire.

Eduquer, c'est mettre en œuvre les moyens propres à développer l'intelligence et la personnalité dans le sens voulu par le milieu culturel ; sans oublier que la santé physique exige aussi l'attention de l'éducateur.

Instruire signifie mettre quelqu'un en possession de connaissances nouvelles.

Les *objectifs généraux* relèvent de l'éducation, les *objectifs spéciaux*, de l'instruction. Mais les deux sont inséparables ; du moins, ils devraient l'être.

Evidemment, c'est *avant le début de l'année scolaire* que le professeur doit s'interroger sur les objectifs à atteindre (sinon il agit en aveugle) *et donc aussi définir la matière des examens.*

Certes, les plans d'études et les programmes scolaires officiels apportent une partie des réponses. Mais il est clair qu'elles n'apparaîtront qu'après une étude minutieuse et une longue méditation de ces documents de base. Il n'est d'ailleurs pas rare, surtout dans l'enseignement secondaire, que chaque branche fasse l'objet d'un programme séparé et que seuls les objectifs privilégiés soient explicitement traités. Il appartient alors au professeur d'ajouter les objectifs plus généraux qu'il poursuit en commun avec ses collègues des autres disciplines. Si le maître n'a pas une conscience vive des objectifs et ne les a pas faits siens, il a bien peu de chances de les atteindre.

Dresser un tableau des objectifs à poursuivre aide beaucoup le professeur au moment où il prépare sa matière de l'année. Ce tableau (voir exemple p. 66) contiendra autant de colonnes que d'objectifs généraux. Pour chaque chapitre ou

chaque étape du programme souhaite atteindre et (échelle à 3 ou 5 degrés)

L'appréciation de l'impact du programme échappe-t-elle d'une matière de base à en aborder d'autres ? Les compétences indispensables pour l'insertion dans la vie ?

Si subjective l'évaluation doit être, l'effort de réflexion est presque toujours, une constante.

A. LES

Au fond, l'objectif est de former l'adulte. P.A. Osterrieth sur ce thème¹.

Plus spécialement, il ne faut pas nuire à la santé physique, autant que possible, et poursuivre systématiquement ce que nous allons essayer.

Séparer les domaines de la théorie fait ou irrité ; aimer ou ne pas aimer la théorie. Peut-on penser sans penser ? On le sait, mais qui répond à une stimulation ? une analyse subtile de ses composantes cognitives ?

Notre propos est de contribuer à la recherche de l'enseignement et de l'apprentissage par arbitrairement le donner à ce dernier moment.

1. Les objectifs cognitifs

On doit à deux chercheurs, Guilford³, des classifications de types qui, malgré certaines

1 P.A. OSTERRIETH, *Faire des adultes*, premier chapitre : « Le but de l'éducation ».

2 B. BLOOM et al., *Taxonomie de Bloom*, traduit par M. Lavallée, Montréal, Éditions Érudition.

3 J.-P. GUILFORD, *Modèle de l'intelligence*, trad. de J.-P. Guilford, graphies ont été publiées sur ce thème en Californie, Los Angeles).

4 Voir spécialement la critique de...

ifs.

eux de constater que lorsqu'on
personnes ayant charge d'édu-
nts de préciser les fins qu'elles
on les plonge souvent dans la

errieth, *Faire des adultes*, p. 9.

non seulement *instruire*,
uisme. Encore importe-t-il
ieux, de préoccupations
on est : notre personnalité,
fluencent directement, et

er notre action, une défi-
vient nécessaire.

es moyens propres à déve-
é dans le sens voulu par
la santé physique exige

n en possession de con-

de l'éducation, les *objec-*
s deux sont inséparables ;

nt de l'année scolaire que
objectifs à atteindre (sinon
ir la matière des examens.

es programmes scolaires
éponses. Mais il est clair
étude minutieuse et une
de base. Il n'est d'ailleurs
t secondaire, que chaque
e séparé et que seuls les
ment traités. Il appartient
ectifs plus généraux qu'il
ues des autres disciplines.
ive des objectifs et ne les
chances de les atteindre.
s à poursuivre aide beau-
il prépare sa matière de
66) contiendra autant de
Pour chaque chapitre ou

chaque étape du programme, on indique les objectifs que l'on souhaite atteindre et l'importance relative à leur réserver (échelle à 3 ou 5 degrés).

L'appréciation de l'importance relative des différents points du programme échappe à la quantification rigoureuse : s'agit-il d'une matière de base qu'il faut nécessairement connaître pour en aborder d'autres ? s'agit-il de connaissances ou de capacités indispensables pour l'exercice d'une profession ou pour l'insertion dans la vie ?

Si subjective l'évaluation de l'importance puisse-t-elle parfois être, l'effort de réflexion qu'elle exige entraîne néanmoins, presque toujours, une clarification utile.

A. LES OBJECTIFS GENERAUX.

Au fond, l'objectif de l'enseignement est unique : faire un adulte. P.A. Osterrieth a consacré des pages pénétrantes à ce thème¹.

Plus spécialement, les maîtres doivent, d'une part, au moins ne pas nuire à la santé physique et mentale de leurs élèves et, autant que possible, aider à la développer, et, d'autre part, poursuivre systématiquement des objectifs cognitifs et affectifs que nous allons essayer de préciser.

Séparer les domaines cognitif (penser), affectif (être satisfait ou irrité ; aimer ou rejeter) et conatif (vouloir, désirer) est théorique. Peut-on penser sans éprouver de sentiment, agir sans penser ? On le sait bien, c'est toujours l'organisme total qui répond à une stimulation et, pour chaque comportement, une analyse subtile devrait révéler l'importance relative des composantes cognitive, affective et psycho-motrice.

Notre propos est bien plus modeste. Nous sommes actuellement à la recherche de points de repère qui serviront à jaloner enseignement et examens. C'est pourquoi nous allons couper arbitrairement le domaine cognitif du domaine affectif et donner à ce dernier mot une acception très large.

1. Les objectifs cognitifs.

On doit à deux chercheurs américains, B. Bloom² et J.-P. Guilford³, des classifications hiérarchisées des objectifs cognitifs qui, malgré certaines faiblesses⁴, se révèlent d'utiles outils.

1 P.A. OSTERRIETH, *Faire des adultes*, Bruxelles, Dessart, 1964. Voir spécialement le premier chapitre : « Le but de l'éducation ».

2 B. BLOOM et al., *Taxonomie des objectifs pédagogiques. I. Domaine cognitif*, traduit par M. Lavallée, Montréal, Education Nouvelle, 1969.

3 J.-P. GUILFORD, *Modèle tridimensionnel de l'intellect*. Une quarantaine de monographies ont été publiées sur ce sujet par Guilford et son équipe (University of South California, Los Angeles).

4 Voir spécialement la critique de J. CARROLL.

a) La taxonomie de Bloom.

En voici d'abord les grandes articulations ; un exemple d'application pratique est proposé page 68.

1. Connaître de mémoire.
2. Comprendre.
3. Appliquer.
4. Analyser.
5. Synthétiser.
6. Evaluer.

Chacun des échelons doit être accepté dans le sens retenu par les auteurs de la taxonomie¹. Par ailleurs, même si connaître de mémoire et évaluer apparaissent comme les comportements cognitifs, le moins et le plus nobles, cela n'implique en rien que la mémorisation doive disparaître de notre enseignement. Théoriquement au moins, il faut avoir franchi chaque échelon pour accéder au niveau supérieur. La taxonomie a été créée pour *aider l'éducateur à ne pas oublier* certaines étapes et pour l'inciter à élever graduellement le niveau de son enseignement.

1. CONNAITRE DE MEMOIRE.

- 1.1. Des données particulières : appellations, faits, dates, symboles.
- 1.2. Des façons de traiter des données particulières (sans les appliquer) : conventions, classifications, critères, méthodes.
- 1.3. Des données universelles : principes, lois, théories, ...

La différence entre les trois sous-catégories est plus quantitative que qualitative : savoir par cœur, sans plus, les dates de naissance des rois de France ou tel exposé de la philosophie de Kant nécessite surtout des efforts de mémorisation différents. Le niveau de pensée reste, dans les deux cas, très bas. Or, combien de questions d'examen, du primaire à l'enseignement supérieur, se situent à ce niveau (Que savez-vous de ... ? Quelles sont les clauses de... ? Comment prépare-t-on le... ?).

¹ Appliquer, par exemple, appelle une analyse plus ou moins fine. Ne faudrait-il donc pas placer l'analyse avant l'application ? Bloom et ses collaborateurs ont rétréci le sens d'*application* pour échapper à cette difficulté. Ils indiquent d'ailleurs à plusieurs endroits de leur ouvrage les recouvrements entre certains échelons et analysent, avec beaucoup de finesse, les difficultés d'utilisation de la *Taxonomie*. Notre but n'est pas d'en restituer toutes les nuances, mais bien d'attirer l'attention sur l'instrument.

2. COMPRENDRE.

Il s'agit ici de la compréhension d'une formulation nouvelle qui a dépassé le psittacisme, la répétition. Toutefois, il n'est pas facile de percevoir les rapports de cette compréhension avec d'autres situations.

Bloom distingue deux types de compréhension :

2.1. Traduire, transposer.

Le contenu de la situation est le même, son ordre soit modifié.

Exemples : Paraphraser un texte de la *taxonomie* es

Exprimer verbalement

A > B signifie

2.2. Interpréter.

Expliquer ou résoudre un problème. L'interprétation consiste à donner une nouvelle vue du contenu de la situation de reconnaître les liens de communication entre les éléments de la situation. « A » est le synonyme de l'analyse de l'évaluation¹.

Exemple : Interpréter une forme de table à deux dimensions, ou déduire des données, ou aller au-delà de ce qui est fourni : extrapoler.

2.3. Extrapoler.

3. APPLIQUER.

L'application suppose que la situation est connue, que la situation s'est donc produite.

Exemple : Emploi de la taxonomie pour résoudre des problèmes de

¹ B. BLOOM et al., *op. cit.*, p. 10.

2. COMPRENDRE.

Il s'agit ici de la compréhension au niveau le plus bas. Par une formulation nouvelle du donné, l'individu montre qu'il a dépassé le psittacisme, que le message a, pour lui, une signification. Toutefois, il n'est pas encore question d'appliquer, donc de percevoir les rapports du donné avec d'autres matériaux, d'autres situations.

Bloom distingue deux échelons :

2.1. Traduire, transposer.

Le contenu de la communication est conservé sans que son ordre soit modifié, mais la forme est changée.

Exemples : Paraphraser un récit, une proposition : « Une *taxonomie* est une *classification* ».

Exprimer verbalement des symboles mathématiques :

$A > B$ signifie que A est plus grand que B.

2.2. Interpréter.

Expliquer ou résumer une communication.

L'interprétation implique un nouvel arrangement, une nouvelle vue du matériel. Elle suppose donc la capacité de reconnaître et de saisir les idées maîtresses d'une communication et de comprendre les rapports existant entre elles. « A cet égard, l'interprétation devient synonyme de l'analyse et possède certaines caractéristiques de l'évaluation¹ ».

Exemple : Interpréter des données représentées sous forme de tableaux ou de graphiques en en tirant des déductions tenant compte des relations entre données, ou de leur signification d'ensemble.

Aller au-delà des données et des renseignements fournis : extension des tendances, généralisation.

2.3. Extrapoler.

3. APPLIQUER.

L'application suppose que le sujet distingue les traits communs à deux situations, à deux problèmes ; une abstraction s'est donc produite.

Exemple : Emploi de procédés expérimentaux pour résoudre des problèmes de travaux ménagers.

¹ B. BLOOM et al., o.c., p. 104.

4. ANALYSER.

4.1. Rechercher des éléments.

Exemple : Distinguer les faits des hypothèses dans une communication.

4.2. Rechercher des relations.

Exemple : Les hypothèses sont-elles logiques par rapport aux renseignements dont on dispose ?

4.3. Rechercher des principes d'organisation.

Exemple : Identifier les techniques de propagande utilisées dans des tracts.

5. SYNTHÉTISER.

5.1. Produire une œuvre personnelle.

Exemple : Narration captivante d'une expérience vécue.

5.2. Elaborer un plan d'action répondant aux exigences fixées.

5.3. Dériver un ensemble de relations abstraites. Induire une règle.

6. EVALUER.

Des jugements qualitatifs ou quantitatifs établissent dans quelle mesure le matériel et les méthodes répondent aux critères (internes ou externes).

Exemples : Déceler les sophismes dans une discussion. Apprécier un travail par comparaison à un modèle.

b) Le modèle de J.-P. Guilford.

Moins utilisé, jusqu'à présent, que la taxonomie de Bloom pour la construction des examens, le modèle de J.-P. Guilford offre peut-être de plus grandes possibilités encore, en raison de sa rigueur.

Voici d'abord comment J.-P. Guilford et R. Marrifels définissent les trois dimensions de l'intellect et leurs composantes¹ :

¹ Voir les *Reports from the Psychological Laboratory of the University of Southern California*, publiés depuis juin 1950.

LES OPERATIONS.

Ce sont les activités pures ; c'est ce que fait l'informationnelle, à

1. Cognition.

Conscience, appréhension, connaissance, compréhension.

2. Mémoire.

Rétention d'information.

3. Production convergente.

Génération d'information acceptées, à partir d'informations respectées.

4. Production divergente.

Génération d'information. Originalité, créativité.

5. Evaluation.

Prise de décisions ou l'exactitude, l'adéquation des critères, à des id

LES CONTENUS.

1. Figuratifs.

Information dans sa forme images.

Un minimum d'organisation. Intelligence pratique.

2. Symboliques.

Informations sous forme de symboles par et en eux-mêmes.

Intelligence théorique.

3. Sémantiques.

Informations sous forme de mots.

Intelligence verbale.

LES OPERATIONS.

Ce sont les activités ou les processus intellectuels principaux ; c'est ce que fait l'organisme à partir de la matière première informationnelle, à partir de ce qu'il discrimine.

1. Cognition.
Conscience, appréhension, découverte ou redécouverte, reconnaissance, compréhension d'informations sous diverses formes.
2. Mémoire.
Rétention d'informations.
3. Production convergente.
Génération d'informations uniques, conventionnellement acceptées, à partir d'un donné. L'usage, la coutume, la règle sont respectés.
4. Production divergente.
Génération d'informations variées à partir d'un même donné. Originalité, créativité.
5. Evaluation.
Prise de décisions ou formulation de jugements concernant l'exactitude, l'adéquation, la désirabilité, ... conformément à des critères, à des idéaux, à des objectifs adoptés.

LES CONTENUS.

1. Figuratifs.
Information dans sa forme concrète, perçue ou rappelée en images.
Un minimum d'organisation, de structuration est nécessaire.
Intelligence pratique.
2. Symboliques.
Informations sous forme de signes dépourvus de signification par et en eux-mêmes : lettres, nombres, notes de musique.
Intelligence théorique.
3. Sémantiques.
Informations sous forme de significations attachées à des mots.
Intelligence verbale.

4. Comportementaux.

Informations, essentiellement non verbales, intervenant dans les interactions humaines, où la perception d'attitudes, de besoins, de désirs, d'intentions, de pensées d'autrui et de soi-même jouent un rôle.

Intelligence sociale.

LES PRODUITS.

Ce sont les résultats du traitement des informations par l'organisme.

1. Unités.

Portions d'information relativement isolées ou circonscrites.

2. Classes.

Unités groupées en raison de leurs propriétés communes.

3. Relations.

Connexions reconnues entre des unités.

4. Systèmes.

Groupements d'unités organisées ou structurées ; complexes de parties se trouvant en interrelation ou en interaction.

5. Transformations.

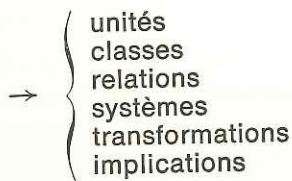
Changements apportés dans des informations ou dans leur utilisation.

6. Implications.

Extrapolation d'informations : prédiction, conséquences, antécédents.

Chacune des composantes des trois dimensions se combine avec toutes les autres. Ainsi :

Mémoire des



soit 24 combinaisons.

Comme on distingue en tout, cent vingt combi

OPERATIONS Jugement _____

Production divergen

Production converg

Mémoire _____

Cognition _____

PRODUITS Unités _____

Classes _____

Relations _____

Systèmes _____

Transformations _____

Implications _____

CONTENUS Figurax _____

Symboliques _____

Sémantiques _____

Comportementau

L'expérience montr les combinaisons en te

Dans un premier retiendront surtout l'att la scolarité, le maître d toutes les activités.

verbales, intervenant dans perception d'attitudes, de de pensées d'autrui et de

ment des informations par

nt isolées ou circonscrites.

urs propriétés communes.

unités.

es ou structurées; com-interrelation ou en interac-

informations ou dans leur

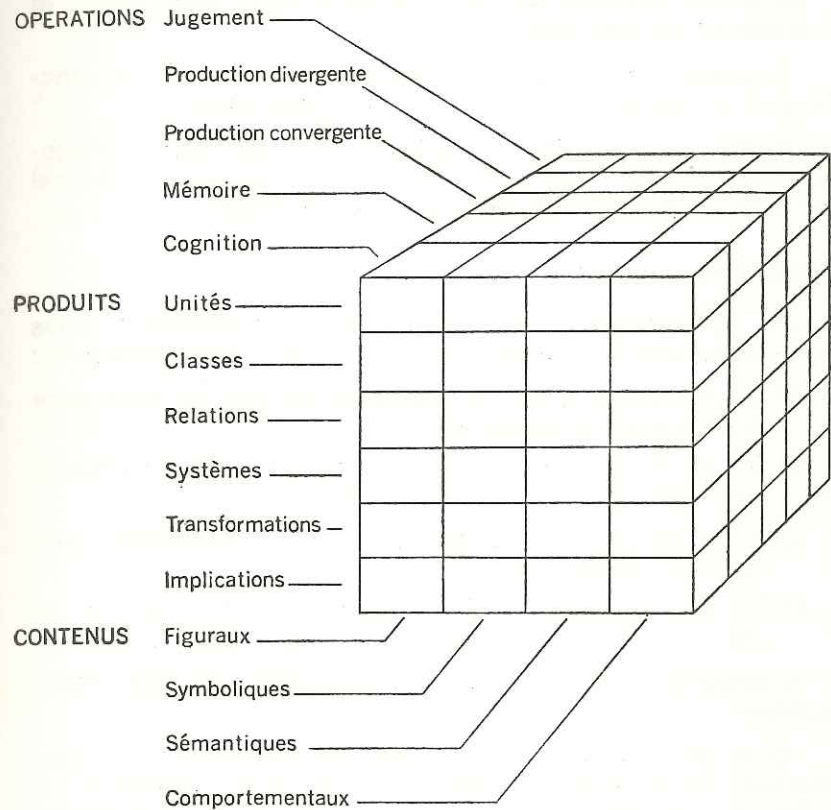
prédiction, conséquences,

trois dimensions se com-

→ { figuratifs
symboliques
sémantiques
comportementaux

soit 24 combinaisons.

Comme on distingue cinq types d'opérations, il existe donc, en tout, cent vingt combinaisons dans le modèle.



L'expérience montre qu'il est difficile de traduire toutes les combinaisons en termes utilisables pour l'enseignement.

Dans un premier temps, les quatre types d'opérations retiendront surtout l'attention. Idéalement, et dès le début de la scolarité, le maître devrait s'efforcer de les introduire dans toutes les activités.

1. La restitution de mémoire témoigne de la fixation de la notion.
2. La production convergente témoigne de la capacité d'appliquer les acquisitions en conformité avec les règles, les usages.
3. La production divergente témoigne de la capacité de découvrir des solutions ou des usages nouveaux.
4. Enfin, l'évaluation témoigne de la capacité de critiquer, de porter un jugement en fonction de critères bien définis et d'agir en conséquence.

Trop souvent, l'activité scolaire et les examens qui la sanctionnent se limitent aux deux premières opérations.

S'efforcer de garder un équilibre entre les quatre catégories des contenus importe autant dans l'enseignement général que dans l'enseignement technique ou artistique.

2. Domaine affectif.

En schématisant l'analyse fine de P. Osterrieth¹, nous dirons qu'un individu accède pleinement à la condition d'adulte :

- 1° a) Si son comportement a trouvé sa logique, sa cohérence et échappe à la versatilité ;
 - b) S'il a, par ailleurs, acquis une saine tolérance au changement, à la contradiction, à l'échec ;
- 2° a) S'il a conquis son indépendance, son autonomie intellectuelle et affective ;
 - b) S'il est cependant capable de faire le don de soi, de rester fidèle à ses engagements et à ses sentiments.

L'éducation est un long acheminement vers cet ultime équilibre.

Nous proposons ci-dessous une adaptation et une interprétation de la taxonomie des objectifs de D. Kratochvil et B. Bloom². On obtient de nouveau cinq échelons, cinq étapes qui acheminent du comportement le plus passif au plus actif.

¹ P. OSTERRIETH, *Faire des adultes*, Bruxelles, Dessart, 1964.

² D.R. KRATOWHL, B.S. BLOOM, B. MASIA, *Taxonomie des objectifs pédagogiques, Domaine affectif*, Montréal, Éducation nouvelle, 1970.

I. L'individu répo

1. Est simplement réce

C'est une sorte d'é la beauté ou la laideur, peu comme un miroir d

Ce comportement e simple cognition qui p certain éveil de l'attenti musique, écoute parler l

2. Reçoit et réagit.

L'individu réagit ne festant du plaisir, par ce stade, on n'observe gnerait d'un choix délibé

Pour le professeur élèves n'ont pas enco choix personnel, où le raffinée pour leur perm mais où, mis en présen à en sentir la grandeur.

3. Reçoit et réagit en

Maintenant, l'indiv condition d'être mis en p il s'engage.

II. L'inc

4. Essaie spontanément

L'individu éprouve struire sans y être invité initiative sentimentale o le sens des valeurs po religion.

5. Agit selon ses optio

C'est le stade psych P. Osterrieth.

I. L'individu répond à une stimulation extérieure.

1. Est simplement réceptif.

C'est une sorte d'état affectif amorphe où le sujet perçoit la beauté ou la laideur, les sentiments divers, sans réagir, un peu comme un miroir qui ne renverrait pas l'image.

Ce comportement est d'ailleurs difficile à distinguer de la simple cognition qui précède la mise en mémoire. Seul, un certain éveil de l'attention est observable. Exemple : Ecoute la musique, écoute parler les autres.

2. Reçoit et réagit.

L'individu réagit nettement soit en obéissant, soit en manifestant du plaisir, par la parole, par le geste ou l'attitude. A ce stade, on n'observe pas encore de rejet explicite qui témoignerait d'un choix délibéré.

Pour le professeur de littérature, c'est le moment où les élèves n'ont pas encore le goût assez formé pour faire un choix personnel, où leur sensibilité n'est pas encore assez raffinée pour leur permettre de partir seuls à la découverte, mais où, mis en présence de belles œuvres, ils commencent à en sentir la grandeur.

3. Reçoit et réagit en acceptant ou en refusant.

Maintenant, l'individu sait ce qu'il veut ou ce qu'il aime, à condition d'être mis en présence des personnes ou des choses ; il s'engage.

II. L'individu prend l'initiative.

4. Essaie spontanément de comprendre, de juger, de ressentir.

L'individu éprouve assez d'intérêt, de curiosité, pour s'instruire sans y être invité, assez de sensibilité pour prendre une initiative sentimentale ou, encore, il a suffisamment découvert le sens des valeurs pour se choisir une philosophie ou une religion.

5. Agit selon ses options.

C'est le stade psychologiquement adulte, comme l'a défini P. Osterrieth.

Par exemple, l'individu vit en fonction de ses options morales, sentimentales, esthétiques, mais il est aussi capable de changer de conduite à la lumière de preuves, d'arguments convaincants.

Cette ultime étape de l'ascension affective correspond à l'évaluation dans le domaine cognitif.

B. LES OBJECTIFS SPECIAUX.

Théoriquement, toutes les matières, tous les points des programmes scolaires offrent l'occasion de se rapprocher des objectifs généraux et de vérifier s'ils ont été atteints.

Trois remarques s'imposent toutefois :

- 1° La pratique montre qu'une même épreuve ne peut porter que sur un nombre limité d'objectifs ;
- 2° Certaines matières se prêtent mieux que d'autres à la poursuite de certains objectifs ;
- 3° Le grave problème du transfert des apprentissages continue à se poser avec acuité. Par exemple, cultiver la divergence à l'occasion des activités artistiques ne garantit en rien que ce trait se manifesterait dans les domaines scientifiques ou dans la vie pratique.

Quoi qu'il en soit, la première démarche visant à l'identification des objectifs spéciaux reste fondamentalement toujours la même et consiste en l'élaboration d'un tableau à double entrée. En haut, on porte les objectifs généraux ; sur le côté, on inscrit les matières du cours. Chaque intersection de colonne et de rangée indique, en principe, un objectif spécial.

Exemple : CHIMIE ¹.

	Connaissance des faits, méthodes et techniques	Application	Évaluation
Équilibres ioniques ¹			
1. Généralités			
1. Degré d'ionisation ou fraction ionisée			
2. Mesures du degré d'ionisation			
3. Application de la loi d'action des masses aux équilibres ioniques			
4. Loi de la dilution d'Ostwald			
2. Produit de solubilité			
1. Définition			
2. Calcul du produit de solubilité en fonction de la solubilité			

¹ Extrait de : BERGER et DIGHAYE, *Chimie IV*, Liège, Sciences et Lettres, 1967.

- 3. Diminution de la solubilité d'un électrolyte peu soluble
- 4. Précipitation d'un électrolyte par adjonction, à sa solubilité, d'un électrolyte à ion commun
- 5. Dissolution des précipités

Appendice : Application du produit de solubilité à la précipitation

- 3. Produit ionique de l'eau
 - 1. Produit ionique de l'eau
 - 2. Le pH
 - a La notation pH
 - b pH de l'eau pure
 - c L'échelle des pH
 - d Détermination du pH
 - e Les indicateurs colorés
 - f Méthodes expérimentales de détermination du pH
 - g Distinction entre acides et bases faibles, par détermination de solutions diluées de concentrations connues
 - h Calcul du pH de solutions d'acides et de bases de leurs concentrations et — éventuellement — de leurs constantes d'équilibre
- 4. Solutions tamponnées
 - 1. Introduction expérimentale
 - 2. Calcul du pH d'une solution tampon
- 5. Hydrolyse
 - 1. Généralités
 - 2. Cas de sels dérivant d'un acide fort et d'une base forte
 - 3. Cas des sels dérivant d'un acide fort et d'une base faible
 - 4. Cas des sels dérivant d'un acide faible et d'une base forte
 - 5. Cas des sels dérivant d'un acide faible et d'une base faible
 - 6. Facteurs influençant l'hydrolyse
- 6. Neutralisation mutuelle d'un acide et d'une base
 - 1. Chaleur de neutralisation
 - 2. Courbes de neutralisation
 - a Acide fort - base forte
 - b Acide faible - base forte
 - c Base faible - acide fort
- 7. Les méthodes d'analyse chimique par voie chimique
 - 1. Méthodes gravimétriques
 - 2. Méthodes volumétriques
 - Acides - bases
 - Choix des indicateurs
 - Titrages rédox

ction de ses options mora-
s il est aussi capable de
de preuves, d'arguments

ion affective correspond à

SPECIAUX.

ières, tous les points des
sion de se rapprocher des
'ils ont été atteints.

efois :

ne épreuve ne peut porter
tifs ;

ieux que d'autres à la pour-

des apprentissages continue
mple, cultiver la divergence
ques ne garantit en rien que
les domaines scientifiques

démarche visant à l'identifi-
fondamentalement toujours
ion d'un tableau à double
fs généraux ; sur le côté, on
que intersection de colonne
un objectif spécial.

CHIMIE¹.

Connaissan- ce des faits, méthodes et techniques	Appli- cation	Evalua- tion

	Connaissance des faits, méthodes et techniques	Appli- cation	Evaluation
3. Diminution de la solubilité d'un électrolyte peu soluble			
4. Précipitation d'un électrolyte par adjonction, à sa solution saturée, d'un électrolyte apportant un ion commun			
5. Dissolution des précipités			
Appendice : Application du produit de solubilité à la précipitation des sulfures			
3. Produit ionique de l'eau et pH			
1. Produit ionique de l'eau			
2. Le pH			
a La notation pH			
b pH de l'eau pure			
c L'échelle des pH			
d Détermination du pH			
e Les indicateurs colorés			
f Méthodes expérimentales de détermination du pH			
g Distinction entre électrolytes (acides et bases) forts et faibles, par détermination du pH de solutions diluées de concentrations connues			
h Calcul du pH des solutions d'acides et de bases à partir de leurs concentrations molaires et — éventuellement — de leurs constantes d'ionisation			
4. Solutions tamponnées			
1. Introduction expérimentale			
2. Calcul du pH d'un mélange-tampon			
5. Hydrolyse			
1. Généralités			
2. Cas de sels dérivant d'un acide fort et d'une base forte			
3. Cas des sels dérivant d'un acide fort et d'une base faible			
4. Cas des sels dérivant d'un acide faible et d'une base forte			
5. Cas des sels dérivant d'un acide faible et d'une base faible			
6. Facteurs influençant l'hydrolyse			
6. Neutralisation mutuelle des acides et des bases			
1. Chaleur de neutralisation			
2. Courbes de neutralisation			
a Acide fort - base forte			
b Acide faible - base forte			
c Base faible - acide fort			
7. Les méthodes d'analyse quantitative par voie chimique			
1. Méthodes gravimétriques			
2. Méthodes volumétriques			
Acides - bases			
Choix des indicateurs			
Titrages rédox			

La matière de chimie figurant dans le tableau ci-dessus constitue un des quatre chapitres d'un cours destiné à l'année supérieure de l'enseignement secondaire. Parmi bien d'autres objectifs généraux possibles, nous n'en avons retenu que trois.

Un simple coup d'œil sur l'ensemble révèle que, même dans ces conditions très simplifiées, le professeur se trouve devant un choix difficile. Sur quels points va-t-il faire porter l'examen pour obtenir un échantillonnage suffisant de la matière ? Quelle importance relative va-t-il réserver aux matières retenues et aux divers objectifs spéciaux qui y correspondent ? En quoi consiste exactement l'évaluation dans un cours comme celui-ci ?

On le voit, le problème est loin d'être simple. Il dépasse presque toujours la compétence d'un seul homme et débouche sur des questions restées sans réponse. Il est souhaitable de créer des commissions de définition d'objectifs où enseignants, psychologues et spécialistes de la recherche en éducation unissent leurs efforts.

Voici maintenant une série d'*items* de sciences illustrant les principaux niveaux de la hiérarchie de B. Bloom.

TAXONOMIE DE BLOOM

Illustration : Sciences naturelles¹

1. Connaissance.

(a) Connaissance de données spécifiques

Les glucides et les lipides sont utilisées par l'organisme :

- A. Comme aliments énergétiques
- B. Pour former du nouveau cytoplasme
- C. Comme source de sels minéraux
- D. Comme aliments plastiques
- E. Comme source de vitamines

¹ Adapté d'après D.G. LEWIS, Objectives in the Teaching of Science, in *Educational Research*, June 1965, pp. 186-199.

(b) Connaissance de données spécifiques

Équilibrez l'équation et comptez le nombre de molécules.

- A. 4
- B. 18
- C. 5
- D. 9
- E. autre nombre

(c) Connaissance de principes

Une machine simple

- A. Changer la direction
- B. Substituer un mouvement
- C. Offrir un gain
- D. Vaincre une résistance appliquée
- E. Augmenter l'efficacité

2. Compréhension (l'élève doit être capable de mentalement établir des rapports)

Si, par temps très chaud, on chauffe un poteau en acier, il se dilatera. Pourquoi ne se dilaterait-il pas si le poteau était en bois ?

C'est parce que l'acier

- A. A un point de fusion plus élevé
- B. Rayonne plus
- C. A une masse plus
- D. Est meilleur conducteur de la chaleur
- E. Possède des molécules plus serrées

3. Application (abstraire l'essence de la situation, transférer l'abstraction d'une situation à une autre)

Quand le magnésium est chauffé, il se combine avec l'oxygène. Si on chauffe 3 grammes de magnésium dans un récipient fermé, combien de grammes de magnésium oxyde se formeront ?

dans le tableau ci-dessus
l'un cours destiné à l'année
ndaire. Parmi bien d'autres
n'en avons retenu que trois.

nsemble révèle que, même
es, le professeur se trouve
s points va-t-il faire porter
onnage suffisant de la ma-
ya-t-il réserver aux matières
ciaux qui y correspondent ?
ation dans un cours comme

in d'être simple. Il dépasse
un seul homme et débouche
ponse. Il est souhaitable de
n d'objectifs où enseignants,
la recherche en éducation

items de sciences illustrant
chie de B. Bloom.

E BLOOM

es naturelles¹

écifiques

nt utilisées par l'organisme :

iques

cytoplasme

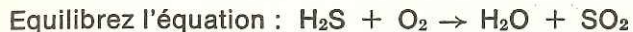
minéraux

es

es

n the Teaching of Science, in Educa-

(b) Connaissance de la manière de traiter des informations
spécifiques



et comptez le nombre de molécules pour l'ensemble.

- A. 4
- B. 18
- C. 5
- D. 9
- E. autre nombre que les précédents

(c) Connaissance des principes généraux

Une machine simple ne peut :

- A. Changer la direction d'une force appliquée
- B. Substituer une grande force à une petite force
- C. Offrir un gain de 100 entre les forces mises en jeu
- D. Vaincre une résistance plus grande que la force appliquée
- E. Augmenter l'énergie du système

2. *Compréhension* (l'élève connaît et sait faire usage de la
matière qui lui a été communiquée sans pouvoir nécessaire-
ment établir des rapports avec d'autres matières).

Si, par temps très froid, un enfant porte la main sur un
poteau en acier, il se peut que la main adhère, ce qui n'arri-
verait pas si le poteau était en bois.

C'est parce que l'acier :

- A. A un point de fusion plus élevé que le bois
- B. Rayonne plus la chaleur que le bois
- C. A une masse spécifique plus élevée que le bois
- D. Est meilleur conducteur de la chaleur que le bois
- E. Possède des propriétés magnétiques alors que le bois pas.

3. *Application* (abstraire d'une situation particulière et appli-
quer l'abstraction d'une autre façon).

Quand le magnésium brûle pour former de l'oxyde de ma-
gnésium, MgO, deux grammes d'oxygène se combinent à
3 grammes de magnésium. Si l'on brûle 6 grammes de
magnésium dans un récipient fermé qui contient 3 g d'oxy-
gène :

- A. Il n'y aura pas de réaction
- B. Il se formera du Mg_2O
- C. Il se formera du MgO_2
- D. Une certaine quantité de magnésium restera
- E. L'oxyde de magnésium contiendra un moindre pourcentage d'oxygène

4. *Analyse.*

Si deux objets, X et Y, ont le même moment (momentum), l'énergie cinétique de Y ne peut être supérieure à celle de X que si :

- A. Y est plus lourd que X
- B. Y se déplace plus vite que X
- C. Y a le même poids que X
- D. Y se déplace plus lentement que X
- E. Y est situé à une distance plus grande de l'axe que X

5. *Synthèse* (combiner les éléments pour former un tout structuré).¹

Un certain nombre de problèmes qui ont retenu l'attention des hommes de science peuvent être classés sous la rubrique « structure de la matière ». En ce domaine, les premières recherches portèrent sur la structure discontinue de la matière, hypothèse émise par Dalton en 1805.

Un autre groupe de recherches concerne la théorie cinétique des gaz. Elle apparaît déjà au XVII^e siècle, puis est reprise de façon plus complète par Maxwell (1859) et Boltzmann (1896).

Montrez comment la théorie cinétique des gaz constitua un argument favorable à l'hypothèse de la représentation atomistique de la matière.

Votre discussion doit comprendre :

1. Un exposé des hypothèses fondamentales de cette théorie ;
2. Quelques déductions importantes de cette théorie, appliquée au cas des gaz parfaits ;
3. L'interprétation des lois expérimentales de Gay-Lussac et de l'hypothèse d'Avogadro.

¹ Cet item est dû à H. Guillaume.

6. *Evaluation* (jugements à une démonstration ou

La théorie atomique de

- A. Un bon exemple
- B. Un bon exemple de science
- C. Exacte et a permis
- D. Partiellement fautive
- E. Ni correcte, ni occupé.

Voici enfin quelques à trois aspects caractérisés

6. *Evaluation* (jugements sur la valeur des arguments servant à une démonstration ou sur des critères).

La théorie atomique de Dalton est :

- A. Un bon exemple de vérité absolue
- B. Un bon exemple d'une conception statique de la science
- C. Exacte et a permis d'autres découvertes scientifiques
- D. Partiellement fausse, mais a permis d'autres découvertes scientifiques
- E. Ni correcte, ni incorrecte, et on ne s'en est plus occupé.

Voici enfin quelques exemples de questions se rapportant à trois aspects caractéristiques du modèle de J.-P. Guilford :

magnésium restera
contiendra un moindre pour-

ème moment (momentum),
être supérieure à celle de

t que X
plus grande de l'axe que X

pour former un tout struc-

s qui ont retenu l'attention
t être classés sous la ru-
. En ce domaine, les pre-
a structure discontinue de
Dalton en 1805.

concerne la théorie ciné-
au XVII^e siècle, puis est re-
r Maxwell (1859) et Boltz-

ique des gaz constitua un
e de la représentation ato-

:
fondamentales de cette théo-

ntes de cette théorie, ap-
its ;

riminentales de Gay-Lussac

Branche	Production convergente	Production divergente	Evaluation
Sciences	Expliquez pourquoi il ne peut pas y avoir de vie sur Mercure.	En quoi la vie sur Mars pourrait-elle différer de la nôtre ?	Pensez-vous qu'il existe une vie sur Mars ?
Géographie	En quoi le détroit de Bering a-t-il influencé le développement de l'Amérique du Nord ?	Que serait-il arrivé si le détroit de Bering n'avait pas existé ?	Actuellement, quel est, à votre avis, le principal rôle du détroit de Bering ?
Histoire	Expliquez l'importance de la découverte de l'Amérique par C. Colomb sur la vie économique de l'Europe.	Que serait-il arrivé si C. Colomb avait découvert la route des Indes au lieu de l'Amérique ?	Quelles sont, à votre avis, les deux conséquences les plus importantes du voyage de C. Colomb ?
Langue maternelle	Expliquez pourquoi la nouvelle s'est beaucoup plus développée aux Etats-Unis qu'en Europe ?	Voici le début d'une nouvelle. Imaginez autant de dénouements que vous le pouvez.	Qu'est-ce qui importe le plus dans une nouvelle : les caractères ou l'intrigue ?

1 D'après J.-R. VERDUIN Jr., *Conceptual Models in Teacher Education*, Washington, A.A.C.T.E., 1967, p. 93.

REDACTI

I. Obs

Pour être équitable toujours comporter un domaine à couvrir est étérger que sur une fraction sur une exigence difficile système d'enseignement élèves sont censés avoir apprentissages est suppo arbitrairement sur une pporter un jugement valable

En réalité, si, par m une des rares parties qu tième de l'ensemble, pa quemment aussi, une sim expliquent une mauvaise bien connu.

Nous avons déjà vu teurs essaient, en général question, sous des formes l'usage dans les examens

Par ailleurs, la mu ainsi que le modèle de C ment concevable un peti

Pourtant, il serait irr savante pour le moindre

Savoir poser des que plus nécessaire au profé erreurs dans le choix du

1 Voir 5^e partie : Le mythe de la

Langue maternelle	Expliquez pourquoi la nouvelle s'est beaucoup plus développée aux Etats-Unis qu'en Europe ?	Voici le début d'une nouvelle. Imaginez autant de dénouements que vous le pouvez.	Qu'est-ce qui importe le plus dans une nouvelle : les caractères ou l'intrigue ?
	Expliquez pourquoi la découverte de l'Amérique par C. Colomb sur la vie économique de l'Europe.	Colomb avait découvert la route des Indes au lieu de l'Amérique ?	les deux conséquences les plus importantes du voyage de C. Colomb ?

1 D'après J.-R. VERDUIN Jr., *Conceptual Models in Teacher Education*, Washington, A.A.C.T.E., 1967, p. 93.

CHAPITRE 2.

REDACTION DES QUESTIONS.

I. Observations générales.

Pour être équitable et valide, un examen doit presque toujours comporter un grand nombre de questions. Si le domaine à couvrir est étendu, la chose semble évidente. N'interroger que sur une fraction, parfois minime, de la matière repose sur une exigence difficilement justifiable, surtout dans notre système d'enseignement actuel, non individualisé¹ : tous les élèves sont censés avoir tout appris et la qualité de tous leurs apprentissages est supposée homogène. Des questions portant arbitrairement sur une partie du tout permettraient donc de porter un jugement valable.

En réalité, si, par malchance, le professeur interroge sur une des rares parties qu'un étudiant n'a pas étudiée, un vingtième de l'ensemble, par exemple, c'est la catastrophe. Fréquemment aussi, une simple distraction, une fatigue passagère expliquent une mauvaise réponse dans un domaine pourtant bien connu.

Nous avons déjà vu que, pour prévenir ce danger, les testeurs essaient, en général, de poser au moins trois fois la même question, sous des formes différentes, ce qui n'est certes guère l'usage dans les examens traditionnels.

Par ailleurs, la multiplicité des objectifs à poursuivre, ainsi que le modèle de Guilford nous l'a révélé, rend difficilement concevable un petit nombre de questions.

Pourtant, il serait irréaliste de préconiser une construction savante pour le moindre exercice de contrôle ou de diagnostic.

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

¹ Voir 5^e partie : Le mythe de la courbe de Gauss.

fréquemment rencontrées dans le libellé des exercices d'application et les questions d'examen en témoignent.

A. Des questions compréhensibles.

Bien des réponses erronées ne sont pas dues à l'ignorance de la matière, mais bien à des malentendus, à une mauvaise compréhension des questions. Le danger apparaît spécialement en un domaine comme les mathématiques parce que trois difficultés s'y combinent aisément : la difficulté inhérente au problème, la difficulté du langage abstrait des mathématiques, et les faiblesses linguistiques des mathématiciens eux-mêmes.

Mais, on le sait, l'abstrus n'épargne pas non plus les historiens, les géographes, les professeurs de sciences... ou de langues.

En ne s'assurant pas de la clarté des questions, les maîtres risquent, en particulier, de commettre une injustice sociale. La recherche contemporaine confirme que beaucoup d'enfants issus de milieux socialement défavorisés souffrent de handicaps graves dans le domaine du langage. Davis et Haggard ont, par exemple, montré qu'il suffit de modifier la forme d'un problème, sans en changer le sens, pour que la différence de réussite entre enfants provenant de milieux socio-économiques favorisés ou non passe de 12 à 32 %¹.

B. Tenir compte du niveau d'information.

Pour vérifier la capacité à transférer les apprentissages, il est souvent nécessaire de situer les problèmes dans un contexte non encore évoqué en classe. Encore faut-il que ce contexte ait un sens pour l'élève.

Tels petits problèmes, récemment rencontrés dans un manuel, supposaient la connaissance des règles du jeu de tennis ou de la manière de remplir des bulletins de pronostic de football. Que signifieront ici les échecs si les élèves n'ont pas la possibilité de s'informer du sens réel des questions avant d'essayer d'y répondre ?

C. Essayer ou prétester les questions.

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 A. DAVIS, Education for the Conservation of Human Resources, in *Progressive Education*, 1950, 27, 221-224.

Pareil essai est difficile ; de nouveau, le travail est isolé ; de nouveau, les faiblesses en ce domaine, à l'avance.

D. Calcul de la facilité

S'il est possible de faire ce calcul se fera avant l'examen de procéder à l'opération de détermination des réponses, soit

Le *pourcentage de réussite*. Rappelons, toutefois, que le *pourcentage de réussite* à partir de la somme de sujets devrait excéder

Une fois l'indice de facilité comment doser l'examen.

E. Calcul de l'efficacité

L'indice de facilité se calcule parfois de facteurs accidentels. L'indice d'efficacité d'une question déterminée est faible.

1. Méthode simple.

Pidgeon et Yates proposent une méthode fort utile pour les enseignants.

- Diviser la classe en trois groupes, sur la base des capacités.
- Pour chaque *item*, voir le nombre de groupes qui a réussi.
- Pour chaque *item*, la réussite du groupe supérieur est une bonne estimation de la difficulté. Si la différence est faible, on ne distingue pas bien les

Plus tard, nous verrons comment calculer plus la distribution des réponses (s'aplatit).

ellé des exercices d'appli-
témoinent.

ont pas dues à l'ignorance
entendus, à une mauvaise
nger apparaît spécialement
iques parce que trois diffi-
difficulté inhérente au pro-
ait des mathématiques, et
athématiciens eux-mêmes.

gne pas non plus les histo-
urs de sciences... ou de

des questions, les maîtres
e une injustice sociale. La
ue beaucoup d'enfants is-
és souffrent de handicaps
Davis et Haggard ont, par
ifier la forme d'un problè-
ue la différence de réussite
ocio-économiques favorisés

ation.

féérer les apprentissages, il
problèmes dans un contexte
re faut-il que ce contexte

ment rencontrés dans un
e des règles du jeu de ten-
s bulletins de pronostic de
ecs si les élèves n'ont pas
réel des questions avant

ns.

oyer les questions avant de
uvre ainsi les imprécisions,
s matérielles et, aussi, le

of Human Ressources, in *Progressive*

Pareil essai est difficilement réalisable par un enseignant isolé ; de nouveau, le travail en groupe offre bien des possibilités en ce domaine, l'examen étant préparé longtemps à l'avance.

D. Calcul de la facilité des questions.

S'il est possible de prétester sur un grand nombre de sujets, ce calcul se fera avant l'examen. Sinon, il est néanmoins utile de procéder à l'opération, soit pour mieux percevoir la physiologie des réponses, soit pour un usage ultérieur.

Le *pourcentage de réussite* est l'indice le plus simple. Rappelons, toutefois, qu'il n'est pas correct de calculer des pourcentages à partir de petits nombres. Normalement, le nombre de sujets devrait excéder cent.

Une fois l'indice de facilité connu, le professeur sait mieux comment doser l'examen.

E. Calcul de l'efficacité - Pouvoir discriminatif.

L'indice de facilité seul peut induire en erreur, car il résulte parfois de facteurs accidentels. Il est donc prudent de l'accompagner de l'*indice d'efficacité*, qui révèle dans quelle mesure une question déterminée *discrimine* les élèves forts des élèves faibles.

1. Méthode simple.

Pidgeon et Yates proposent une méthode de calcul simple, fort utile pour les enseignants :

- Diviser la classe en trois groupes : supérieur, moyen et inférieur, sur la base des scores totaux à l'examen considéré.
- Pour chaque *item*, voir quel pourcentage de chacun des trois groupes a réussi.
- Pour chaque *item*, la différence entre le pourcentage de réussite du groupe supérieur et celui du groupe inférieur donne une bonne estimation de l'efficacité.
Si la différence est faible ou nulle, c'est que la question ne distingue pas bien les élèves forts des élèves faibles.

Plus tard, nous verrons que, plus la discrimination est fine, plus la distribution des résultats est large (la courbe de Gauss s'aplatit).

2. Méthode plus fine.

Voici une autre méthode, utilisée pour l'examen de fin du secondaire en Angleterre¹.

Quatre groupes A, B, C, D sont constitués.

On donne les consignes suivantes :

- 1) Déterminer, pour chacun des *items* et dans chacun des quatre groupes, le nombre d'élèves qui répondent correctement.
- 2) Calculer la moyenne obtenue par chacun des quatre groupes pour chaque *item* ou question.
- 3) Réunir toutes ces données en un tableau d'ensemble et comparer les moyennes pour chaque *item* dans les quatre groupes.

Si les quatre moyennes s'ordonnent de la même façon que les scores moyens pour le test entier, dans les quatre groupes, on peut affirmer que l'*item* considéré contribue à la discrimination totale du test. Sinon, l'*item* est suspect.

Exemple :

Question n°	1	2	3	4	5	6	7	Total
Maximum	6	6	6	6	10	8	8	50
Groupe A	4,9	5,1	5,3	5,2	4,3	5,8	4,3	34,9
Groupe B	3,2	4,9	5,8	4,1	4,7	5,9	3,8	32,3
Groupe C	3,4	4,6	5,4	3	3,4	3,6	3,9	27,3
Groupe D	2	3,4	5	2,8	2,3	4,1	2,3	21,8
Ordre des scores moyens	A	A	B	A	B	B	A	A
	C	B	C	B	A	A	C	B
	B	C	A	C	C	D	B	C
	D	D	D	D	D	C	D	D
Discrimine	Non	Oui	Non	Oui	Non	Non	Non	Oui

La réponse que l'on obtient ainsi manque de finesse. Pour les *items* considérés comme acceptables, nous ne pouvons pas dire s'ils discriminent finement ou si, au contraire, ils sont tout

¹ D. MATHER et al., *The C.S.E., A Handbook for Moderators*, Londres, Collins, 1965, p. 108.

² A partir d'un ensemble de 80 copies, D. Mather constitue quatre groupes de 20 : A = les 20 premiers ; B = les 20 suivants ; etc.

juste acceptables. De améliorer ou la rédaction

Pour obtenir des in utilisent des techniques suivante est simple et a

Considérons les ré élèves de chaque grou questions 2, 3 et 4. A nous les pointons dans U

Question 2 — Maximum			
Scor. obt.	Groupes		
	A	B	C
6			
5			
4			
3			
2			
1			
0			

Question	
Scor. obt.	
6	
5	
4	
3	
2	
1	
0	

Scor. obt : scores obtenus.

1 Encore que l'utilisation de

sée pour l'examen de fin du

nt constitués.

tes :

items et dans chacun des élèves qui répondent correcte-

par chacun des quatre grou-

un tableau d'ensemble et chaque item dans les quatre

onnent de la même façon test entier, dans les quatre l'item considéré contribue à t. Sinon, l'item est suspect.

4	5	6	7	Total
6	10	8	8	50
5,2	4,3	5,8	4,3	34,9
4,1	4,7	5,9	3,8	32,3
3	3,4	3,6	3,9	27,3
2,8	2,3	4,1	2,3	21,8

A	B	B	A	A
B	A	A	C	B
C	C	D	B	C
D	D	C	D	D

Oui	Non	Non	Non	Oui
-----	-----	-----	-----	-----

nsi manque de finesse. Pour tables, nous ne pouvons pas si, au contraire, ils sont tout

for Moderators, Londres, Collins, 1965,

Mather constitue quatre groupes de 20 :

juste acceptables. De plus, nous ne savons trop comment améliorer ou la rédaction de la question, ou la façon de l'évaluer.

Pour obtenir des informations plus fines, les statisticiens utilisent des techniques compliquées¹. La solution graphique suivante est simple et apporte les nuances souhaitées.

Considérons les résultats individuels de chacun des 20 élèves de chaque groupe. Nous prendrons pour exemple les questions 2, 3 et 4. Au lieu d'écrire les résultats en chiffres, nous les pointons dans un tableau.

Question 2 — Maximum 6				
Scor. obt.	Groupes			
	A	B	C	D
6				
5				
4				
3				
2				
1				
0				

Question 3 — Maximum 6				
Scor. obt.	Groupes			
	A	B	C	D
6				
5				
4				
3				
2				
1				
0				

Question 4 — Maximum 6				
Scor. obt.	Groupes			
	A	B	C	D
6				
5				
4				
3				
2				
1				
0				

Scor. obt : scores obtenus.

1 Encore que l'utilisation de l'ordinateur les mette aujourd'hui à la portée de chacun.

Si, sur des tableaux semblables, nous indiquons le quartile supérieur, le médian et le quartile inférieur¹, le comportement des élèves apparaît clairement.

Question 2				
Points	A	B	C	D
6	QS	QS		
5	M	M	QS	QS
4	QI		M	
3		QI	QI	
2				M
1				QI
0				

Question 3				
Points	A	B	C	D
6	QS	QS	QS	QS
5	M	M	M	M
4	QI	QI	QI	QI
3				
2				
1				
0				

Question 4				
Points	A	B	C	D
6	QS	QS		QS
5	M	M	QS	
4	QI		M	
3		QI		M
2				
1			QI	
0				QI

Idéalement, pour chacune des mesures, nous devrions observer une descente continue de A en D. La descente est plus ou moins accusée, selon le pouvoir discriminatif. Une

¹ Le MEDIAN : note du milieu dans une série de notes ordonnées ; si le nombre de notes est pair, on calcule la moyenne arithmétique entre les deux notes médianes

Dans une échelle de mesure, le médian est le point au-dessus duquel se trouvent exactement la moitié des cas.

Le QUARTILE SUPERIEUR ou Q_3 est la note du milieu de la moitié supérieure de la série.

Le QUARTILE INFÉRIEUR ou Q_1 est la note du milieu de la moitié inférieure.

montée indique qu'un g l'ensemble de l'examen) groupe supérieur.

On observe :

Question 2 :

— Il n'y a aucune remon mais pas très nette.

Item faiblement discriminatif.

Question 3 :

— Aucune descente pour

— Remontée pour le méd

— Remontée pour le qua

Item non discriminatif. A

Question 4 :

— Aucune remontée.

— Descente bien marquée

— Descente très bien marquée

Bonne discrimination. U groupe D : les correcteurs leurs élèves du groupe fa

II. Réponses

A une question à réponse fermée, en utilisant son propre jugement, le sujet est tenu de donner une réponse proposée.

Traditionnellement, le premier type de question a une grande utilité.

A. Réponses

Ce sont les questions qui nous posons à tout instant dans

es, nous indiquons le quartile inférieur¹, le comportement

Question 3				
Points	A	B	C	D
6	QS	QS	QS	QS
5	M	M	M	M
4	QI	QI	QI	QI
3				
2				
1				
0				

Question 4		
B	C	D
QS		QS
M	QS	M
QI	M	QI
	QI	
		QI

les mesures, nous devrions de A en D. La descente est le pouvoir discriminatif. Une

série de notes ordonnées ; si le nombre symétrique entre les deux notes médianes 11 13

est le point au-dessus duquel se trouvent du milieu de la moitié supérieure de la du milieu de la moitié inférieure.

montée indique qu'un groupe inférieur (d'après le résultat à l'ensemble de l'examen) obtient de meilleurs points que le groupe supérieur.

On observe :

Question 2 :

— Il n'y a aucune remontée et une tendance à la descente, mais pas très nette.

Item faiblement discriminatif.

Question 3 :

— Aucune descente pour le quartile supérieur.

— Remontée pour le médian.

— Remontée pour le quartile inférieur.

Item non discriminatif. A rejeter.

Question 4 :

— Aucune remontée.

— Descente bien marquée pour le médian.

— Descente très bien marquée pour le quartile inférieur.

Bonne discrimination. Une réserve, cependant, pour le QS du groupe D : les correcteurs ont été trop généreux pour les meilleurs élèves du groupe faible (D).

II. Réponses ouvertes ou fermées ?

A une question à *réponse ouverte*, l'élève répond spontanément, en utilisant son propre vocabulaire. La *réponse* est dite *fermée* si le sujet est tenu d'opérer un choix parmi plusieurs réponses proposées.

Traditionnellement, l'école s'en est exclusivement tenue au premier type de questions ; elle négligeait ainsi un outil de grande utilité.

A. Réponses ouvertes (orales ou écrites).

Ce sont les questions les plus naturelles, celles que nous posons à tout instant dans la vie.

Elles conviennent spécialement, soit pour des épreuves de contrôle faites rapidement, en toute spontanéité par les maîtres, en cours d'enseignement, soit pour la vérification d'apprentissages tellement complexes qu'ils échappent à l'analyse rigoureuse.

L'évaluation des capacités supérieures (créativité, jugement, esprit critique,...) semble de leur ressort. Pas exclusivement, toutefois. D'abord parce que, comme nous allons le voir dans un instant, les questions à choix multiple bien construites permettent des explorations beaucoup plus subtiles qu'il n'y paraît. Ensuite, parce que, par souci d'objectivité, de rigueur, les notateurs tendent peut-être, de façon inconsciente, à ne retenir des réponses ouvertes que les éléments les plus concrets, les plus factuels qu'elles contiennent. La divergence des évaluations n'est-elle pas fonction de la subtilité de l'objet sur lequel elles portent ?

« Nous devons admettre, » écrit Vernon¹, « que les élèves excellents ont l'occasion de montrer certaines qualités exceptionnelles dans l'examen traditionnel, et qu'un notateur perspicace peut, parfois, s'en apercevoir, alors que les autres peuvent très bien pénaliser la réponse à cause de son anticonformisme. » Le débat est loin d'être clos.

En construisant entièrement la réponse, l'élève tente de prouver deux choses : sa connaissance de la matière et sa capacité à l'exprimer verbalement.

Longtemps, on a cru que les deux allaient de pair. Le vieil adage : « Ce que l'on conçoit bien s'énonce clairement, Et les mots pour le dire arrivent aisément », fut vérité reçue par bien des maîtres jusqu'aujourd'hui. Pourtant, les choses sont loin d'être aussi simples. La traduction verbale de la pensée n'est qu'une forme d'expression parmi d'autres, dédaignées pendant des siècles par les classes sociales privilégiées auxquelles tout travail manuel, tout traitement du réel répugnaient.

Renoncer entièrement aux questions à réponses ouvertes serait une erreur dans une civilisation où la communication verbale reste dominante. Mais il est parfois bon d'isoler, au moins partiellement, les connaissances et les capacités d'expression verbale, ce que permettent les réponses fermées. Les professeurs doivent d'autant moins hésiter à y recourir que, comme le remarquent Pidgeon et Yates, on n'a jamais démontré que leur utilisation, même fréquente, nuisait au développement verbal.

¹ Bull. 4, p. 7.

Une réponse formulée de cette façon est un inconconvient grave : son évaluation par comparaison est impossible. C'est assurément une des raisons pour lesquelles, entre correcteurs, si sou-

B. Réponses fermées

1. Utilité.

L'examen intensif vise à tester les acquisitions ; il a souvent une portée sur une matière vaste.

Dans les deux cas, d'ailleurs, d'où le recours à des questions surtout du type « à choix multiples ».

Exemple : Le premier exemple de la pratique courante

- a) Binet
- b) Galton
- c) Goddard
- d) Spearman
- e) Terman.

2. Constituer une provision.

Trouver plusieurs di- verses questions pour un seul examen, met l'imagination à l'épreuve. Le fichier d'items que l'on constitue est ainsi considérablement la tâche.

On n'inscrit qu'une seule question, mais on porte aussi les « distracteurs », mais vraisemblablement les distracteurs sont notamment fréquemment par les élèves, leur « logique » et la vraie.

Les professeurs d'université ont unifié leurs efforts en cette matière.

On semble aussi s'orienter vers des d'items, offices centraux de questions, maîtres plusieurs centaines de questions mises au point pour des

t, soit pour des épreuves de
spontanéité par les maîtres,
pour la vérification d'apprentis-
sage échappent à l'analyse rigou-

supérieures (créativité, juge-
ment leur ressort. Pas exclusi-
ve, comme nous allons le
voir, les questions à choix multiple bien
construites sont beaucoup plus subtiles
que les questions à choix unique, par souci d'objectivité, de
ne pas être, de façon inconsciente,
biaisées en faveur des éléments les plus
connus qui contiennent. La divergence
de la subtilité de l'objet

dit Vernon¹, « que les élèves
possèdent certaines qualités excep-
tionnelles, et qu'un notateur perspi-
cace, alors que les autres peuvent
être affectés par la cause de son anticonformis-

à la réponse, l'élève tente de
maîtriser l'essence de la matière et sa

deux allaient de pair. Le vieil
s'énonce clairement, Et les
t », fut vérité reçue par bien
d'autant, les choses sont loin
de la verbale de la pensée n'est
d'autres, dédaignées pendant
des siècles privilégiées auxquelles
le monde du réel répugnaient.

questions à réponses ouvertes
situation où la communication
est parfois bon d'isoler, au
niveau des capacités d'ex-
ercice les réponses fermées. Les
hésiter à y recourir que, com-
me, on n'a jamais démontré que
nuisait au développement

Une réponse formulée en toute liberté présente un autre
inconvenient grave : son caractère unique se prête mal à l'éva-
luation par comparaison avec les réponses d'autres individus.
C'est assurément une des principales sources des désaccords
entre correcteurs, si souvent dénoncés par la docimologie.

B. Réponses fermées. Questions (*items*) à choix multiple.

1. Utilité.

L'examen intensif vise à vérifier en détail la qualité des
acquisitions ; il a souvent un but diagnostique. L'examen extensif
porte sur une matière vaste.

Dans les deux cas, de nombreuses vérifications sont néces-
saires. D'où le recours aux questions à « réponses fermées »,
surtout du type « à choix multiple ».

Exemple : Le premier test d'intelligence utilisable dans
la pratique courante a été construit par :

- a) Binet
- b) Galton
- c) Goddard
- d) Spearman
- e) Terman.

2. Constituer une provision de questions.

Trouver plusieurs dizaines de questions précises, pour un
seul examen, met l'imagination à rude épreuve. Constituer un
fichier *d'items* que l'on enrichit à chaque occasion facilite
considérablement la tâche.

On n'inscrit qu'une question par fiche afin de pouvoir
porter aussi les « distracteurs », c'est-à-dire les réponses faus-
ses, mais vraisemblables, à mesure de leur découverte. Ces
distracteurs sont notamment fournis par les erreurs commises
fréquemment par les élèves ; on est ainsi assuré de respecter
leur « logique » et la vraisemblance est garantie.

Les professeurs d'une même branche peuvent aisément
unir leurs efforts en cette matière.

On semble aussi s'acheminer vers la création de *banques*
d'items, offices centraux pouvant mettre à la disposition des
maîtres plusieurs centaines de questions à choix multiple, bien
mises au point pour des populations déterminées.

3. Exploiter la gamme des possibilités logiques.

Répondre à un *item* à choix multiple n'est pas nécessairement un simple exercice de mémoire, loin s'en faut.¹

Les recherches entreprises par les services de l'enseignement supérieur en vue de la réforme des examens de médecine, en France, ont bien mis en lumière la richesse du système². Nous choisissons à dessein des exemples dans le domaine des examens universitaires, dans l'espoir de convaincre les maîtres que, quelle que soit la matière qu'ils enseignent, le recours à l'examen objectif n'est pas exclu *a priori*.

Dans le présent examen, on distingue huit types de questions :

a) Question à complément simple.

L'*item* se présente, dans son esprit sinon toujours dans sa forme, comme une phrase à compléter :

Ex. : Parmi les caractères suivants, celui qui s'applique à tous les enzymes est :

- A. Ils contiennent toujours un coenzyme dissociable.
- B. Ils sont thermostables.
- C. Ils contiennent toujours de l'azote dans leur molécule.
- D. Ils contiennent toujours du phosphore dans leur molécule.
- E. Ils sont dialysables.

Remarquons qu'il n'est pas toujours nécessaire que la réponse correcte figure parmi les choix proposés. C aurait pu prendre la place de E et devenir : « Aucune des réponses précédentes. »

L'*item* peut prendre la forme négative :

Ex. : Un hydrosol métallique a tous les caractères suivants, sauf un. Indiquez lequel.

A — B — C — D — E

Pour répondre, le candidat doit connaître tous les caractères de l'hydrosol. Toutefois, la forme négative contraint à une gymnastique intellectuelle qui se superpose à la difficulté inhérente à la matière.

¹ La technique a même été appliquée avec succès à l'analyse littéraire - voir B. CHOPPIN et A. PURVES, *A comparison of open-ended and multiple choice items dealing with literary understanding in Research in the Teaching of English*, 3, 1, 1969, 15-24.

² Paris, Ministère de l'Education nationale. Enseignement supérieur, Examens et concours, section médicale 1961. Tous les exemples médicaux suivants sont empruntés à cette publication.

b) Association simple

Elle sert à vérifier l'association d'entités qui peuvent être

Il faut ici faire correspondre une lettre à un élément précédent (par exemple, par exemple).

Ex. : Branche postérieure

- A. Segment d'origine
- B. Groupe des rameaux
- C. Groupe des rameaux
- D. Nerf interosseux
- 5. est appliqué sur les osseux
- 6. contourne le col
- 7. innerve les muscles de la région antibrachiale
- 8. passe entre les deux osseux
- 9. innerve les muscles de la région antibrachiale postérieure

c) Association complexe

C'est une simple variante

Exemple :

- A. Paludisme à plasmodium
- B. Paludisme à plasmodium
- C. Les deux à la fois
- D. Aucun des deux

Question 1 :

- 1. Le développement de la maladie est moins grave chez l'homme de race blanche
- 2. Une association de symptômes est le traitement de choix
- 3. Les épisodes cliniques sont traités par la prise de chloroquine unique (réponse C)
- 4. Guérit d'une manière définitive par la prise de chloroquine (réponse C)
- 5. La contamination est due à la prise de chloroquine une fois par semaine

Remarquons que le développement de la maladie est augmenté ou diminué selon

tés logiques.

Multiple n'est pas nécessaire-
e, loin s'en faut.¹

les services de l'enseigne-
des examens de médecine,
e la richesse du système².
exemples dans le domaine
l'espoir de convaincre les
matière qu'ils enseignent, le
pas exclu a priori.

istingue huit types de ques-

ple.

prit sinon toujours dans sa
er :

ants, celui qui s'applique à

coenzyme dissociable.

azote dans leur molécule.
phosphore dans leur molé-

ours nécessaire que la ré-
oix proposés. C aurait pu
aucune des réponses précé-

gative :

ous les caractères suivants,

D — E

connaître tous les carac-
ne négative contraint à une
erpose à la difficulté inhé-

uccès à l'analyse littéraire - voir B.
led and multiple choice items dealing
ching of English, 3, 1, 1969, 15-24.
Enseignement supérieur, Examens et
s médicaux suivants sont empruntés à

b) Association simple.

Elle sert à vérifier la connaissance « d'un certain nombre d'entités qui peuvent être ou ne pas être en relation ».

Il faut ici faire correspondre un élément précédé d'une lettre à un élément précédé d'un chiffre (association ou appariement).

Ex. : *Branche postérieure du nerf radial*

- A. Segment d'origine
- B. Groupe des rameaux postérieurs
- C. Groupe des rameaux antérieurs
- D. Nerf interosseux postérieur
5. est appliqué sur la face postérieure du ligament interosseux
6. contourne le col du radius
7. innerve les muscles de la couche superficielle de la région antibrachiale postérieure
8. passe entre les deux chefs du court supinateur
9. innerve les muscles de la couche profonde de la région antibrachiale postérieure

c) Association composée.

C'est une simple variante de la précédente.

Exemple :

- A. Paludisme à plasmodium vivax
- B. Paludisme à plasmodium falciparum
- C. Les deux à la fois (A et B)
- D. Aucun des deux

Question 1 :

1. Le développement clinique a toutes chances d'être moins grave chez un homme de race noire que chez un homme de race blanche (réponse B)
2. Une association de primaquine et de chloroquine est le traitement de choix pour une attaque aiguë (réponse A)
3. Les épisodes cliniques sont supprimés par l'ingestion de chloroquine une fois par semaine en zone endémique (réponse C)
4. Guérit d'une manière définitive par le traitement avec la chloroquine (réponse B)
5. La contamination est évitée par l'ingestion de chloroquine une fois par semaine (réponse D)

Remarquons que le nombre de questions peut être ici augmenté ou diminué selon l'importance accordée à la matière.

d) Association à terme exclu.

Dans l'*item* suivant, 4 des 5 phénomènes numérotés sont communs à un des troubles A.B.C. Il faut indiquer le trouble (A) et le phénomène qui n'y correspond pas (2).

Exemple :

- A. *eosinophilie d'importance diagnostique*
- B. plasmocytose d'importance diagnostique
- C. lymphocytose d'importance diagnostique
- 1° trichynose
- 2° myélome multiple
- 3° syndrome de Loeffler
- 4° maladie de Hodgkin
- 5° schistosomiase

e) Analyse de relation de cause à effet.

Ex. : L'articulation radio-cubitale supérieure permet des mouvements de rotation limités PARCE QUE la tête du radius est entourée par le ligament annulaire.

- A. La constatation et la raison proposée sont toutes les deux vraies et il existe une relation de cause à effet entre les deux.
- B. La constatation et la raison proposée sont toutes les deux vraies et il n'y a pas de relation de cause à effet entre elles.
- C. La constatation est vraie, mais la raison proposée est fausse.
- D. La constatation est fausse, mais la raison proposée est un fait ou un principe accepté.
- E. La constatation et la raison proposée sont toutes les deux fausses.

f) Analyse d'observation.

L'ensemble complexe suivant met le candidat dans une situation comparable à l'expérience réelle.

Exposé de la maladie : Le malade est un homme de 21 ans qui se plaint de malaises, d'une toux et de fièvre.

La maladie a débuté dix jours avant l'admission par un malaise et une toux sans expectoration, suivis dans les 24 heures d'une température variant de 37.8 à 38.3 qui a persisté jusqu'au moment de l'admission.

Le quatrième jour de la maladie, la toux s'accroît, produisant de petites quantités d'expectoration blanche et visqueuse.

Trois jours avant l'admission, la toux commença, avec des sensations de frissons et des frissons véritables, avec une fièvre antérieure à l'apparition de la maladie.

A l'examen physique, la température à 110, le rythme cardiaque 10 1/2, minima 8.

Le malade est bien orienté, il est alerte, semble aiguë, il est

L'examen physique : les vibrations vocales, le murmure vésiculaire, on entend quelques râles, la son est augmentée,

La formule sanguine : polynucléaires 30%, 5% d'éosinophiles. La radio du thorax montre une opacité de la région périhilaire, la répartition est inégale, nuages pulmonaires supérieurs.

Questions :

1. *Quel est le diagnostic ?*
 - a) tuberculose
 - b) pneumonie à pneumocoque
 - c) pneumonie (primaire)
 - d) coccidiomycose
 - e) broncho-pneumonie
2. *Quel est le signe pathognomonique ?*
 - a) spénomégalie
 - b) signe de souffrance
 - c) bruit de frottement
 - d) changements fréquents des signes thoraciques
 - e) signe de condensation
3. *Lequel des examens suivants est le plus utile pour le diagnostic ?*
 - a) l'élévation et l'au

phénomènes numérotés sont
C. Il faut indiquer le trouble
fond pas (2).

diagnostique
diagnostique
diagnostique

se à effet.

ale supérieure permet des
és PARCE QUE la tête du
ment annulaire.

n proposée sont toutes les
e relation de cause à effet

n proposée sont toutes les
de relation de cause à effet

mais la raison proposée est

mais la raison proposée est
é.

n proposée sont toutes les

met le candidat dans une
réelle.

de est un homme de 21 ans
de toux et de fièvre.

s avant l'admission par un
toration, suivis dans les 24
nt de 37.8 à 38.3 qui a per-
mission.

e, la toux s'accroît, pro-
xpectoration blanche et vis-

Trois jours avant l'admission, des accès paroxystiques de toux commencèrent, parfois suivis de vomissements. Des sensations de frissons furent notées, mais non pas de véritables frissons avec tremblements. Une douleur parasternale antérieure à la toux existe depuis le cinquième jour de maladie.

A l'examen physique, la température est à 38.3, le pouls à 110, le rythme respiratoire 32, la tension maxima 10 1/2, minima 8.

Le malade est bien développé, sans maigreur, sa maladie semble aiguë, il est dyspnéique, mais non cyanosé.

L'examen physique de la cage thoracique montre des vibrations vocales à la palpation et à l'auscultation. Le murmure vésiculaire est normal. Dans l'aisselle gauche, on entend quelques râles fins et la qualité bronchique du son est augmentée, bien que d'intensité normale.

La formule sanguine est la suivante : globules blancs 3 400 (polynucléaires 30 %, lymphocytes 62 %, monocytes 5 %, éosinophiles 3 %).

La radio du thorax révèle une augmentation de la densité de la région périciliaire avec des aires mal définies de densité inégale, nuageuse aux deux bases et dans un champ pulmonaire supérieur gauche.

Questions :

1. *Quel est le diagnostic le plus probable ?*
 - a) tuberculose
 - b) pneumonie à pneumocoques
 - c) pneumonie (primaire atypique) à virus
 - d) coccidiomycose
 - e) broncho-pneumonie
2. *Quel est le signe physique qui s'y ajoute probablement ?*
 - a) spénomégalie
 - b) signe de souffrance méningée
 - c) bruit de frottements pleuraux
 - d) changements fréquents dans la distribution des symptômes thoraciques
 - e) signe de condensation lobaire gauche
3. *Lequel des examens de laboratoire suivants va de pair avec le diagnostic ?*
 - a) l'élévation et l'augmentation des agglutines froides

- b) hémoculture positive
- c) leucocytose marquée au début de la convalescence
- d) examen des expectorations
- e) cuti-réaction positive

4. *Quelle thérapeutique devra être employée ?*

- a) repos au lit et streptomycine
- b) repos au lit et pénicilline
- c) streptomycine et acide para-amino salicylique
- d) repos au lit et auréomycine
- e) psychothérapie et rééducation physique

5. *Quelle est l'issue probable de cette maladie sans traitement?*

- a) la fièvre va disparaître spontanément par une crise terminale
- b) la convalescence va être progressive avec une rechute prévisible
- c) un empyème résiduel va se développer
- d) une fibrose résiduelle va apparaître après guérison
- e) une caverne pulmonaire peut apparaître

g) **Comparaisons quantitatives.**

Comparez X à Y :

X — pression mécanique dans le capillaire veineux
 Y — pression oncotique dans le capillaire veineux

Et dites si :

- A. X est plus grand que Y
- B. Y est plus grand que X
- C. X est égal à Y

h) **Relations.**

Soit :

1. Le débit circulaire cutané — ET
2. La quantité de chaleur perdue par unité de temps
- A. L'augmentation du premier est accompagnée d'une augmentation du second, ou, la diminution du premier est accompagnée d'une diminution du second.
- B. L'augmentation du premier est accompagnée d'une diminution du second, ou, la diminution du premier est accompagnée d'une augmentation du second.
- C. Les variations du second sont indépendantes des variations du premier.

i) **Compléments graphiques.**
 Procédé utilisé quand la réponse est unique.

Ex. : Cinq conscrits
 1,63 m et 1,61 m.

1. La moyenne des hauteurs est 1,62 m.
2. L'écart type est 0,01 m.
3. L'écart type est 0,02 m.
4. L'échantillon a été tiré d'une population des hauteurs de conscrits.

- A. 1, 2 et 3 sont correctes.
- B. 1 et 3 sont correctes.
- C. 2 et 4 sont correctes.
- D. 4 est correcte.
- E. Une seule des propositions est correcte.

On trouvera, en annexe, un exemple de question conduite selon la méthode du choix multiple.

4. **Calcul de l'efficacité.**

Dans une question à choix multiples, les distracteurs jouent un rôle important. Pour chaque question, il y a une seule possibilité de réponse a) et quatre autres possibilités b, c, d, e.

a) **Situation idéale.**

A	B	C	D	E
10%	10%	10%	10%	10%

D est la bonne réponse. Les autres choix ont une chance sur cinq d'être choisis.

b) **Situation à corriger.**

A	B	C
3%	0%	35%

Seul, le distracteur C est choisi. Les autres sont ignorés.

ut de la convalescence

employée ?

mino salicylique

n physique

tte maladie sans traitement?

anément par une crise ter-

ogressive avec une rechute

velopper

raître après guérison

apparaître

e capillaire veineux

e capillaire veineux

ET
e par unité de temps
est accompagnée d'une
, la diminution du premier
minution du second.
est accompagnée d'une di-
diminution du premier est
tation du second.
nt indépendantes des varia-

i) Compléments groupés.

Procédé utilisé quand une question peut avoir plus d'une réponse correcte.

Ex. : Cinq conscrits mesurent : 1,65 m - 1,67 m - 1,69 m - 1,63 m et 1,61 m.

1. La moyenne des tailles de l'échantillon est 1,65 m
 2. L'écart type est proche de 8
 3. L'écart type est proche de 2,8
 4. L'échantillon a de bonnes chances de renseigner sur la population des tailles des individus en général
- A. 1, 2 et 3 sont corrects
B. 1 et 3 sont corrects
C. 2 et 4 sont corrects
D. 4 est correct
E. Une seule des propositions 1, 2 ou 3 est correcte

On trouvera, en annexe, une comparaison entre un examen conduit selon la méthode traditionnelle et selon la méthode à choix multiple.

4. Calcul de l'efficacité des distracteurs.

Dans une question à choix multiple, il importe de vérifier si les distracteurs jouent bien leur rôle. A cet effet, on calcule, pour chaque question, quel pourcentage des choix chaque possibilité de réponse a recueilli.

a) Situation idéale.

A	B	C	D	E
10 %	10 %	10 %	60 %	10 %

D est la bonne réponse.

Les autres choix ont été également attractifs.

b) Situation à corriger.

A	B	C	D	E
3 %	0 %	35 %	60 %	2 %

Seul, le distracteur C a réellement joué. Les autres se révèlent sans pouvoir.

5. Critiques et réfutation partielle.

Bien que sa valeur soit prouvée depuis longtemps, l'examen par questions à choix multiple continue à susciter de vives critiques. Certaines peuvent être aisément réfutées ; d'autres paraissent justifiées ; d'autres encore ne peuvent être ni infirmées, ni confirmées, faute de critères scientifiques.

a) Une objectivité trompeuse.

La notation des questions à choix multiple est indiscutablement objective : les réponses correctes sont définies d'avance ; l'élève les a trouvées ou non. Les avantages de cette méthode sont évidents.

Pourtant, 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.

Dans la rédaction des questions d'abord. Elles sont le fruit de la réflexion, de l'invention des examinateurs qui, en dernière analyse, agissent subjectivement. A côté des questions auxquelles ils ont pensé, combien d'autres, peut-être plus valides, n'aurait-on pas pu imaginer ?

Dans la rédaction des réponses considérées comme idéales, ensuite. On a rarement affaire à des réponses aussi évidemment correctes que $2 \times 2 = 4$. Au cours de la construction de tests avec nos étudiants, nous avons plus d'une fois rencontré des réponses proposées comme correctes, passées victorieusement au crible de l'analyse mathématique, et qui, pourtant, n'étaient, au mieux, que grossières approximations exprimées dans un langage douteux.

Comme Vernon le note, non sans malice, la supériorité indiscutable des examens « objectifs » sur les épreuves traditionnelles est sans doute plus souvent due à la préparation très soignée et à l'union des efforts et de la compétence de nombreux enseignants et psychopédagogues, qu'à la nature même de l'examen¹.

On oublie aussi qu'entre l'utopie de l'objectivité et la subjectivité totale, bien des stades intermédiaires existent.

¹ VERNON, The C.S.E.: An Introduction to Objective-Type Examinations, Londres, H.M.S.O., 1964, pp. 4-5.

b) Choix « corrects

Ce point aussi a été écrit :

« Il est fréquent que accueillent les questions telles que : « Cet item est considéré comme « fautive », sinon plus, que la vient principalement du sur les mêmes capacités demment, des examens contenir pas mal de mauvaises épreuves construites par exemple, car les items insatisfaisants, par un contrôle préliminaire, par une cernée, ou bien éliminées des réponses (...). Les critiques et les réactions sophistiquées des élèves sur lesquels les questions peuvent lire dans les examens ne tiennent même pas à l'esprit ».

c) Un jeu de hasard

Dans une question fautive et l'autre fautive correcte et l'autre fautive qui travaille au pur hasard, les constructeurs de tests le savent.

Aujourd'hui, on a le choix de cinq choix, dont l'un est la réponse correcte et quatre distracteurs² considérablement réduits, leurs selon leur niveau de difficulté. La loi de la probabilité de la chance minimale de tomber sur la réponse correcte est de 1/5. Par contre, une connaissance de la réponse correcte dans le choix aveugle présente une chance beaucoup plus grande de tomber sur la réponse correcte.

¹ VERNON, o.c., p. 6.

² Parfois aussi, plusieurs choix sont proposés pour sélectionner le plus adéquat.

b) Choix « corrects » contestables.

Ce point aussi a été excellemment discuté par Vernon¹ qui écrit :

« Il est fréquent que des personnes hautement éduquées accueillent les questions à choix multiple par des critiques telles que : « Cet *item* est stupide » ou prétendent que les choix, considérés comme « faux » par le testeur, sont tout aussi admissibles, sinon plus, que la réponse dite « bonne ». Cette critique vient principalement du fait que l'*item* objectif ne s'appuie pas sur les mêmes capacités que les questions traditionnelles. Evidemment, des examens construits par des amateurs peuvent contenir pas mal de mauvais *items*. On en trouve aussi dans les épreuves construites par des professionnels, mais plus rarement, car les *items* insatisfaisants sont ou bien rejetés, lors du contrôle préliminaire, par des spécialistes de la discipline concernée, ou bien éliminés lors de l'analyse mathématique des réponses (...). Les critiques risquent aussi d'oublier que leurs réactions sophistiquées peuvent différer très fort de celles des élèves sur lesquels les *items* ont été essayés. Enfin, les critiques peuvent lire dans les questions des choses qui ne viennent même pas à l'esprit d'élèves intelligents. »

c) Un jeu de hasard.

Dans une question fermée à deux choix de réponses, l'une correcte et l'autre fausse, la probabilité de réussite de l'élève qui travaille au pur hasard est, théoriquement, de 50 %. Les constructeurs de tests le savent depuis longtemps.

Aujourd'hui, on a le plus souvent recours aux questions à cinq choix, dont l'un est correct et les autres seulement vraisemblables (distracteurs)². Dans ce cas, le jeu du hasard est considérablement réduit. La démarche des élèves variera d'ailleurs selon leur niveau de connaissance de la matière. En vertu de la loi de la probabilité, l'étudiant qui ignore tout n'a qu'une chance minime de tomber aveuglément sur la solution correcte. Par contre, une connaissance partielle permet d'éliminer sciemment un certain nombre de distracteurs ; la chance de réussite dans le choix aveugle parmi les possibilités restantes est alors beaucoup plus grande que dans le cas précédent, et c'est justice.

¹ VERNON, o.c., p. 6.

² Parfois aussi, plusieurs choix conviennent, mais à des degrés divers. Il faut alors sélectionner le plus adéquat.

On peut d'ailleurs réduire largement le rôle de la chance en corrigeant les résultats obtenus à l'aide d'une formule simple¹, laquelle, il est vrai, pénalise généralement de façon exagérée, puisque la part du choix aveugle varie selon les élèves.

d) Acrobatie mentale.

On reproche aux examens traditionnels de faire la part trop belle à la facilité d'expression. On peut toutefois se demander si les *items* à choix multiple un peu compliqués (nous venons d'en rencontrer des exemples) ne présentent pas un inconvénient au moins aussi grave en ajoutant, à la difficulté inhérente à la matière de la question, l'obligation de démêler des doubles négations, de saisir des subtilités logiques ou, plus généralement, en accordant une prime à une certaine aptitude à l'abstraction à partir de données verbales. P. Vernon constate, en tout cas, que la compréhension de la lecture joue un rôle important dans la réussite de ces épreuves².

Il importe, non seulement de connaître le niveau de développement mental des élèves, mais aussi de savoir clairement ce que l'on veut : vérifier la connaissance de la matière, la capacité en compréhension de la lecture et en raisonnement, ou la combinaison des deux.

On ne voit pas comment le sondage des qualités intellectuelles supérieures irait sans un accroissement de la complexité des *items*. C'est pourquoi plusieurs auteurs estiment que les questions ouvertes continueront à jouer un rôle important dans les examens de niveau avancé.

En cas de recours à la correction automatique, une autre difficulté vient encore s'ajouter : l'utilisation de cartes de réponses où ne figurent généralement que les numéros des questions et les lettres A, B, C, D, E, représentant les cinq choix. Ce système exige un supplément d'attention de la part de l'élève, en particulier quand il ne peut répondre à certaines questions et décide de les sauter. S'il oublie de sauter aussi la ligne correspondante sur la carte de réponses, les conséquences peuvent être désastreuses. La difficulté ne doit toutefois pas être exagérée : Remmers, Gage et Rummel ont montré expérimentalement que le système est utilisable avec des enfants, à partir de 9 ou 10 ans³.

1 $\frac{\text{Nombre de réponses correctes}}{\text{Nombre de réponses incorrectes}}$

2 P. VERNON, The Determinants of Reading Comprehension, in *Educational Psychological Measurement*, 1962, 22, 269-286.

3 H. REMMERS, N. GAGE et RUMMEL, *Educational Measurement and Evaluation*, New York, Harper, 1955, p. 246.

De toute façon, les entraînés à la technique de subir une épreuve décisive de construction rigoureuse, les bons résultats, leur utilité, depuis de nombreuses années, répandue ailleurs, le prouvent.

e) Inconvénients inconvénients

Nous qualifions les réponses que, à notre connaissance, aucune n'en a établi le bien ou le mal ne se justifient que par le construit.

- Choisir la bonne réponse et construire. La mémoire intellectuelle, une répétition claire de la pensée.
- Sachant que, pour subir une épreuve, il suffirait de souligner ou de cocher une croix, les maîtres préfèrent la solution.
- La préférence irait à des questions à un temps très court ; on exige une très longue épreuve et la vie réelle ne nous

C. En guise de

Un compromis nous est proposé.

Puisque les questions ouvertes possèdent chacune des avantages, cause certains apprentissages, raison de revendiquer un rôle, peuvent parfaitement coexister, estiment cependant qu'il faut tester les deux types de questions.

Par ailleurs, une forme de questions ouvertes et fermées existe, plus, selon les progrès de

L'examen consisterait à tester des questions ouvertes, d'une portée limitée, testées. L'éventail des réponses

ement le rôle de la chance
à l'aide d'une formule sim-
généralement de façon exa-
mple varie selon les élèves.

tionnels de faire la part trop
peut toutefois se demander si
complicqués (nous venons
présentent pas un inconvé-
nant, à la difficulté inhérente
de démêler des doubles
logiques ou, plus générale-
ment, une certaine aptitude à l'ab-
solutisme. P. Vernon constate, en
la lecture joue un rôle im-
portant².

connaître le niveau de déve-
loppement aussi de savoir clairement
la matière, la capa-
cité et en raisonnement, ou la

évaluation des qualités intellec-
tuelles et de la complexité
des auteurs estiment que les
examinés jouent un rôle important dans

évaluation automatique, une autre
l'utilisation de cartes de
évaluation que les numéros des
représentant les cinq choix.
d'attention de la part de
peut répondre à certaines
S'il oublie de sauter aussi
de réponses, les consé-
quences. La difficulté ne doit toute-
fois pas être négligée. Gage et Rummel ont montré
qu'elle est utilisable avec des en-

de réponses incorrectes

nombre de choix — 1
Comprehension, in *Educational Psycho-
logical Measurement and Evaluation*,

De toute façon, les élèves doivent être soigneusement
entraînés à la technique des examens à choix multiple avant de
subir une épreuve décisive. Moyennant cette précaution, et une
construction rigoureuse, les questions à choix multiple donnent
de bons résultats, leur utilisation intensive dans les pays anglo-
saxons, depuis de nombreuses années, et de plus en plus
répandue ailleurs, le prouve à suffisance.

e) Inconvénients incertains.

Nous qualifions les reproches suivants d'incertains parce
que, à notre connaissance, aucune recherche scientifique rigou-
reuse n'en a établi le bien-fondé. Il semble que certaines criti-
ques ne se justifient que dans la mesure où l'examen est mal
construit.

- Choisir la bonne réponse parmi d'autres est plus facile que la
construire. La mémoire intervient trop. Une certaine paresse
intellectuelle, une répugnance à l'effort nécessaire, à la for-
mulation claire de la pensée risquent de s'installer.
- Sachant que, pour subir l'examen en langue maternelle, il
suffirait de souligner ou d'indiquer la réponse choisie par
une croix, les maîtres négligeraient les exercices d'expres-
sion.
- La préférence irait à de nombreux petits exercices à faire en
un temps très court ; on négligerait ainsi les problèmes qui
exigent une très longue réflexion. Or, les études supérieures
et la vie réelle ne nous épargnent pas ce genre d'épreuves.

C. En guise de conclusion : un compromis.

Un compromis nous paraît s'imposer à deux niveaux au
moins.

Puisque questions ouvertes et questions fermées semblent
posséder chacune des avantages particuliers et mettre en
cause certains apprentissages différents, on ne voit aucune
raison de revendiquer un monopole pour l'une des deux. Elles
peuvent parfaitement coexister. Les meilleurs docimologistes
estiment cependant qu'il n'est pas souhaitable de mélanger
les deux types de questions dans une même épreuve.

Par ailleurs, une formule intermédiaire entre questions
ouvertes et fermées existe et s'imposera peut-être de plus en
plus, selon les progrès de la recherche pédagogique.

L'examen consisterait en un nombre assez élevé de ques-
tions *ouvertes*, d'une portée relativement limitée, ayant été pré-
testées. L'éventail des réponses probables serait donc connu

d'avance, ce qui permettrait de proposer un schéma de notation s'approchant très fort de la rigueur de la question à réponse fermée.

Cette possibilité est déjà confirmée par plusieurs travaux expérimentaux. Ainsi, des corrélations presque parfaites (.98 et .99) ont pu être obtenues entre sept groupes de notateurs représentant chacun une commission d'examens de Grande-Bretagne. L'expérience a été faite sur des épreuves d'histoire de niveau supérieur du G.G.E. (fin du secondaire)¹.

Sans aller aussi loin, des professeurs peuvent parfaitement s'entendre sur des points à exiger dans les réponses à des questions ouvertes.

Par ailleurs, la façon de rédiger les questions peut fermer partiellement des réponses et donc diminuer le jeu de la subjectivité. Une question comme : Ecrivez un début approprié à la phrase : « ... quand il commença à pleuvoir. » est, en quelque sorte, à mi-chemin entre l'épreuve subjective et l'épreuve objective. Car ici, les cas de désaccord entre notateurs invités à se prononcer sur la correction de la réponse seront rares.

¹ JOINT MATRICULATION BOARD, *The Marking of Scripts in Advanced Level History*. Universities of Manchester, Liverpool, Leeds, Sheffield and Birmingham, 1964.

I. Un préambule ind

Il est impossible de
sans s'appuyer sur que
n'oublions pas la prome
l'arithmétique élémentaire

A. La courbe de Gauss,

Un sac contient deu
sinon que l'une est rouge

Dans ces conditions,
de sortir à chaque tirage a

Si nous nous livrons
il est possible que nous
même couleur. A mesure
tirages, cette probabilité d

Imaginons 100 tirage
le rouge sortira à peu pr
par contre, fort peu prob
vement la boule rouge o
petit peu plus probable q
fois blanc, plus probable
etc. Bref, la probabilité
blanc, puis diminuer prog
99 blanc.

poser un schéma de notation
ur de la question à réponse

firmée par plusieurs travaux
ons presque parfaites (.98 et
t groupes de notateurs repré-
xamens de Grande-Bretagne.
preuves d'histoire de niveau
aire) ¹.

esseurs peuvent parfaitement
er dans les réponses à des

er les questions peut fermer
c diminuer le jeu de la sub-
rivez un début approprié à la
a pleuvir. » est, en quelque
subjective et l'épreuve objec-
entre notateurs invités à se
ponse seront rares.

CHAPITRE 3.

LA NOTATION.

I. Un préambule indispensable : la courbe de Gauss.

Il est impossible de discuter de l'évaluation des résultats sans s'appuyer sur quelques notions mathématiques. Nous n'oublions pas la promesse faite au début de cet ouvrage : l'arithmétique élémentaire nous suffira !

A. La courbe de Gauss, image de la probabilité.

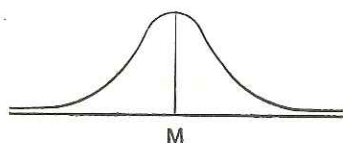
Un sac contient deux boules, en tout point semblables, sinon que l'une est rouge et l'autre blanche.

Dans ces conditions, chaque boule a une chance sur deux de sortir à chaque tirage aveugle.

Si nous nous livrons à ce jeu, un très petit nombre de fois, il est possible que nous tirions plusieurs fois consécutives la même couleur. A mesure que nous augmentons le nombre de tirages, cette probabilité diminue.

Imaginons 100 tirages consécutifs. Il est fort probable que le rouge sortira à peu près autant de fois que le blanc. Il est, par contre, fort peu probable que l'on tirera 100 fois consécutivement la boule rouge ou la boule blanche. Il est déjà un tout petit peu plus probable que l'on obtiendra 99 fois rouge et une fois blanc, plus probable déjà 98 fois rouge et 2 fois blanc, etc. Bref, la probabilité va augmenter jusqu'à 50 rouge et 50 blanc, puis diminuer progressivement pour arriver à 1 rouge et 99 blanc.

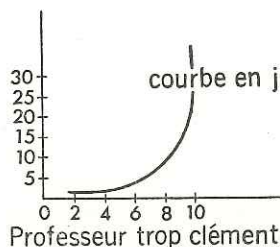
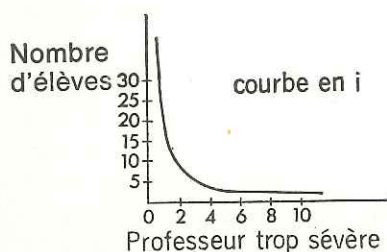
Dans un très grand nombre de tirages, ce mouvement ascendant-descendant correspond à une courbe revêtant la forme d'une cloche : c'est la fameuse courbe de Gauss, dont les deux moitiés sont symétriques par rapport à la moyenne arithmétique et dont les extrémités ne touchent jamais la ligne du zéro, la probabilité nulle n'existant qu'à l'infini.



Cette *distribution*, dite *normale*, est à l'image de beaucoup de qualités humaines, telles qu'elles se répartissent dans des groupes nombreux, *pris au hasard*. Ainsi, parmi les habitants d'une grande ville, les hommes de taille moyenne sont les plus nombreux, tandis que les géants et les nains sont très rares ; entre ces deux extrêmes, la population se distribue selon la courbe de Gauss.

Même l'erreur se soumet souvent à la loi normale : si l'on fait un très grand nombre de mesures, on verra fort probablement apparaître une erreur de grandeur moyenne, et une distribution allant de l'erreur infime à l'erreur maximum. Nous verrons, par la suite, que cette observation est importante.

Dès que le hasard ne joue plus, la distribution se modifie. Un professeur peut, par exemple, donner une très grande majorité de mauvaises notes et fort peu de bonnes, ou le contraire. Il est possible que l'on obtienne alors une des deux courbes caractéristiques suivantes :



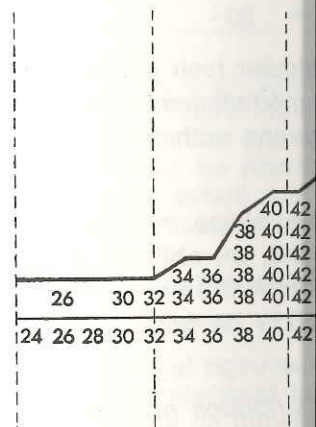
B. La courbe de Gauss, de l'enseignement normal

Un professeur qui dans une classe où les sélectionnés, donne non difficulté adapté à la m correct, il y aura donc b très bons et peu de très tats s'approchera de la

Pareil phénomène s dans les exercices écha parce qu'ils mettent en j

Ci-dessous, on a di nus à un examen, lui-mêm ves ont obtenu 50 sur obtient la note la plus b

On observe aussi q tion des résultats n'est p seulement son ébauche donner quelques précis le nombre de sujets av polie et on considère la



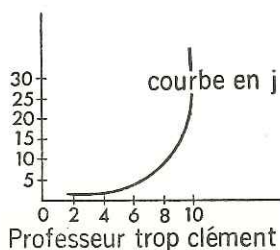
le tirages, ce mouvement
à une courbe revêtant la
courbe de Gauss, dont les
apport à la moyenne arith-
touchent jamais la ligne du
à l'infini.



est à l'image de beaucoup
s se répartissent dans des
Ainsi, parmi les habitants
de taille moyenne sont les
ants et les nains sont très
population se distribue selon

nt à la loi normale : si l'on
es, on verra fort probable-
leur moyenne, et une distri-
erreur maximum. Nous ver-
tion est importante.

, la distribution se modifie.
donner une très grande
ort peu de bonnes, ou le
obtienne alors une des deux



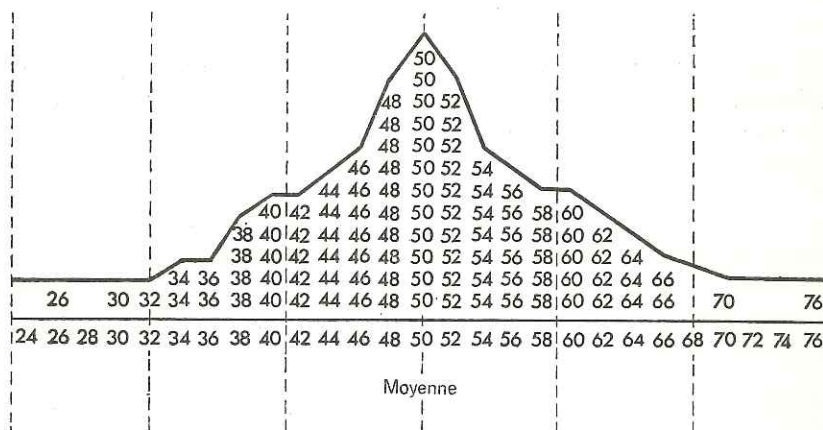
B. La courbe de Gauss, image des résultats de l'enseignement non individualisé.

Un professeur qui enseigne de façon non individualisée dans une classe où les élèves ne sont pas spécialement sélectionnés, donne normalement à son cours un degré de difficulté adapté à la majorité du groupe. Si l'ajustement est correct, il y aura donc beaucoup de résultats moyens, peu de très bons et peu de très mauvais. La distribution de ces résultats s'approchera de la courbe gaussienne.

Pareil phénomène se produit, plus spontanément encore, dans les exercices échappant à la quantification rigoureuse, parce qu'ils mettent en jeu un ensemble complexe de facteurs.

Ci-dessous, on a disposé 100 résultats (imaginaires) obtenus à un examen, lui-même noté sur 100. On constate que 12 élèves ont obtenu 50 sur 100 (note moyenne), alors qu'un seul obtient la note la plus basse (26) et un seul la plus haute (76).

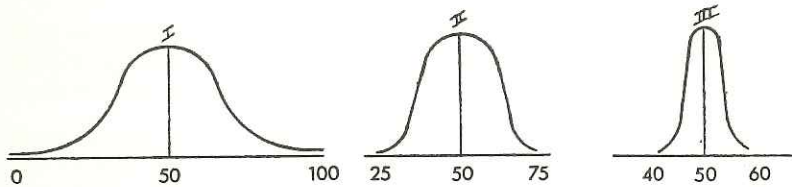
On observe aussi que la ligne correspondant à la répartition des résultats n'est pas une courbe en cloche parfaite, mais seulement son ébauche. Dans pareil cas (nous allons bientôt donner quelques précisions à ce propos), on suppose que, si le nombre de sujets avait été plus élevé, la courbe se serait polie et on considère la répartition comme « normale ».



C. L'écart type ou sigma, indice précieux.

1. Signification.

La variation des 100 résultats que nous venons d'examiner aurait pu être soit plus grande, par exemple de 0 à 100, soit plus petite, par exemple de 40 à 60, tout en se distribuant toujours en courbe de Gauss. Nous aurions pu avoir :



Dans les trois cas, la moyenne est 50. Pourtant, il s'agit de situations très différentes.

La marge de variation des résultats est :

$$\text{courbe I : } 100 - 0 = 100$$

$$\text{courbe II : } 75 - 25 = 50$$

$$\text{courbe III : } 60 - 40 = 20$$

Sigma ou σ , est un indice facile à calculer (voir plus loin la méthode de calcul), qui nous indique immédiatement comment les résultats varient par rapport à la moyenne arithmétique des notes obtenues.

Pour les 100 notes prises en exemple, $\sigma = 9$. Si l'on prend un σ avant et après la moyenne, on obtient : 41 et 59. Comptez le nombre de résultats qui tombent entre ces deux limites, vous obtenez 68 notes, soit ici 68 % de notes de tout le groupe.

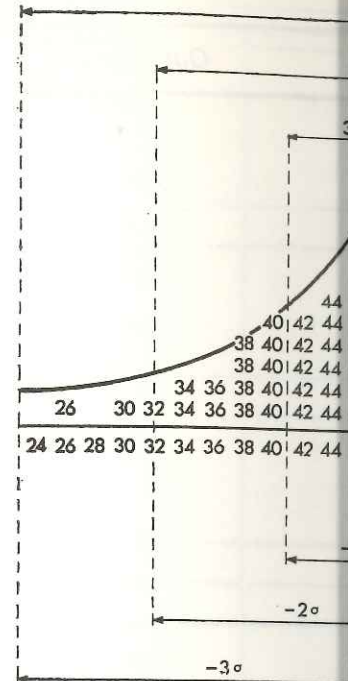
Il en est toujours ainsi, dans une *distribution normale* : le σ indique toujours la même proportion des résultats par rapport à l'ensemble :

1 σ de part et d'autre de la moyenne = 68 % des notes

2 σ de part et d'autre de la moyenne = 95 % des notes

3 σ de part et d'autre de la moyenne = 99 % des notes.

Voyons maintenant la
yeux nouveaux :



Comme, à un σ de part et d'autre de la surface de l'aire déterminée, la même proportion des résultats est obtenue. *de comparer les résultats obtenus, pour autant que ce soit la même. Nous montrerons bien*

2. Estimation rapide d'un σ

a) Problème.

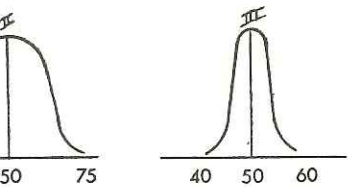
Le calcul rigoureux d'un σ (par la méthode de Ma) est relativement lourd. Si la distribution n'est pas normale, la moyenne obtenue est souvent éloignée de la tendance centrale.

Un procédé simple permet de fournir une estimation de σ à partir de quelques résultats obtenus dans presque tous les cas.

1 D'après D.R. MATHER et al., o.c.

précieux.

que nous venons d'examiner
 par exemple de 0 à 100, soit
 tout en se distribuant tou-
 jours pu avoir :



est 50. Pourtant, il s'agit de

ats est :

- 0 = 100
- 25 = 50
- 40 = 20

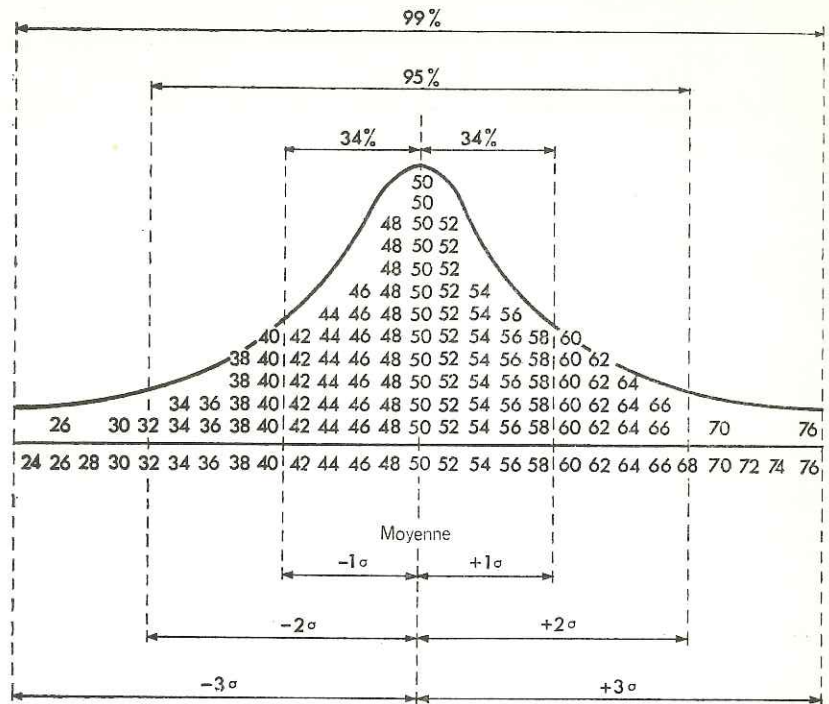
à calculer (voir plus loin la
 que immédiatement comment
 la moyenne arithmétique des

exemple, sigma = 9. Si l'on
 moyenne, on obtient : 41 et
 qui tombent entre ces deux
 t ici 68 % de notes de tout

une *distribution normale* : le
 proportion des résultats par

- moyenne = 68 % des notes
- moyenne = 95 % des notes
- moyenne = 99 % des notes.

Voyons maintenant la distribution des résultats avec des
 yeux nouveaux :



Comme, à un sigma donné, correspond toujours la même
 surface de l'aire déterminée par la courbe, c'est-à-dire la
 même proportion des résultats, *cet indice nous fournit le moyen
 de comparer les résultats attribués par des professeurs diffé-
 rents, pour autant que ces résultats se répartissent normale-
 ment*. Nous montrerons bientôt comment.

2. Estimation rapide de la moyenne et du sigma.

a) Problème.

Le calcul rigoureux de la moyenne et de l'écart type (sig-
 ma) est relativement lourd. De plus, quand la distribution n'est
 pas normale, la moyenne arithmétique donne une idée fautive
 de la tendance centrale.

Un procédé simple permet d'éviter de longs calculs. Il ne
 fournit que des résultats *approximatifs*. Ils suffisent néanmoins
 dans presque tous les cas de la pratique courante à l'école¹.

¹ D'après D.R. MATHER et al., o.c.

Soit deux questions d'examen auxquelles 10 élèves ont répondu. Voici les notes attribuées, les moyennes et les écarts types calculés par la méthode classique.

Elève	Question 1	Question 2
1	39	32
2	33	28
3	25	32
4	22	28
5	26	27
6	18	31
7	23	33
8	13	27
9	57	35
10	45	36

301

309

Moyenne arithmétique : 30

Moyenne arithmétique : 31

Ecart type : $\sigma =$

$$\sqrt{\frac{\sum x^2}{N}} = \pm 13$$

Ecart type : $\sigma = \pm 3.$

b) Méthode simple
1° Représentation graph

Notes	Question 1	Question 2
60		
58		
56		
54		
52		
50		
48		
46		
44		
42		
40		
38		
36		
34		
32		
30		
28		
26	:	:
24		
22	:	:
20		
18		
16		
14		
12		
10		

Remarquons que, male, la moyenne et le

2° Estimation de l'écart

$\sigma = 3/4$ de l'écart inte

Donc :

- a. Chercher le quar notes au-dessus d

Question 1 = 39

Question 2 = 33

auxquelles 10 élèves ont les moyennes et les écarts que.

1	Question 2
	32
	28
	32
	28
	27
	31
	33
	27
	35
	36

309

Moyenne arithmétique : 31

Ecart type : $\sigma = \pm 3$.

b) *Méthode simple de calcul.*

1° Représentation graphique des résultats.

Notes	Question 1	Question 2
60		
58		
56		
54		
52		
50		
48		
46		
44		
42		
40		
38		
36		
34		
32		
30		
28		
26	:	:
24	:	:
22	:	:
20		
18		
16		
14		
12		
10		

Marges de variation :

Question 1 : $57 - 13 = 44$

Question 2 : $36 - 27 = 9$

Les notes de la question 1 sont beaucoup plus dispersées. Elles influencent plus le classement final que celles de la question 2. Il est souvent utile d'ignorer les deux notes extrêmes : dans ce cas :

Marge Question 1 = 27

Question 2 = 8

Médians

= notes du milieu.

Ici, chiffre pair, donc moyenne entre 5° et 6°.

Question 1 :

$(25 + 26) : 2 = 25,5$

Question 2 :

$(31 + 32) : 2 = 31,5$

Pour la question 1, la différence assez nette entre la moyenne arithmétique et le médian (30 et 25,5) indique une distribution asymétrique des notes (ici : majorité sous le médian).

Pour la question 2, la symétrie est bonne (Moy. 31 et Méd. 31,5).

Remarquons que, dans une distribution parfaitement normale, la moyenne et le médian coïncident : voir notre exemple.

2° Estimation de l'écart type (σ).

$\sigma = 3/4$ de l'écart interquartile.

Donc :

a. Chercher le quartile supérieur, c'est-à-dire le milieu des notes au-dessus du médian.

Question 1 = 39

Question 2 = 33

b. Chercher le quartile inférieur.

$$\text{Question 1} = 22$$

$$\text{Question 2} = 28$$

c. Ecart interquartile.

$$\text{Question 1} = 39 - 22 = 17$$

$$\text{Question 2} = 33 - 28 = 5$$

d. Estimation du σ :

$$\text{Question 1} = 17 \times \frac{3}{4} = 12 \frac{3}{4}$$

$$\text{Question 2} = 5 \times \frac{3}{4} = 3 \frac{3}{4}$$

D. La concentration des résultats autour de la moyenne.

Reportons-nous de nouveau à la distribution des 100 notes prises comme exemple (p. 97). Nous observons que, plus nous nous rapprochons de la moyenne, plus les notes sont nombreuses :

12 notes sur 100 à la moyenne exacte

22 notes sur 100 entre 48 et 52

68 notes sur 100 entre 41 (-1σ) et 59 ($+1\sigma$)

Imaginons qu'il s'agisse de compositions françaises. Alors que le professeur pouvait noter de 0 à 100, presque un quart des notes se situent entre 48 et 52, soit à 4 points de distance.

Que signifie une différence de 4 points sur 100 en dissertation française ? Il est à peu près certain que, si nous faisons juger le même travail par dix professeurs différents, les écarts seront presque toujours supérieurs à 2. Or, si c'est la moyenne qui constitue la limite de l'échec (ici, elle coïncide avec la moitié de l'échelle totale), on peut dire, sans crainte de se tromper que, pour 22% du groupe d'élèves pris comme exemple, l'échec ou la réussite dépendra du pur hasard.

Voici, en dégradé, une image de la densité de population par rapport à la moyenne :



E. Courbe de Gauss vou

La courbe de Gauss ment puisqu'elle permet entre 25 et 35) et les moind

Dans le cas d'un postulent un petit nombre choisir parmi ceux qui s qui échappe à la forte co même, si l'on souhaite c trop exclusivement dans utile pour permettre à c normale est certainement

Comme il ne s'agit p hasard, mais de sujets crée artificiellement les c gaussienne. Comment ? E de difficulté (et d'efficacité

F. Comment savoir si u

Tout ce que nous ve de Gauss et du sigma n'e buent *normalement*. Avan donc de vérifier si cette

Il existe des procé atteindre ce but, mais, de d'une simple approximat

E. Courbe de Gauss voulue par les maîtres.

La courbe de Gauss est un excellent instrument de classement puisqu'elle permet d'identifier les meilleurs (par exemple entre 25 et 35) et les moins bons.

Dans le cas d'un concours où de nombreux candidats postulent un petit nombre de places, on peut, par exemple, choisir parmi ceux qui se situent entre $+1\sigma$ et $+3\sigma$ (partie qui échappe à la forte concentration autour de la moyenne). De même, si l'on souhaite classer les élèves — ce que l'on a fait trop exclusivement dans le passé, mais qui reste tout de même utile pour permettre à chacun de se situer —, la distribution normale est certainement utile.

Comme il ne s'agit plus, cette fois, de groupes choisis au hasard, mais de sujets spécialement entraînés, le professeur crée artificiellement les conditions nécessaires à une répartition gaussienne. Comment ? En dosant des *items* selon leurs indices de difficulté (et d'efficacité).

F. Comment savoir si une distribution est normale ?

Tout ce que nous venons de dire à propos de la courbe de Gauss et du sigma n'est valable que si les résultats se distribuent *normalement*. Avant d'engager toute opération, il importe donc de vérifier si cette condition est remplie.

Il existe des procédés mathématiques rigoureux pour atteindre ce but, mais, de nouveau, nous nous contenterons ici d'une simple approximation graphique : l'*histogramme*.

Le procédé est simple :

1. *Classer les résultats.*

On peut les ordonner du plus grand au plus petit, ou inversement.

Voici ce qu'on obtient en classant les 100 résultats que nous avons déjà pris comme exemples :

Notes	Nombre d'élèves ayant obtenu ces notes (fréquences) (f)
26	1
30	1
32	1
34	2
36	2
38	4
40	5
42	5
44	6
46	7
48	10
50	12
52	10
54	7
56	6
58	5
60	5
62	4
64	3
66	2
70	1
76	1

$N = 100$

Mais, si les notes sont nombreuses, ce procédé n'est ni rapide, ni pratique. Il est bien plus aisé de constituer des classes.

- 1° Calculer la marge de variation entre les deux notes extrêmes : $76 - 26 = 50$.
- 2° Diviser cette marge par 15^1 : $50 : 15 = 3,33$.
- 3° Choisir comme intervalle de classe un des deux nombres impairs les plus proches : 3 ou 5. On choisit ici 5, vu le petit nombre de notes.

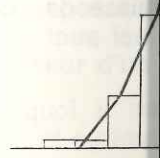
¹ Ce nombre est arbitraire. La pratique montre que, dans la plupart des cas, il conduit à une bonne répartition.

4° Placer la note supérieure : note supérieure : 74-75-76-77-78

	Classes
1	74 - 78
2	69 - 73
3	64 - 68
4	59 - 63
5	54 - 58
6	49 - 53
7	44 - 48
8	39 - 43
9	34 - 38
10	29 - 33
11	24 - 28
11 classes	Intervalle (i) =

2. *Dessiner l'histogramme*

Dans la partie « fréquence » a été fait avec page horizontalement graphique appelé histogrammes.

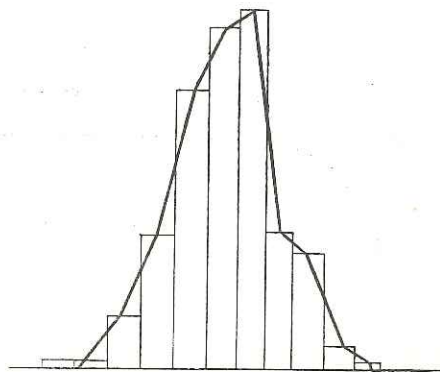


4° Placer la note supérieure au milieu de l'intervalle supérieur : note supérieure : 76 ; la classe supérieure est donc : 74-75-76-77-78.

	Classes	Fréquences (f)
1	74 - 78	I 1
2	69 - 73	I 1
3	64 - 68	III 5
4	59 - 63	III III 9
5	54 - 58	III III III III 18
6	49 - 53	III III III III II 22
7	44 - 48	III III III III III 23
8	39 - 43	III III 10
9	34 - 38	III III 8
10	29 - 33	II 2
11	24 - 28	I 1
11 classes	Intervalle (i) = 5	Nombre de notes (N) = 100

2. Dessiner l'histogramme.

Dans la partie « fréquences » du tableau ci-dessus, le pointage a été fait avec beaucoup de soin. Il suffit de tenir la page horizontalement pour voir apparaître le squelette du graphique appelé histogramme : il reste à tracer les rectangles.



B. Nature et faiblesse des échelles d'évaluation.

L'échelle d'évaluation est *une échelle ordinale*.

- 1) *Elle n'a ni zéro ni maximum naturels ou absolus ; elle commence et finit n'importe où, au choix de l'utilisateur. Par conséquent, même s'ils évaluent un même phénomène, du même point de vue, les juges n'ont jamais des échelles rigoureusement les mêmes ; les différences peuvent être considérables, et le sont souvent.*

Exemple : *Connaissance de l'anglais deuxième langue.*

Théoriquement, le zéro pourrait être le moment où l'élève ne connaît pas encore un seul mot de la langue. Toutefois, on juge rarement à partir de ce point à l'école.

Quant à la connaissance totale, absolue, elle n'existe évidemment pas.

Imaginons deux professeurs d'anglais jugeant, sans se concerter, deux classes où chaque élève a fait un court exposé oral. Fort probablement, les deux extrémités des échelles d'évaluation utilisées indépendamment par les deux professeurs seront plus définies par l'élève jugé le plus fort et l'élève jugé le plus faible, dans son groupe, que par une évaluation abstraite de la quantité d'anglais qui devrait être connue au niveau pédagogique considéré.

- 2) *Les degrés ne sont pas de même grandeur à l'intérieur d'une même échelle.*

Il est, par exemple, impossible de démontrer que la distance séparant la bonne de la très bonne connaissance de l'anglais est la même que la distance séparant la connaissance moyenne de la bonne.

Si les degrés sont : A, B, C, D, E, on ne peut donc pas dire que $A - B = B - C = 0$

- 3) *Des degrés correspondants d'une échelle à l'autre (pour une même branche) ne sont pas de même grandeur.*

Bon en anglais, pour un professeur, n'est pas quantitativement égal à bon en anglais pour un autre : rien ne peut établir pareille égalité de façon mathématique.

- 4) *Des échelles portant pas comparables.*

Pour les élèves d'une seule classe, le seul maître enseignant, la notation ne recouvre pas la même *bon* en français. Les échelles ne sont pas comparables sur le plan de la notation.

De tout ceci, il résulte que les échelles ne mettent aucune opération de comparaison. La notation française et l'anglais ne permettent rien de plus. C'est un savoir total. C'est un savoir décennal, dans nos écoles, d'utiliser des échelles différentes branches, différentes branches : de 0 à 120 points en sciences, etc.

Si le système a permis de permettre d'identifier les élèves (tout dans les premières années).

C. Utilité.

Malgré ses limitations, dans des cas, le seul instrument pour évaluer notre jugement sur des cas complexes.

En fait, elle permet de classer des éléments *qualitatifs*, à l'aide de *sables complémentaires* des examens objectifs ou les tests.

D. Construction.

Beaucoup d'échelles d'évaluation apparente avec laquelle on construit.

En réalité, il n'est pas possible de construire un instrument exige de longue clarification théorique, et des méthodes raffinées comme l'analyse de contenu à déterminer dans quelle mesure les mêmes choses sous

'évaluation.

helle ordinale.

ls ou *absolus* ; elle com-
hoix de l'utilisateur. Par
n même phénomène, du
ont jamais des échelles
différences peuvent être

ais deuxième langue.

être le moment où l'élève
mot de la langue. Toute-
de ce point à l'école.
absolue, elle n'existe évi-

nglais jugeant, sans se
e élève a fait un court
es deux extrémités des
pendamment par les deux
r l'élève jugé le plus fort
son groupe, que par une
d'anglais qui devrait être
sidéré.

randeur à l'intérieur d'une

démontrer que la distance
ne connaissance de l'an-
séparant la connaissance

on ne peut donc pas dire

chelle à l'autre (pour une
me grandeur.

ar, n'est pas quantitative-
un autre : rien ne peut
mathématique.

4) *Des échelles portant sur des branches différentes ne sont pas comparables.*

Pour les élèves d'une même classe d'école primaire où un seul maître enseigne toutes les branches, être *bon* en mathématique ne recouvre pas les mêmes phénomènes qu'être *bon* en français. Les échelons ne sont certainement pas comparables sur le plan quantitatif.

De tout ceci, il résulte que les échelles d'évaluation ne permettent aucune opération arithmétique. Si on note la composition française et l'anglais deuxième langue de 0 à 10, rien ne permet d'additionner les deux résultats pour évaluer un savoir total. C'est pourtant ce que l'on fait, depuis des décennies, dans nos écoles. Il est tout aussi suspect, sinon plus, d'utiliser des échelles numériques différentes pour les différentes branches, en fonction d'une pondération empirique : de 0 à 120 pour le français, de 0 à 80 pour les sciences, etc.

Si le système a pu fonctionner, c'est qu'il a, de toute façon, permis d'identifier les meilleurs (ceux qui se classent *en tout* dans les premiers) et les plus faibles (derniers *en tout*).

C. Utilité.

Malgré ses limitations, l'échelle d'évaluation est, dans bien des cas, le seul instrument dont nous disposons pour concrétiser notre jugement sur des comportements humains complexes.

En fait, elle permet d'apporter, dans un dossier scolaire, des éléments *qualitatifs*, aussi objectifs que possible, *indispensables compléments* des résultats *quantitatifs* apportés par les examens objectifs ou les tests.

D. Construction.

Beaucoup d'éducateurs se laissent abuser par la facilité apparente avec laquelle on construit une échelle.

En réalité, il n'est pas rare que la mise au point de pareil instrument exige de longs mois de travail, nécessaires à la clarification théorique, et appelle l'utilisation de techniques raffinées comme l'analyse factorielle. Celle-ci aide notamment à déterminer dans quelle mesure plusieurs échelles évaluent les mêmes choses sous des appellations différentes.

Dans ce domaine aussi, il importe donc que, le plus rapidement possible, les enseignants puissent s'appuyer sur des services de recherche bien équipés en spécialistes et en ordinateurs.

Cela ne signifie toutefois pas que les maîtres ne peuvent rien faire en attendant.

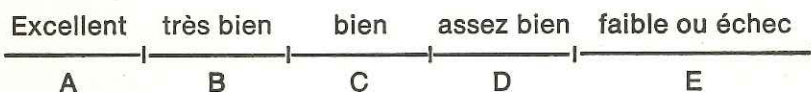
1. Combien de degrés ?

Théoriquement, une échelle d'évaluation peut compter une quantité infinie d'échelons. En pratique, les juges ne classent avec quelque sûreté que si l'on n'utilise qu'un petit nombre de degrés : 3, 5, 7 ou 9.

En nous référant à la littérature expérimentale et à de multiples observations que nous avons pu faire directement, nous conseillons que l'on s'en tienne le plus souvent à 5 échelons. Pareilles échelles permettent des classements relativement sûrs et fidèles, à condition que chaque échelon soit très clairement défini et les notateurs bien formés.

Pratiquement,

- 1° On classe d'abord en trois catégories : les meilleurs, les plus faibles, les autres ;
- 2° On répartit ensuite les « autres » en bons, moyens et faibles.



Remarquons que l'échelon C est à cheval sur le centre de l'échelle, c'est-à-dire sur la moyenne arithmétique. Comme nous l'avons vu, à propos de la courbe de Gauss, une forte concentration se produit souvent autour de ce point. Choisir un nombre pair d'échelons créerait un centre dans l'échelle, devant ou derrière lequel plusieurs élèves basculeraient au hasard.

Dans le cas d'une distribution normale, il n'est pas exclu que ce verdict du hasard affecte 20, 30, voire 40 % des élèves.

2. Définir l'objet de l'évaluation.

Demander, par exemple, d'apprécier le « courage » des membres d'un groupe donné — sans autre précision — conduirait à des réponses presque entièrement dépourvues d'intérêt. Quel ou quels sens les observateurs auront-ils donnés au mot « courage » : ardeur, volonté, zèle, persévérance, bravoure, fermeté, stoïcisme ?

Et même si nous précisons la fermeté devant le distinction entre l'intrépidité et

Pour pallier les imprécisions, *bien, moyen, etc.*, il est nécessaire une description aussi précise que possible pour apprécier, et d'illustrer la valeur. Définir non par des formules mais par des jugements précis est une condition

Voici comment Schonegger a défini « soi » :

Extrêmement confiant en soi. Presque trop sûr de lui.	Très confiant en ses propres forces.
---	--------------------------------------

Description du trait.

« Sous sa forme positive, il se manifeste par des manifestations suivantes : capable de faire face aux difficultés pendant et prêt à assumer la responsabilité »

» L'enfant qui a confiance avec le minimum d'assistance doit être aidé constamment. Il est capable de construire des instructions claires ; le second degré qu'on l'aide durant toute la durée de l'activité »

Situations types.

- 1° A-t-il peur de l'obscurité ?
- 2° Peut-il prendre soin de lui-même faut-il que quelqu'un s'occupe de lui ?
- 3° Voyage-t-il seul en train (à l'âge de 9 ans).

1 F.J. SCHONELL, *Backwardness Measurement of Personality* (« Educational Psychology »)

... donc que, le plus rapide-
ment s'appuyer sur des ser-
spécialistes et en ordina-

... les maîtres ne peuvent

... évaluation peut compter
... ratique, les juges ne clas-
... n'utilise qu'un petit nom-

... expérimentale et à de
... ons pu faire directement,
... le plus souvent à 5 éche-
... des classements relative-
... chaque échelon soit très
... formés.

... gories : les meilleurs, les

... en bons, moyens et faibles.

... ez bien faible ou échec
D |-----| E

... à cheval sur le centre de
... ne arithmétique. Comme
... urbe de Gauss, une forte
... ur de ce point. Choisir un
... entre dans l'échelle, devant
... asculeraient au hasard.

... normale, il n'est pas exclu
... 30, voire 40 % des élèves.

... évaluer le « courage » des
... s autre précision — con-
... rement dépourvues d'inté-
... eurs auront-ils donnés au
... e, persévérance, bravoure,

Et même si nous précisons que par « courage » nous enten-
dons la fermeté devant le danger, ferons-nous aisément la dis-
tinction entre l'intrépidité et la témérité ?

Pour pallier les imprécisions de notations telles que *très
bien, moyen, etc.*, il est nécessaire d'ajouter à l'échelle une
description aussi précise que possible du trait ou de l'objet à
apprécier, et d'illustrer la définition par des situations types.
Définir non par des formules abstraites, mais par des comporte-
ments précis est une condition *sine qua non* de validité.

Voici comment Schonell¹ présente le trait « Confiance en
soi » :

Confiance en soi

Extrêmement confiant en soi. Presque trop sûr de lui.	Très confiant en ses propres forces.	Confiant.	Manque de confiance. Timide.	Manque extrême de confiance. Dépend des autres. Décline les responsa- bilités.
--	---	-----------	------------------------------------	---

Description du trait.

« Sous sa forme positive, cette qualité est marquée par les
manifestations suivantes : l'individu compte sur lui-même, est
capable de faire face aux difficultés, a de l'assurance, est indé-
pendant et prêt à assumer des responsabilités.

» L'enfant qui a confiance en lui-même essaie d'avancer
avec le minimum d'assistance ; celui qui manque de confiance
doit être aidé constamment. Le premier aime de voir ce qu'il
est capable de construire et de produire quand il a reçu des
instructions claires ; le second veut qu'on lui mâche la besogne,
qu'on l'aide durant toute la phase de la réalisation. »

Situations types.

- 1° A-t-il peur de l'obscurité ?
- 2° Peut-il prendre soin de lui-même et de ce qu'il possède ou
faut-il que quelqu'un soit tout le temps à ses côtés ?
- 3° Voyage-t-il seul en tram ou en bus ? (pour enfants de plus
de 9 ans).

¹ F.J. SCHONELL, *Backwardness in the Backward Subjects*, cité par F. WARBURNE,
Measurement of Personality (« Educational Research », novembre 1961, p. 9).

- 4° Parle-t-il librement à des visiteurs inconnus ?
- 5° Est-il bon dans les jeux ? Sait-il nager ?
- 6° Est-il à l'aise et répond-il avec assurance aux examens oraux ?
- 7° Lit-il bien, dramatise-t-il bien un texte devant la classe ?
- 8° S'attaque-t-il bien à des tâches nouvelles ou pose-t-il constamment des questions à ses compagnons et à ses maîtres ?

Définir les degrés de l'échelle.

Exemple : Organisation de l'enseignement de la lecture.

0	1	2	3	4	5
Médiocre En lecture, tous les élèves sui- vent la même progression. Pas de travail par groupes.	Assez bien Cf. 1. Mais parfois un élève très lent reçoit un peu moins de travail que les autres.	Moyen Constitution de 2 ou 3 groupes, selon les aptitudes en lecture. Peu de flexibilité dans le groupement.	Très bien Groupement selon les aptitudes. Flexibilité.	Excellent Groupement après étude approfondie des aptitudes et des difficultés rencontrées. Grande flexibilité.	

On aura remarqué que ce dernier exemple combine les échelles graphique, numérique et descriptive.

Le danger d'une définition insuffisante des cinq degrés a été montré expérimentalement par R. Weiss¹. Adoptant les échelles descriptives imposées par le Ministère autrichien de l'Education nationale, Weiss a invité 92 instituteurs à noter deux compositions d'élèves du niveau où ils enseignaient (4^e primaire) et 272 instituteurs à noter deux travaux d'arithmétique (problèmes ; 4^e et 5^e années).

Les diagrammes ci-dessous montrent combien les notations sont divergentes, pour un même travail (dispersion des notes sur toute la longueur de l'échelle).

¹ R. WEISS, o.c.

Notation de

I

Orthographe

Style

Fond

Note globale

1	2
1	
1	
1	

II

Orthographe

Style

Fond

Note globale

1	2
1	
1	
1	

Notation de deu

I

II

1	
2	

On le voit, il ne suffit en remplacement du système tous les problèmes soient r

E. Utilisation.

1) Combien d'élèves par é

Premier cas : Elève co

Les échelons marquent sion vers un pouvoir ou un à atteindre. Dans un enseig tif est commun à la majori la classe. Il doit évidemmer les élèves puissent se dé large, ce qui risque de nuire inopérant.

inconnus ?
ager ?
assurance aux examens

texte devant la classe ?
nouvelles ou pose-t-il con-
compagnons et à ses maî-

seignement de la lecture.

3 4 5

	Très bien Groupement selon les aptitudes. Flexibilité.	Excellent Groupement après étude approfondie des aptitudes et des difficultés rencontrées. Grande flexibilité.
--	--	---

nièr exemple combine les
escriptive.

ffisante des cinq degrés a
R. Weiss¹. Adoptant les
le Ministère autrichien de
92 instituteurs à noter deux
ils enseignaient (4^e primai-
travaux d'arithmétique (pro-

ntrent combien les notations
avail (dispersion des notes

NOTATION DE DEUX COMPOSITIONS

Notation de deux rédactions

		Note moyenne	Sigma
I	Orthographe	2.89	1.04
	Style	2.29	0.53
	Fond	2.08	0.79
	Note globale	2.45	0.80

		Note moyenne	Sigma
II	Orthographe	2.99	0.96
	Style	2.12	1.05
	Fond	1.78	0.85
	Note globale	2.54	0.80

Notation de deux travaux d'arithmétique

		Note moyenne	Sigma
I		2.56	0.77
II		3.55	0.80

On le voit, il ne suffit pas d'adopter une nouvelle échelle en remplacement du système de notes traditionnel pour que tous les problèmes soient résolus. Loin s'en faut !

E. Utilisation.

1) Combien d'élèves par échelon ?

Premier cas : Elève comparé à lui-même.

Les échelons marquent les degrés à franchir dans l'ascension vers un pouvoir ou un savoir. Le maître a défini un objectif à atteindre. Dans un enseignement non individualisé, cet objectif est commun à la majorité, sinon à la totalité des élèves de la classe. Il doit évidemment être choisi de telle sorte que *tous* les élèves puissent se dépasser. Ce sera donc un objectif large, ce qui risque de nuire à sa précision au point de le rendre inopérant.

Il est clair que nous agissons actuellement comme si, au début de l'année scolaire, tous les élèves se trouvaient approximativement à un même endroit de l'échelle du savoir. On suppose donc que, l'année précédente, tous ont atteint le degré supérieur de l'échelle précédente.

Or, nous le sentons bien, la réalité est différente et la recherche confirme notre sentiment :

a) *Marges de variation de l'âge mental.*

Lors de la révision du test d'intelligence de Binet-Simon par Terman et Merrill (1937), on a constaté, aux Etats-Unis¹ :

- qu'en première année primaire, l'âge mental variait de 4 à 8 ans ;
- qu'en sixième année primaire, l'âge mental allait de 8 à 16 ans ;
- que, dans l'enseignement secondaire, une marge de variation de 8 à 10 ans n'était pas exceptionnelle.

b) *Marges de variation du rendement scolaire.*

Il suffit d'observer combien les normes des tests de connaissances, étalonnés par année scolaire, se chevauchent, pour prendre conscience de l'ampleur des marges de variation.

S'appuyant sur ses propres recherches et sur celles de Lindquist, Cornell, Learned et Wood, W.-W. Cook² fournit les indications suivantes :

- Pour la compréhension de la lecture, pour le vocabulaire, les sciences, la géographie et l'histoire, la marge de variation du rendement est :
 - En 1^{re} primaire : de 3 à 4 ans.
 - En 4^e primaire : de 5 à 6 ans.
 - En 6^e primaire : de 7 à 8 ans.
- Pour le raisonnement arithmétique et le calcul :
 - En 6^e primaire : de 6 à 7 ans.
- En culture générale (sciences, littérature, arts, histoire, géographie et langue maternelle), au niveau universitaire :

¹ Cf. Q. McNEMAR, *The Revision of the Stanford-Binet Scale*, Boston, Houghton-M., 1942.

² W.W. COOK, *The Functions of Measurement in the Facilitation of Learning*, in E.F. LINDQUIST, Ed., *Educational Measurement*, Washington, A.C.E., 1961. 4^e éd. pp. 3-47.

- 28 % des étudiants d'âge moyen des étudiants
- 10 % des étudiants de 4^e année ne des élèves de fin d'année

Semblables observations sur le pied de la lettre, mais dont il ne faudrait être multipliées à volonté.

Elles prouvent qu'aussi bien qu'ils progresseront selon le système de classement allant de pair avec le souci de progresser à son allure propre, ils iront aussi loin qu'il le peut, sera.

Toutefois, comme pour les problèmes économiques, l'enseignement individualisé offre une solution pas avant de nombreuses années. Dans une école semi-individualisée, par branches principales. Ainsi, un élève doué en langue, pour ce niveau A, pour cette branche, les problèmes d'arithmétique.

Afin de ne pas alourdir l'annexe une description détaillée puis cinquante ans selon ce p...

Deuxième cas : Elèves c...

Même si l'on essaie d'atténuer des différences suffisantes sur de nouveau apparaître une fois moins, son ébauche) où la m... cée vers le haut.

Il semble donc justifié d'expliquer selon des pourcentages qui s... male.

¹ Ce n'est pas par hasard que nous voyons d'arithmétique, en général. Dans une branche bien doué en calcul, mais faible en probabilité si l'on ne groupe pas assez finement, considérables.

ctuellement comme si, au
èves se trouvaient approxi-
e l'échelle du savoir. On
e, tous ont atteint le degré

éalité est différente et la
:

mental.

Intelligence de Binet-Simon
constaté, aux Etats-Unis¹ :

l'âge mental variait de 4 à

l'âge mental allait de 8 à

laire, une marge de varia-
ceptionnelle.

ement scolaire.

normes des tests de con-
laire, se chevauchent, pour
marges de variation.

cherches et sur celles de
J. W.-W. Cook² fournit les

cture, pour le vocabulaire,
histoire, la marge de varia-

e et le calcul :

littérature, arts, histoire,
), au niveau universitaire :

ford-Binet Scale, Boston, Houghton-M.,

in the Facilitation of Learning, in E.F.
ngton, A.C.E., 1961. 4^e éd. pp. 3-47.

- 28 % des étudiants de 4^e année sont inférieurs à la moyenne des étudiants de 2^e année ;
- 10 % des étudiants de 4^e année sont inférieurs à la moyenne des élèves de fin d'enseignement secondaire.

Semblables observations — qu'il ne faut pas prendre au pied de la lettre, mais dont il faut retenir la tendance — pourraient être multipliées à volonté.

Elles prouvent qu'aussi longtemps que nos écoles fonctionneront selon le système de classes rigides, l'évaluation continue allant de pair avec le souci de permettre à chaque élève de progresser à son allure propre, de se dépasser toujours, d'aller aussi loin qu'il le peut, sera un leurre.

Toutefois, comme pour des raisons technologiques et économiques, l'enseignement individualisé ne se généralisera pas avant de nombreuses années, seul l'enseignement semi-individualisé offre une solution réaliste, immédiatement applicable. Dans une école semi-individualisée, des groupes homogènes, par branches principales, se substituent à la classe. Ainsi, un élève doué en langue étrangère peut travailler à un niveau A, pour cette branche, mais se joindre au niveau D pour les problèmes d'arithmétique¹.

Afin de ne pas alourdir ce texte, nous avons reporté en annexe une description détaillée d'une école fonctionnant depuis cinquante ans selon ce principe.

Deuxième cas : Elèves comparés entre eux.

Même si l'on essaie d'amener chacun à un niveau élevé, des différences suffisantes subsisteront probablement pour faire de nouveau apparaître une distribution gaussienne (ou, du moins, son ébauche) où la moyenne se sera simplement déplacée vers le haut.

Il semble donc justifié d'essayer de peupler chaque échelon selon des pourcentages qui s'approchent de la distribution normale.

¹ Ce n'est pas par hasard que nous parlons de *problèmes d'arithmétique* et non d'arithmétique, en général. Dans une branche complexe comme celle-ci, un élève peut être bien doué en calcul, mais faible en problèmes. La recherche contemporaine montre que si l'on ne groupe pas assez finement, les marges de variation entre élèves restent considérables.

Dans une classe habituelle, il n'est évidemment pas possible de respecter strictement cette proportion, en raison du petit nombre d'élèves. On aura, par exemple :

La distribution normale serait :

Excellent :	1	5 %	2,5 %
Très bons :	4	20 %	13,5 %
Bons (moyens) :	10	50 %	68 %
Faibles :	4	20 %	13,5 %
Très faible :	1	5 %	2,5 %
	<hr/>	<hr/>	<hr/>
	20	100 %	100 %

Il faut y insister, se forcer à respecter pareille répartition (*distribution forcée*) n'implique pas nécessairement un échec pour un certain nombre d'élèves. Toutes les notes sont *relatives* les unes par rapport aux autres : « très faible » peut, dans certains groupes forts, se situer au-dessus de la note d'échec.

C'est d'ailleurs pourquoi *il vaut mieux d'adopter des lettres de classement* plutôt que les formules traditionnelles.

2) Lutter contre la contamination et la tendance centrale.

Outre les dangers de stéréotypie et d'effets de halo que nous décrivons, page 33, un autre phénomène de contamination, de nature plus mécanique, guette aussi les juges. S'ils utilisent consécutivement, pour un même élève, une série d'échelles orientées dans le même sens (par exemple du positif au négatif ou du meilleur au plus faible) et peut-être même présentées sur une même page, les notations tendent à se grouper dans un même côté.

On conseille de faire figurer chaque échelle sur une page différente, de tirer au sort le sens de présentation de chacune.

Enfin, il est bon de ménager un intervalle assez long après l'évaluation de chaque qualité d'un même sujet et de faire évaluer une même qualité par autant de juges que possible.

Beaucoup de juges ont aussi tendance à grouper leurs notes vers le centre de l'échelle : crainte de surévaluer ou de sous-évaluer un élève, peur de supporter la responsabilité d'un échec. La règle de la distribution forcée rend ici d'utiles services. Mais, pour être réellement efficace, elle doit pouvoir s'appuyer sur une description précise et nuancée des différents échelons.

Quand les évaluations combinées, les professeurs ont tendance à donner une note de facilité qu'est la note la plus élevée. On a vu que souvent une large partie de les collègues qui osent utiliser l'échelle.

F. Comment synthétiser les notes.

Au terme d'une période d'évaluations (que l'on souhaite synthétiser), faute de quoi maîtriser la tâche à accomplir.

En revenant au simple classement des notes, on risque de ruiner les efforts de rigueur des ordinales, théoriquement du moins, par la soustraction.

Voici une méthode simple pour synthétiser les notes dans chaque branche, puis les comparer.

Au départ, on dispose de :

Note	
A	
B	
C	
D	
E	
	Dates

Un professeur utilise au moins deux aspects de la notation : le contenu, le style, l'orthographe, la rédaction.

Au fur et à mesure des évaluations, on inscrit les notes dans le tableau, en indiquant :

est évidemment pas possible proportion, en raison du exemple :

La distribution normale serait :

5 %	2,5 %
10 %	13,5 %
10 %	68 %
10 %	13,5 %
5 %	2,5 %
<hr/>		
100 %		100 %

respecter pareille répartition n'est pas nécessairement un échec. Toutes les notes sont relatives : « très faible » peut, dans certains cas, être au-dessus de la note d'échec.

Il est préférable d'adopter des lettres de notation traditionnelles.

Il faut éviter la tendance centrale.

l'influence de l'habitude et d'effets de halo que l'on observe dans le phénomène de contamination aussi les juges. S'ils évaluent le même élève, une série de notes faibles (par exemple du positif au négatif) et peut-être même des notations tendent à se grouper.

Chaque échelle sur une page doit être présentée de façon claire.

Il faut laisser un intervalle assez long après l'évaluation d'un même sujet et de faire intervenir un grand nombre de juges que possible.

Il faut éviter la tendance à grouper leurs notes par crainte de surévaluer ou de sous-évaluer. Porter la responsabilité d'un jugement forcé rend ici d'utiles services. Si elle est efficace, elle doit pouvoir être précise et nuancée des différents aspects.

Quand les évaluations par plusieurs juges doivent être combinées, les professeurs qui se laissent séduire par la solution de facilité qu'est la note centrale doivent savoir qu'ils abdiquent une large partie de leur influence en faveur de leurs collègues qui osent utiliser l'ensemble de l'échelle.

F. Comment synthétiser les évaluations.

Au terme d'une période scolaire, il importe de synthétiser les évaluations (que l'on souhaite aussi nombreuses que possible), faute de quoi maîtres et élèves ne sauront pas les interpréter.

En revenir au simple calcul de la moyenne risquerait de ruiner les efforts de rigueur précédents. Car des évaluations ordinales, théoriquement du moins, ne permettent ni addition, ni soustraction.

Voici une méthode simple permettant de situer un élève dans chaque branche, puis pour l'ensemble des branches.

Au départ, on dispose du tableau suivant :

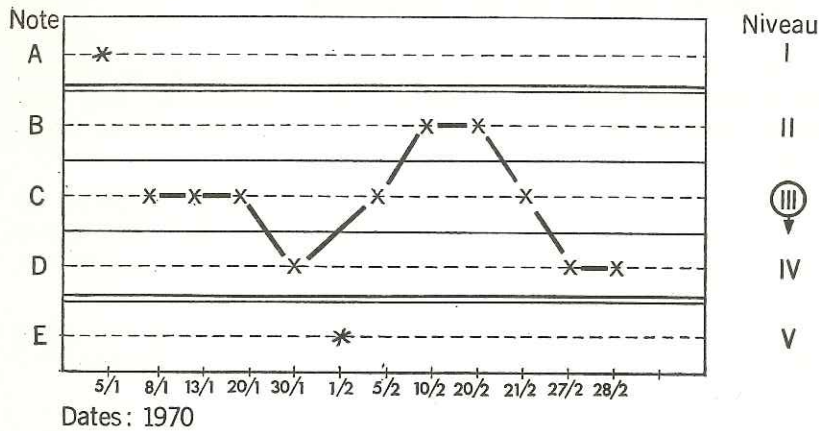
Note		Niveau
A		I
B		II
C		III
D		IV
E		V

Un professeur utilise autant de tableaux semblables qu'il souhaite distinguer d'aspects dans son cours : dictée, grammaire, orthographe, rédaction,...

Au fur et à mesure des évaluations, on porte une croix dans le tableau, en indiquant la date au bas.

Au moment de faire la synthèse, on commence par barrer (sans l'effacer !) la croix correspondant à la note la plus haute et celle qui représente la note la plus basse, afin d'éliminer des notes qui pourraient être accidentelles.

Supposons que l'on obtienne le tableau ci-dessous.



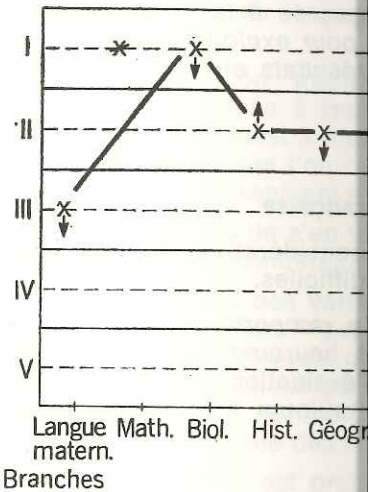
Interprétation.

- 1) Les résultats se concentrent sur trois bandes : II, III, IV. Ils ne sont donc ni excellents, ni faibles ou insuffisants. La zone de concentration est marquée par des traits doubles.
- 2) Le plus grand nombre de notes (le mode) se situent en bande III. C'est la tendance dominante.
- 3) Comme la bande la plus peuplée, après la III, est la IV, une flèche sera placée, dans cette direction, sous la III.
- 4) En réunissant les croix, on obtient un profil qui fait apparaître :
 - a) Une assez bonne homogénéité des résultats. Si l'élève se promenait anarchiquement sur les 5 bandes, il faudrait voir s'il est seul dans ce cas (cause à déterminer au niveau de l'élève), ou si la majorité de la classe se trouve dans la même situation (voir si les travaux du professeur ne sont pas mal ajustés).
 - b) Une descente en fin de période. Elle n'est pas très accusée : elle peut s'expliquer par la fatigue, par une maladie bénigne,...

On le voit, il ne suffit pas par le profil, il faut tâcher de commentaires, oraux ou écrits et guider l'élève et ses parents.

Remarquons, par ailleurs, que si le professeur se borne à noter moins avec une seule note de certains maîtres obnubilés par l'enseignement... ou allergiques à l'usage de la notation, que l'introduction d'un système de notation plus riche pourra heureusement améliorer.

TABLEAU POUR L'ÉVALUATION DE LA PERIODE DU



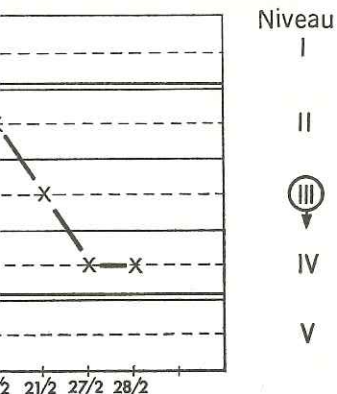
Interprétation.

Quatre notes d'ensemble se situent en bande III. C'est la tendance dominante. Elle sera donc marquée par deux cercles par un trait.

En cas d'égalité pratique (deux notes d'ensemble dans deux niveaux contigus), soit par un trait.

on commence par barrer
ant à la note la plus haute
basse, afin d'éliminer des
es.

tableau ci-dessous.



trois bandes : II, III, IV. Ils
les ou insuffisants. La zone
r des traits doubles.

s (le mode) se situent en
ante.

ée, après la III, est la IV,
ette direction, sous la III.

nt un profil qui fait apparaî-

é des résultats. Si l'élève
sur les 5 bandes, il faudrait
s (cause à déterminer au
orité de la classe se trouve
si les travaux du professeur

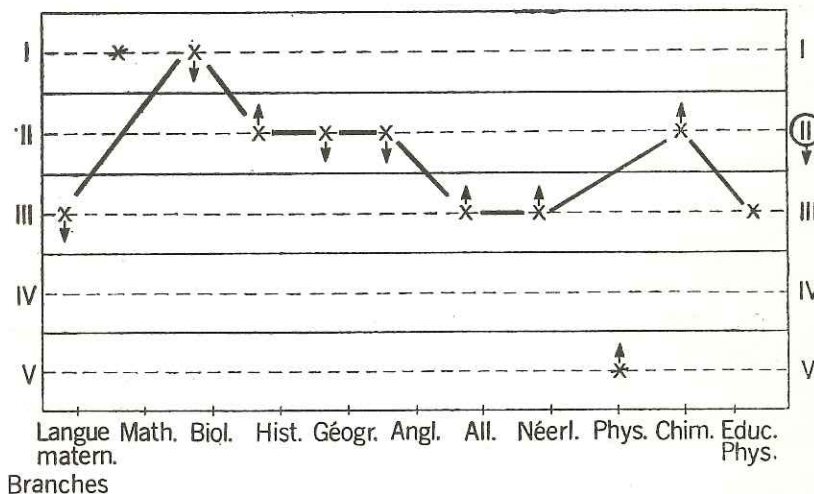
e. Elle n'est pas très accu-
r la fatigue, par une mala-

On le voit, il ne suffit pas de constater une situation traduite par le profil, il faut tâcher d'en découvrir le pourquoi. Ainsi, les commentaires, oraux ou écrits, des maîtres viennent-ils éclairer et guider l'élève et ses parents.

Remarquons, par ailleurs, que ce système ne peut fonctionner si le professeur se borne à deux ou trois évaluations ; encore moins avec une seule ! Cette fois, c'est le comportement de certains maîtres obnubilés par la quantité de matière à enseigner... ou allergiques au long travail de notation des travaux, que l'introduction d'une nouvelle forme d'évaluation viendra heureusement améliorer.

TABLEAU POUR L'ENSEMBLE DES BRANCHES

PERIODE DU AU 19 ..



Interprétation.

Quatre notes d'ensemble se trouvent en bande II et quatre en bande III. C'est la tendance du profil (ici vers le haut) qui décidera de la dominante : ici II, avec flèche vers le bas.

En cas d'égalité pratique entre deux niveaux, on peut soit décider que le niveau supérieur l'emporte (surtout s'il s'agit de deux niveaux contigus), soit entourer deux chiffres et réunir les deux cercles par un trait.

Remarques.

1° Ce système de synthèse des évaluations ne détient aucun pouvoir magique ! Si l'on commet des erreurs signalées dans le présent chapitre, la note de synthèse sera probablement dépourvue de sens.

Un phénomène de concentration dans les bandes médianes étonne souvent les utilisateurs. Ils oublient que, dans le système traditionnel, aboutissant à un pourcentage, les notes se concentraient aussi dans la même zone : entre 65 et 75 ou 80 %, tout le reste de l'échelle (de 100 échelons !) étant bien souvent presque vide.

En réalité, le phénomène de concentration est plus frappant parce que l'échelle à cinq degrés est beaucoup mieux visualisée que l'ancienne.

2° Rien n'empêche de pondérer les notes, c'est-à-dire d'accorder, par exemple, deux fois plus d'importance au travail en classe qu'aux travaux à domicile, à tel aspect d'une branche plutôt qu'à tel autre, etc.

3° *On trouvera, dans la partie consacrée à la « modération », des indications supplémentaires pour exploiter les notes, en particulier pour déterminer les résultats et les classements en fin d'année (voir page 171).*

G. Un cas particulier :

la notation de la composition française.

Ce problème mérite quelques considérations particulières, parce qu'il compte parmi les plus difficiles.

S'il était résolu, la docimologie remporterait peut-être sa plus belle victoire. Hélas ! ou, plutôt, heureusement pour l'homme, la plus noble de ses activités, l'évaluation du beau, du vrai et du bien, échappera sans doute toujours à la quantification objective et donc automatisable.

1) Quatre méthodes d'évaluation.

Les principales méthodes se ramènent à quatre : la méthode de l'impression générale, l'échelle de spécimens, la méthode analytique et les comptages de fréquences.

a) La méthode de l'impression générale.

C'est la plus subjective. On lui reproche le manque de fidélité d'un même correcteur, les désaccords entre plusieurs notateurs, spécialistes ou non, et son inefficacité fréquente : maint professeur s'est usé à noter avec soin des centaines de travaux par an sans obtenir des résultats convaincants.

Il y aurait beaucoup à dire sur la composition française, et sur les habitudes profondes de l'écrivain en prose, pas d'en traiter. Retenons deux faiblesses souvent méthodologiques, grève l'évaluation que tellement lourde que...

Malgré ses faiblesses, le système est la mieux en harmonie avec la composition. En utilisant le maximum, à cinq degrés, obtenu entre des notateurs et appartenant à un même niveau, loin les précautions à prendre.

Une composition corrigée par nous, ambitionner une simple correction du langage à cet effet.

La composition traditionnelle, d'exception, destinée à la maturation, et aussi à la correction des maladresses que l'on évite par des exercices méthodiques d'enrichissement.

Si l'on décide de s'en tenir à la technique de l'évaluation.

Elle comporte des variantes, décrites par S. Wiseman¹ et qui s'évaluent indépendamment, à une allure d'environ cinquante pour les obliger à se décider sur la moyenne arithmétique des notes.

Cette méthode est précieuse pour les travaux narratifs ou descriptifs, recouvrant toutes les compétences, la supériorité de la méthode, pourtant valide aussi en ce qui concerne le haut niveau intellectuel.

La validité générale de la « dissertation » est confirmée par les travaux de J. Britton.

¹ S. WISEMAN, The Marking of Essays, *British Journal of Educational Psychology*, 1958, 28, 1-10.

évaluations ne détient aucun
net des erreurs signalées
te de synthèse sera pro-

dans les bandes médianes
oublie que, dans le sys-
un pourcentage, les notes
même zone : entre 65 et 75
e (de 100 échelons !) étant

centration est plus frappant
est beaucoup mieux visua-

notes, c'est-à-dire d'accor-
d'importance au travail en
à tel aspect d'une branche

acrée à la « modération »,
pour exploiter les notes, en
résultats et les classements

nçaise.

onsidérations particulières,
fficiles.

remporterait peut-être sa
heureusement pour l'hom-
évaluation du beau, du vrai
oujours à la quantification

èment à quatre : la métho-
de spécimens, la méthode
nces.

générale.

proche le manque de fidé-
saccords entre plusieurs
on inefficacité fréquente :
avec soin des centaines de
ultats convaincants.

Il y aurait beaucoup à dire sur la méthodologie de la composition française, et surtout sur la nécessité d'une réforme profonde des habitudes scolaires qui traitent chaque élève comme un écrivain en puissance, mais il ne nous appartient pas d'en traiter. Retenons simplement que la conjugaison de deux faiblesses souvent accusées, l'une docimologique, l'autre méthodologique, grève l'exercice de rédaction d'une hypothèse tellement lourde que certains déclarent cet exercice inutile. Nous ne sommes pas de cet avis.

Malgré ses faiblesses, la méthode de l'impression générale est la mieux en harmonie avec la complexité essentielle de la composition. En utilisant une échelle courte, à trois ou, au maximum, à cinq degrés, un accord assez élevé peut être obtenu entre des notateurs expérimentés, de même formation et appartenant à un même milieu scolaire. Nous verrons plus loin les précautions à prendre avant de décider d'un échec.

Une composition corrigée par cette méthode ne doit pas, croyons-nous, ambitionner d'améliorer beaucoup le style ou la simple correction du langage écrit : d'autres exercices existent à cet effet.

La composition traditionnelle doit être une prestation d'exception, destinée à repérer des talents ou à en contrôler la maturation, et aussi à identifier des faiblesses, des carences, des maladresses que l'on combattrait dans des exercices systématiques d'enrichissement et de remédiation.

Si l'on décide de s'en tenir à la seule appréciation globale, *la technique de l'évaluation en temps imposé* mérite l'attention.

Elle comporte des variantes mineures. La méthode proposée par S. Wiseman¹ est caractéristique. Quatre notateurs évaluent indépendamment des travaux faits en trente minutes, à une allure d'environ cinquante copies à l'heure (taux imposé pour les obliger à se décider très rapidement). On calcule la moyenne arithmétique des notes.

Cette méthode est probablement la plus efficace pour les travaux narratifs ou descriptifs. Pour les « dissertations », mot recouvrant toutes les compositions où le sujet doit argumenter, la supériorité de la méthode est plus discutée. Elle semble pourtant valide aussi en ce domaine si les notateurs sont de haut niveau intellectuel.

La validité générale de la méthode et son applicabilité à la « dissertation » est confirmée par deux recherches expérimentales de J. Britton.

1 S. WISEMAN, The Marking of English Composition in Grammar School Selection, in *British Journal of Educational Psychology*, XIX, nov. 1949, 208.

1^{re} expérience ¹.

Un ensemble de compositions écrites par 168 élèves de 15 ans a d'abord été noté, selon la méthode analytique, par un examinateur expérimenté. Les mêmes travaux ont été ensuite évalués indépendamment, selon la méthode de l'impression générale en temps imposé, par huit notateurs.

Pour disposer d'un critère, cinq autres travaux de chaque élève ont été notés par deux examinateurs.

Les notes analytiques et la note moyenne de trois impressions rapides tirées au hasard parmi les huit dont on disposait pour chaque étudiant, ont alors été comparées aux notes de critère :

corrélation entre notes analytiques et critère : .71

corrélation entre notes rapides et critère : .76

En outre, la fidélité de la notation rapide a été calculée en comparant les notes de deux groupes de trois examinateurs, choisis au hasard parmi les huit. On obtient un coefficient très élevé : .87.

2^e expérience ².

Cette seconde expérience, beaucoup plus fine que la précédente, a porté sur un échantillon de 500 élèves de 17 ans, stratifié selon le sexe, le type d'école secondaire et le lieu d'habitation (Londres, centre et faubourgs, grandes villes, petites villes). Tous les types de sujets de compositions, de la simple narration à la dissertation philosophique et à l'amplification poétique, étaient aussi représentés.

De nouveau, la notation multiple selon l'impression rapide s'est montrée plus fidèle et plus valide que la notation analytique.

b) Les échelles de spécimens.

Un petit nombre de compositions, souvent cinq, sont choisies de façon à constituer des modèles représentatifs des différents niveaux de qualité : de la médiocrité à l'excellence.

Le notateur évalue les travaux par comparaison aux cinq textes-étalons de l'échelle.

¹ J. BRITTON, *Experimental Marking of English Composition Written by Fifteen-Year-Olds*, in *Educational Review* (Birmingham), vol. 16, 1, 1963, 17-23.

² J. BRITTON, N. MARTIN et H. ROSEN, *Multiple Marking of English Composition*, Londres, H.M.S.O., 1966.

Après avoir connu un pays ¹, les échelles de plusieurs raisons ² :

1. Il est rare que la comparaison aux textes-étalons ;
2. Les notateurs ont tendent à donner des notes communes qui lésent les autres ;
3. Des échelles différentes par sujet et selon le niveau.

D. Pidgeon nous a signalé récemment en Grande-Bretagne à 5 degrés, s'est soldée par un accord des correcteurs. Le résultat n'est que de notations purement subjectives.

Notons cependant que les échelles de spécimens des chercheurs.

c) La méthode analytique

Deux ou trois notateurs sont assignés à des points pour chacune des compositions à observer. On totalise et on divise.

Cette méthode est la plus souvent critiquée parce qu'elle tend à échapper à toute remarque d'ailleurs que, dans le travail estimé unanimement par les experts, se classe rarement.

Pourtant, la recherche sur l'importance à leur attribuer et la méthode analytique est la seule méthode qui n'est pas le seul correcteur. La méthode que nous l'avons déjà vu, d'ailleurs, est la participation de plusieurs notateurs.

¹ Pour une étude d'ensemble, voir *Rhetoric Found in English Composition*.

² R. BRADDOCK et al., *Research in English Composition*, 1963, p. 12.

³ On trouvera aussi une bonne étude sur *Different Methods of Marking English Composition*, IX, Nov. 1939, 257-269 et X, F.

Après avoir connu une vogue considérable dans certains pays¹, les échelles de spécimens ont été abandonnées pour plusieurs raisons² :

1. Il est rare que la composition à noter ressemble étroitement aux textes-étalons ;
2. Les notateurs ont tendance à n'identifier que les caractéristiques communes qui les intéressent spécialement et négligent les autres ;
3. Des échelles différentes sont nécessaires selon le type de sujet et selon le niveau des élèves.

D. Pidgeon nous a signalé qu'une vaste expérience, menée récemment en Grande-Bretagne, avec une échelle de spécimens à 5 degrés, s'est soldée par un écart moyen de 2 degrés entre correcteurs. Le résultat n'est donc guère meilleur qu'en conditions purement subjectives.

Notons cependant que, depuis 1903 au moins, la méthode des échelles de spécimens retrouve périodiquement la faveur des chercheurs.

c) La méthode analytique.

Deux ou trois notateurs accordent un certain nombre de points pour chacune des différentes qualités qu'il est convenu d'observer. On totalise et on calcule la moyenne entre notateurs.

Cette méthode est la plus lourde des quatre ; elle est souvent critiquée parce qu'elle dissèque un tout qui semble précisément échapper à toute dissection systématique. S. Wiseman remarque d'ailleurs que, dans un ensemble noté analytiquement, le travail estimé unanimement comme le meilleur, par les experts, se classe rarement en tête.

Pourtant, la recherche montre que, si les qualités à observer et l'importance à leur accorder sont définies avec précision, la méthode analytique est la plus sûre, si l'on doit se fier à un seul correcteur. La méthode de l'impression générale donne, nous l'avons déjà vu, d'aussi bons résultats, mais elle exige la participation de plusieurs notateurs³.

¹ Pour une étude d'ensemble, voir : E. HINTON, *Study of the Qualities of Style and Rhetoric Found in English Compositions*, New York, 1940.

² R. BRADDOCK et al., *Research in Written Composition*, Champaign, Ill, N.C.T.E., 1963, p. 12.

³ On trouvera aussi une bonne étude comparative in B. CAST, *The Efficiency of Different Methods of Marking English Composition*, in *British Journal of Educational Psychology*, IX, Nov. 1939, 257-269 et X, Feb. 1940, 49-60, cité par R. BRADDOCK, o.c., p. 13.

Pour être efficace et praticable, la méthode analytique semble devoir répondre à deux exigences au moins :

- 1° Les correcteurs doivent accepter le plan d'analyse. L'idéal est qu'ils participent à son élaboration ;
- 2° Le plan ne doit pas être trop détaillé. Ainsi, la méthode très fine utilisée par E. Burton¹ n'offre d'intérêt que pour la recherche, en raison du temps très long qu'elle exige.

Quelles qualités observer ?

C. Remondino² a montré que les attitudes fondamentales des correcteurs cultivés (professeurs ou non) ne diffèrent pas essentiellement : les mêmes qualités retiennent leur attention. Mais ils leur accordent des poids très variables, dans l'appréciation d'ensemble, et là gît réellement la source des divergences.

Elaborer un tableau très détaillé d'analyse du contenu est un leurre, car, même si les nuances distinguées existent, les notateurs les refondent inconsciemment. L'analyse factorielle met ce processus en évidence. Le travail de Remondino apporte d'intéressants résultats à ce propos.

Remondino a d'abord interrogé longuement vingt professeurs, de branches littéraires (enseignement secondaire) et il a aussi dressé la liste des qualités relevées dans des compositions scolaires. Il aboutit à l'ensemble suivant :

- | |
|--|
| 1. Lisibilité |
| 2. Esthétique |
| 3. Présentation |
| 4. Exactitude de l'orthographe |
| 5. Exactitude morphologique |
| 6. Exactitude syntaxique |
| 7. Structure de l'exposé |
| 8. Richesse d'idées |
| 9. Pertinence des idées |
| 10. Précision d'information |
| 11. Exhaustivité |
| 12. Concision |
| 13. Propriété du langage |
| 14. Style |
| 15. Originalité |
| 16. Maturité |
| 17. Imagination |

Une analyse factorielle des qualités relevées ci-dessus a permis de constater que les qualités relevées se regroupaient en quatre groupes :

1. Présentation graphique
2. Usage de la langue (orthographe, morphologie, syntaxe)
3. Contenu et organisation
4. Aspects personnels (richesse, pertinence, précision, exhaustivité, concision, maturité, imagination)

¹ E. BUXTON, An Experiment to Test the Effects of Writing Frequency..., in *Alberta Journal of Educational Research*, V, Juin 1959, 91-99.

² Etude factorielle sur la notation des compositions scolaires portant sur la langue maternelle, in *Le Travail Humain*, XXII, Janv.-Juin 1959, 27-40.

la méthode analytique
nces au moins :

le plan d'analyse. L'idéal
on ;

lé. Ainsi, la méthode très
e d'intérêt que pour la
s long qu'elle exige.

attitudes fondamentales
(ou non) ne diffèrent pas
retiennent leur attention.
variables, dans l'appré-
t la source des divergen-

d'analyse du contenu est
distinguées existent, les
ent. L'analyse factorielle
ail de Remondino apporte

onguement vingt profes-
ement secondaire) et il a
vées dans des composi-
e suivant :

1. Lisibilité	Qualité d'une écriture qui se prête à une lecture facile, rapide, sans équivoque.
2. Esthétique	Ligne harmonieuse et agréable des lettres et bon goût dans la mise en pages.
3. Présentation	Propreté, soin, bonne présentation de la copie.
4. Exactitude de l'orthographe	Densité des erreurs d'orthographe.
5. Exactitude morphologique	Densité des erreurs morphologiques.
6. Exactitude syntaxique	Densité des erreurs de syntaxe.
7. Structure de l'exposé	Qualité d'un exposé fait avec ordre, dans les proportions voulues, et selon un plan.
8. Richesse d'idées	Quantité d'idées ; ressources utilisées.
9. Pertinence des idées	Qualité par laquelle les idées exposées sont en juste rapport avec le thème traité.
10. Précision d'information	Véracité et exactitude objective des affirmations et des faits exposés.
11. Exhaustivité	Qualité qui consiste à ne rien laisser de côté de tout ce qui devait se dire.
12. Concision	Qualité par laquelle les choses à dire le sont avec le minimum indispensable de termes sans répétitions, redondance ou tortuosité.
13. Propriété du langage	Juste emploi des termes.
14. Style	Facilité, exactitude, maîtrise de la langue au point de vue de la « construction des phrases ».
15. Originalité	Qualité par laquelle, à travers le travail, transparaît et s'affirme quelque chose de la personnalité.
16. Maturité	Capacité de jugement, profondeur critique, acuité des raisonnements exposés.
17. Imagination	Capacité de création, de transfiguration, de « projection » révélée par le travail.

Une analyse factorielle a ensuite montré que ces 17 qualités relevaient de quatre groupes seulement :

1. Présentation graphique (1, 2 et 3) ;
2. Usage de la langue (4, 5, 6, 13, 14) ;
3. Contenu et organisation de l'exposé (7, 8, 9, 10, 11) ;
4. Aspects personnels du fond (15, 16 et 17).

of Writing Frequency..., in *Alberta*
ons scolaires portant sur la langue
27-40.

Un exemple.

Les conclusions de Remondino recourent largement celles des autres chercheurs. Les échelles suivantes, construites par l'*Educational Testing Service* (E.T.S.) de Princeton¹, et légèrement remaniées par E. Page², témoignent de la similitude des vues.

Nous traduisons le texte complet parce qu'il apporte un des meilleurs exemples que l'on puisse donner actuellement pour indiquer la voie à suivre.

CRITERES POUR NOTER LES COMPOSITIONS.

I. Définition des traits à évaluer.

- A. Idées ou contenu : la quantité et la qualité du matériel utilisé pour traiter du sujet.
- B. Organisation : la relation entre les parties de la composition et l'ensemble.
- C. Style : utilisation du langage au-delà de la simple correction grammaticale.
- D. Mécanique : orthographe, grammaire, ponctuation.
- E. Créativité.

II. Guide pour l'évaluation de ces cinq traits.

A. Idées ou contenu.

Niveau élevé.

L'étudiant omet beaucoup d'aspects importants du sujet. le plan de travail. Il comprend bien le sujet et utilise des définitions claires. Il sait considérer le sujet dans une perspective plus large que celle des autres élèves de la classe. Autrement dit, il témoigne d'une expérience plus riche.

Niveau moyen.

Les idées sont appropriées, mais conventionnelles et peu nombreuses. Certains aspects du sujet sont négligés. L'élève ne semble pas avoir un esprit richement meublé.

¹ E.T.S., *Definitions of Ratings on the E.T.S. Compositions Scale*, cité par E. PAGE, o.c., pp. 76-77.

² O.c., pp. 78-80.

Niveau bas.

L'étudiant omet beaucoup. Il semble ne pas disposer de idées au sujet et, par conséquent, idées simples.

B. Organisation.

Niveau élevé.

L'étudiant suit un plan contre, il avance des raisons. S'il décrit quelque chose, sommet à la base, par ordre de complexité, etc.). Si l'étudiant utilise un plan cohérent, il utilise un plan cohérent. L'étudiant sent bien l'importance des répétitions. Il témoigne de différentes parties de son travail.

Niveau moyen.

L'étudiant ne s'en tient pas à des idées sans rapport avec le sujet. Les choses peu importantes sont associées librement (qu'est-ce que pense à Hawaï ?) plutôt qu'avec un ordre de présentation est.

Niveau bas.

L'étudiant ne semble pas avoir de plan avant de commencer à écrire. Il prend une direction, puis change de direction, jusqu'à ce que le lecteur ne soit pas clairement sûr de l'ordre de présentation est.

C. Style.

(Plusieurs aspects du style sont évalués : individualité, vivacité, nous intéressons ici à trois aspects : répétition, variation et éventail de mots.)

Niveau élevé.

L'étudiant utilise un langage varié dans la composition du travail. Il utilise un langage habituel. Les mots sont pertinents.

Niveau bas.

L'étudiant omet beaucoup d'aspects importants du sujet. Il semble ne pas disposer d'une réserve de connaissances relatives au sujet et, par conséquent, répète sans cesse quelques idées simples.

B. Organisation.

Niveau élevé.

L'étudiant suit un plan défini. S'il présente le pour et le contre, il avance des raisons pertinentes, dans un ordre efficace. S'il décrit quelque chose, il le fait de façon ordonnée (du sommet à la base, par ordre d'importance, par ordre de complexité, etc.). Si l'étudiant explique un concept ou un processus, il utilise un plan cohérent d'analyse, de définition ou d'illustration. L'étudiant sent bien ce qui se rapporte à son plan et évite des répétitions. Il témoigne du sens de la mesure en traitant les différentes parties de son travail.

Niveau moyen.

L'étudiant ne s'en tient pas à son plan ou introduit des idées sans rapport avec le sujet. Il consacre trop de temps à des choses peu importantes ou se répète. Il traite le sujet par association libre (qu'est-ce qui me vient à l'esprit quand je pense à Hawaï ?) plutôt qu'en poursuivant un but bien défini.

Niveau bas.

L'étudiant ne semble pas s'être demandé ce qu'il allait faire avant de commencer à écrire. Il ne suit pas de plan. Le travail prend une direction, puis en change, en change encore et encore, jusqu'à ce que le lecteur soit perdu. Les points principaux ne sont pas clairement séparés les uns des autres, et leur ordre de présentation est laissé au hasard.

C. Style.

(Plusieurs aspects du style peuvent intervenir dans l'évaluation : individualité, vivacité, élégance, etc. Toutefois, nous nous intéressons ici à trois aspects stylistiques seulement : clarté, variation et éventail des ressources linguistiques.)

Niveau élevé.

L'étudiant utilise un langage qui rend aisé la compréhension du travail. Il utilise des mots adéquats, dans leur sens habituel. Les mots sont présentés dans un ordre normal. Les

transitions sont bien ménagées. L'élève évite les ambiguïtés et ne trompe pas l'attente du lecteur. En même temps, l'étudiant évite la répétition monotone de mots, de compléments ou de structures de phrases. Finalement, il témoigne de la connaissance d'un large éventail de ressources linguistiques. Son vocabulaire est bon. Il utilise des structures parallèles ou fait un usage subtil de la subordination.

Niveau moyen.

L'étudiant égare parfois le lecteur en utilisant un mot inapproprié ou une tournure bizarre ; ou bien en utilisant une métaphore peu claire, ou en déplaçant de façon inopportune un complément ou une subordonnée, ou encore en pratiquant des transitions abruptes. La répétition de mots, de tournures et de structures de phrases devient monotone. Les ressources linguistiques sont limitées. L'élève utilise volontiers des clichés et des tournures éculées.

Niveau bas.

Les mots sont utilisés de façon vague. Tournures ambiguës, constructions boiteuses, vocabulaire et structures de phrases enfantins.

D. Mécanique.

Niveau élevé.

La structure des phrases est habituellement correcte, même lorsqu'il s'agit de modèles variés et compliqués. Les règles de l'orthographe sont respectées, même les mots difficiles sont généralement écrits sans faute. Pas de violation grave des règles de ponctuation, de majuscules, d'abréviations, d'écriture des nombres.

Niveau moyen.

Défauts de syntaxe occasionnels. Les mots difficiles sont parfois mal orthographiés. Quelques violations des règles de ponctuation, etc.

Niveau bas.

Très grand nombre de fautes.

E. Créativité.

Niveau élevé.

L'étudiant surprend par la façon de considérer le problème. Il traite son traitement du sujet. Il trouve des manières sages d'utiliser le langage.

Niveau moyen.

L'étudiant pense à ce qu'il veut dire et traite les choses comme à propos. Il utilise des expressions et des structures variées.

Niveau bas.

L'étudiant utilise des clichés. Il traite le sujet de façon superficielle. Il ne réussit pas réellement en comprenant la tâche.

d) La méthode des comptes.

Nous la signalons ici, bien qu'elle ne soit pas dans la catégorie des évaluations. Elle est présentée dans le tableau des méthodes.

Elle consiste à définir (à l'aide d'une étude diagnostique), à compter (à l'aide d'un tableau) et à exprimer le nombre obtenu.

Au lieu d'adopter un procédé traditionnel, on fait faire l'inventaire (relativement simple) des méthodes utilisées dans la composition. Une méthode est fournie par les *Éléments de l'analyse littéraire* de A. Purves¹.

Une dernière méthode présente des aspects qui se révèlent particulièrement intéressants, normalement attribuée par défaut à l'élève, programmé en conséquence, la notation ; c'est ce qu'a fait Purves.

Les comptages de fréquence sont une source d'espoir de notation objective, mais il faut toutefois devoir rester du domaine de la pratique scolaire, à moins que l'élève ne soit capable de faire un inventaire de ses méthodes.

¹ A. PURVES et V. RIPPÈRE, *Éléments de l'analyse littéraire* (traduit de l'anglais par J. Dubois), in *Studia Neoscholastica*, 229-334.

évitent les ambiguïtés et
au même temps, l'étudiant
utilise des compléments ou de
des témoignages de la connais-
sance linguistiques. Son voca-
bulaire est parallèle ou fait un

en utilisant un mot inap-
proprié ou bien en utilisant une
construction de façon inopportune
ou encore en pratiquant
l'usage de mots, de tournures
monotone. Les ressources
sont volontiers des clichés

que. Tournures ambiguës,
et structures de phrases

généralement correcte, même
compliqués. Les règles de
des mots difficiles sont
de violation grave des
d'abréviations, d'écriture

Les mots difficiles sont
violations des règles de

E. Créativité.

Niveau élevé.

L'étudiant surprend par des façons neuves et efficaces de considérer le problème. Il introduit des idées nouvelles dans son traitement du sujet. Il trouve des façons fraîches et intéressantes d'utiliser le langage pour faire ressortir ses idées.

Niveau moyen.

L'étudiant pense à ce que l'on s'attendait qu'il penserait. Il traite les choses comme à peu près tout le monde. Il utilise des expressions et des structures de phrases ordinaires.

Niveau bas.

L'étudiant utilise des clichés de pensée et d'expression. Il traite le sujet de façon superficielle. Il répète des formules sans réellement en comprendre la signification.

d) La méthode des comptages de fréquences.

Nous la signalons ici, bien qu'elle n'ait pas de place dans la catégorie des évaluations subjectives, afin de compléter le tableau des méthodes.

Elle consiste à définir des types de fautes (après une étude diagnostique), à compter celles-ci dans la composition et à exprimer le nombre obtenu pour 100 ou pour 1 000 mots.

Au lieu d'adopter un point de vue négatif, on peut aussi faire l'inventaire (relativement exhaustif) des divers processus utilisés dans la composition. Le plus bel exemple de pareille méthode est fourni par les *Eléments pour servir à l'étude d'analyses littéraires* de A. Purves¹.

Une dernière méthode positive consiste à isoler certains aspects qui se révèlent prédictifs, à un seuil donné, de la note normalement attribuée par des correcteurs qualifiés. Un ordinateur, programmé en conséquence, peut alors se charger de la notation ; c'est ce qu'a fait E. Page.

Les comptages de fréquences offrent probablement le seul espoir de notation objective des compositions. Ils semblent toutefois devoir rester du domaine de la recherche étrangère à la pratique scolaire, à moins que l'ordinateur ne devienne d'un

¹ A. PURVES et V. RIPPERE, *Eléments pour servir à l'étude d'analyses littéraires* (traduit de l'anglais par J. Dubois), in *Scientia Paedagogica Experimentalis*, VI, 2, 1969, 229-334.

accès tellement facile que l'étude des fréquences se fasse de façon économique et constitue un des aspects de l'évaluation. Ceci n'est nullement exclu.

2) Plusieurs sujets au choix ?

J. Britton remarque avec raison que la rédaction d'examen sur un sujet unique implique souvent, chez le professeur, l'hypothèse qu'un sujet comme « Nuages » révélera aussi bien l'aptitude d'un étudiant que « Exposez vos vues sur la peine capitale ».

On a montré expérimentalement qu'il n'en est rien, nous allons y revenir. Toutefois, il est douteux que la justice soit beaucoup mieux respectée si l'on offre trois sujets au choix quand les points d'examens sont attribués pour une rédaction unique. Il faudrait d'abord prouver qu'avec trois, voire cinq sujets, on recouvre tout le spectre des stimuli nécessaires pour donner à chacun une chance égale ; et même si cela était établi, il importerait encore que le professeur sélectionne toujours les sujets à bon escient.

L'expérience, disions-nous, nous montre que le rendement de l'élève varie selon le thème proposé. Finlayson¹ a établi, pour la fin de l'école primaire, que si, à une semaine d'intervalle, les élèves sont invités à faire une rédaction à partir d'un même choix de quatre sujets, ceux qui changent de sujets (plus de la moitié dans l'expérience) obtiennent des notes significativement différentes.

Wiseman et Wrigley² arrivent à la même conclusion (choix parmi le même ensemble de cinq sujets à quatre mois d'intervalle), tout en prouvant que la différence de note n'est pas imputable au manque de fidélité du professeur. Vernon et Millican³ confirment la conclusion au niveau universitaire.

Que faire ? Si l'on veut vérifier la capacité de s'exprimer par écrit dans une perspective étroitement définie (par exemple : capacité de faire un rapport scientifique ou d'écrire une lettre commerciale), la sensibilité particulière des candidats n'a guère à être considérée et un sujet suffit.

Par contre, s'il s'agit ou bien de déceler une aptitude à l'expression, où qu'elle soit, ou d'amener les élèves à discuter de problèmes où la tournure d'esprit et le niveau d'information

1 D.S. FINLAYSON, The Reliability of the Marking of Essays, in *British Journal of Educational Psychology*, XXI, 2, 1951, cité par Britton (1966).

2 S. WISEMAN et J. WRIGLEY, Essay-Reliability : the Effect of Choice of Essay-Title, in *Educational and Psychological Measurement*, 18, 1, 1958.

3 P. VERNON et G. MILLICAN, A Further Study of the Reliability of English Essays, in *British Journal of Statistical Psychology*, VII, 2, 1954.

peuvent jouer un grand rôle dans le jugement à partir d'une seule rédaction. On se réfère aussi au travail qui a été pu choisir parmi plusieurs sujets. L'élève réagit à plusieurs dizaines de sujets par mois et l'on peut supposer que la chance. Laisser inventer un sujet, l'on pouvait se prémunir contre cela. C'est bien difficile...

3) Conclusion.

Le sort à réserver à la composition est un but poursuivi. Que veut-on ? Dans ce cas, il vaut mieux donner plusieurs travaux ou l'attaquer dans une seule rédaction.

Connaître la capacité de l'élève à la prose conforme au bon usage de la langue avec réponses à choix multiples.

Faire acquérir la capacité de l'élève. Dans ce cas, des thèmes très limités sont proposés sur l'objet du travail. « Décrivez un vélo de course ».

Déceler des talents d'expression à l'expression de l'élève sur son même sujet.

Mais que l'on ne s'y trompe pas : les capacités différentes. Dans ce cas, on décrit plus haut, l'auteur a

1° Que la corrélation entre la note et un résumé de texte va

2° Que la corrélation de la note et la compréhension de texte est un

Les écarts sont donc

Nous l'avons déjà dit, la composition traditionnelle n'a pas existé. L'écriture avait été supprimée à cause de la nouveauté. On l'a d'ailleurs réintroduite où la Commission des Examinations Supérieures a renoncé à l'examen en Angleterre, aux deux époques. le 16 + Examination.

fréquences se fasse de
aspects de l'évaluation.

ue la rédaction d'examen
chez le professeur, l'hypo-
révélera aussi bien l'apti-
vues sur la peine capi-

qu'il n'en est rien, nous
uteux que la justice soit
fre trois sujets au choix
bués pour une rédaction
qu'avec trois, voire cinq
stimuli nécessaires pour
même si cela était établi,
r sélectionne toujours les

montre que le rendement
osé. Finlayson¹ a établi,
i, à une semaine d'inter-
ne rédaction à partir d'un
i changent de sujets (plus
ment des notes significati-

a même conclusion (choix
ets à quatre mois d'inter-
rence de note n'est pas
professeur. Vernon et Malli-
eau universitaire.

la capacité de s'exprimer
ement définie (par exem-
scientifique ou d'écrire une
iculaire des candidats n'a
ffit.

e déceler une aptitude à
ener les élèves à discuter
et le niveau d'information

king of Essays, in *British Journal of*
(1966).

ty: the Effect of Choice of Essay-
18, 1, 1958.

of the Reliability of English Essays,

peuvent jouer un grand rôle, il n'est guère possible de porter un jugement à partir d'une seule composition. L'idéal est alors de se référer aussi au travail de l'année ; si, chaque fois, l'élève a pu choisir parmi plusieurs sujets (cinq, par exemple), il aura réagi à plusieurs dizaines de thèmes différents en quelques mois et l'on peut supposer que chacun aura vraiment eu sa chance. Laisser inventer le sujet à traiter serait tentant, si l'on pouvait se prémunir contre la préparation frauduleuse. C'est bien difficile...

3) Conclusion.

Le sort à réserver à la composition française dépendra du but poursuivi. Que veut-on ? Corriger un défaut particulier ? Dans ce cas, il vaut mieux n'examiner que celui-là dans les travaux ou l'attaquer dans des exercices spéciaux.

Connaître la capacité des élèves à se corriger, à identifier la prose conforme au bon usage ? Dans ce cas, un test objectif, avec réponses à choix multiple, peut faire l'affaire.

Faire acquérir la capacité de s'exprimer clairement ? Dans ce cas, des thèmes très limités permettent de concentrer l'attention sur l'objet du travail. Pourquoi bannir les sujets tels que : « Décrivez un vélo de course » ?

Déceler des talents d'écrivains ? Alors, il faut laisser libre cours à l'expression de l'élève, l'invitant souvent à choisir lui-même son sujet.

Mais que l'on ne s'y trompe pas, on testera ainsi des capacités différentes. Dans la seconde expérience de J. Britton, décrite plus haut, l'auteur a montré :

- 1° Que la corrélation entre les résultats à la rédaction créative et un résumé de texte varie de .30 à .40 ;
- 2° Que la corrélation de la rédaction avec un test de compréhension de texte est un peu plus élevée : de .35 à .45.

Les écarts sont donc considérables.

Nous l'avons déjà dit, nous ne pensons pas qu'il faille bannir la composition traditionnelle. Dans certains pays où elle avait été supprimée à cause de la difficulté d'évaluation rigoureuse, on l'a d'ailleurs réintroduite, notamment aux Etats-Unis où la Commission des Examens d'Admission dans l'Enseignement Supérieur a renoncé au « Test objectif de composition » et en Angleterre, aux deux épreuves-clés de sélection : le **11 +** et le **16 + Examination**.

à jouer un rôle important
p répandue de noter doit
position de R. Braddock :

nombre d'étudiants pour
échec, il importe de per-
un second travail (...). Si
s travaux au moins sont
meilleurs dont on fera la
être notés par au moins
de notation, bien compris
nés. »¹

jective.

intervenir l'avis personnel
simple est fourni par une
s à raison d'un point par
d'un test composé d'*items*
comptage du nombre de

appelle notation objective
s à des échantillons com-
définis et contrôlés pour
es juges ou notateurs »².

déjà, l'objectivité de la
té de l'examen dans son
este, en dernière analyse,
isément un examen dont
éellement la capacité que
que puisse être la correc-
cas, aucune validité. Les
fourmillent.

la position relative.

soulève un problème de
s normes, c'est-à-dire des
de situer les performances

Selon le but poursuivi, les normes sont établies pour un groupe restreint ou pour de plus vastes ensembles pouvant comprendre jusqu'à la population d'un pays. Pratiquement, leur validité se limite toujours à un ensemble déterminé et pour les utiliser, il importe de savoir lequel.

Longtemps, le pourcentage a été l'alpha et l'oméga de la norme scolaire. Pour étalonner certains tests, on se contentait parfois aussi d'exprimer le nombre de réponses correctes en pour cent. On sait toutes les faiblesses de pareil système.

Aujourd'hui, la différence entre des individus et non la situation sur une échelle du savoir considérée comme absolue¹ joue un rôle déterminant. La moyenne ou le médian² constituent les points de repère à partir desquels on construit des échelles « universelles », qui, moyennant certaines conditions sur lesquelles nous reviendrons, permettent d'exprimer des scores bruts, obtenus à différentes épreuves, en unités comparables.

Les systèmes d'étalonnage les plus connus et les plus accessibles sont les suivants :

1. Le centilage ;
2. Les notes standards ou notes Z ;
3. L'échelle normalisée à 5 classes ;
4. L'échelle normalisée à 9 classes (Stanines).

Nous les étudions successivement.

A. Le centilage.

1) Définition.

Le centile, ou rang occupé sur une échelle à cent degrés, jouit encore d'une certaine faveur à cause de sa ressemblance extérieure avec le pourcentage, dont il diffère pourtant de façon fondamentale. Le pourcentage indique quelle proportion du total des points attribués à un examen l'élève a obtenu. Le centile indique combien d'élèves se classeraient après un élève donné si la classe comptait cent élèves. Donc un score équivalant au 90^e centile est supérieur à 90% de la population considérée.

Le médian correspond au 50^e centile.

¹ Voir nature et faiblesse des échelles d'évaluation. Calculer le pourcentage équivaut à considérer une échelle à cent degrés.

² Rappelons que le médian est le point qui divise une série de notes ordonnées en deux parties égales : 1 — 2 — 3 — 4 — 5 ou 1 — 2 — 3 . 4 — 5 — 6.

2) Calcul.

En principe, on ne calcule pas les centiles à partir de moins de cent notes. L'exemple suivant utilise les cent notes qui nous ont déjà servi lors de l'étude de la courbe de Gauss.

Démarche générale :

- Dresser le tableau des classes (en commençant par la classe inférieure) ;
- Indiquer les fréquences (f) ;
- Calculer les fréquences cumulées (f_c) : chaque fréquence est additionnée à celle qui la précède ;
- Appliquer la formule :

$$C_x = l + \frac{(N.C)/100 - f_c}{f_s} \times i.$$

C_x = valeur de la note (x) en centile.

$(N.C)/100$ = le rang de la note correspondant à un centile quelconque.

Exemple : Le 50^e centile correspond à la note médiane, soit $N/2$. Pour les autres, il s'agit d'une simple règle de trois. Le 75^e centile sera donc : $(100 \times 75) : 100 = 75$ (les chiffres sont les mêmes parce que, dans notre exemple, il y a précisément 100 notes).

l = la moyenne entre la limite supérieure de la classe où se trouve le centile cherché et la limite inférieure de la classe qui la précède.

f_c = fréquence cumulée de la classe qui précède celle où se trouve le centile cherché.

f_s = la fréquence simple de l'intervalle où se trouve le centile cherché.

i = intervalle de classe.

Exemple :

Classes	
24-28	
29-33	
34-38	
39-43	
44-48	
49-53	
54-58	
59-63	
64-68	
69-73	
74-78	

Calcul du 50 ^e centile	
$N.C$	100×50
100	100
l	$\frac{49+48}{2} = 48,5$
f_c	44
f_s	22
C_x	$48,5 + \frac{50-44}{22} \times 5 = 49$

N.B. Décilage : le 10^e centile =
Ecart semi-interquartile :

3) Critique.

L'échelle en centiles vaste échelle d'évaluation rang occupé, mais non s distribution est normale, l'on se rapproche du méd normalité parfaite).

Le tableau suivant m

	- 2σ	- 1σ	1σ	2σ
Centiles	2 ^e	16 ^e		

Exemple :

Classes	Fréquences (f)	Fréquences cumulées (f _c)
24-28	1	1
29-33	2	3
34-38	8	11
39-43	10	21
44-48	23	44
49-53	22	66
54-58	18	84
59-63	9	93
64-68	5	98
69-73	1	99
74-78	1	100

	Calcul du 50 ^e centile	Calcul du 16 ^e centile	Calcul du 84 ^e centile
N.C	$\frac{100 \times 50}{100}$	$\frac{100 \times 16}{100}$	$\frac{100 \times 84}{100}$
100	100	100	100
l	$\frac{49+48}{2} = 48,5$	$\frac{39+38}{3} = 38,5$	$\frac{54+53}{2} = 53,5$
f _c	44	11	66
f _s	22	10	18
C _x	$48,5 + \frac{50-44}{22} \times 5 = 49,9$	$38,5 + \frac{16-11}{10} \times 5 = 41$	$53,5 + \frac{84-66}{18} \times 5 = 58,5$

N.B. *Décilage* : le 10^e centile = le 1^{er} décile, etc.
Ecart semi-interquartile : (75^e C — 25^e C)/2.

3) Critique.

L'échelle en centiles n'est, nous l'avons déjà dit, qu'une vaste échelle d'évaluation à 100 degrés. Elle informe sur le rang occupé, mais non sur la distance entre rangs. Or, si la distribution est normale, cette distance rétrécit à mesure que l'on se rapproche du médian (égal à la moyenne dans le cas de normalité parfaite).

Le tableau suivant montre clairement ce phénomène.

	- 2 σ	- 1 σ	0	+ 1 σ	+ 2 σ
Centiles	2 ^e	16 ^e	50 ^e	84 ^e	98 ^e

On constate que 68 centiles sont agglomérés autour de la moyenne, ce qui donne un classement trop grossier pour la majorité du groupe. On risque, par exemple, de considérer qu'une différence importante sépare le 75^e du 25^e centile. L'ancienne norme du pourcentage subsiste tenacement dans les esprits et incline à penser que le 75^e C vaut trois fois le 25^e. Le tableau ci-dessus nous montre qu'il s'agit, en réalité, de deux résultats relativement proches et moyens.

En fait, si les normes en centiles restent utilisées, notamment pour la sélection dans certaines universités américaines, c'est parce que celles-ci ne recrutent que dans la bande des dix ou vingt centiles supérieurs, zone où la discrimination est satisfaisante.

B. Les notes standard ou notes Z.

Un élève a obtenu les notes brutes suivantes¹. Calcul : 22 sur 25 ; Lecture : 72 sur 100 ; Sciences : 26 sur 50.

L'examen de ces notes n'apprend pas grand-chose. Si le professeur est sévère en lecture et beaucoup moins en calcul, le 72/100 en lecture est peut-être plus méritoire que le 22/25 en calcul. De plus, des notes brutes ne permettent pas la comparaison avec d'autres élèves de même niveau pédagogique, par exemple.

Pour rendre la comparaison possible, on exprime ces notes en fonction des écarts types, ce qui permet de les situer sur une même courbe (notes étalonnées ou scores standard).

Supposons que l'on obtienne les résultats suivants :

	Moyenne	Ecart type
Calcul	15	2,5
Lecture	50	10
Sciences	29	5

Calcul des scores standard : $z = (x - M) / \sigma$.

Exemple :

$$(22 - 15) : 2,5 = 2,8 \sigma.$$

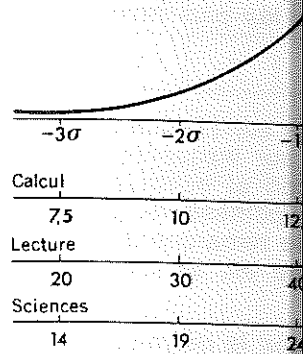
Connaissant z, on peut retrouver la note brute :

$$x = M + z \sigma$$

$$x = 15 + (2,8 \times 2,5) = 22.$$

¹ Adapté d'après R. THOMAS, o.c.

Représentation gra



Application.

Les étudiants I et II
Comparer les résultats

TEST	Moyenne M	Ecart type σ
Anglais	155,7	26,4
Lecture	33,7	8,2
Information	54,5	9,2
Aptitudes scolaires	87,1	25,8
Attention	24,8	6,8
Totaux Moyennes		

Si l'on s'en tenait à l'étudiant I serait considéré le calcul des pourcentages

¹ Chiffres empruntés à J.-P.

agglomérés autour de la
 nt trop grossier pour la
 exemple, de considérer
 75^e du 25^e centile. L'an-
 te tenacement dans les
 vaut trois fois le 25^e. Le
 l s'agit, en réalité, de
 moyens.

restent utilisées, notam-
 universités américaines,
 que dans la bande des
 où la discrimination est

utes suivantes¹. Calcul :
 nces : 26 sur 50.

l pas grand-chose. Si le
 beaucoup moins en calcul,
 méritoire que le 22/25 en
 permettent pas la compa-
 niveau pédagogique, par

ble, on exprime ces notes
 permet de les situer sur
 u scores standard).

résultats suivants :

Ecart type
2,5
10
5

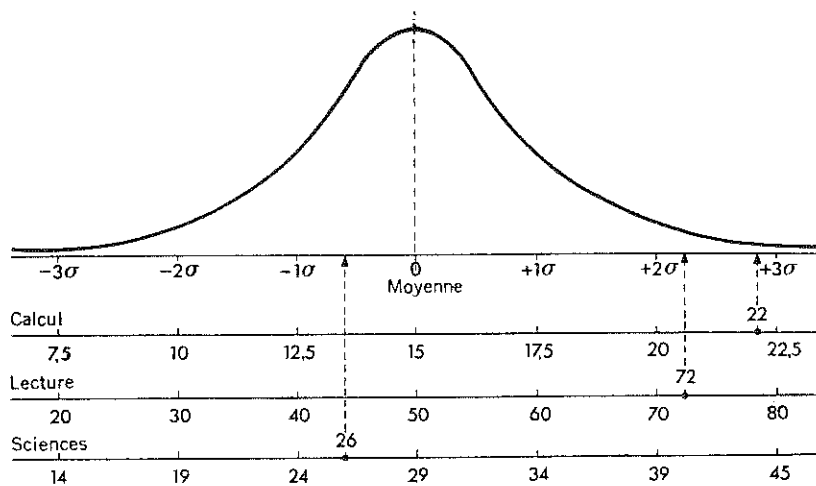
$(x - M) / \sigma$.

2,8 σ .

la note brute :

5) = 22.

Représentation graphique :



Application.

Les étudiants I et II ont passé une série de tests.

Comparer les résultats¹.

TEST	Moyen- ne M	Ecart type σ	X Scores bruts		x Déviations (M-X)		Scores standards z	
			I	II	I	II	I	II
			Anglais	155,7	26,4	195	162	+ 39,3
Lecture	33,7	8,2	20	54	- 13,7	+ 20,3	- 1,67	+ 2,48
Information	54,5	9,2	39	72	- 15,5	+ 17,5	- 1,67	+ 1,88
Aptitudes scolaires	87,1	25,8	139	84	+ 51,9	- 3,1	+ 2,01	- 0,12
Attention	24,8	6,8	41	25	+ 16,2	+ 0,2	+ 2,38	+ 0,03
Totaux Moyennes			434	397			+ 2,54 + 0,51	+ 4,51 + 0,90

Si l'on s'en tenait aux simples totaux (scores bruts I et II), l'étudiant I serait considéré comme supérieur à l'étudiant II. Le calcul des pourcentages confirmerait cette impression.

¹ Chiffres empruntés à J.-P. GUILFORD, *Fundamental Statistics*, o.c., p. 513.

Or si, par le calcul des scores Z, on ramène tous les résultats à la même moyenne et à la même unité de déviation par rapport à celle-ci, on constate :

- 1° Qu'en moyenne, l'étudiant II (0,90) obtient un résultat presque deux fois aussi favorable que l'étudiant I (0,51) ;
- 2° Que l'étudiant II obtient les résultats les plus homogènes.

Critique.

La transformation en scores Z repose sur l'hypothèse que, pour le groupe considéré, les aptitudes sur lesquelles portent les tests se distribuent de la même façon et conduisent aux mêmes moyennes et dispersions. Cette hypothèse est pratiquement impossible à vérifier, mais, comme le remarque J.-P. Guilford, on peut être assuré que ce système est de toute façon meilleur que le calcul du pourcentage.

Par ailleurs, le centrage de l'échelle sur la moyenne donne de ce point une image discriminative fautive. En raison des imperfections inévitables de la notation, le pur hasard fera basculer un nombre plus ou moins élevé de sujets d'un côté ou de l'autre.

Une transformation.

Les scores Z prennent zéro pour moyenne et un écart type égal à un comme unité.

Certains considèrent ce système peu parlant et préfèrent exprimer les résultats à la façon des quotients intellectuels où 100 est traditionnellement pris comme moyenne et 145 considéré comme un quotient de « génie ».

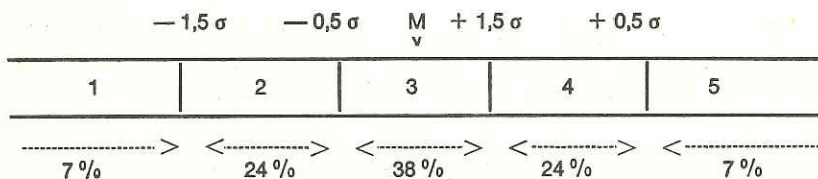
La transformation est aisée. Cent est donc pris comme moyenne et chaque score Z multiplié par 15.

C. L'échelle à 5 classes (N entre 50 et 100).

Les 2^e, 3^e et 4^e classes comportent un écart type.

La 3^e classe est centrée sur la moyenne.

Les 1^{re} et 5^e classes sont d'étendue illimitée.



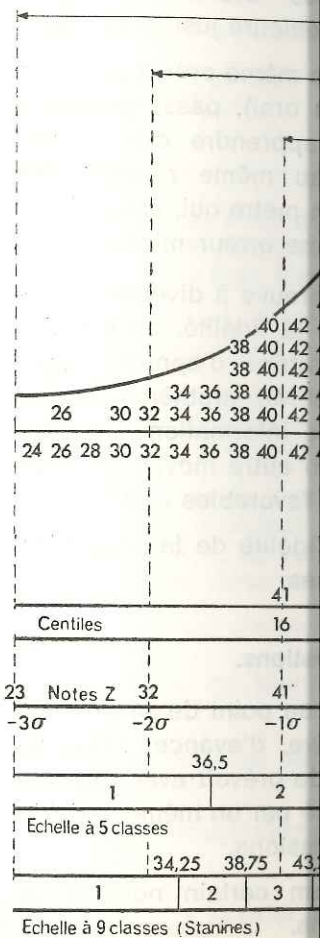
D. L'échelle à 9 classe

Les 1^{re} et 9^e classes

Les autres classes étant centrée sur la moy

$-1,75\sigma$	$-1,25\sigma$	$-0,75\sigma$	
1	2	3	
>	<	<	<
4%	7%	12%	17%

Le graphique ci-dessous illustre des quatre types de norm



on ramène tous les résultats à une unité de déviation par

obtient un résultat prescel de l'étudiant I (0,51) ;

ats les plus homogènes.

pose sur l'hypothèse que, les sur lesquelles portent façon et conduisent aux hypothèse est pratique- me le remarque J.-P. Guil- tème est de toute façon

lle sur la moyenne donne e fausse. En raison des tion, le pur hasard fera levé de sujets d'un côté

moyenne et un écart type

peu parlant et préfèrent quotients intellectuels où e moyenne et 145 consi-

nt est donc pris comme par 15.

et 100).

nt un écart type.

enne.

illimitée.

5σ + $0,5\sigma$

4	5
---	---

<----->	<----->
24%	7%

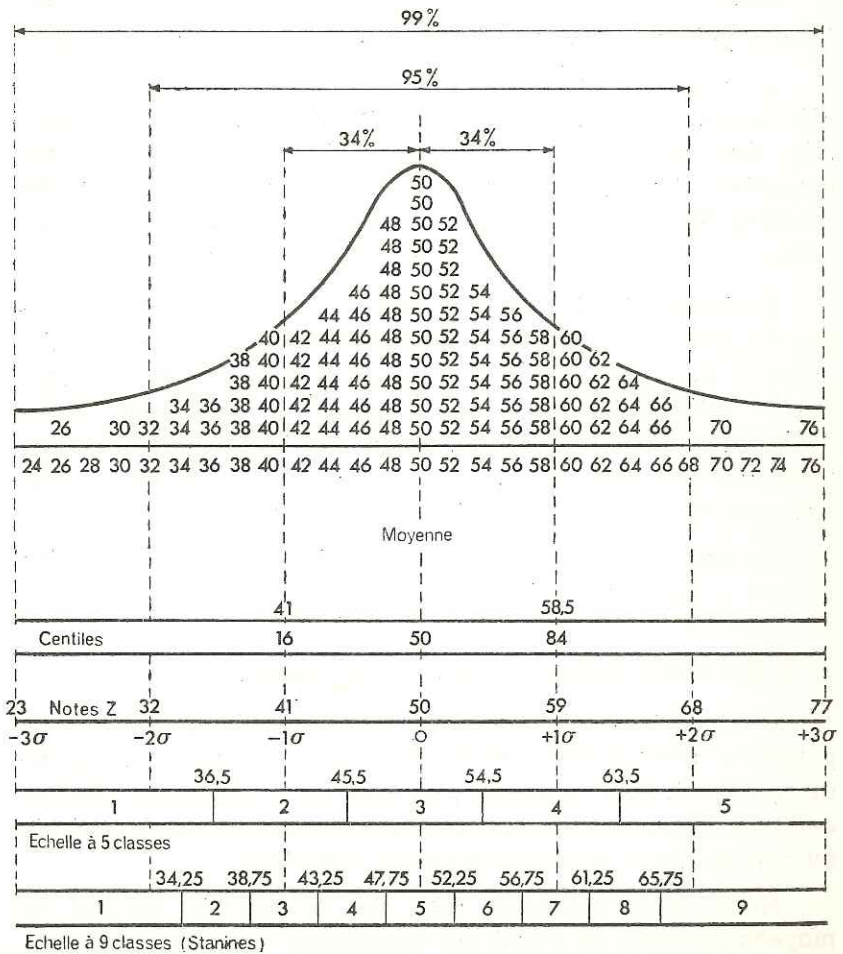
D. L'échelle à 9 classes (Stanines) (N>100).

Les 1^{re} et 9^e classes sont d'étendue illimitée.

Les autres classes comportent un demi-écart type, la 5^e étant centrée sur la moyenne.

M								
-1,75 σ	-1,25 σ	-0,75 σ	-0,25 σ	v + 0,25 σ	+ 0,75 σ	+ 1,25 σ	+ 1,75 σ	
1	2	3	4	5	6	7	8	9
<----->	<----->	<----->	<----->	<----->	<----->	<----->	<----->	<----->
4%	7%	12%	17%	20%	17%	12%	7%	4%

Le graphique ci-dessous permet une comparaison aisée des quatre types de normes.



CHAPITRE 4.

CONTROLE DE LA FIDELITE DE L'EXAMEN.

Beaucoup de mesures que nous avons envisagées jusqu'à présent tendent à assurer une meilleure justice scolaire.

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. On souhaite donc qu'il soit aussi *fidèle* qu'un mètre qui, bien utilisé, mesure toujours la même longueur, à une erreur minime près.

Evidemment, faire subir la même épreuve à diverses reprises ne permet pas un bon contrôle de la fidélité. Le premier passage provoque une réflexion génératrice d'apprentissage. Par la suite, l'élève continue à penser au problème, vérifie l'exactitude de ses réponses, prend des informations complémentaires, etc. Bref, il faut recourir à un autre moyen ou, plus exactement, à un ensemble de mesures favorables à la fidélité.

En supposant que soit acquise la fidélité de la correction il reste à garantir la fidélité des réponses.

1. Eviter toute ambiguïté dans les questions.

Une question qui manque de clarté au point de se prêter à plusieurs interprétations différentes prive, d'avance, l'examen de sa fidélité. Rien ne permet, en effet, de prévoir avec quelque certitude lequel des sens aurait été donné par un même individu ou par un même groupe, à diverses occasions.

Nous avons envisagé, page 74, un certain nombre de moyens d'assurer la clarté des questions.

2. Des questions en no

Si l'on ne pose qu'une question, on laisse de vastes zones d'ombre que l'élève n'ait pas comprises. Les parties du cours en référence accusée du professeur. Si on pose plusieurs questions, on enregistre

Le seul moyen d'évaluer la maîtrise de toute la matière. En d'autre termes, c'est d'appréhender l'ensemble.

En principe, plus on pose de questions, plus on augmente.

3. Un contrôle mathéma

Pour des examens il faut des garanties encore.

a) La méthode pairs-imp

Si le nombre de questions est pair, la suivante simule une situation équivalente à la précédente, mais effectuée en même temps.

a) Les questions étant réparties en groupes : paires et impaires.

b) On dresse deux tableaux de questions et on calcule leur corrélation. On vérifie ainsi que la relation est la même.

c) Les deux examens sont effectués à la fois. On compare les résultats. Si le résultat est inférieur à celui du paragraphe 2, ils sont considérés comme une sous-estime donc la

Une formule simple

saire : l'indice de fidélité

Plus le résultat sera élevé.

On exige généralement

1 La méthode du calcul de la statistique élémentaires.

2. Des questions en nombre suffisant.

Si l'on ne pose qu'un très petit nombre de questions, on laisse de vastes zones de matières inexplorées. Or, il se peut que l'élève n'ait pas compris toutes les matières ou ait négligé des parties du cours en spéculant sur la chance ou sur la préférence accusée du professeur. Selon l'endroit où tomberont les questions, on enregistrera un brillant résultat ou un échec.

Le seul moyen d'éviter pareille aventure est de couvrir toute la matière. En d'autres termes, l'examen doit *bien échantillonner* l'ensemble.

En principe, plus on ajoute de questions, plus la fidélité augmente.

3. Un contrôle mathématique.

Pour des examens importants, on souhaitera acquérir plus de garanties encore.

a) La méthode pairs-impairs.

Si le nombre de questions est assez élevé, la démarche suivante simule une situation où un même élève subirait, pratiquement en même temps, deux fois le même examen.

- a) Les questions étant numérotées, on les divise en deux groupes : paires et impaires.
- b) On dresse deux tableaux parallèles des résultats et l'on calcule leur corrélation (r) ; autrement dit, on évalue numériquement la relation qui existe entre les deux groupes¹.
- c) Les deux examens artificiellement créés sont de moitié plus courts que l'original. En vertu de ce que nous avons vu au paragraphe 2, ils sont donc moins fidèles et la corrélation sous-estime donc la situation réelle.

Une formule simple permet d'opérer la correction nécessaire : l'indice de fidélité est finalement obtenu par $\frac{2r}{1+r}$.

Plus le résultat sera proche de 1, plus la fidélité sera élevée.

On exige généralement au moins 0,80.

¹ La méthode du calcul de la corrélation est expliquée dans tous les manuels de statistique élémentaires.

Quand on prépare l'examen par un prétest, une fidélité insuffisante est, notamment, combattue par une augmentation du nombre de questions.

Si le contrôle de fidélité se fait après l'examen effectif, une fidélité trop faible incitera les maîtres à un supplément de prudence dans l'interprétation des résultats et à plus de soin dans la construction des examens futurs.

b) *Deux formes parallèles.*

Le contrôle mathématique est le même que dans la méthode *pairs-impairs*, mais ici les examinateurs préparent deux séries complètes de questions, unanimement jugées équivalentes, à partir de la même grille d'objectifs.

Si la population est assez nombreuse, une moitié, choisie au hasard (par exemple, selon l'ordre alphabétique des noms) reçoit la première forme et l'autre, la seconde. Pour des populations peu nombreuses, on invite les élèves à répondre aux deux séries de questions après quelques jours. Cette seconde solution est évidemment beaucoup moins rigoureuse.

4. Répétition de la notation.

La fidélité des réponses ne suffit pas. Elle doit être accompagnée de la fidélité de la notation. On contrôle celle-ci soit en invitant un même professeur à noter deux fois les mêmes travaux à un intervalle de quelques jours ou de quelques semaines, soit en comparant les notes accordées par plusieurs correcteurs à un même travail.

CONTRÔLE

La fidélité d'un examen mal construit, qui indiquerait la même longueur néanmoins à une conclusion...

En bref, valider un examen, c'est valider ce pour quoi il est capital : car si la conclusion est déployée pour l'organisation vaine ou, au moins, détournée...

Or, la validation n'est que notre savoir sur la validité de la recherche a à peine touché le problème mérite d'autant plus de professeurs.

Selon la nature rétroactive on distingue la validité de...

I. La validité du contenu.

Elle intéresse principalement le bilan des acquisitions. Tel but avait été assigné, atteints ?

Dans les examens postérieurs, on ne peut évaluer la matière. Non seulement les points principaux, mais l'importance relative. Si les questions pourront susciter l'ensemble des acquisitions...

1 R. COX, Reliability and Validity of Examinations, o.c., p. 43.

r un prétest, une fidélité
ue par une augmentation

près l'examen effectif, une
res à un supplément de
ultats et à plus de soin
rs.

même que dans la méthode
rs préparent deux séries
nt jugées équivalentes, à

reuse, une moitié, choisie
e alphabétique des noms)
seconde. Pour des popu-
s élèves à répondre aux
ques jours. Cette seconde
oins rigoureuse.

pas. Elle doit être accom-
n contrôle celle-ci soit en
deux fois les mêmes tra-
ou de quelques semaines,
ées par plusieurs correc-

CHAPITRE 5.

CONTROLE DE LA VALIDITE.

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. Acquérir cette certitude est capital : car si la conclusion est négative, tous les efforts déployés pour l'organisation et le passage des examens sont vains ou, au moins, détournés de leur objet.

Or, la validation n'est pas chose aisée. Faisant le total de notre savoir sur la validité, R. Cox conclut, en 1969, que la recherche a à peine touché aux questions fondamentales¹. Le problème mérite d'autant plus l'attention et la prudence des professeurs.

Selon la nature rétrospective ou prospective de l'examen, on distingue la validité du contenu et la validité prédictive.

I. La validité du contenu.

Elle intéresse principalement les examens destinés à dresser le bilan des acquisitions et donc aussi d'un enseignement. Tel but avait été assigné, tels objectifs choisis. Ont-ils été atteints ?

Dans les examens portant sur de longues étapes d'apprentissage, on ne peut évidemment interroger en détail sur toute la matière. Non seulement, il importe donc de localiser les points principaux, mais encore faut-il tenir compte de leur importance relative. Si l'on satisfait à ces deux critères, les questions pourront susciter des réponses *représentatives* de l'ensemble des acquisitions, de la compétence totale.

¹ R. COX, Reliability and Validity of Examinations, in J. LAUWERYS et D. SCANLON, *Examinations*, o.c., p. 43.

En d'autres mots, la validité du contenu dépend de la qualité et de l'adéquation de l'échantillonnage des connaissances et des capacités.

Nous avons vu que, dans un enseignement bien conçu, les objectifs à atteindre, et donc à contrôler lors de l'examen, ont été définis au début de l'année.

Mais, à ce moment, il s'agissait d'intentions que les circonstances peuvent avoir modifiées. Avant d'essayer de dresser le bilan du travail de ses élèves, le professeur fera donc d'abord le sien. Faute de ce retour sur soi-même, les examens n'ont qu'une validité illusoire ou, pire encore, une validité de façade, destinée à dédouaner le professeur aux yeux de ses supérieurs. Ainsi s'explique que des questions portent sur des matières à peine effleurées, mais figurant au programme officiel, ou sur des capacités intellectuelles élevées que l'on a omis d'installer patiemment, systématiquement pendant l'année. Combien d'échecs scolaires ne sont-ils pas dus aux belles questions « d'intelligence », posées au terme d'un enseignement qui ne l'a pas cultivée ?

Si la définition précise des objectifs permet une validation de contenu beaucoup meilleure que dans le passé, celle-ci n'en reste pas moins encore très limitée dans l'état actuel de nos connaissances. D'abord, parce qu'il est bien difficile d'acquérir la certitude que tel comportement réel est bien représentatif de tel trait que l'on ambitionnait de cultiver. Ensuite, parce que, même si la vision est correcte au départ, les vicissitudes de la notation peuvent toujours trahir les intentions.

Un professeur d'histoire ou de géographie, écrit P. Vernon¹ qui, par un examen écrit, essaie de juger la compréhension profonde de ces disciplines, attribue souvent une bonne partie des points pour la reproduction exacte de faits détaillés, la restitution de ses théories favorites, la longueur des réponses, leur intelligibilité et l'élégance du style dans lequel elles sont exprimées. Dans un examen oral, combien les points ne dépendent-ils pas de l'intelligence sociale, du contrôle de soi-même et, spécialement, de l'aptitude de l'élève à créer, par la parole et l'attitude, une bonne relation avec l'examineur ?

Que faire, en pratique ?

D'abord, exiger que, pour chaque question d'examen qu'il rédige, le professeur indique l'objectif poursuivi sur une fiche spéciale.

¹ P. VERNON, Types of Examination, in J. LAUWERYS et D. SCANLON, *Examinations*, p. 43.

Ensuite, soumettre à plusieurs éducateurs (tiendra à un groupe de cerné) qui jugeront indépendamment des questions. En cas d'être rejetée¹.

L'expérience montre un petit nombre d'objets large. Sinon, les juges Nous l'avons dit plus validation sont encore fr

Pour la même raison ne sera obtenu sur les doute pas un motif suffisant matière ou de la brève d'accord doit alors être

On risque, par exemple propos du contrôle systématiquement assignés aux langues mathématique, etc.

De nouveau, nous rigueur scientifique de mis en cause. Le fétichisme ce prix.

II. La validité prédictive

Elle semble, de loin n'exige pas de compétences. Même si l'on ne s'explique succès à tel test et une minimisé, le fait peut être ais

Qu'il s'agisse d'un cycle d'études ou autorisés de suivre les individus si le pronostic formulé ou non.

On constate avec est rarement faite. J.-C. est moins destiné à m

¹ Un coefficient de concordance Kendall et un seuil d'acceptation d

tenu dépend de la qualité
ge des connaissances et

ignement bien conçu, les
bler lors de l'examen, ont

d'intentions que les cir-
vant d'essayer de dresser
fesseur fera donc d'abord
même, les examens n'ont
e, une validité de façade,
x yeux de ses supérieurs.
rtent sur des matières à
ramme officiel, ou sur des
e l'on a omis d'installer
dant l'année. Combien
us aux belles questions
n enseignement qui ne l'a

tifs permet une validation
ans le passé, celle-ci n'en
dans l'état actuel de nos
st bien difficile d'acquérir
l est bien représentatif de
ttiver. Ensuite, parce que,
part, les vicissitudes de la
intentions.

ographie, écrit P. Vernon ¹
e juger la compréhension
souvent une bonne partie
acte de faits détaillés, la
la longueur des réponses,
yle dans lequel elles sont
mbien les points ne dépen-
du contrôle de soi-même
ève à créer, par la parole
ec l'examineur ?

ne question d'examen qu'il
tif poursuivi sur une fiche

Ensuite, soumettre les questions et le schéma de notation à plusieurs éducateurs expérimentés (dont un au moins appartiendra à un groupe de disciplines différent du domaine concerné) qui jugeront indépendamment de la validité du contenu des questions. En cas de désaccord grave, la question doit être rejetée ¹.

L'expérience montre qu'il faut actuellement s'en tenir à un petit nombre d'objectifs généraux conçus de façon assez large. Sinon, les juges ne réussissent plus à se mettre d'accord. Nous l'avons dit plus haut, nos connaissances en matière de validation sont encore frustes.

Pour la même raison, dans certains cas, aucun accord ne sera obtenu sur les questions fondamentales. Ce n'est sans doute pas un motif suffisant pour cesser l'enseignement de la matière ou de la branche incriminée, mais l'impossibilité d'accord doit alors être clairement indiquée.

On risque, par exemple, de rencontrer pareille difficulté à propos du contrôle systématique de la réalisation des objectifs assignés aux langues anciennes, à certaines matières de la mathématique, etc.

De nouveau, nous constatons qu'en tentant d'accroître la rigueur scientifique des examens, des enseignements seront mis en cause. Le fétichisme pédagogique ne disparaîtra qu'à ce prix.

II. La validité prédictive.

Elle semble, de loin, la plus simple à contrôler, car elle n'exige pas de compréhension profonde des phénomènes. Même si l'on ne s'explique pas le rapport existant entre le succès à tel test et une réussite brillante dans un domaine déterminé, le fait peut être aisément observé.

Qu'il s'agisse d'un examen donnant accès à un nouveau cycle d'études ou autorisant l'exercice d'une profession, il suffit de suivre les individus pendant un certain temps pour savoir si le pronostic formulé à partir des notes attribuées se vérifie ou non.

On constate avec étonnement que semblable vérification est rarement faite. J.-C. Passeron y voit le signe que l'examen est moins destiné à mesurer objectivement la capacité qu'à

¹ Un coefficient de concordance entre les avis peut être calculé par la formule de Kendall et un seuil d'acceptation choisi de commun accord.

servir les besoins sociaux des classes privilégiées¹. Cette affirmation mérite certes des réserves, mais nous avons vu, dans la première partie, qu'elle n'est pas entièrement dénuée de fondement.

Pour n'être pas un leurre, la vérification systématique de la validité prédictive exige le contrôle rigoureux des variables, faute de quoi les variables cachées fausseront les conclusions. Par exemple, on tirait encore récemment argument en faveur du latin en faisant la statistique du nombre d'étudiants brillants qui avaient fait des études secondaires classiques. Or, cette observation ne prouve rien, sinon que, traditionnellement dans nos pays, les élèves les plus intelligents sont orientés dès douze ans vers ce type d'études.

En guise de synthèse, nous empruntons la récapitulation suivante à l'*Examination Bulletin* n° 3².

FACTEURS DE VALIDITE D'UN EXAMEN

1. Identification adéquate des objectifs dont la réalisation doit être vérifiée par l'examen écrit, l'évaluation du travail de l'année, l'épreuve pratique, etc.
2. Parmi toute la gamme des objectifs ainsi identifiés, sélection de ceux sur lesquels la vérification se concentrera.
3. Evaluation efficace de l'adéquation du contenu et de la structure de l'examen au but poursuivi.
4. Relation claire entre chaque question et les objectifs de l'enseignement.
5. Elaboration d'un schéma de notation et rédaction de directives aux correcteurs en fonction des objectifs. Un niveau suffisant de fidélité doit aussi être assuré.
6. Bonne connaissance des capacités des candidats.
7. Disposition des notateurs à tenir compte des jugements indépendants du leur et des données objectives qui leur seraient fournies.
8. Comparaison avec des épreuves antérieures dont la validité a été prouvée.

1 J.-C. PASSERON, *Sociologie des examens*, o.c., p. 7.

2 Londres, H.M.S.O., 1964, pp. 19-20.

classes privilégiées¹. Cette
rives, mais nous avons vu,
pas entièrement dénuée de

ification systématique de la
e rigoureux des variables,
fausseront les conclusions.
mmment argument en faveur
u nombre d'étudiants bril-
econdaires classiques. Or,
non que, traditionnellement
telligents sont orientés dès

empruntons la récapitulation
3².

E D'UN EXAMEN

ectifs dont la réalisation doit
l'évaluation du travail de

ifs ainsi identifiés, sélection
on se concentrera.

tion du contenu et de la
suivi.

estion et les objectifs de

ation et rédaction de direc-
n des objectifs. Un niveau
tre assuré.

ités des candidats.

air compte des jugements
onnées objectives qui leur

antérieures dont la validité

QUATRIEME PARTIE

LES PROCEDURES DE MODERATION

POSITI

1. Définition.

La *modération*, en d
de tempérer les excès de
tains examinateurs.

Au sens large, elle
mesures prises pour rend
internes et donc pour
des différentes classes d
un groupe d'écoles, dans
région ou d'un pays.

Les recherches les p
tainement dues aux docim
tème complètement déce
Bretagne, les programmes
ou secondaires peuvent c
dant très longtemps, des
nisés en dehors des école
dû sanctionner les études
On connaît le fameux 11 -
le secondaire, décidant,
l'enseignement général c
Secondary Education), av
rieur. Ce dernier examen,
d'un document signalant l
par rapport à des normes r

Pour admettre un élè
universités annoncent leu
notes d'excellence (A) so
biologie et en langue mat
sont aussi acceptées, etc.

CHAPITRE 1.

POSITION DU PROBLEME.

1. Définition.

La *modération*, en docimologie, a d'abord eu pour objet de tempérer les excès de sévérité ou de générosité chez certains examinateurs.

Au sens large, elle désigne aujourd'hui l'ensemble des mesures prises pour rendre comparables les notes d'examens internes et donc pour unifier leur signification au niveau des différentes classes de même type dans une école, dans un groupe d'écoles, dans des établissements similaires d'une région ou d'un pays.

Les recherches les plus fines sur la modération sont certainement dues aux docimologistes britanniques. Dans un système complètement décentralisé comme celui de la Grande-Bretagne, les programmes et les méthodes des écoles primaires ou secondaires peuvent différer considérablement. Aussi, pendant très longtemps, des examens externes, c'est-à-dire organisés en dehors des écoles, par des commissions spéciales, ont dû sanctionner les études aux moments cruciaux de la scolarité. On connaît le fameux 11 + *Examination*, examen d'entrée dans le secondaire, décidant, en particulier, de l'admission dans l'enseignement général classique, et le C.S.E. (*Certificate of Secondary Education*), avant l'entrée dans l'enseignement supérieur. Ce dernier examen, par exemple, pourvoit chaque élève d'un document signalant le niveau dans chacune des branches, par rapport à des normes nationales.

Pour admettre un élève dans une section déterminée, les universités annoncent leurs exigences particulières : ici, des notes d'excellence (A) sont demandées en mathématiques, en biologie et en langue maternelle ; là, les notes « très bien » (B) sont aussi acceptées, etc.

Ce système d'examens externes est critiqué, parfois très sévèrement, parce qu'il limite gravement la liberté pédagogique des maîtres et des autorités locales.

Aussi, depuis quelques années, on tente de les remplacer par des examens internes dont les résultats sont rendus comparables par des procédures de modération locale, régionale, puis nationale, expérimentalement mises au point.

Dans des pays comme la Belgique, où un programme d'études unique est imposé par une autorité centrale, il est beaucoup plus facile, si on le veut, de modérer les examens. Avant de voir comment, il n'est peut-être pas inutile de rappeler pourquoi une modération est souhaitable.

En gros, la raison est double. La première concerne les individus, élèves et parents, qui ont le droit de connaître le niveau « réel » des performances scolaires, avant de décider de l'orientation ultérieure dans les domaines des études ou de la profession. Qui se croirait qualifié pour les jeux olympiques parce qu'il a gagné une course organisée entre quelques amis ? Un constat défavorable n'est d'ailleurs pas plus une condamnation sans appel qu'un diagnostic de faiblesse ou de maladie. Il faut savoir à temps qu'un problème se pose pour en chercher les causes, puis les remèdes, s'ils existent.

La seconde raison concerne la communauté. De même qu'un consommateur ne peut être trompé sur la marchandise, de même la société ne peut être tenue ou à payer fort cher les échecs d'un étudiant entré à l'université sur la foi d'un certificat invalide, ou à confier une fonction à quelqu'un qui ne la mérite pas. Le jour où, par exemple, les examens de l'enseignement normal seront rigoureusement contrôlés, on peut espérer ne plus voir confier des générations d'enfants à certains instituteurs dont la connaissance de la langue maternelle, notamment de l'orthographe¹, est très insuffisante.

2. Modérer n'est pas caporaliser.

On pourrait craindre que la volonté de rendre les résultats scolaires comparables ne provoque une résurgence des vieilles contraintes. On se souvient de cet inspecteur, de tradition napoléonienne, qui, consultant sa montre, croyait pouvoir dire : « A cet instant, on enseigne telle leçon dans toutes les cinquièmes années primaires de France. » Nous savons les faiblesses d'un enseignement caporalisé.

¹ Nous n'avons nullement le fétichisme de l'orthographe, mais tenons pour évident que celui qui doit l'enseigner doit la connaître.

Mais la nécessité d'un pas qu'au-delà des années cycle d'études poursuit de à tous : acquisition de tech et de capacités jugées ess vent s'interroger et s'enter

3. Modération volontaire

Normaliser tous les es des ne semble ni souhaitable, parce que les élèves très différents et s'accom en périodes et en années modération, même par le être présentées, reste un dans son exécution.

La modération doit é avant tout lors de l'attribu fin de cycle.

Pour le reste, la décis tres individuellement, soit autorités locales.

4. La modération comm

Nous venons d'y faire a modérateurs aient pu trou l'enseignement et sur que pour que les examens soien

Des affirmations vague s'exprimer correctement pa désirée doit être traduite en que l'on aura d'ailleurs int d'évaluation descriptives. S des questions se posent :

Comment graduer les élèves et du type d'éco

L'originalité sera-t-elle ment sera-t-elle identifi

Quelle importance attr

La longueur des trava dérée ? Exige-t-on une laquelle des points se

es est critiqué, parfois très
ment la liberté pédagogique

, on tente de les remplacer
résultats sont rendus compa-
modération locale, régionale,
nises au point.

que, où un programme d'étu-
rité centrale, il est beaucoup
érer les examens. Avant de
inutile de rappeler pourquoi

La première concerne les
nt le droit de connaître le
scolaires, avant de décider
domaines des études ou de
é pour les jeux olympiques
nisée entre quelques amis ?
urs pas plus une condamna-
faiblesse ou de maladie. Il
se pose pour en chercher
istent.

la communauté. De même
trompé sur la marchandise,
due ou à payer fort cher les
ersité sur la foi d'un certifi-
tion à quelqu'un qui ne la
les examens de l'enseigne-
contrôlés, on peut espérer
s d'enfants à certains insti-
langue maternelle, notam-
uffisante.

onté de rendre les résultats
une résurgence des vieilles
specteur, de tradition napo-
e, croyait pouvoir dire : « A
dans toutes les cinquièmes
s savons les faiblesses d'un

orthographe, mais tenons pour évident

Mais la nécessité d'une plasticité pédagogique n'empêche pas qu'au-delà des aménagements circonstanciels, chaque cycle d'études poursuit des objectifs fondamentaux, communs à tous : acquisition de techniques de base ou de connaissances et de capacités jugées essentielles. Maîtres et modérateurs doivent s'interroger et s'entendre sur ces apprentissages cruciaux.

3. Modération volontaire ou imposée ?

Normaliser tous les examens à tous les moments des études ne semble ni souhaitable, ni d'ailleurs possible. Peu souhaitable, parce que les élèves progressent à des rythmes parfois très différents et s'accommodent mal du découpage rigoureux en périodes et en années scolaires. Impossible parce que la modération, même par les méthodes économiques qui vont être présentées, reste un travail lourd, dans sa préparation et dans son exécution.

La modération doit être imposée aux moments décisifs, avant tout lors de l'attribution de diplômes ou de certificats de fin de cycle.

Pour le reste, la décision devrait être laissée soit aux maîtres individuellement, soit aux chefs d'établissement ou aux autorités locales.

4. La modération commence au début de l'année scolaire.

Nous venons d'y faire allusion, il importe que professeurs et modérateurs aient pu trouver un accord sur les objectifs de l'enseignement et sur quelques grands principes de notation pour que les examens soient comparables.

Des affirmations vagues telles que : « L'élève doit pouvoir s'exprimer correctement par écrit » sont sans utilité. La capacité désirée doit être traduite en termes de comportements concrets, que l'on aura d'ailleurs intérêt à incorporer dans les échelles d'évaluation descriptives. S'il s'agit de l'expression écrite, bien des questions se posent :

Comment graduer les exigences en fonction de l'âge des élèves et du type d'école ?

L'originalité sera-t-elle récompensée ? Comment ? Comment sera-t-elle identifiée ?

Quelle importance attribuer à l'orthographe ?

La longueur des travaux de composition sera-t-elle considérée ? Exige-t-on une longueur minimum, en dessous de laquelle des points seront décomptés ? Combien ?

Quelle importance accorder à la richesse du vocabulaire ?

Comment la définir ? Quelles seront les exigences syntaxiques ?

Les mathématiciens peuvent aussi accorder leurs violons. Si, pour nous limiter à un seul exemple, certains professeurs pénalisent gravement les fautes d'opérations, voire les fautes d'orthographe dans la solution de problèmes, alors que leurs collègues estiment que ces aspects sont secondaires, des notes égales risquent de recouvrir des réalités très différentes.

Les programmes officiels peuvent aussi jouer un rôle important en précisant au maximum les buts à atteindre à chaque échelon. L'efficacité de ces données de base est d'autant plus grande que le système scolaire est centralisé.

Inutile de se faire des illusions : les objectifs comportementaux des différentes branches ne se définissent pas en quelques heures de méditation. Des commissions où enseignants, inspecteurs et chercheurs uniront leurs efforts devront travailler longtemps avant d'arriver à un résultat satisfaisant. Beaucoup de questions risquent d'ailleurs de rester provisoirement sans réponse.

La définition des objectifs fera l'objet d'une recherche permanente, non seulement parce qu'ils pourront se préciser en fonction de l'avancement de la psychologie de l'apprentissage, mais aussi parce que les objectifs mêmes évoluent.

Les instructions relatives aux examens se préciseront parallèlement. A côté des grilles d'objectifs généraux et spéciaux, les programmes scolaires de l'avenir détermineront nettement la quantité et les types de questions à poser aux examens et fourniront des indications sur les lignes de force des schémas de notation.

5. Pas de comparabilité sans fidélité élevée.

Il n'est pas concevable de comparer entre eux des résultats d'examens qui, pris isolément, seraient éminemment fluctuants.

Le problème de la fidélité a été discuté dans la partie consacrée à la construction de l'examen. Nous n'y revenons pas.

6. Peut-on se fier aux tests

Dans les systèmes de notation, les tests d'intelligence occupent une place importante. En ce qui concerne les points de repère sur lesquels on se base, en d'autres cas, on leur attribue une importance qu'au travail de l'année.

Maintenant que nous avons pris les précautions les plus sérieuses, la confiance accordée à ces tests diminue. Par ailleurs, dès que nous avons aussi vu que si le test verbal, administré par un maître, est un simple test verbal, administré par un maître, la prédiction presque aussi bonne que celle du maître.

Pronostic du maître
Test verbal

Voici encore quelques observations.

A moyen terme, une combinaison de scores de tests est un meilleur prédicteur du succès scolaire que les examens. Cette supériorité est surtout par des chercheurs.

- 1) Emmet¹ montre qu'il est possible de mieux prévoir les résultats scolaires à long terme (niveau secondaire maternelle et d'arithmétique) que les résultats à court terme (niveau primaire).
- 2) Emmet et Wilmut² ont montré que tout aussi convaincant que les tests d'intelligence, cinq ans avant l'examen.

1 W. EMMET, *An Inquiry into the Validity of Intelligence Tests*, Univ. of London Press, 1942.

2 W. EMMET and F. WILMUT, *Intelligence Tests and Specific Subjects*, in *British Journal of Educational Psychology*, 1942.

la richesse du vocabulaire ?

seront les exigences syn-

aussi accorder leurs violons.
exemple, certains professeurs
opérations, voire les fautes
problèmes, alors que leurs
sont secondaires, des notes
qualités très différentes.

uvent aussi jouer un rôle
n les buts à atteindre à cha-
années de base est d'autant
re est centralisé.

: les objectifs comporte-
définissent pas en quelques
ions où enseignants, inspec-
forts devront travailler long-
t satisfaisant. Beaucoup de
rester provisoirement sans

l'objet d'une recherche per-
ils pourront se préciser en
chologie de l'apprentissage,
mêmes évoluent.

examens se préciseront paral-
ectifs généraux et spéciaux,
nir détermineront nettement
ns à poser aux examens et
ignes de force des schémas

é élevée.

parer entre eux des résultats
ient éminemment fluctuants.

été discuté dans la partie
examen. Nous n'y revenons

6. Peut-on se fier aux tests ?

Dans les systèmes de modération que nous allons examiner, les tests d'intelligence ou de connaissances occupent une place importante. En certaines occasions, ils apportent les points de repère sur lesquels les autres résultats sont ajustés ; en d'autres cas, on leur accorde, à eux seuls, autant de valeur qu'au travail de l'année et qu'à l'examen final.

Maintenant que nous avons vu avec quelle rigueur et quelles précautions les bons tests de connaissances sont construits, la confiance accordée à ces instruments ne doit plus nous étonner. Par ailleurs, dès la première partie de ce livre, nous avons aussi vu que si les maîtres jugent bien leurs élèves, un simple test verbal, administré en moins d'une heure, permet une prédiction presque aussi sûre :

	Corrélation avec résultats de l'élève :	
	2 ans après	3 ans après
Pronostic du maître	. 821	. 748
Test verbal	. 796	. 722

Voici encore quelques résultats de recherches confirmant ces observations.

A moyen terme, un test d'intelligence ou, mieux encore, une combinaison de scores à différents tests d'intelligence, sont meilleurs prédicteurs du succès scolaire que les résultats d'examens. Cette supériorité a été démontrée à plusieurs reprises, surtout par des chercheurs anglo-saxons :

- 1) Emmet¹ montre qu'un test d'intelligence verbale permet de mieux prévoir les résultats, après deux ou trois ans d'enseignement secondaire général, que les examens de langue maternelle et d'arithmétique évalués par les maîtres.
- 2) Emmet et Wilmut² ont, par la suite, fait une démonstration tout aussi convaincante de la valeur prédictive du test d'intelligence, cinq ans après.

1 W. EMMET, *An Inquiry into the Prediction of Secondary School Success*, London, Univ. of London Press, 1942.

2 W. EMMET and F. WILMUT, *The Prediction of School Certificates Performance in Specific Subjects*, in *British Journal of Educ. Psychol.*, 22, 1952, 52-62.

- 3) Wrigley¹ a confirmé ces résultats et montré que la prédiction peut être meilleure encore si les résultats à un test d'intelligence et à des tests de connaissances standardisés peuvent être combinés.
- 4) En Belgique, tous les orienteurs connaissent la valeur prédictive des scores verbaux et de raisonnement du Test P.M.A. de Thurstone, le fameux 2 V + R, pour l'enseignement secondaire général.

Nous ne disposons pas de recherches similaires pour l'enseignement technique ou pour des structures d'enseignement peut-être mieux adaptées à la civilisation de l'an 2000. Il est très possible que les scores d'autres types de tests devraient être utilisés.

¹ J. WRIGHLEY, The Relative Efficiency of Intelligence and Attainment Tests as Predictors of Success in Grammar Schools, in *British Journal of Educ. Psychol.*, 25, 1955, 107-116.

QUELQUES SYSTEMES

I. La forme

Le système suédois à partir d'un

Dès la fin de la seconde, un système de modération est appliqué à chaque professeur, dans sa

Toutes les notes des élèves sont sur une échelle de 7 degrés¹.

On suggère que le pourcentage soit déterminé soit, en gros, en tenant compte des caractéristiques de tête très forte, ou presque

Notes

Pourcentage

Par ailleurs, une batterie de tests est appliquée sur les branches principales chaque année par un officier spécialisé et des professeurs. Les résultats sont répartis en 7 classes pour les notes des professeurs.

¹ Nous nous référons à l'exposition "System of Equalising Marks, in *Educational Research*, 1955, p. 107-116.
² Distribution normale (voir courbe)

ultats et montré que la prédic-
ore si les résultats à un test
e connaissances standardisés

rs connaissent la valeur pré-
et de raisonnement du Test
ux 2 V + R, pour l'enseigne-

cherches similaires pour l'en-
es structures d'enseignement
lisation de l'an 2000. Il est très
types de tests devraient être

CHAPITRE 2.

QUELQUES SYSTEMES DE MODERATION DES EXAMENS.

I. La formule la plus libérale :

Le système suédois de modération par branche à partir d'un test de connaissances.

Dès la fin de la seconde guerre mondiale, la Suède a adopté un système de modération, simple et facultatif, applicable par chaque professeur, dans sa classe, tout au long de la scolarité.

Toutes les notes des professeurs sont attribuées selon une échelle de 7 degrés ¹.

On suggère que le pourcentage d'élèves recevant une note déterminée soit, en gros, le suivant ², le professeur restant libre de tenir compte des caractéristiques de sa classe (par exemple : tête très forte, ou presque tous élèves moyens, etc.) :

Notes	1	2	3	4	5	6	7
Pourcentage	1	6	24	38	24	6	1

Par ailleurs, une batterie de tests de connaissances portant sur les branches principales du programme est construite chaque année par un office central où collaborent des spécialistes et des professeurs expérimentés. Cette batterie est étalonnée sur un échantillon national représentatif ; les résultats sont répartis en 7 classes, dans la proportion conseillée pour les notes des professeurs.

¹ Intelligence and Attainment Tests as
British Journal of Educ. Psychol., 25, 1955,

¹ Nous nous référons à l'exposé très clair de S. HENRYSSON, The Swedish System of Equalising Marks, in *Educational Research*, VI, 2, Feb. 1964, 156-160.

² Distribution normale (voir courbe de Gauss).

Si le professeur administre le test national dans sa classe, il dispose donc de normes directement comparables aux notes qu'il a attribuées. Il lui est donc facile d'ajuster ces dernières.

Voici un exemple concret proposé par S. Henrysson :

24 élèves			
Notes	Distribution des notes préliminaires attribuées par le professeur	Distribution des notes obtenues au test	Distribution des notes après ajustement
7	—	1	—
6	—	4	4
5	4	4	4
4	4	9	10
3	14	5	4
2	2	1	2
1	—	—	—
Moyenne	3,42	4,33	4,17

Le professeur inscrit d'abord, dans la 2^e colonne, les notes qu'il a attribuées. La moyenne (3,42) fournit une première indication, si la moyenne nationale est connue. (En Suède, elle est de 4 pour la 6^e primaire. Il semble donc que le professeur soit ici trop sévère).

Voici maintenant les résultats obtenus au test par les mêmes élèves :

Score brut	Note correspondante	Nombre d'élèves obtenant cette note
94-100	7	1
84-93	6	4
67-83	5	4
46-66	4	9
30-45	3	5
22-29	2	1
0-21	1	—

Ces résultats sont reportés sur le tableau n° 3.

La moyenne est 4,33, ce qui indique une distribution de sévérité.

Sans changer l'ordre des notes, le professeur rectifie ses notes (colonne 4).

Le système suédois s'explique ainsi :

1° Les maîtres gardent leur liberté de jugement :

- a) de faire passer ou non un élève à l'étape suivante ;
- b) de tenir compte des progrès réalisés ;
- c) de communiquer les résultats à d'autres professeurs, à d'autres parents.

Ainsi, le test devient une aide à la sélection des praticiens. La tension est maintenue au minimum.

2° Les constructeurs des tests ont défini un ensemble de objectifs échantillonnés.

Il est évident que les praticiens ne sont pas attachés à la fidélité aux matières traitées, mais à l'efficacité ainsi d'un moyen efficace pour promouvoir des innovations.

Remarquons que, depuis 1969, ce système a été expérimenté dans des écoles primaires (10-12 ans) pour évaluer les progrès.

Parmi les critiques formulées, on retiendra :

1. Si un professeur a jugé qu'un élève a une sévérité excessive, il ne réparera pas l'injustice.
2. L'épreuve de référence ne prend pas en compte les dimensions que les notes ne reflètent pas.
3. La construction de l'épreuve est parfois défectueuse : fidélité trop faible.

1 Voir S. ROLLER, Le problème de la normalisation des notes, Genève, in *Docimologie et éducation*, avril-septembre 1969, pp. 1-10.

2 W. ANGOFF, Can Usual Grade Be Determined from Different College Admission Tests, in *Journal of Educational Measurement*, Washington, ACE, 1966, pp. 251-264.

F. BACHER, La normalisation des notes, Paris, 1969.

test national dans sa classe, ment comparables aux notes facile d'ajuster ces dernières.

posé par S. Henrysson :

Notes	
Distribution des notes obtenues au test	Distribution des notes après ajustement
1	—
4	4
4	4
9	10
5	4
1	2
—	—
4,33	4,17

dans la 2^e colonne, les notes (42) fournit une première indication connue. (En Suède, elle est donc que le professeur soit

ts obtenus au test par les

Indicateur	Nombre d'élèves obtenant cette note
	1
	4
	4
	9
	5
	1
	—

Ces résultats sont reportés dans le premier tableau, colonne 3.

La moyenne est 4,33, ce qui confirme la première impression de sévérité.

Sans changer l'ordre du classement initial, le professeur rectifie ses notes (colonne 4).

Le système suédois séduit pour plusieurs raisons :

1^o Les maîtres gardent leur entière liberté :

- a) de faire passer ou non le test (presque tous le font) ;
- b) de tenir compte des résultats ;
- c) de communiquer les résultats des tests aux élèves, aux autres professeurs, au directeur, aux parents.

Ainsi, le test devient simplement un outil mis à la disposition des praticiens. La tentation de « bachotage » est donc réduite au minimum.

2^o Les constructeurs des tests veillent à couvrir une large gamme d'objectifs échantillonnant bien tout le programme.

Il est évident que les professeurs prêtent une attention particulière aux matières traitées dans les tests. On dispose donc ainsi d'un moyen efficace pour sensibiliser le maître à certaines innovations.

Remarquons que, depuis 1965, un système fort semblable est expérimenté dans des classes genevoises de 5^e et de 6^e primaire (10-12 ans) pour l'orthographe et l'arithmétique¹.

Parmi les critiques formulées contre ce système d'ajustement², on retiendra :

1. Si un professeur a jugé un des élèves de sa classe avec une sévérité excessive, la référence à l'épreuve nationale ne réparera pas l'injustice.
2. L'épreuve de référence peut ne pas mesurer les mêmes dimensions que les notes.
3. La construction de l'épreuve de référence peut aussi être défectueuse : fidélité trop basse,...

¹ Voir S. ROLLER, Le problème de l'attribution des notes scolaires. Essai de solution, Genève, in *Docimologie et éducation*, numéro spécial de la revue *Les sciences de l'éducation*, avril-septembre 1969, pp. 66 sq.

² W. ANGOFF, Can Usual General Purpose Equivalency Tables be Prepared for Different College Admission Tests, in A. ANASTASI, Ed., *Testing Problems in Perspective*, Washington, ACE, 1966, pp. 251-264.

F. BACHER, *La normalisation des notes*, o.c., p. 63.

II. Système imposé de modération par branche à partir d'un test de connaissances.

A partir des mêmes données de base que celles de la Suède (examens internes et tests de connaissances, pour les branches principales, étalonnés nationalement ou régionalement), le système directif suivant peut donner de bons résultats.

Une commission nationale ou régionale de modération prend un certain nombre d'écoles en charge.

Ces écoles lui envoient les résultats aux épreuves préparées librement par les professeurs et les scores obtenus aux tests de connaissances.

En cas de différence, en plus ou en moins, égale ou supérieure à deux écarts types (par exemple) entre les moyennes aux examens et aux tests, l'école reçoit la visite des modérateurs.

Leur mission n'est pas de dire au directeur comment il doit conduire son école, ni au professeur comment faire son cours, mais bien d'attirer l'attention sur un fait et de tâcher de trouver, en pleine collaboration avec l'école, l'explication et, si possible, le remède.

Si le désaccord subsiste, le droit d'ajuster d'office les notes peut être donné aux autorités régionales, l'école ayant, de son côté, le droit d'interjeter appel.

III. Modération volontaire par appel à une banque d'items.

Dans le cas où les maîtres utilisent des questions sous forme d'items à choix multiple, la démarche suivante, actuellement expérimentée en Grande-Bretagne semble pleine de promesses¹. Toutefois, elle requiert l'existence d'un service de recherche, nécessité dont un éducateur averti ne peut d'ailleurs plus douter aujourd'hui.

1. Les maîtres indiquent sur une grille les objectifs poursuivis.
2. Ils envoient cette grille au service de recherche, en même temps que les items qu'ils ont rédigés et, éventuellement, déjà prétestés localement.
3. Le service examine ces items et, selon les possibilités, met certains d'entre eux à l'épreuve ; leur difficulté et leur pouvoir discriminatif sont calculés.

¹ Voir D. PIDGEON et A. YATES, o.c

4. Le service renvoie au maître

- a) les items examinés
- b) des items compléments d'efficacité sont compilés par ville, canton, pays,.... Ces items fourniront duquel le résultat est ajusté.

Il est intéressant de noter que, dans la rédaction de questions à choix multiple, les Britanniques ont adopté des questions semi-ouvertes. Il s'agit, en fait, de questions généralement formulées, pour lesquelles les grands types de réponses sont mentalement les grands types de réponses auxquelles des barèmes de correction sont établis.

L'organisation d'une banque d'items pour les premières années. Par la suite, elle devient telle que le travail est plus souple et une grande souplesse et un

IV. Un système similaire au service de recherche

Nous avons déjà observé que les maîtres ont bien leurs élèves, mais qu'ils ont des difficultés par rapport au niveau global de l'enseignement.

Le système suivant, mis en œuvre en Angleterre pour la Recherche dans l'enseignement secondaire, sélectionne en tenant compte de la difficulté et du pouvoir discriminatif par les maîtres, d

1. L'instituteur classe ses élèves selon son évaluation et les classements dans l'enseignement secondaire peuvent être classés ex
2. Les élèves subissent un test
3. Les scores au test sont comparés (1) et (2)
4. Les classements (1) et (2) sont comparés

¹ Voir A. YATES et D. PIDGEON, o.c

modération par branche connaissances.

de base que celles de la
de connaissances, pour les
ationalement ou régionale-
peut donner de bons résul-

u régionale de modération
n charge.

sultats aux épreuves prépa-
et les scores obtenus aux

ou en moins, égale ou supé-
(exemple) entre les moyennes
reçoit la visite des modéra-

re au directeur comment il
professeur comment faire son
on sur un fait et de tâcher
n avec l'école, l'explication

droit d'ajuster d'office les
régionales, l'école ayant, de

appel à une banque d'items.

utilisent des questions sous
démarche suivante, actuelle-
tagne semble pleine de pro-
l'existence d'un service de
cateur averti ne peut d'ail-

rille les objectifs poursuivis.

ice de recherche, en même
rédigés et, éventuellement,

t, selon les possibilités, met
; leur difficulté et leur pou-

4. Le service renvoie au maître :

- a) les *items* examinés ;
 - b) des *items* complémentaires dont les indices de facilité et d'efficacité sont connus pour une population donnée : ville, canton, pays,...
- Ces *items* fourniront un point de comparaison à partir duquel le résultat enregistré pour les autres pourra être ajusté.

Il est intéressant de noter qu'après s'être étroitement cantonnés dans la rédaction d'*items* à réponse fermée, par choix multiple, les Britanniques commencent à proposer des questions semi-ouvertes. Il s'agit, en fait, de questions ouvertes, très soigneusement formulées, pour lesquelles on a identifié expérimentalement les grands types de réponses, en fonction desquelles des barèmes de correction sont proposés.

L'organisation d'une banque d'*items* est lourde pendant les premières années. Par la suite, la provision d'*items* étalonnés devient telle que le travail s'allège et que le système acquiert une grande souplesse et une grande rapidité de fonctionnement.

IV. Un système simple de modération globale au service de la sélection.

Nous avons déjà observé qu'en général, les maîtres jugent bien leurs élèves, mais qu'ils tendent à relativiser leurs jugements par rapport au niveau global de la classe.

Le système suivant, mis au point par la Fondation nationale anglaise pour la Recherche en Education¹ pour l'admission dans l'enseignement secondaire général classique, permet de *sélectionner* en tenant compte des jugements portés indépendamment par les maîtres, dans leur classe.

1. L'instituteur classe ses élèves par ordre de mérite, c'est-à-dire selon son évaluation globale de la chance de réussite dans l'enseignement secondaire général. Plusieurs élèves peuvent être classés *ex aequo* (Classement 1).
2. Les élèves subissent un test d'intelligence verbale (2).
3. Les scores au test sont classés par ordre décroissant (Classement 3).
4. Les classements (1) et (3) sont placés côte à côte.

¹ Voir A. YATES et D. PIDGEON, *Admission to Grammar Schools*, o.c.

5. Le score d'intelligence qui tombe en face du nom de l'élève est considéré comme score d'évaluation par l'instituteur (Jugement ajusté).

Classement par l'instituteur (1)	Score obtenu par l'élève au test verbal (2)	Classement des scores par ordre décroissant (3)	Jugement ajusté (4)
A	121	132	132
B	120	128	128
C	132	121	121
D	128	120	120
E	100	106	106
F	106	100	100
G	94	100	97
H	82	96	97
I	96	94	97
J	100	86	86
K	78	82	78
L	79	79	78
M	86	78	78
N	73	73	78
O	65	65	65
		Total 290	
		Total 312	

On voit que le classement auquel on aboutit finalement ne modifie en rien l'ordre initialement choisi par l'instituteur. Mais, cette fois, nous disposons d'un moyen de comparaison entre écoles.

Utilisation des jugements

Exemple : Dans un lycée latin, 25 élèves provenant

	Ecole primaire I	
	132	
130	-----	
	128	
125	-----	
	121	
120	-----	
	120	
115	-----	115 x 3
	112	
110	-----	
105	-----	106
100	-----	100

Pour sélectionner 25 élèves, à la coupure de 110, 25 élèves dans la catégorie élémentaire aurait exigé qu

Tracer une ligne de détermination est une source d'injustice, seul n'est-il pas responsable de sujets juste au-dessus

Le psychométricien est té : il prend une marge de dard qu'une formule simple coefficient de fidélité du test

1 Erreur standard = $\sigma \sqrt{1-r}$
 où σ = écart type des scores.
 r = coefficient de fidélité. Ce manuel qui accompagne le

en face du nom de l'élève
l'évaluation par l'instituteur

Classement des scores par ordre décroissant (3)		Jugement ajusté (4)
132		132
128		128
121		121
120		120
106		106
100		100
100	Total 290	97
96		97
94		97
86		86
82	Total 312	78
79		78
78		78
73		78
65		65

quel on aboutit finalement ne
choisi par l'instituteur. Mais,
moyen de comparaison entre

Utilisation des jugements ajustés.

Exemple : Dans un lycée, on souhaite recruter, pour la 6^e latine, 25 élèves provenant de trois écoles primaires différentes.

Jugements ajustés

	Ecole primaire I	Ecole primaire II	Ecole primaire III
	132		
130	----- 130		----- 130 x 2
	128		
125	----- 125	125	----- 126
		124	----- 125
120	121	121	123
	----- 120	119	----- 120
		117	
115	115 x 3	115	116
	----- 115		
	112		114
110	----- 110	110	----- 113
			----- 110
105	106	107	
	----- 105	104	
100	100		
	----- 100		

Pour sélectionner 25 élèves, on coupe, dans ce cas, à 110. Si, à la coupure de 110, des *ex aequo* avaient placé plus de 25 élèves dans la catégorie des sélectionnés, la justice la plus élémentaire aurait exigé qu'on les admette tous.

Tracer une ligne de démarcation comme on vient de le faire est une source d'injustice pour une autre raison : le hasard seul n'est-il pas responsable du placement d'un certain nombre de sujets juste au-dessus ou en dessous de la limite ?

Le psychométricien est armé pour surmonter cette difficulté : il prend une marge de sécurité de trois fois l'erreur standard qu'une formule simple permet de calculer à partir du coefficient de fidélité du test utilisé¹.

¹ Erreur standard = $\sigma \sqrt{1-r}$

où σ = écart type des scores.

r = coefficient de fidélité. Ce coefficient est généralement indiqué dans le manuel qui accompagne le test.

V. En Angleterre, un système de modération complet.

Le système que nous allons décrire maintenant est, lui aussi, dû à la Grande-Bretagne. A notre connaissance, il n'en existe pas de plus complet : il porte sur la préparation de l'examen en collaboration avec les écoles, la notation, l'ajustement des notes et des grades finaux.

Le but poursuivi est de *perfectionner les examens internes*, au point de pouvoir leur confier le rôle joué jusqu'à présent par les grandes épreuves externes.

Pour en arriver à une modération nationale, on procède par paliers : à l'échelon local où un certain nombre d'écoles poursuivant une même finalité se groupent ; à l'échelon régional ensuite, selon une technique qui ne diffère pas fondamentalement de celle appliquée localement ; pour le passage au niveau national, s'ajoute, à la technique de modération proprement dite, un échantillonnage fin dont l'étude technique serait hors de propos ici.

Pour notre pays, nous suggérons que quelques écoles de même esprit commencent par faire une expérience volontaire du système. C'est pourquoi nous concentrons notre présentation sur le processus de base.

La modération inter-écoles ne serait faite qu'aux moments cruciaux de la scolarité et pour quelques branches principales. Rien n'empêche, toutefois, les professeurs d'une même école, enseignant les mêmes cours, d'utiliser spontanément la même méthode de correction.

A. Préliminaires.

Nous avons déjà fait allusion aux accords à prendre sur les objectifs et sur la construction de l'examen. Nous n'y revenons plus.

Chaque école choisit en son sein un professeur qui jouera deux rôles :

- 1° Coordonnateur des examens dans son école ;
- 2° Membre de la commission de modération inter-écoles.

Un portrait idéal du modérateur n'existe pas. Les qualités suivantes paraissent souhaitables¹ :

1. Maturité générale et bonne expérience pédagogique ;

¹ MATHER, o.c., p. 67.

2. Contacts fréquents avec les autres écoles concernées par l'examen ;
3. Capacité de formulations précises ;
4. Bonté mais fermeté ;
5. Capacité de discuter ;
6. Aptitude à comprendre les statistiques.

Dans chaque école, avec une grande liberté d'action, les professeurs ont des objectifs principaux, le nombre de candidats à respecter selon la convention.

Au début au moins, les questions à la commission de modération sont à l'unité générale.

B. Les professeurs notent les copies.

C. Nouvelle correction des copies par les modérateurs.

Trois principes dominent :

- 1° L'intervention dans la notation doit être discrète que possible ; les corrections éventuelles méritent l'attention ;
- 2° Les correcteurs doivent apprécier les mêmes qualités ; la notation des grades finaux doit être la même ;
- 3° Les échantillons à rediger pour la commission statistique doit être au nombre de 20.

Dans chaque école, on prélève, au hasard, 20 copies.

Soit le cas de 12 lycées, on envoie 240 copies à la commission, laquelle en choisit 20.

La première opération est la notation par les modérateurs à trois points.

- 1° *Sévérité* : son degré est noté sur 3 ;
- 2° *Discrimination* : notation sur 3 ; Révélée par la dispersion des notes.

¹ Exemple emprunté à l'*Examen*.

de modération complet.

À décrire maintenant est, lui-même. À notre connaissance, il n'en existe pas sur la préparation de l'examen, la notation, l'ajustement

Correction des examens internes, le rôle joué jusqu'à présent par

Commission nationale, on procède à un certain nombre d'écoles groupées ; à l'échelon régional qui ne diffère pas fondamentalement ; pour le passage au système de modération proprement dit l'étude technique serait

Il nous faut que quelques écoles de la région aient une expérience volontaire et nous concentrons notre présentation

La correction serait faite qu'aux moments clés de quelques branches principales. Les professeurs d'une même école, devraient utiliser spontanément la même

Il faut aux accords à prendre sur la notation de l'examen. Nous n'y revenons

Il faut à un professeur qui jouera

Il faut dans son école ; la modération inter-écoles.

Il faut pour n'exister pas. Les qualités de la notation :

Il faut l'expérience pédagogique ;

2. Contacts fréquents avec des élèves du niveau et du type d'école concernés par l'examen ;
3. Capacité de formuler clairement ses critères et ses jugements ;
4. Bonté mais fermeté ;
5. Capacité de discuter sans passion, sans créer de tension ;
6. Aptitude à comprendre quelques techniques d'analyse statistique.

Dans chaque école, les questions d'examens sont rédigées avec une grande liberté de mouvement : seuls la grille des objectifs principaux, le nombre et le volume des questions sont à respecter selon la convention prise.

Au début au moins, il est souhaitable de soumettre les questions à la commission de modération qui s'assure de l'unité générale.

B. Les professeurs notent leurs examens.

C. Nouvelle correction d'un échantillon de copies par les modérateurs.

Trois principes dominent le travail :

- 1° L'intervention dans les examens des écoles doit être aussi discrète que possible. Seuls les cas de divergences importantes méritent l'attention.
- 2° Les correcteurs doivent, en gros, être de même sévérité, apprécier les mêmes qualités et être d'accord sur la signification des grades finaux.
- 3° Les échantillons à recorriger doivent être petits, et le travail statistique doit être aussi simple que possible.

Dans chaque école et pour un même type d'examen, on prélève, au hasard, 20 copies corrigées.

Soit le cas de 12 lycées¹. Chacun a délégué un modérateur à la commission, laquelle reçoit donc 12 paquets de 20 copies.

La première opération vise à déterminer l'équivalence des modérateurs à trois points de vue :

- 1° *Sévérité* : son degré est révélé par la moyenne.
- 2° *Discrimination* : notation trop désinvolte ou trop prudente. Révélée par la dispersion ou marge de variation des notes.

¹ Exemple emprunté à l'*Examinations Bulletin* n° 5, Londres, HMSO, 1965.

3° *Conformité* : un même élève est-il classé de la même façon par tous les correcteurs ? Révélée par la corrélation entre deux séries de notes.

Pour vérifier l'accord entre modérateurs, à ces trois points de vue, les douze recorrigent d'abord un même paquet de 20 copies.

Les 20 copies sont réparties, toujours au hasard, en cinq groupes de quatre. Dans chaque groupe, les élèves sont classés par ordre alphabétique.

On fait ensuite les simples opérations suivantes, dans l'ordre où elles figurent dans le tableau :

ETUDE DE L'ACCORD ENTRE LES MODERATEURS

Notes accordées par 12 modérateurs

Nom du candidat	MODERATEURS												Moyenne
	a	b	c	d	e	f	g	h	i	j	k	l	
A	4	4	4	3	4	4	3	1	4	4	4	4	4
B	2	2	2	2	3	2	1	1	2	2	2	2	2
C	6	5	5	5	5	5	5	5	5	5	5	5	5
D	2	1	1	2	2	3	3	2	2	3	2	3	2
Total	14	12	12	12	14	14	12	9	13	14	13	14	13
Marge de variat.	4	4	4	3	3	3	4	4	3	3	3	3	3
E	3	3	3	4	3	5	3	3	3	3	3	3	3
F	1	1	1	1	1	1	1	1	1	1	1	1	1
G	5	6	6	6	5	3	6	6	6	6	6	6	6
H	3	4	4	4	3	4	3	4	4	2	4	4	4
Total	12	14	14	15	12	15	13	14	14	12	14	14	14
Marge de variat.	4	5	5	5	4	4	5	5	5	5	5	5	5
I	2	2	2	2	4	3	2	3	3	3	3	3	3
J	2	1	2	1	3	2	2	2	2	2	2	2	2
K	5	5	5	6	5	5	5	6	5	5	5	4	5
L	1	2	2	1	2	2	2	3	2	2	3	2	2
Total	10	10	11	10	14	12	11	14	12	12	13	11	12
Marge de variat.	4	4	3	5	3	3	3	4	3	3	3	2	3
M	4	2	3	4	3	4	5	5	5	4	5	4	4
N	4	3	4	4	4	5	4	4	4	4	4	4	4
O	5	4	4	3	4	3	4	4	4	5	4	5	4
P	5	6	5	5	5	5	5	5	5	5	5	5	5
Total	18	15	16	16	16	17	18	18	18	18	18	18	17
Marge de variat.	1	4	2	2	2	2	1	1	1	1	1	1	1
Q	1	2	1	1	1	1	1	1	1	1	2	1	1
R	1	1	2	1	1	2	1	1	1	1	1	1	1
S	3	3	2	3	3	2	2	3	3	3	3	3	3
T	2	1	2	1	2	2	1	2	2	2	2	2	2
Total	7	7	7	6	7	7	5	7	7	7	8	7	7
Marge de variat.	2	2	1	2	2	1	1	2	2	2	2	2	2
Total général	61 ²	58	60	59	63	65	59	62	64	63	66	64	63 ¹
Somme des marges	15 ⁴	19	15	17	14	13	14	16	14	14	14	13	14 ³

Différences entre les notes et la moyenne des notes à un

Nom du candidat	a	b	c
A	0	0	0
B	0	0	0
C	1	0	0
D	0	-1	-1
Marge de variat.	1	1	1
E	0	0	0
F	0	0	0
G	-1	0	0
H	-1	0	0
Marge de variat.	1	0	0
I	-1	-1	-1
J	0	-1	0
K	0	0	0
L	-1	0	0
Marge de variat.	1	1	1
M	0	-2	-1
N	0	-1	0
O	1	0	0
P	0	1	0
Marge de variat.	1	3	1
Q	0	1	0
R	0	0	1
S	0	0	-1
T	0	-1	0
Marge de variat.	0	2	2
Somme des marges ⁵	4	7	5

Contrôles.

Les règles suivantes n'ont pas été utilisées dans l'industrie, G. Peaker s'est inspiré de la qualité utilisées dans l'industrie dans le domaine des examens.

Le contrôle ainsi réalisé est très grossier mais non très fin. Aussi, si on ne prend pas en compte un aspect de la correction, on se pose un problème !

1° Sévérité.

L'expérience révèle que la solution est heureusement trouvée, vent porter surtout sur ce point, la solution est heureusement trouvée.

est-il classé de la même façon
révélée par la corrélation entre

modérateurs, à ces trois points
abord un même paquet de 20

, toujours au hasard, en cinq
groupe, les élèves sont classés

s opérations suivantes, dans
bleau :

PRE LES MODERATEURS

r 12 modérateurs

MODERATEURS							Moyenne
f	g	h	i	j	k	l	
4	3	1	4	4	4	4	4
2	1	1	2	2	2	2	2
5	5	5	5	5	5	5	5
3	3	2	2	3	2	3	2
4	12	9	13	14	13	14	13
3	4	4	3	3	3	3	3
5	3	3	3	3	3	3	3
1	1	1	1	1	1	1	1
3	6	6	6	6	6	6	6
4	3	4	4	2	4	4	4
5	13	14	14	12	14	14	14
4	5	5	5	5	5	5	5
3	2	3	3	3	3	3	3
2	2	2	2	2	2	2	2
5	5	6	5	5	5	4	5
2	2	3	2	2	3	2	2
2	11	14	12	12	13	11	12
3	3	4	3	3	3	2	3
4	5	5	5	4	5	4	4
5	4	4	4	4	4	4	4
3	4	4	4	5	4	5	4
5	5	5	5	5	5	5	5
7	18	18	18	18	18	18	17
2	1	1	1	1	1	1	1
	1	1	1	1	2	1	1
	1	1	1	1	1	1	1
	2	3	3	3	3	3	3
	1	2	2	2	2	2	2
	5	7	7	7	8	7	7
	1	2	2	2	2	2	2
5	59	62	64	63	66	64	63 ¹
3	14	16	14	14	14	13	14 ³

Différences entre les notes accordées par chaque modérateur et la moyenne des notes accordées par tous les modérateurs à un même candidat.

Nom du candidat	MODERATEURS											
	a	b	c	d	e	f	g	h	i	j	k	l
A	0	0	0	-1	0	0	-1	-3	0	0	0	0
B	0	0	0	0	1	0	-1	-1	0	0	0	0
C	1	0	0	0	0	0	0	0	0	0	0	0
D	0	-1	-1	0	0	1	1	0	0	1	0	1
Marge de variat.	1	1	1	1	1	1	2	3	0	1	0	1
E	0	0	0	1	0	2	0	0	0	0	0	0
F	0	0	0	0	0	0	0	0	0	0	0	0
G	-1	0	0	0	-1	-1	0	0	0	0	0	0
H	-1	0	0	0	-1	0	-1	0	0	-2	0	0
Marge de variat.	1	0	0	1	1	3	1	0	0	2	0	0
I	-1	-1	-1	-1	1	0	-1	0	0	0	0	0
J	0	-1	0	-1	1	0	0	0	0	0	0	0
K	0	0	0	1	0	0	0	1	0	0	0	-1
L	-1	0	0	-1	0	0	0	1	0	0	1	0
Marge de variat.	1	1	1	2	1	0	1	1	0	0	1	1
M	0	-2	-1	0	-1	0	1	1	1	0	1	0
N	0	-1	0	0	0	1	0	0	0	0	0	0
O	1	0	0	-1	0	-1	0	0	0	1	0	1
P	0	1	0	0	0	0	0	0	0	0	0	0
Marge de variat.	1	3	1	1	1	2	1	1	1	1	1	1
Q	0	1	0	0	0	0	0	0	0	0	1	0
R	0	0	1	0	0	1	0	0	0	0	0	0
S	0	0	-1	0	0	-1	-1	0	0	0	0	0
T	0	-1	0	-1	0	0	-1	0	0	0	0	0
Marge de variat.	0	2	2	1	0	2	1	0	0	0	1	0
Somme des marges ⁵	4	7	5	6	4	8	6	5	1	4	3	3

Contrôles.

Les règles suivantes n'ont rien de magique. Pour les formuler, G. Peaker s'est inspiré des techniques de contrôle de qualité utilisées dans l'industrie. Des expériences très poussées dans le domaine des examens ont permis de les adapter.

Le contrôle ainsi réalisé est raisonnable, expéditif, efficace, mais non très fin. Aussi, si malgré la grossièreté des critères, un aspect de la correction paraît inacceptable, il y a certainement un problème !

1° Sévérité.

L'expérience révèle que les efforts d'harmonisation doivent porter surtout sur ce point. On va le voir, en cas de problème, la solution est heureusement facile.

Règle :

- a) Partir de la moyenne¹ des totaux généraux (ici 63) ;
- b) Ne pas tolérer des écarts au-delà d'une marge centrale de 10 points¹ par rapport à cette moyenne, soit 5 points en plus ou en moins (ici : 58-68).

Constatation :

Tous les totaux généraux sont compris dans cette marge.
Tous les correcteurs sont donc d'une sévérité acceptable.

Solution en cas d'excès de sévérité ou de générosité :
Supposons que l'on se trouve devant les totaux suivants :
75-72-68-66-65-64-60-49-46-44-41-40.

On constate que si on réunit le plus sévère et le plus généreux, puis le deuxième plus sévère et le deuxième plus généreux, etc., les moyennes des couples sont fort proches de la moyenne générale.

Soit :			Moyenne
(1)	40	75	57,5
(2)	41	72	56,5
(3)	44	68	56
(4)	46	66	56
(5)	49	65	57
(6)	60	64	62
Moyenne générale		57,5
Marge de 10 points		52,5-62,5

Seuls, les deux correcteurs du couple n° 6 se situent entre 52,5 et 62,5. Ces deux correcteurs seront autorisés à corriger seuls. Dans les autres cas, les copies devront être vues par les deux membres d'un des groupes constitués, et la moyenne des points sera faite.

2° Discrimination.

Règle :

Le total des marges moyennes (3) ne peut pas être supérieur au double du total des marges d'un correcteur (4) et réciproquement.

Constatation :

Aucun problème.

¹ Les correcteurs notent 20 compositions de 1 à 5. Dix points équivalent à une différence moyenne d'un demi-point par copie pour l'ensemble des 20 copies.

3° Conformité.

Règle :

Le total des marges des différentes épreuves évaluées par un modérateur et par l'ensemble ne doit pas être...

Constatation :

Aucun problème.

Conclusion.

Dans le cas présent, toutes les épreuves de contrôle...

Où en sommes-nous ? De cet échantillon est maintenant notateurs l'ont vu !).

Comme chaque modérateur des opérations ira vite.

D. Nouvelle correction des épreuves.

Les opérations sont pratiquées par le contrôle des modérateurs. Nous ne détaillons pas la présentation de la copie par le modérateur et un seul professeur semble...

¹ Nous montrons p. 167 que l'on peut estimer la corrélation entre deux séries...

aux généraux (ici 63) ;
au-delà d'une marge centrale de
la moyenne, soit 5 points en

pris dans cette marge.
une sévérité acceptable.

de sévérité ou de générosité :
devant les totaux suivants :

le plus sévère et le plus généreux,
le deuxième plus généreux,
et fort proches de la moyenne

	Moyenne
	57,5
	56,5
	56
	56
	57
	62
...	57,5
...	52,5-62,5

l'exemple n° 6 se situent entre 52,5
et sont autorisés à corriger seuls.
Ils devront être vus par les deux
notés, et la moyenne des

ne peut pas être supérieur
à un correcteur (4) et récipro-

1 à 5. Dix points équivalent à une
note pour l'ensemble des 20 copies.

3° Conformité.

Règle :

Le total des marges des différences (5) entre les notes attri-
buées par un modérateur et la moyenne des notes attribuées
par l'ensemble ne doit pas être supérieur à 12¹.

Constatation :

Aucun problème.

Conclusion.

Dans le cas présent, tous les modérateurs ont surmonté
les trois épreuves de contrôle. Ils pourront donc travailler seuls.

Où en sommes-nous ? Des douze échantillons de 20 copies,
un échantillon est maintenant corrigé définitivement (douze
notés l'ont vu !).

Comme chaque modérateur peut travailler seul, la suite
des opérations ira vite.

D. Nouvelle correction des échantillons restants et contrôle.

Les opérations sont pratiquement les mêmes que pour le
contrôle des modérateurs. Nous donnons, toutefois, un exemple
détaillé parce que la présentation plus concise (un seul modé-
rateur et un seul professeur) donne une meilleure vue d'en-
semble.

¹ Nous montrons p. 167 que l'on obtient avec ce repère empirique une bonne
estimation de la corrélation entre deux séries de notes.

**CALCULS POUR LA COMPARAISON
ENTRE UN PROFESSEUR ET UN MODERATEUR
(1 est la meilleure note ; 5 la moins bonne !)**

5 groupes de 4	Elèves choisis au hasard	Note attribuée par				Différence (Modérateur-Professeur)
		Modérateur		Professeur		
1 ^{er} groupe	Henri	Min.	5	5	Min.	0
	Jean		4	4		0 Min.
	Paul	Max.	2	1	Max.	1
	Pierre		3	2		1 Max.
	Marge : Min-Max	3	14	12	4	Marge des différences = 1
2 ^e groupe	André	Max.	2	1	Max.	1 Max.
	Edouard		1	1		0 Min.
	Jules	Min.	1	1	Min.	0
	René		4	4		0
	Marge : Min-Max	3	8	7	3	Marge des différences = 1
3 ^e groupe	Antoine		3	3	Max.	0 Max.
	Camille		5	5		0
	Eugène	Max.	2	3		-1 ¹ Min.
	Jérôme	Min.	6	6	Min.	0
	Marge : Min-Max	4	16	17	3	Marge des différences = 1
4 ^e groupe	Jacques	Min.	5	5	Min.	0
	Laurent	Max.	1	1	Max.	0 Min.
	Martin		4	3		1 Max.
	Victor		4	3		1
	Marge : Min-Max	4	14	12	4	Marge des différences = 1
5 ^e groupe	Bruno	Max.	3	2	Max.	1
	Hugues	Min.	4	2		2 Max.
	Léon		4	5	Min.	-1 Min.
	Simon		3	2		1
	Marge : Min-Max	1	14	11	3	Marge des différences = 3
Les cinq groupes réunis	Total des grades réunies		66 (1)	59 (2)		Total des marges de différences = 7 (5)
		15 (3)		17 (4)		

* On considère que (0) est plus grand que (-1).
166

Contrôles.

1° *Sévérité des correctifs.*

La différence entre totaux est supérieure à 10.

Ici, 66 — 59 = 7 : Acceptable

2° *Discrimination.*

Le total (3) ne peut pas être supérieur à (4) et réciproquement.

Ici, 15 et 17 : Acceptable

3° *Conformité.*

Le total des marges de différences est supérieur à 12.

Ici, = 7 : acceptable¹.

4° *Conclusion.*

Dans ce cas, aucun problème. Les notes du professeur sont donc terminées.

Si les notes du professeur ne sont pas terminées, l'idéal est d'engager une discussion avec le modérateur.

Ce n'est pas toujours facile.

La solution suivante est proposée. On constate un problème, il invite le modérateur à agir indépendamment le même problème. On calcule la moyenne entre les deux séries de notes. On est invité à ajuster toutes les notes du modérateur.

Selon G. Peaker, les notes du professeur deviennent de plus en plus précises. La modération fonctionne, en particulier dans les cas où les notes du professeur sont faibles qu'ils provoquent chaque fois un problème.

¹ Cette façon simple de procéder est valable pour deux séries de notes. On verra d'ailleurs que l'on obtient le même coefficient calculé par la méthode classique.

COMPARAISON
ET UN MODÉRATEUR
(5 la moins bonne !)

attribuée par		Différence (Modérateur- Professeur)	
Professeur			
5 4 1 2	Min. Max.	0 0 1 1	Min. Max.
12	4	Marge des différences = 1	
1 1 1 4	Max. Min.	1 0 0 0	Max. Min.
7	3	Marge des différences = 1	
3 5 3 6	Max. Min.	0 0 -1 ¹ 0	Max. Min.
17	3	Marge des différences = 1	
5 1 3 3	Min. Max.	0 0 1 1	Min. Max.
12	4	Marge des différences = 1	
2 2 5 2	Max. Min.	1 2 -1 1	Max. Min.
11	3	Marge des différences = 3	
59 (2) (4)	17	Total des marges de différences = 7 (5)	

Contrôles.

1° *Sévérité des correcteurs.*

La différence entre total (1) et total (2) ne doit pas être supérieure à 10.

Ici, $66 - 59 = 7$: Acceptable.

2° *Discrimination.*

Le total (3) ne peut pas être supérieur au double du total (4) et réciproquement.

Ici, 15 et 17 : Acceptable.

3° *Conformité.*

Le total des marges des différences (5) ne doit pas être supérieur à 12.

Ici, = 7 : acceptable¹.

4° *Conclusion.*

Dans ce cas, aucun problème ne semble se poser. Les notes du professeur sont donc acceptées et les opérations sont terminées.

Si les notes du professeur ne semblent pas acceptables, l'idéal est d'engager une discussion avec lui afin de rechercher un accord.

Ce n'est pas toujours possible, en particulier faute de temps.

La solution suivante est plus expéditive. Si un modérateur constate un problème, il invite quatre de ses collègues à recorriger indépendamment le même échantillon de vingt copies. On calcule la moyenne entre les cinq modérateurs, et le professeur est invité à ajuster toutes ses notes en fonction de celles des modérateurs.

Selon G. Peaker, les cas nécessitant pareille correction deviennent de plus en plus rares, à mesure que le système de modération fonctionne, en raison des échanges de vues fructueux qu'ils provoquent chaque fois.

¹ Cette façon simple de procéder évite le long calcul de la corrélation entre les deux séries de notes.

On verra d'ailleurs que l'on obtient un coefficient de corrélation très proche du coefficient calculé par la méthode classique en faisant l'opération :

$$\frac{15^2 + 17^2 - 7^2}{2 \times 15 \times 17} = .90$$

cordantes ?

tement de la moyenne (ou du placement des notes de tous les excessives), soit aussi d'une érale, pour améliorer la discri-

uation est la suivante :

Median	Quartile inférieur	Ecart type approximatif
13	11	$\sigma = 4$
13	6	$\sigma = 4$

que le modérateur, mais il leur ajoute 5 points à chaque

l'écart type.

Median	Quartile inférieur	Ecart type approximatif
13	11	$\sigma = 4$
12	9	$\sigma = 5$

er, on corrige d'abord pour le s les élèves.

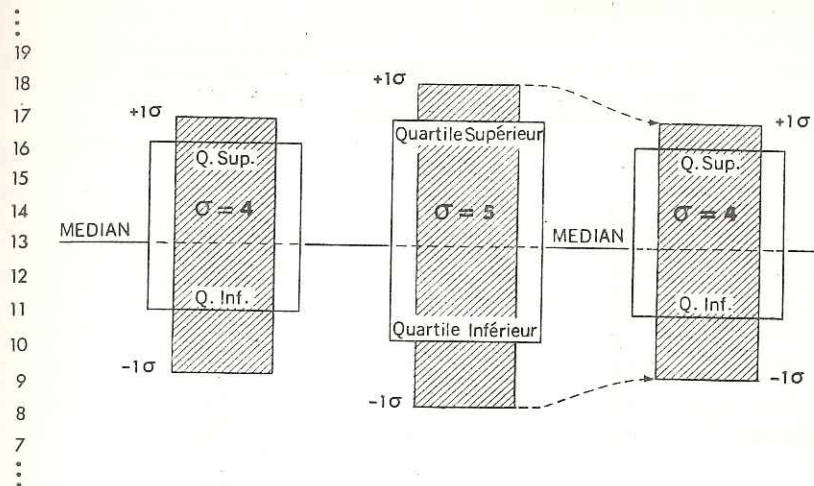
départ du professeur sont :

2 11 10 9 8 7 6 5 ...

lian

183.

Nous allons d'abord montrer par le dessin en quoi consiste l'ajustement de l'écart type.



I. Notes du modérateur.

II. Notes du professeur, augmentées d'un point. Le médian est devenu le même que celui du modérateur. Reste à ajuster σ .

III. Notes du professeur diminuées une seconde fois pour aligner l'écart type sur celui du modérateur.

Pour l'opération III, l'ajustement en fonction du nouvel écart type est un peu plus compliqué que pour le médian. La confection d'une table de conversion facilitera les opérations.

Pour trouver la nouvelle note correspondant à $+ 1 \sigma$, il suffit d'ajouter le nouveau sigma au nouveau médian, soit $13 + 4 = 17$. La nouvelle note correspondant à $- 1 \sigma = 13 - 4 = 9$, etc.

Pour les notes correspondant à des fractions de sigma, on calcule d'abord l'écart par rapport au médian, on multiplie par 4/5 et on ajoute le résultat au médian. L'exemple suivant va éclairer cette phase.

	Notes de départ du professeur	Notes ajustées en fonction du nouveau médian	Second ajustement en fonction du nouvel écart type (σ)
	.	.	.
	19	.	.
	18	.	.
+ 1 σ	17	18	17 *
	16	17	16 **
	15	16	15 ***
	14	15	15 ****
	13	14	14
Médian	12	13	13
	11	12	12
	10	11	11
	9	10	11
	8	9	10
- 1 σ	7	8	9
	6	.	.
	5	.	.
	.	.	.
	.	.	.

Comment a-t-on trouvé ces chiffres ?

* Nouveau médian : 13

Nouvel écart type : 4

$$13 + 4 = 17$$

** Dans la 2^e colonne, 17 est à 4 points du médian.

$$\text{Multiplier 4 par } \frac{4}{5} = 3,2$$

$$\text{Médian} + 3,2 = 16,2 \text{ arrondi à } 16$$

$$*** 16 - 13 = 3 ; 3 \times \frac{4}{5} = 2,4 ; 13 + 2,4 = 15,4 \text{ arrondi à } 15$$

$$**** 15 - 13 = 2 ; 2 \times \frac{4}{5} = 1,6 ; 13 + 1,6 = 14,6 \text{ arrondi à } 15$$

Problème ¹

On veut classer les élèves à la fin de leurs études. A sera le grade supérieur et B sera le grade inférieur et va...

On veut tenir compte de ces notes selon les mêmes échelles à...

— travail de l'année = T.A.

— travaux pratiques = T.P.

— test régional = T.R.

Le test régional se voit attribuer des notes sur les deux autres éléments. D'où...

T.A.
T.P.
T.R.

¹ D'après D. MATHER *et al.*, o.c., pp.

à des fractions de sigma,
port au médian, on multiplie
le médian. L'exemple suivant

Notes en du médian	Second ajustement en fonction du nouvel écart type (σ)
.	.
.	.
.	.
.	.
.	.
.	17 *
.	16 **
.	15 ***
.	15 ****
.	14
.	13
.	12
.	11
.	11
.	10
.	9
.	.
.	.
.	.

notes ?

$$13 + 4 = 17$$

4 points du médian.

16

$$13 + 2,4 = 15,4 \text{ arrondi à } 15$$

$$13 + 1,6 = 14,6 \text{ arrondi à } 15$$

VI. LA NOTE DE FIN D'ANNEE.

Travail de l'année + travaux pratiques + test.

Problème ¹

On veut classer les élèves en cinq groupes de mérite à la fin de leurs études. A sera le grade supérieur et vaudra 1, E sera le grade inférieur et vaudra 5.

On veut tenir compte de trois éléments, évalués chacun selon les mêmes échelles à cinq degrés :

- travail de l'année = T.A.
- travaux pratiques = T.P.
- test régional = T.R.

Le test régional se voit attribuer autant d'importance que les deux autres éléments. D'où la pondération :

$$T.A. = \frac{1}{4}$$

$$T.P. = \frac{1}{4}$$

$$T.R. = \frac{1}{2}$$

¹ D'après D. MATHER *et al.*, o.c., pp. 149-154.

Voici le tableau général des grades et le tableau des résultats après pondération et ajustement final.

ATTRIBUTION DU GRADE FINAL

Elève N°	Test régional T.R		Evaluation par l'école		Pondération			Total	Grade final
	Score	Note	T.A.	T.P.	T.R $\times \frac{1}{2}$	T.A $\times \frac{1}{4}$	T.P $\times \frac{1}{4}$		
1	59	2	3	3	1	1	<i>$\frac{3}{4}$</i>	<i>$2\frac{3}{4}$</i>	3
2	77	1	1	1	<i>$\frac{1}{2}$</i>	<i>$\frac{1}{4}$</i>	<i>$\frac{1}{4}$</i>	1	1
3	51	3	4	3	<i>$1\frac{1}{2}$</i>	1	<i>$\frac{3}{4}$</i>	<i>$3\frac{1}{4}$</i>	3
4	12	6	5	5	3	<i>$1\frac{1}{4}$</i>	<i>$1\frac{1}{4}$</i>	<i>$5\frac{1}{2}$</i>	6
5	53	3	3	4	<i>$1\frac{1}{2}$</i>	1	1	<i>$3\frac{1}{2}$</i>	4
6	40	4	2	5	2	<i>$\frac{3}{4}$</i>	<i>$1\frac{1}{4}$</i>	4	4
7	66	1	2	2	<i>$\frac{1}{2}$</i>	<i>$\frac{1}{2}$</i>	<i>$\frac{1}{2}$</i>	<i>$1\frac{1}{2}$</i>	1
8	60	2	3	4	1	<i>$\frac{3}{4}$</i>	1	<i>$2\frac{3}{4}$</i>	3
9	38	5	4	5	<i>$2\frac{1}{2}$</i>	<i>$1\frac{1}{4}$</i>	<i>$1\frac{1}{4}$</i>	5	5
10	70	1	2	2	<i>$\frac{1}{2}$</i>	<i>$\frac{1}{2}$</i>	<i>$\frac{1}{2}$</i>	<i>$1\frac{1}{2}$</i>	1
11	56	2	2	3	1	<i>$\frac{1}{2}$</i>	<i>$\frac{3}{4}$</i>	<i>$2\frac{1}{4}$</i>	2
12	69	1	2	2	<i>$\frac{1}{2}$</i>	<i>$\frac{1}{2}$</i>	<i>$\frac{1}{2}$</i>	<i>$1\frac{1}{2}$</i>	1
13	44	4	4	5	2	1	<i>$1\frac{1}{4}$</i>	<i>$4\frac{1}{4}$</i>	4
14	64	2	1	2	1	<i>$\frac{1}{4}$</i>	<i>$\frac{1}{2}$</i>	<i>$1\frac{3}{4}$</i>	2
15	19	6	5	5	3	<i>$1\frac{1}{4}$</i>	<i>$1\frac{1}{4}$</i>	<i>$5\frac{1}{2}$</i>	6
16	49	3	2	4	<i>$1\frac{1}{2}$</i>	<i>$\frac{3}{4}$</i>	1	<i>$3\frac{1}{4}$</i>	3
17	54	3	2	1	<i>$1\frac{1}{2}$</i>	<i>$\frac{1}{2}$</i>	<i>$\frac{1}{4}$</i>	<i>$2\frac{1}{4}$</i>	2
18	47	4	1	2	2	<i>$\frac{1}{4}$</i>	<i>$\frac{1}{2}$</i>	<i>$2\frac{3}{4}$</i>	3
19	52	3	3	1	<i>$1\frac{1}{2}$</i>	<i>$\frac{3}{4}$</i>	<i>$\frac{1}{4}$</i>	<i>$2\frac{1}{2}$</i>	2
20	48	3	3	3	<i>$1\frac{1}{2}$</i>	<i>$\frac{3}{4}$</i>	<i>$\frac{3}{4}$</i>	3	3
21	50	3	1	2	<i>$1\frac{1}{2}$</i>	<i>$\frac{1}{4}$</i>	<i>$\frac{1}{2}$</i>	<i>$2\frac{1}{4}$</i>	2
22	24	6	3	4	3	<i>$\frac{3}{4}$</i>	1	<i>$4\frac{3}{4}$</i>	5
23	61	2	4	2	1	1	<i>$\frac{1}{2}$</i>	<i>$2\frac{1}{2}$</i>	2
24	57	2	1	3	1	<i>$\frac{1}{4}$</i>	<i>$\frac{3}{4}$</i>	2	2
25	42	4	3	5	2	<i>$\frac{3}{4}$</i>	<i>$1\frac{1}{4}$</i>	4	4
26	35	5	5	2	<i>$2\frac{1}{2}$</i>	<i>$1\frac{1}{4}$</i>	<i>$\frac{3}{4}$</i>	<i>$4\frac{1}{2}$</i>	4
27	45	4	4	1	2	1	<i>$\frac{1}{4}$</i>	<i>$3\frac{1}{4}$</i>	3
28	41	4	2	4	2	<i>$\frac{3}{4}$</i>	1	<i>$3\frac{3}{4}$</i>	4
29	27	5	5	5	<i>$2\frac{1}{2}$</i>	<i>$1\frac{1}{4}$</i>	<i>$1\frac{1}{4}$</i>	5	5
30	43	4	3	2	2	1	<i>$\frac{1}{2}$</i>	<i>$3\frac{1}{2}$</i>	4
31	67	1	1	1	<i>$\frac{1}{2}$</i>	<i>$\frac{1}{4}$</i>	<i>$\frac{1}{4}$</i>	1	1
32	31	5	5	3	<i>$2\frac{1}{2}$</i>	<i>$1\frac{1}{4}$</i>	1	<i>$4\frac{3}{4}$</i>	5
33	72	1	2	2	<i>$\frac{1}{2}$</i>	<i>$\frac{1}{2}$</i>	<i>$\frac{1}{2}$</i>	<i>$1\frac{1}{2}$</i>	1

Les notes pondérées en italiques seront ajustées en fonction du test régional (T.R.).

La lecture du tableau
trois notes pondérées : 1,
final 3 *. L'élève 3 : $1\frac{1}{2}$ +

Pourquoi un dernier
grade final ? Parce que
nouvelle concentration ven
tel que, dans une échelle
risque de se retrouver au m

L'exemple fictif suivant

E L E V E S	
Pierre	
Paul	
Jean	
Roger	
Henri	
Dispersion	

Si l'on se reporte au ta
les grades pondérés, on ren
tout), alors qu'on n'en retro
Total. Les notes d'excellent
grades.

Le procédé suivant as
tive entre les élèves d'une
ment sur le niveau régional
étant reconnu comme l'étal

* Les chiffres en italiques dans
l'accord de l'école, ajusté la note du
dans la région.

ades et le tableau des résultats final.

GRADE FINAL

Pondération			Total	Grade final
T.R × 1/2	T.A × 1/4	T.P × 1/4		
1	1	3/4	2 ³ / ₄	3
1/2	1/4	1/4	1	1
1 1/2	1	3/4	3 ¹ / ₄	3
3	1 1/4	1 1/4	5 ¹ / ₂	6
1 1/2	1	1	3 ¹ / ₂	4
2	3/4	1 1/4	4	4
1/2	1/2	1/2	1 1/2	1
1	3/4	1	2 ³ / ₄	3
2 1/2	1 1/4	1 1/4	5	5
1/2	1/2	1/2	1 1/2	1
1	1/2	3/4	2 1/4	2
1/2	1/2	1/2	1 1/2	1
2	1	1 1/4	4 1/4	4
1	1/4	1/2	1 3/4	2
3	1 1/4	1 1/4	5 1/2	6
1 1/2	3/4	1	3 1/4	3
1 1/2	1/2	1/4	2 1/4	2
2	1/4	1/2	2 3/4	3
1 1/2	3/4	1/4	2 1/2	2
1 1/2	3/4	3/4	3	3
1 1/2	1/4	1/2	2 1/4	2
3	3/4	1	4 3/4	5
1	1	1/2	2 1/2	2
1	1/4	3/4	2	2
2	3/4	1 1/4	4	4
2 1/2	1 1/4	3/4	4 1/2	4
2	1	1/4	3 1/4	3
2	3/4	1	3 3/4	4
2 1/2	1 1/4	1 1/4	5	5
2	1	1/2	3 1/2	4
1/2	1/4	1/4	1	1
2 1/2	1 1/4	1	4 3/4	5
1/2	1/2	1/2	1 1/2	1

es seront ajustées en fonc-

La lecture du tableau nous montre que l'élève 1 obtient les trois notes pondérées : 1, 1 et 3/4. Soit au total 2³/₄ ; grade final 3*. L'élève 3 : 1 1/2 + 1 + 3/4 = 3 1/4 ; grade final 3.

Pourquoi un dernier ajustement au moment d'attribuer le grade final ? Parce que l'addition des grades provoque une nouvelle concentration vers la moyenne ; le rétrécissement est tel que, dans une échelle finale à 5 degrés, toute la population risque de se retrouver au milieu.

L'exemple fictif suivant illustre ce phénomène.

ELEVES	GRADE		
	Test R.	Trav. année	Moyenne
Pierre	1	5	3
Paul	2	4	3
Jean	3	3	3
Roger	4	2	3
Henri	5	1	3
Dispersion	4	4	0

Si l'on se reporte au tableau p. 172, on observe que, parmi les grades pondérés, on rencontre un grand nombre de 1 (17 en tout), alors qu'on n'en retrouve plus que deux dans la colonne Total. Les notes d'excellence ont été noyées par l'addition des grades.

Le procédé suivant assure une meilleure justice distributive entre les élèves d'une même classe et un meilleur alignement sur le niveau régional révélé par le test régional, celui-ci étant reconnu comme l'étalon le plus sûr.

* Les chiffres en italiques dans le tableau indiquent que le modérateur a, avec l'accord de l'école, ajusté la note du T.A. afin de l'harmoniser avec le niveau moyen dans la région.

On part du tableau suivant.

TEST REGIONAL		TOTAL DES GRADES PONDERES		GRADE FINAL	
Grade (1)	N. d'élèves ayant obtenu ce grade (2)	Grade (3)	N. d'élèves ayant obtenu ce grade (4)	N. d'élèves auxquels il est attribué (5)	Grade (9)
1		1			1
		1 1/4			
		1 1/2			
		1 3/4			
2		2			2
		2 1/4			
		2 1/2			
		2 3/4			
3		3			3
		3 1/4			
		3 1/2			
		3 3/4			
4		4			4
		4 1/4			
		4 1/2			
		4 3/4			
5		5			5
Au-delà		Au-delà			Non classé

Opérations.

1. Pointer dans la colonne (2) le nombre d'élèves et faire le total.
2. Commencer à pointer dans la colonne (4) et s'arrêter quand le total égale celui de la colonne (2). On pratique la première coupure à cet endroit. Afin de ne pas désavantager certains élèves, on dépasse, au besoin, le nombre de la colonne (3), de façon à épuiser le niveau où l'on s'est arrêté. On verra, par exemple, dans le tableau ci-dessous, qu'on accorde le grade 2 à 7 élèves et non à 6, parce que 2 d'entre eux ont obtenu 2 1/2. On ne pouvait évidemment attribuer un grade 2 à l'un et un grade 3 à l'autre.

TEST REGIONAL			DES GRAD
GRADE (1)	N. D'ÉLÈVES AYANT OBTENU CE GRADE (2)	GRADE (3)	N. O
1	ИИ I 6	1	
		1 1/4	
		1 1/2	
		1 3/4	
2	ИИ I 6	2	
		2 1/4	
		2 1/2	
		2 3/4	
3	ИИ II 7	3	
		3 1/4	
		3 1/2	
		3 3/4	
4	ИИ II 7	4	
		4 1/4	
		4 1/2	
		4 3/4	
5	IIII 4	5	
AU-DELÀ	III 3	AU-DELÀ	

Conclusion.

Aucun des systèmes décrits ne mettraient d'améliorer notre système.

Un premier choix s'opère entre la sélection ou comparaison. Ce choix attire notre attention.

Nous ne nous prononçons pas sur le choix. La décision appartient aux pédagogues et aux enseignants.

Un effort intense d'expérimentation est en cours à différents niveaux : petits groupes de projets régionaux et nationales en collaboration avec les laboratoires sociaux et les laboratoires de recherche des universités.

A mesure que ces travaux se poursuivent, il est probable qu'émergeront de nouvelles méthodes et aussi des objectifs pédagogiques, toutefois, la recherche pédagogique régionale nous y reviendrons encore !

GRADES ES	GRADE FINAL	
	N. d'élèves auxquels il est attribué (5)	Grade (9)
N. d'élèves obtenus grade (4)		1
		2
		3
		4
		5
		Non classé

nombre d'élèves et faire le

colonne (4) et s'arrêter quand
(2). On pratique la première
ne pas désavantager certains
le nombre de la colonne (3),
on s'est arrêté.

le tableau ci-dessous, qu'on
s et non à 6, parce que 2
ne pouvait évidemment attri-
de 3 à l'autre.

TEST REGIONAL			TOTAL DES GRADES PONDÉRÉS		GRADE FINAL	
GRADE (1)	N. D'ÉLÈVES AYANT OBTENU CE GRADE (2)	GRADE (3)	N. D'ÉLÈVES AYANT OBTENU CE GRADE (4)	N. D'ÉLÈVES AUXQUELS IL EST ATTRIBUÉ (5)	GRADE (6)	
1	JHT I 6	1	// 2	6	1	
		1 1/4 1 1/2 1 3/4	//// 4			
2	JHT I 6	2	/ 1 / 1			7
		2 1/4 2 1/2 2 3/4	/// 3 // 2 /// 3			
3	JHT II 7	3	/ 1 /// 3	7	3	
		3 1/4 3 1/2 3 3/4	// 2 / 1 // 2			
4	JHT II 7	4	/ 1 / 1			7
		4 1/4 4 1/2 4 3/4	// 2			
5	IIII 4	5	// 2	4	5	
AU-DELÀ	III 3	AU-DELÀ				
				2	NON CLASSÉS	

Conclusion.

Aucun des systèmes décrits n'est parfait, mais tous permettraient d'améliorer notre système traditionnel de notation.

Un premier choix s'opérera en fonction du but poursuivi : sélection ou comparaison. Ce second aspect a surtout retenu notre attention.

Nous ne nous prononçons pas en faveur d'un système particulier. La décision appartient aux autorités politiques et pédagogiques et aux enseignants.

Un effort intense d'expérimentation doit être fait à tous les niveaux : petits groupes de professeurs, recherches régionales et nationales en collaboration avec les centres psycho-médico-sociaux et les laboratoires de pédagogie expérimentale des universités.

A mesure que ces travaux avanceront, on verra probablement émerger un nouveau système tenant compte de nos traditions et aussi des objectifs particuliers à notre pays. Presque immédiatement, toutefois, la nécessité de puissants centres de recherche pédagogique régionaux se fera de nouveau sentir. Nous y reviendrons encore !

CINQUIEM

LE MYTHE DE
DE C

CINQUIEME PARTIE

LE MYTHE DE LA COURBE
DE GAUSS

LE DANGEREUX MYTHE

Dans les sciences humaines joue un rôle considérable, par la répartition de bien des aptitudes, les moyens abondent, mais les grands et les nains sont rares.

La courbe de Gauss est celle qui préside à notre naissance, elle est le résultat d'un grand nombre de facteurs indépendants sur un individu.

Comme les tests de mesure de personnalité ou des performances, il est naturel qu'ils soient étalonnés : en gros, 70 % de moyens, 2 % d'excellents, 2 % de

Au cours de la construction de la classe, ce sont les questions qui seraient posées à nos sujets. Le but poursuivi est d'attribuer la place qui leur revient. Bref, il s'agit d'organiser une situation qui occupera nécessairement la place

C'est pourquoi beaucoup de connaissances sont d'actualité.

Dans sa classe, l'enseignant a un idéal différent. Son idéal n'est pas de lire, à calculer et, de fait, de transmettre toutes les connaissances par la société. *Instruire n'est pas transmettre.* C'est s'efforcer que tous réussissent, c'est la courbe de Gauss prise comme

Les conséquences pédagogiques sont particulièrement importantes.

LE DANGEREUX MYTHE DE LA COURBE DE GAUSS.

Dans les sciences humaines, la courbe en cloche de Gauss joue un rôle considérable, parce qu'elle est l'image même de la répartition de bien des aptitudes et des qualités : les individus moyens abondent, mais les génies et les idiots, les géants et les nains sont rares.

La courbe de Gauss est soit le reflet de la loi du hasard qui préside à notre naissance, soit la résultante de l'influence d'un grand nombre de facteurs agissant de façon plus ou moins indépendante sur un individu ou un objet.

Comme les tests mesurent souvent des aptitudes, des traits de personnalité ou des performances de vastes populations, il est naturel qu'ils soient étalonnés selon la répartition gaussienne : en gros, 70 % de moyens, 13 % de bons, 13 % de médiocres, 2 % d'excellents, 2 % de très mauvais.

Au cours de la construction de tels tests, on élimine notamment les questions qui seraient réussies par trop ou trop peu de sujets. Le but poursuivi est de classer chacun, de lui attribuer la place qui lui revient dans un groupe nombreux. Bref, il s'agit d'organiser une sorte de concours, où le plus fort occupera nécessairement la première place.

C'est pourquoi beaucoup de tests d'aptitudes ou d'inventaires de connaissances sont d'excellents *instruments de sélection*.

Dans sa classe, l'enseignant poursuit un objectif totalement différent. Son idéal n'est-il pas que *tous* les élèves apprennent à lire, à calculer et, de façon générale, à maîtriser parfaitement toutes les connaissances jugées nécessaires ou utiles par la société. *Instruire n'est pas sélectionner*. Au contraire ! C'est s'efforcer que *tous* réussissent. *C'est donc lutter contre la courbe de Gauss prise comme modèle de sélection*.

Les conséquences pédagogiques de ces observations sont particulièrement importantes.

CHAPITRE 1.

EVOLUTION DE LA COURBE DES CONNAISSANCES.

Quand, dans notre système de classes rigides, un maître reçoit, le jour de la rentrée scolaire, un groupe d'élèves qu'il ne connaît pas, il se trouve normalement devant... deux courbes : l'une représente la distribution des aptitudes et l'autre celle du savoir.

I. La courbe des aptitudes.

Dans son acception habituelle — que nous retenons provisoirement — le mot *aptitude* désigne des caractéristiques, innées ou acquises, considérées comme symptomatiques de la capacité d'un individu à acquérir un niveau de compétence plus ou moins élevé, dans un domaine déterminé.

Dans l'enseignement non étroitement spécialisé (il va jusqu'à passé vingt ans pour bon nombre de nos élèves), la largeur même de l'éventail des connaissances et des capacités à faire acquérir (mathématiques, langues, sciences naturelles, arts, ...) rend impossible la sélection très rigoureuse selon une aptitude particulière.

Aussi, jusqu'à un niveau fort avancé de la scolarité, les aptitudes des élèves restent-elles distribuées au hasard.

Dans ces conditions, un professeur de mathématiques, par exemple, qui mesurerait les aptitudes spéciales pour sa branche obtiendrait normalement une distribution gaussienne, ou — vu le nombre probablement peu élevé d'élèves — une ébauche de celle-ci.

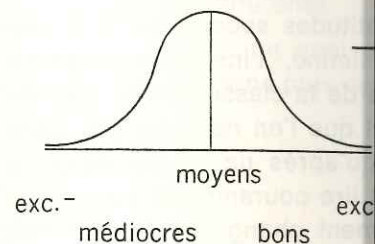
Les professeurs n'ont d'ailleurs guère besoin de recourir à des tests pour connaître cette situation ; ils savent d'expérience que les moyens seront majorité et que les excellents sont rares...

Toutefois, une singularité se manifeste dans les esprits. On considère les aptitudes comme pronostique de la réussite et on fixe le niveau de l'enseignement ; le pronostic se vérifie : il sera permis aux meilleurs de s'élever et de laisser la chance aux « médiocres ».

Insistons-y, la difficulté est fixée, la moyenne des aptitudes du groupe est la même, la moyenne de difficulté objective est la même.

Dans ces conditions, si on considère à toute la classe, il est normal que les connaissances acquises en fin d'année ressemblent à une courbe gaussienne.

Courbe des aptitudes en début d'année



La vocation de l'enseignant

II. La courbe des connaissances

Revenons au premier jour de l'année pour considérer les aptitudes des élèves et leurs connaissances.

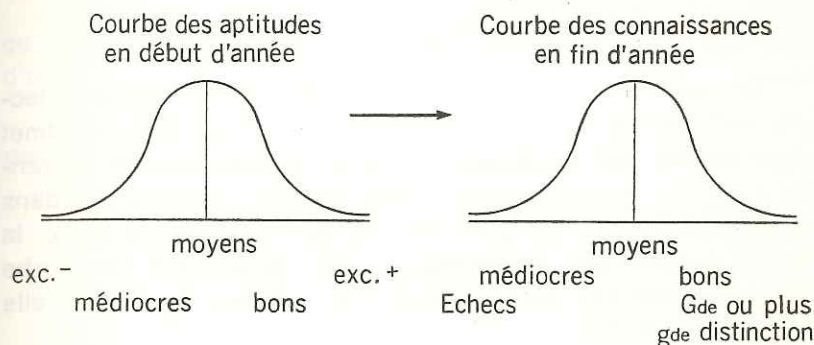
En toute logique, le rôle de l'enseignement est de faire acquérir à tous les individus formés la même quantité et qualité de connaissances. Néanmoins, le système de l'enseignement repose sur l'hypothèse que tous les élèves ne sont pas au même niveau. Sinon, comment expliquer le même enseignement à

¹ Ainsi s'expliquent les différences de résultats entre les classes et les régions, dont nous avons

Toutefois, une singulière distorsion se produit généralement dans les esprits. On considère cette répartition des aptitudes comme pronostique des résultats scolaires en fin d'année et on fixe le niveau de l'enseignement de telle façon que ce pronostic se vérifie : il sera « moyennement » difficile tout en permettant aux meilleurs de s'épanouir et en laissant une mince chance aux « médiocres ».

Insistons-y, la difficulté « moyenne » est déterminée par la moyenne des aptitudes du groupe considéré et non par une moyenne de difficulté *objective* des notions à enseigner¹.

Dans ces conditions, si le professeur fait *le même cours* à toute la classe, il est normal que la *courbe des connaissances* acquises en fin d'année respecte, à son tour, la distribution gaussienne.



La vocation de l'enseignement est-elle ainsi respectée ?

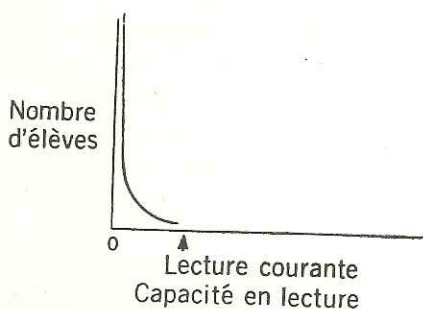
II. La courbe des connaissances.

Revenons au premier jour de l'année scolaire et, au lieu de considérer les aptitudes spécifiques des élèves, examinons leurs connaissances.

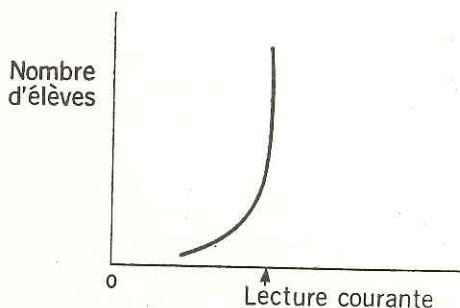
En toute logique, le rôle du professeur est de susciter l'apprentissage de connaissances *nouvelles*. Certes, imaginer que tous les individus formant une classe possèdent exactement la même quantité et qualité de connaissances est utopique. Néanmoins, le système de classes fixes que nous pratiquons repose sur l'hypothèse que *tous* se trouvent approximativement au même niveau. Sinon, comment oserions-nous encore dispenser le même enseignement à chacun ?

¹ Ainsi s'expliquent les différences considérables du niveau moyen selon les classes et les régions, dont nous avons parlé dans la première partie de ce livre.

Et en réalité ? Prenons le cas de l'entrée en première primaire. La majorité des enfants ne savent pas lire ; quelques-uns sont en bonne voie ; deux ou trois lisent déjà couramment. A ce moment, la courbe de la capacité en lecture épouse, en gros, la forme de la lettre *i*.

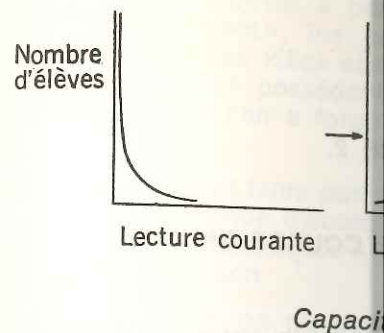


Or, bien que la courbe des aptitudes spécifiques à la lecture soit fort probablement gaussienne, l'instituteur n'admet pas d'emblée que, seule, une partie de la classe pourra apprendre à lire, au contraire. Pour autant que l'on ne verse pas dans le perfectionnisme, on peut dire qu'après un an ou deux, la grande majorité des élèves sauront lire couramment. La courbe des connaissances aura complètement changé de forme : elle ressemblera à un *j*.



Entre la courbe en *i* et la courbe en *j*, il est probable qu'un moment a existé où les mieux doués ont avancé le plus vite, où les moins doués ont traîné et où les moyens se sont situés entre les deux.

L'évolution est donc sché

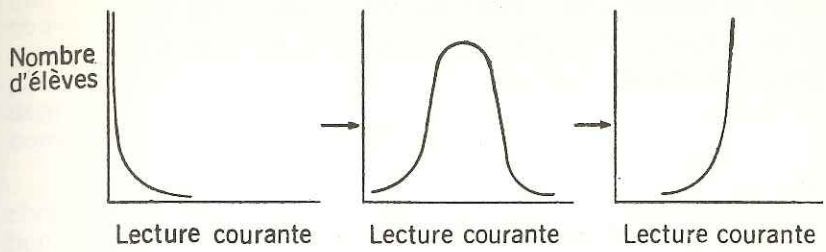


Dans le cas de la lecture, ou les responsables de l'ens d'une évolution différente.

Mais, à partir de quel ni gence n'est-elle donc plus de

de l'entrée en première pri-
vent pas lire ; quelques-uns
is lisent déjà couramment.
acité en lecture épouse, en

L'évolution est donc schématiquement la suivante :



Capacité en lecture.

Dans le cas de la lecture, on ne conçoit pas que les parents ou les responsables de l'enseignement puissent se contenter d'une évolution différente.

Mais, à partir de quel niveau de la scolarité, pareille exigence n'est-elle donc plus de mise ?

tudes spécifiques à la lec-
enne, l'instituteur n'admet
de la classe pourra appren-
que l'on ne verse pas dans
'après un an ou deux, la
lire couramment. La courbe
ent changé de forme : elle

en j, il est probable qu'un
s ont avancé le plus vite,
les moyens se sont situés

CHAPITRE 2.

UNE PEDAGOGIE DE COURBE EN J.

Un examen où les résultats se distribuent selon la courbe en cloche de Gauss se prête bien à la sélection, au concours.

Or, pour trois raisons historiques principales, l'esprit de concours a imprégné notre enseignement pendant des siècles :

- 1° Pour des raisons socio-économiques, une partie seulement de la population scolarisable avait accès à l'école. Il y a cent ans, les familles modestes faisaient la première sélection en ne laissant finir l'école primaire qu'aux plus doués de leurs enfants, c'est-à-dire à ceux qui réussissaient le mieux leurs examens. Les bourses d'études, parcimonieusement distribuées, se gagnaient en concours.
- 2° Jusqu'à ces derniers temps, — disons avant l'ère de l'ordinateur, — tous les pays industrialisés disposaient d'un énorme surplus de matière grise. On exploitait donc celle qui s'offrait à meilleur marché et avec le plus de facilité, c'est-à-dire que l'on se souciait surtout d'identifier les mieux doués. Même pour les enfants fortunés, l'enseignement secondaire jouait un rôle sélectif.
- 3° Les connaissances psychologiques et pédagogiques étaient rudimentaires. Les maîtres n'étaient donc pas en mesure d'appliquer des traitements fins aux élèves éprouvant des difficultés d'apprentissage.
D'ailleurs, aujourd'hui encore, on fait souvent répéter une année à l'élève qui n'a pas été capable des performances minima exigées pour le passage. Autrement dit, au lieu d'appliquer des remèdes particuliers, on se contente de placer de nouveau l'élève dans les conditions (même professeur, même méthode) dans lesquelles l'échec s'est produit.

Chaque année scolaire étant considérée comme un filtre pour la suivante, une certaine quantité d'échecs paraissait donc normale. Bref, la répartition gaussienne semblait satisfaisante.

Mais le souci constant causé, progressivement, une telle situation. On en est arrivé à penser que « comptent vraiment », les choses « véritable jeu des idées abstraites » par une minorité possédant un degré élevé. Et l'on a forgé une conséquence.

Les élèves et leurs parents se débattaient avec ces choses sans grande discussion « latin-mathématiques » dans cette direction.

Assurément, on ne peut pas attendre de chacun un virtuose de la littérature. Mais où se situe le degré l'accès à un savoir essentiel dépend des aptitudes moyennes.

Par nature, chacun de nous a des domaines. Il ne manque plus ou moins réfractaires à l'observation fréquemment qu'à faire redire les explications, de sorte que pour trouver une forme d'expressions de notions de mathématiques d'ailleurs, les plus sibles sont bel et bien maîtrisées.

Pour autant qu'ils y ont accès, les élèves moyens (c'est-à-dire, vu les études antérieures, probablement au-dessous du niveau scolaire d'un niveau donné) qu'on ne l'imagine.

Dans cette perspective, la proposition de J. Carroll prend la quantité de temps demandée pour la matière »¹.

L'implication de cette proposition leur alloue le temps nécessaire normalement dans une classe pour obtenir un très bon résultat.

¹ J. CARROLL, A Model of School, 723-733.

Mais le souci constant de la promotion des plus aptes a causé, progressivement, une déformation pédagogique plus subtile. On en est arrivé à penser que les connaissances « qui comptent vraiment », les connaissances « approfondies », le « véritable jeu des idées abstraites », ne sont assimilables que par une minorité possédant des aptitudes spécifiques à un degré élevé. Et l'on a forgé les méthodes d'enseignement en conséquence.

Les élèves et leurs parents acceptent d'ailleurs cet état de choses sans grande discussion. On s'inscrit rarement en section « latin-mathématiques » sans se sentir spécialement doué dans cette direction.

Assurément, on ne peut ambitionner de faire indifféremment de chacun un virtuose de la mathématique, du piano ou de la littérature. Mais où se situe la limite ? A partir de quel degré l'accès à un savoir est-il impossible à ceux qui possèdent des aptitudes moyennes, voire médiocres ?

Par nature, chacun de nous est médiocrement doué en bien des domaines. Il ne manque, par exemple, pas d'intellectuels plus ou moins réfractaires à la mathématique. Pourtant, on observe fréquemment qu'à force de vouloir, de persévérer, de faire redire les explications, de changer de manuel ou de maître pour trouver une forme d'enseignement qui convienne, des notions de mathématiques d'abord considérées comme inaccessibles sont bel et bien maîtrisées par certains.

Pour autant qu'ils y consacrent le temps nécessaire, les élèves moyens (c'est-à-dire, vu la sélection déjà opérée par les études antérieures, probablement plus de 80 % de la population scolaire d'un niveau donné) peuvent aller beaucoup plus loin qu'on ne l'imagine.

Dans cette perspective, la nouvelle définition de l'aptitude, proposée par J. Carroll prend toute sa valeur : « L'aptitude est la quantité de temps demandée par le *learner* pour dominer une matière »¹.

L'implication de cette conception est considérable : si on leur alloue le temps nécessaire, tous les élèves se trouvant normalement dans une classe pourraient arriver à un bon, voire un très bon résultat.

¹ J. CARROLL, A Model of School Learning, *Teachers College Record*, 1963, 64 : 723-733.

CHAPITRE 3.

LA THEORIE DE L'EVALUATION FORMATIVE.

L'expression *évaluation formative* — il l'oppose à *évaluation sommative* — a été forgée par Michael Scriven¹.

L'*évaluation sommative* (S. Roller parle dans ce cas de « norme psychotechnique ») nous est familière. Pour interpréter le score obtenu à un test classique d'inventaire de connaissances ou d'intelligence, on le situe dans une distribution statistique : la performance d'un individu est jugée par référence à celles d'autrui. De même, on classe souvent encore les élèves entre eux selon l'ordre croissant ou décroissant de leurs résultats scolaires, et c'est d'après la place ainsi occupée que bien des parents apprécient le travail de leurs enfants.

Or, dans les deux cas, le résultat est essentiellement relatif. Que le groupe de référence varie de composition et le résultat apparaît sous un autre jour.

Une simple différence d'âge peut aussi changer considérablement la face des choses. Dans bien des normes de tests de connaissances d'usage courant, un an d'âge ou une année scolaire en plus ou en moins suffisent pour qu'une même performance soit considérée comme médiocre ou bonne.

Que la connaissance soit acquise ou non n'a donc pas été la préoccupation première des constructeurs de ces tests, mais bien à quelle vitesse cette acquisition s'est réalisée. En nous reportant à la définition de J. Carroll, on évalue donc l'aptitude au lieu d'évaluer le contenu de l'apprentissage.

Les partisans de l'*évaluation formative* prennent le contre-pied de cette conception.

¹ M. SCRIVEN, *The Methodology of Evaluation*, in R. STAKER, (Ed.), *Perspectives of Curriculum Evaluation*, Chicago, Rand McNally, 1967.

Voir aussi : B.S. BLOOM, *Learning for Mastery*, in *Evaluation Comment*, 1968, 2. Notre exposé de la méthode de l'évaluation formative s'appuie directement sur cette publication. Plusieurs résultats d'expériences où la théorie du *mastery learning* a été appliquée sont publiés par J. BLOCK, B. BLOOM, HASTINGS (1970).

Puisque l'élève vient à l'...
n'est-il pas de le situer dans

Imaginons qu'une analyse révèle qu'avec toutes ses nuances du participe passé, employé cent cas différents. Pour chaque maîtrise.

Selon le niveau scolaire, alors être fixé. Dans ce contexte de nature.

L'élève est noté en fonction du parcours dans l'acquisition.

En outre, si l'on ambitionne jusqu'à un niveau de connaissance maîtrise totale, de la notion de enseigner est elle-même remise d'appliquer indifféremment à pendant une même durée.

Que deviendrait l'enseignant ?

Trois problèmes cruciaux :

1. Comment jalonner l'ascension ?
2. Comment conduire l'élève ?
3. Où se situe la limite pratique de la courbe en *j* ?

A ma connaissance, il n'y a pas de réponses complètes à ces questions. Peut-être la normale en sciences naturelles nous devons apprendre à l'accepter. Notre discussion aboutit à des recherches à entreprendre ou à

I. Jalonner l'ascension du savoir

En voyage, pour déterminer le but, deux conditions doivent être réunies : savoir où l'on est et où l'on va, la carte indiquant clairement le chemin.

De même, en éducation, il faut définir les objectifs à atteindre et déterminer les conditions particulières qui y conduiront. L'élève a-t-il la matière ou non à des apprentissages ?

Puisque l'élève vient à l'école pour apprendre, l'important n'est-il pas de le situer dans l'ascension du savoir ?

Imaginons qu'une analyse scientifique rigoureuse nous révèle qu'avec toutes ses nuances et ses complications, l'accord du participe passé, employé avec l'auxiliaire avoir, présente cent cas différents. Pour chacun, on peut définir des critères de maîtrise.

Selon le niveau scolaire, le nombre de cas à dominer peut alors être fixé. Dans ce contexte, l'évaluation scolaire change de nature.

L'élève est noté en fonction d'un critère objectif : le chemin parcouru dans l'acquisition.

En outre, si l'on ambitionne de conduire tous les élèves jusqu'à un niveau de connaissances minimum, sinon jusqu'à la maîtrise totale, de la notion ou de la technique, la façon d'enseigner est elle-même remise en cause : il n'est plus possible d'appliquer indifféremment à tous une seule et même méthode pendant une même durée.

Que deviendrait l'enseignement dans cette perspective ?

Trois problèmes cruciaux se posent :

1. Comment jalonner l'ascension du savoir ?
2. Comment conduire l'élève ?
3. Où se situe la limite pratique de cette pédagogie de la courbe en *j* ?

A ma connaissance, il n'existe pas encore de réponses complètes à ces questions. Pareille imprécision du savoir semble normale en sciences naturelles ou en médecine. Nous devons apprendre à l'accepter aussi dans les sciences de l'éducation. Notre discussion aboutit donc maintes fois sur des recherches à entreprendre ou à continuer.

I. Jalonner l'ascension du savoir.

En voyage, pour déterminer à quelle distance on se trouve du but, deux conditions doivent être remplies : d'une part, savoir où l'on est et où l'on va, et, d'autre part, disposer d'une carte indiquant clairement le chemin.

De même, en éducation, nous devons définir les objectifs à atteindre et déterminer avec précision les apprentissages particuliers qui y conduiront. Le problème varie selon que l'on a affaire ou non à des apprentissages de base. Pour ces der-

niers, il faut découvrir l'enchaînement « critique » des matières, c'est-à-dire celui où l'une n'est accessible que si la précédente est assimilée. Pour cette raison, l'apprentissage des connaissances et des techniques de base doit, en dernière analyse, être linéaire, tandis que les acquisitions et les applications qui vont au-delà voient s'ouvrir devant elles des voies de plus en plus nombreuses.

Par exemple, quelle que soit la méthode d'enseignement, il n'est pas possible d'appliquer complètement une règle de trois, sans avoir — notamment — la notion de la multiplication et de la division. A un niveau plus élevé, comment faire du calcul intégral sans savoir ce qu'est une fonction ? Mais pareilles propositions sont encore trop vagues. Quels sont exactement les apprentissages nécessaires et suffisants pour pouvoir assimiler la règle de trois ? Et, parmi eux, lesquels sont critiques par rapport aux autres ?

La question est redoutable. Et il serait naïf de croire que la réflexion et les tâtonnements qui ont présidé à l'élaboration de nos meilleurs manuels scolaires l'ont résolue. Certes, le bon enchaînement des chapitres constitue une amorce de solution, mais il suffit d'essayer de programmer la moindre notion, ne fût-ce qu'en système skinnérien, pour découvrir toutes les ignorances, toute l'imprécision des méthodes d'enseignement actuelles.

Cependant, c'est précisément dans les efforts de programmation entrepris par des chercheurs et des enseignants de plus en plus nombreux que gît une des grandes sources de progrès.

Assurément, les sciences déductives sont le terrain d'élection de la méthode qui s'esquisse. Dussions-nous nous y borner, nous n'aurions pas accompli œuvre négligeable. Mais la limite n'apparaît pas si vite. Là où l'induction domine dans l'élaboration du savoir, la déduction réapparaît pour son application. D'ailleurs, l'apprentissage inductif aussi peut être systématiquement guidé.

II. Guider l'élève.

Une fois clairement défini le chemin qui conduit à un apprentissage, il faut y engager l'élève. Pour les acquisitions fondamentales, nous avons, en outre, décidé que *tous* devraient, en principe, arriver au but.

Dans ce cas, l'évaluation fréquente des progrès est essentielle. D'où la nécessité d'*exercices de maîtrise* et de *tests diagnostics* portant sur des matières très limitées et utilisés par les maîtres eux-mêmes.

Ces instruments — que s
gnants et des chercheurs per
suffisant — indiquent où l'élève
et où il éprouve des difficultés.

Les remèdes sont multipliés
pratique pédagogique quotidienne

- a) Indiquer de façon précise
- b) Le travail par sous-groupe
élèves rencontre la ou les r
avantage à consacrer une
groupe, les autres travailla
- c) Si le groupe compte maxim
gère que chacun d'eux soi
gène de deux ou trois. Co
concurrence, mais luttent
l'entraide joue.
- d) Dans les écoles comptant p
branche, un élève devrait a
explication, voire une cour
professeur que le sien.
- e) Varier la présentation : me
des manuels expliquant la
permettre le recours à l'
moyens audio-visuels ; vari
- f) En général, ne pas fixer à
mêmes choses.

Pareille pratique paie. B.
conduite dans cette ligne où plu
la meilleure note en fin d'anné
gogie de la courbe en *j*.

Peut-il toujours en être ainsi ?

III. Limite pratique du système

Même s'il est vrai que tou
capables d'atteindre un haut
qu'on leur laisse un temps su
d'une certaine limite, la durée

Comment permettre de pas
l'enseignement secondaire ! Ce
fois, pas remettre le système en
élèves en fonction de leurs apt
vation et d'orientation, introdui
gnement secondaire de la plupa
bilité.

ment « critique » des matières, accessible que si la précédente l'apprentissage des connaissances et les applications qui vont des voies de plus en plus

la méthode d'enseignement, il complètement une règle de trois, notion de la multiplication et levé, comment faire du calcul ne fonction ? Mais pareilles es. Quels sont exactement les isants pour pouvoir assimiler, lesquels sont critiques par

il serait naïf de croire que la ont présidé à l'élaboration de ont résolue. Certes, le bon itue une amorce de solution, mmer la moindre notion, ne our découvrir toutes les igno méthodes d'enseignement

dans les efforts de program- s et des enseignants de plus grandes sources de progrès.

actives sont le terrain d'élec- Dussions-nous nous y borner, e négligeable. Mais la limite ction domine dans l'élabora- paraît pour son application. aussi peut être systématique-

chemin qui conduit à un élève. Pour les acquisitions e, décidé que tous devraient,

ente des progrès est essen- es de maîtrise et de tests res très limitées et utilisés

Ces instruments — que seule la collaboration des enseignants et des chercheurs permettra de construire en nombre suffisant — indiquent où l'élève en est (ce qui lui donne sa note) et où il éprouve des difficultés.

Les remèdes sont multiples et doivent entrer dans notre pratique pédagogique quotidienne :

- a) Indiquer de façon précise la partie du cours à réétudier.
- b) Le travail par sous-groupes : si un groupe de plus de trois élèves rencontre la ou les mêmes difficultés, le professeur a avantage à consacrer une partie du temps de la leçon à ce groupe, les autres travaillant indépendamment.
- c) Si le groupe compte maximum trois élèves, B. Bloom suggère que chacun d'eux soit intégré dans un groupe hétérogène de deux ou trois. Comme les élèves ne sont plus en concurrence, mais luttent pour la maîtrise d'une matière, l'entraide joue.
- d) Dans les écoles comptant plusieurs professeurs d'une même branche, un élève devrait avoir la faculté de demander une explication, voire une courte leçon particulière à un autre professeur que le sien.
- e) Varier la présentation : mettre à la disposition des élèves des manuels expliquant la notion de différentes façons ; permettre le recours à l'enseignement programmé, aux moyens audio-visuels ; varier le niveau d'abstraction.
- f) En général, ne pas fixer à tous le même temps pour les mêmes choses.

Pareille pratique paie. B. Bloom rapporte une expérience conduite dans cette ligne où plus de 80 % des élèves ont obtenu la meilleure note en fin d'année. C'est le triomphe de la pédagogie de la courbe en *j*.

Peut-il toujours en être ainsi ? Rien ne permet de l'affirmer.

III. Limite pratique du système.

Même s'il est vrai que tous les élèves de nos écoles sont capables d'atteindre un haut degré de maîtrise pour autant qu'on leur laisse un temps suffisant, il est évident qu'au-delà d'une certaine limite, la durée devient prohibitive.

Comment permettre de passer vingt ans au lieu de six dans l'enseignement secondaire ! Cette constatation ne doit, toutefois, pas remettre le système en cause. Le tout est d'orienter les élèves en fonction de leurs aptitudes réelles. Le cycle d'observation et d'orientation, introduit progressivement dans l'enseignement secondaire de la plupart des pays, en donne la possibilité.

Mais quand, après ce triage, l'étudiant est admis dans une classe, il doit normalement pouvoir apprendre l'essentiel de *tout* ce qu'on y enseigne.

IV. Le système d'enseignement mis en cause.

La pédagogie de la courbe en *j* appelle le nivellement par le haut ; il serait néanmoins chimérique d'imaginer qu'elle permet le nivellement par le... génie.

Nous l'avons vu, moins de cinq pour cent de la population possèdent des aptitudes *exceptionnelles*. Elles doivent être développées au maximum, tant par respect pour la personne que dans l'intérêt de la communauté.

Un nouveau système d'enseignement permet à la fois d'instruire chacun à l'allure convenable et de laisser s'épanouir les surdoués : c'est l'école sans classe, où l'on travaille par groupes homogénéisés selon les aptitudes spécifiques *pour une seule branche*¹.

Dans les conditions actuelles, la classe est un carcan. Pourquoi un élève de six ans qui sait lire en entrant en première primaire ne pourrait-il pas participer aux exercices de lecture de la deuxième année ? Pourquoi tel élève de quatrième de l'enseignement secondaire ne pourrait-il pas suivre un cours de mathématique du niveau de la 3^e ou de la 2^e ?

Dans nos vieilles écoles primaires de campagne, où un seul instituteur se voyait parfois confier les six années, il était commun de voir de tels déplacements. Dans des écoles secondaires qui, actuellement, comptent plusieurs classes de même âge, il est possible de travailler à quatre ou cinq niveaux d'aptitudes différents, au moins dans quelques branches principales. Pareil système fonctionne parfaitement, depuis de nombreuses années, dans des écoles comme la New Trier High School, que nous avons décrite en détail dès 1961².

Prévenons immédiatement l'objection financière. Une école à groupes homogènes spécifiques ne coûte pas plus cher qu'une autre, au contraire. Si l'on ajoute le gain de temps, de talent et la raréfaction des échecs que le système permet, on peut affirmer sans crainte de se tromper que le bénéfice est considérable...

¹ Système à ne pas confondre avec le *streaming* qui consiste à constituer des classes homogènes à l'aide de tests d'aptitudes générales ou, pire encore, en fonction des résultats scolaires globaux. On commet ainsi une double erreur scientifique : dans la grande majorité des cas, un même sujet est inégalement doué pour les diverses branches du programme ; de plus, homogénéiser globalement fait baisser le rendement de l'ensemble. Dans ce système, on aboutit d'ailleurs souvent à une semi-ségrégation sociale. Tous ces inconvénients ont été si bien mis en évidence que le *streaming* est interdit par la loi en Suède.

² *L'individualisation de l'enseignement dans une école multilatérale*, in Education, mai 1961.

étudiant est admis dans une
pour apprendre l'essentiel de

s en cause.

j appelle le nivellement par
rique d'imaginer qu'elle per-

pour cent de la population
elles. Elles doivent être dé-
spect pour la personne que

ement permet à la fois d'in-
et de laisser s'épanouir les
, où l'on travaille par grou-
udes spécifiques pour une

la classe est un carcan.
t lire en entrant en première
er aux exercices de lecture
tel élève de quatrième de
ait-il pas suivre un cours de
u de la 2^e ?

aires de campagne, où un
nfier les six années, il était
nts. Dans des écoles secon-
plusieurs classes de même
atre ou cinq niveaux d'apti-
lques branches principales.
ent, depuis de nombreuses
New Trier High School, que
61².

ection financière. Une école
ne coûte pas plus cher
ajoute le gain de temps, de
que le système permet, on
omper que le bénéfice est

aming qui consiste à constituer des
générales ou, pire encore, en fonction
une double erreur scientifique : dans
inégalement doué pour les diverses
globalement fait baisser le rendement
leurs souvent à une semi-ségrégation
mis en évidence que le streaming est

une école multilatérale, in Education,

CONCLUSIONS ET RECOMMANDATIONS

Les conclusions partielles nous avançons dans cette ét pourtant loin d'avoir épuisé l apporté des solutions définitive

Sortir des routines irréflé les grandes théories pédagogic de l'éducation, découvre tant remplir que l'on risque d'être ragement !

Au risque de laisser nos l pourtant pas voulu voiler la d croire à l'existence d'une do conduire à une évaluation p moyens. Vraisemblablement, e mesure rigoureuse des traits h jours impossible : le sort nous âmes !

Que l'évaluation continue le processus d'enseignement indissociable, n'en supprime d'une docimastique, c'est-à-dir de notation, autrement dit end tion systématique du rendemen

Par ailleurs, l'exigence d'é teur peut regretter pour des ra ne scolaire : elle correspond notre civilisation.

Peut-être en sera-t-il un j continue à s'intellectualiser, e nous ne voyons toutefois pas complètement aux examens et

Notre position docimologi dans son désir de concilier le ques nouvelles, de nature surt niable et fondamental de l'analy

Pour nous limiter à un se nous ne pouvons concevoir que venant en fin de chaque année prises seulement en fonction d choix multiple se prêtant à la n l'introduction de ce type d'épre pour le contrôle objectif et app il ne suffit pas. L'évaluation du et la rencontre ultime du maître les nuances que l'approche qua

Les conclusions partielles se sont imposées à mesure que nous avançons dans cette étude déjà longue. Nous sommes pourtant loin d'avoir épuisé le sujet et, encore plus, d'avoir apporté des solutions définitives.

Sortir des routines irréflechies et ambitionner de traduire les grandes théories pédagogiques dans la pratique quotidienne de l'éducation, découvre tant d'écueils, tant de conditions à remplir que l'on risque d'être envahi par le doute et le découragement !

Au risque de laisser nos lecteurs insatisfaits, nous n'avons pourtant pas voulu voiler la difficulté, et moins encore laisser croire à l'existence d'une docimologie achevée, capable de conduire à une évaluation parfaite, si on lui en donne les moyens. Vraisemblablement, et heureusement sans doute, la mesure rigoureuse des traits humains les plus fins restera toujours impossible : le sort nous garde de la machine à peser les âmes !

Que l'évaluation continue s'insère fonctionnellement dans le processus d'enseignement et d'apprentissage dont elle est indissociable, n'en supprime pas, pour autant, la nécessité d'une docimastique, c'est-à-dire d'une technique d'examen et de notation, autrement dit encore, d'une technique d'observation systématique du rendement immédiat ou différé.

Par ailleurs, l'exigence d'évaluation comparée, que l'éducateur peut regretter pour des raisons idéales, dépasse le domaine scolaire : elle correspond à un caractère fondamental de notre civilisation.

Peut-être en sera-t-il un jour autrement. Si notre culture continue à s'intellectualiser, et donc à croître en complexité, nous ne voyons toutefois pas comment elle pourrait renoncer complètement aux examens et aux concours.

Notre position docimologique est résolument éclectique dans son désir de concilier les avantages certains de techniques nouvelles, de nature surtout statistique, et l'apport indéniable et fondamental de l'analyse qualitative.

Pour nous limiter à un seul exemple de cet éclectisme, nous ne pouvons concevoir que les décisions de passage intervenant en fin de chaque année d'études universitaires soient prises seulement en fonction des résultats obtenus à un test à choix multiple se prêtant à la notation automatique. Assurément, l'introduction de ce type d'épreuves est hautement souhaitable pour le contrôle objectif et approfondi des connaissances, mais il ne suffit pas. L'évaluation du travail de l'année doit intervenir et la rencontre ultime du maître et de son élève apporter toutes les nuances que l'approche quantitative a laissées dans l'ombre.

De même que l'introduction des machines dans l'industrie a permis à l'intelligence de prendre le pas sur la force musculaire et la routine avilissante, de même le contrôle automatique de la connaissance des faits, des méthodes et des techniques rend possible un examen final réellement centré sur les processus mentaux et les traits de personnalité les plus nobles. Quelle importance respective accordera-t-on aux trois types d'évaluation ? La décision appartient, dans ce cas, à la communauté académique, éclairée par les recherches docimologiques.

Un certain nombre de *recommandations* paraissent s'imposer.

En raison même de nos ignorances et de la grande complexité de l'évaluation, il est d'abord urgent qu'un *réseau de centres de recherche en éducation* couvre le pays entier et soit mis au service du système scolaire. Les fonctions de ces centres sont apparues au cours de notre analyse : recherches portant sur les problèmes fondamentaux et développement des méthodes, des techniques et des instruments nécessaires. Ces centres ne doivent pas être greffés en corps étrangers dans l'ensemble de nos institutions éducatives, mais bien travailler en étroite collaboration avec elles.

Plus particulièrement, *chaque université* devrait créer en son sein un *bureau d'études des problèmes d'enseignement et d'examens*. On ne peut concevoir que l'institution scientifique par excellence exclue de la rigueur de son analyse une de ses deux raisons d'être : l'éducation.

La *définition des objectifs* généraux et spéciaux, fondement de tout l'édifice de l'enseignement et de l'évaluation, requiert le travail en commun d'un grand nombre de chercheurs et de praticiens de toutes les disciplines. *Des commissions spécialisées* permettront la coopération nécessaire et la communication, — dans les deux sens, — entre la base et le sommet de la hiérarchie pédagogique.

La réforme des examens et des modes d'évaluation entraînera sans doute la disparition définitive du calcul du pourcentage, presque universellement répandu dans nos écoles jusqu'à ces temps derniers. Quelles que soient ses faiblesses, l'échelle des pourcentages offrait cependant l'avantage d'être familière à tous et chacun avait au moins l'illusion de la comprendre. Actuellement, les échelles les plus diverses sont essayées dans nos écoles. *Une unification du système de notation* et *une campagne d'information* sont nécessaires dans le plus bref délai, faute de quoi la communication deviendra

malaisée, sinon impossible dans différentes écoles, entre le moment de prendre une décision et l'expertise de la docimologie en sciences humaines dev

Un vaste effort d'information des enseignants appelés à participer à des d'évaluation. Le moyen de l'organisation d'expérience des examens, la notation seraient, par exemple, inconnus à l'introduction de l'enseignement, et dans la réforme de l'évaluation, toutes les écoles y

Des commissions de travail seront petit à petit constituées

Une action trop hâtive d'évaluation éminemment louable ne doit pas décevoir ou d'égaler les esprits sous l'autorité de docimolo

Œuvre humaine par excellence, la mesure de la qualité que dans la mesure de la connaissance, fondément les idées nouvelles mêmes.

machines dans l'industrie
le pas sur la force musculaire
le contrôle automatique
méthodes et des techniques
ement centré sur les pro-
personnalité les plus nobles.
rdera-t-on aux trois types
t, dans ce cas, à la com-
les recherches docimolo-

ndations paraissent s'impo-

nces et de la grande com-
d urgent qu'un réseau de
ouvre le pays entier et soit
Les fonctions de ces cen-
e analyse : recherches por-
ux et développement des
truments nécessaires. Ces
en corps étrangers dans
ves, mais bien travailler en

université devrait créer en
problèmes d'enseignement et
ue l'institution scientifique
de son analyse une de ses

aux et spéciaux, fondement
de l'évaluation, requiert le
e de chercheurs et de pra-
commissions spécialisées
e et la communication, —
et le sommet de la hiéar-

modes d'évaluation entraî-
ive du calcul du pourcen-
du dans nos écoles jus-
ue soient ses faiblesses,
pendant l'avantage d'être
moins l'illusion de la com-
les plus diverses sont
ation du système de nota-
sont nécessaires dans le
communication deviendra

malaisée, sinon impossible, entre l'école et la famille, entre les différentes écoles, entre les écoles et les autorités. Pourtant, avant de prendre une décision dont l'importance est patente, les experts de la docimologie et, plus généralement, de la mesure en sciences humaines devraient être entendus.

Un vaste effort d'information sera aussi entrepris auprès des enseignants appelés à changer profondément leurs habitudes d'évaluation. Le moyen le plus efficace semble résider dans l'organisation d'expériences locales portant sur la préparation des examens, la notation et la modération. Ces expériences seraient, par exemple, incorporées dans les travaux nécessaires à l'introduction de l'enseignement secondaire rénové, en Belgique, et dans la réforme de la formation des maîtres ; progressivement, toutes les écoles y participeraient.

Des commissions de modération locales et régionales seront petit à petit constituées.

Une action trop hâtive risque de mettre en péril une rénovation éminemment louable. On se gardera, en particulier, de décevoir ou d'égarer les enseignants en tentant de les placer sous l'autorité de docimologistes improvisés.

Œuvre humaine par excellence, l'éducation ne gagnera en qualité que dans la mesure où les éducateurs accepteront profondément les idées nouvelles et tenteront de se dépasser eux-mêmes.

Faint, illegible text on the left page, possibly bleed-through from the reverse side.

Faint, illegible text on the right page, possibly bleed-through from the reverse side.

AN

ANNEXES

ETUDE COMPAREE D
PRESENTEE SELON LA
ET SELON
PAR QUESTION

*Exposé de la question*¹ (M)

- 1° a) Donner la liste des causes
b) Décrire les lésions anatomiques de l'hypoglycémie.

Pour permettre aux correcteurs de participer dans l'appréciation, la liste de l'ensemble des correcteurs :

- A. Valeur égale pour les deux questions
B. Pour obtenir une note de 10, il faut qu'au moins 50% des correcteurs aient répondu que dans la liste les causes

- 1) Excès d'insuline ;
- 2) Tumeur fonctionnelle de l'pancréas ;
- 3) Influence endocrinienne : hypoparathyroïdisme, hyperadrenocorticisme, hyperadrenophyse et des surrénales ;
- 4) Maladie du foie empêchant la synthèse du glucose (nécrose aiguë du foie, cirrhose, maladie de Von Gierke).

- 2° A quoi reconnaît-on la différence entre une seule crise aiguë d'hypoglycémie et une hypoglycémie chronique ?

¹ MOORE, Robert A., *Methods of Medical Education*, janvier 1954, vol. 29, n° 1, p. 10-11. Traduction empruntée au rapport français.

**ETUDE COMPAREE D'UNE QUESTION D'EXAMEN
PRESENTEE SELON LA METHODE TRADITIONNELLE
ET SELON LA METHODE
PAR QUESTIONS A CHOIX MULTIPLE.**

*Exposé de la question*¹ (Méthode traditionnelle).

- 1° a) Donner la liste des causes d'hypoglycémie ;
b) Décrire les lésions anatomo-pathologiques résultant de l'hypoglycémie.

Pour permettre aux correcteurs un certain degré d'uniformité dans l'appréciation, la grille suivante a été établie par l'ensemble des correcteurs :

- A. Valeur égale pour les deux parties de la question ;
B. Pour obtenir une note de 75/100, le candidat doit avoir indiqué dans la liste les causes :
- 1) Excès d'insuline ;
 - 2) Tumeur fonctionnelle des îlots de Langerhans ;
 - 3) Influence endocrinienne par hypofonctionnement de l'hypophyse et des surrénales ;
 - 4) Maladie du foie empêchant la mise en réserve du glycogène (nécrose aiguë) ou la libération du glycogène (maladie de Von Gierke).
- 2° A quoi reconnaît-on la différence des lésions causées par une seule crise aiguë d'hypoglycémie de celles causées par une hypoglycémie chronique ?

¹ MOORE, Robert A., *Methods of Examining Students in Medicine*, in « Journal of Medical Education », janvier 1954, vol. 29, n° 1.
Traduction empruntée au rapport français sur la réforme des examens de médecine, o.c.

3° Décrire les modifications visibles au niveau du cerveau.

C. Pour obtenir une note allant de 75 à 90/100, le candidat doit avoir indiqué :

- soit une liste montrant qu'il a compris le mécanisme d'action de chacune des causes,
- soit une liste comprenant d'autres causes, mais en indiquant qu'il comprend que les mécanismes d'homéostasie ont habituellement tendance à corriger l'hypoglycémie, dans les cas tels que :

- 1) Ingesta diminués ;
- 2) Absorption perturbée ;
- 3) Métabolisme augmenté comme dans l'hyperthyroïdisme ;
- 4) Infection grave ;
- 5) Surmenage physique ;
- 6) Glycosurie rénale.

D. Pour obtenir une note supérieure à 90/100, le candidat doit avoir indiqué :

- une liste logique des quatre causes majeures et des six causes mineures, en plus d'un type idiopathique,
- une différenciation de l'hypoglycémie aiguë et chronique par une description détaillée des lésions cérébrales.

Exposé de la question (Méthode par questions à choix multiple).

Voici comment une même question pourrait être présentée selon la nouvelle méthode.

1° On peut déterminer si le candidat est capable de reconnaître les quatre causes les plus importantes et les différencier des moins importantes dans la question suivante :

INSTRUCTIONS. Quatre des cinq phrases numérotées sont communes à l'un des trois troubles fonctionnels de la liste alphabétique (A.B.C.).

Indiquez celle qui est l'exception et le trouble fonctionnel commun aux quatre phrases restantes.

- | | |
|---|---|
| 1. Excès d'insuline | A. Hypoglycémie cliniquement décelable |
| 2. Tumeur fonctionnelle des îlots de Langerhans | B. Hyperglycémie cliniquement décelable |
| 3. Glycosurie rénale | C. Glycosurie cliniquement décelable |
| 4. Hypofonctionnement hypophysaire | |
| 5. Maladie de Von Gierke | |

Si le candidat répond correct, 1, 2, 4 et 5 peuvent produire l'hypoglycémie, ce qui n'est pas le cas de la maladie de Von Gierke. La possession de connaissances suffisantes pour reconnaître la relation de cause à effet dans ce cas prouve la possession de connaissances suffisantes pour reconnaître la relation de cause à effet dans les autres cas.

2° Si l'on veut savoir ce que sont les hormones qui commandent la sécrétion de l'insuline, on peut poser la question suivante :

INSTRUCTIONS. Chacune des phrases numérotées est posée de deux parties : la première énonce le fait, la seconde énonce la relation de cause à effet.

Indiquez pour chacune des phrases numérotées A, B, C, D ou E, selon ce qui est vrai.

A. L'exposé du fait est correct, la relation de cause à effet est correcte.

B. L'exposé du fait est correct, la relation de cause à effet est incorrecte.

C. L'exposé du fait est incorrect, la relation de cause à effet est correcte.

D. L'exposé du fait est incorrect, la relation de cause à effet est incorrecte.

E. L'exposé du fait est correct, la relation de cause à effet est correcte.

1. Le taux du sucre est élevé.

(A) L'exposé du fait est correct, la relation de cause à effet est correcte.

2. Le diagnostic anatomique de la maladie de Von Gierke est hyperinsulinisme.

(E) L'exposé du fait est correct, la relation de cause à effet est incorrecte.

3. Les malades atteints de Von Gierke ont de grandes chances d'avoir un hyperinsulinisme.

(C) L'exposé du fait est correct, la relation de cause à effet est correcte.

4. Les malades atteints de Von Gierke sentent une diminution de l'appétit parce que, dans cet état, le glycogène est mis en réserve par le foie.

(E) L'exposé du fait est correct, la relation de cause à effet est incorrecte.

5. Une hypoglycémie de Von Gierke n'est pas suivie de séquestration de lipides cellulaires produites par le foie.

es au niveau du cerveau.

75 à 90/100, le candidat

l a compris le mécanisme
uses,

d'autres causes, mais en
que les mécanismes d'ho-
ent tendance à corriger
as tels que :

comme dans l'hyperthyroï-

re à 90/100, le candidat

causes majeures et des
d'un type idiopathique,
glycémie aiguë et chroni-
ée des lésions cérébrales.

uestions à choix multiple).

on pourrait être présentée

est capable de reconnaître
rtantes et les différencier
estion suivante :

phrases numérotées sont
oubles fonctionnels de la

n et le trouble fonctionnel
stantes.

Hypoglycémie clinique-
ment décelable

Hyperglycémie clinique-
ment décelable

Glycosurie cliniquement
décelable

Si le candidat répond correctement, il montre qu'il sait que 1, 2, 4 et 5 peuvent produire une hypoglycémie cliniquement décelable, que ce n'est pas le cas de 3, et qu'aucune combinaison des quatre cas parmi les cinq ne peut être associée ni à l'hyperglycémie, ni à la glycosurie. En d'autres termes, la possession de connaissances positives et négatives est prouvée.

2° Si l'on veut savoir ce que sait le candidat sur les phénomènes qui commandent l'équilibre du niveau du sucre sanguin et de l'insuline, on peut poser la question suivante :

INSTRUCTIONS. Chacune des phrases suivantes est composée de deux parties : l'exposé d'un fait et la raison de ce fait.

Indiquez pour chacune des phrases numérotées la lettre A, B, C, D ou E, selon que :

- A. L'exposé du fait et sa raison sont vrais et ont une relation de cause à effet.
 - B. L'exposé du fait et sa raison sont vrais, mais n'ont pas de relation de cause à effet.
 - C. L'exposé du fait est vrai, mais la raison est fausse.
 - D. L'exposé du fait est faux, mais la raison est un fait ou un principe accepté.
 - E. L'exposé du fait et la raison sont faux.
1. Le taux du sucre sanguin tombe brutalement après hépatectomie parce que le glycogène contenu dans le foie est la source principale du sucre sanguin.
(A)
 2. Le diagnostic anatomo-pathologique d'un adénome des îlots de Langerhans du pancréas implique que la maladie était hyperinsulinique parce que tous les adénomes des îlots sont fonctionnels et secrètent de l'insuline.
(E)
 3. Les malades atteints d'hyperthyroïdisme ont toutes chances d'avoir une hypoglycémie parce qu'il existe un hyperinsulinisme associé.
(C)
 4. Les malades atteints de la maladie de Von Gierke présentent une diminution du taux du sucre sanguin parce que, dans cette maladie, le glycogène n'est pas mis en réserve par le foie.
(C)
 5. Une hypoglycémie durant depuis plusieurs mois n'est pas suivie de séquelles parce que les modifications cellulaires produites par l'hypoglycémie sont réversibles.
(E)

Si le candidat répond correctement à cette série, cela montre qu'il connaît :

- a) que le glycogène hépatique est la source principale permettant le maintien du taux du sucre sanguin ;
- b) que toutes les tumeurs des îlots de Langerhans ne sont pas fonctionnelles ;
- c) que les malades atteints d'hyperthyroïdie ont une hypoglycémie, mais que la cause n'est pas un hyperinsulinisme associé ;
- d) que les malades atteints de la maladie de Von Gierke ont une hypoglycémie, mais qu'elle n'est pas due au fait qu'il n'y a pas de glycogène dans le foie ;
- e) qu'il y a des séquelles après hypoglycémie chronique et que les modifications cellulaires résultantes ne sont pas réversibles.

3° On peut déterminer si le candidat a quelques notions simples concernant les modifications au niveau du cerveau dans la question suivante :

INSTRUCTIONS. Chacun des exposés incomplets (numérotés) est suivi de cinq compléments au choix. Indiquez, dans chaque cas, le complément qui convient le mieux.

1. Les modifications anatomiques secondaires à une hypoglycémie chronique sont mises en évidence le plus souvent dans :
 - a) la rate
 - b) le rein
 - c) les surrénales
 - d) le cerveau
 - e) la thyroïde
2. Parmi les modifications provoquées par l'hypoglycémie chronique au niveau du cerveau, les plus importantes sont sur :
 - a) les neurones
 - b) les astrocytes
 - c) les cellules de l'épendyme
 - d) les cellules oligodendrogliques
 - e) les cellules microgliales

3. Parmi les altérations provoquées par l'hypoglycémie :
 - a) le déplacement de la courbe de la glycémie
 - b) la disparition de la glycémie
 - c) l'altération de la glycémie
 - d) l'éclatement du foie
 - e) la fusion de mitochondries

4. Parmi les modifications au niveau du cerveau provoquées par l'hypoglycémie chronique :
 - a) l'hydrocéphalie
 - b) l'épaississement de la paroi des vaisseaux
 - c) la destruction de cellules de Nissl
 - d) la prolifération des cellules de Nissl
 - e) la prolifération des cellules de Schwann

Si le candidat répond correctement à cette série, cela montre qu'il sait que les altérations au niveau du cerveau atteintes par l'hypoglycémie chronique provoquent la destruction de cellules de Nissl des neurones et qu'il y a une augmentation de la destruction de cellules de Nissl.

4° On peut déterminer si le candidat a quelques notions simples concernant les troubles, dans la question suivante :

INSTRUCTIONS. Dans la question suivante, cinq mécanismes différents sont proposés. Inscrivez la lettre appropriée à côté de chaque mécanisme numérotés en associant la lettre au mécanisme :

- A. Augmentation de l'activité de l'enzyme de la glycémie.
 - B. Hyperinsulinisme.
 - C. Mise en réserve de glycogène dans le foie.
 - D. Absence de mise en réserve de glycogène dans le foie.
 - E. Hypofonctionnement de l'enzyme de la glycémie.
- (B) 1. Adénomes des îlots de Langerhans
 - (A) 2. Exercice physique
 - (A) 3. Hyperthyroïdie

nt à cette série, cela mon-

t la source principale per-
t sucre sanguin ;

ts de Langerhans ne sont

erthyroïdie ont une hypo-
n'est pas un hyperinsuli-

maladie de Von Gierke ont
le n'est pas due au fait
s le foie ;

ypoglycémie chronique et
s résultantes ne sont pas

a quelques notions sim-
s au niveau du cerveau

sés incomplets (numéro-
ments au choix. Indiquez,
nt qui convient le mieux.

es secondaires à une hy-
mises en évidence le plus

voquées par l'hypoglycé-
cerveau, les plus impor-

giques

3. Parmi les altérations des cellules nerveuses provoquées par l'hypoglycémie aiguë, la plus évidente est :
 - a) le déplacement du noyau
 - b) la disparition de la paroi cellulaire
 - c) l'altération de la substance de Nissl
 - d) l'éclatement du noyau de la cellule
 - e) la fusion de mitochondries.
4. Parmi les modifications provoquées au niveau du cerveau par l'hypoglycémie chronique, la plus significative est :
 - a) l'hydrocéphalie interne
 - b) l'épaississement fibreux de l'arachnoïde
 - c) la destruction des neurones
 - d) la prolifération des astrocytes
 - e) la prolifération des cellules de l'épendyme

Si le candidat répond correctement à cette série, il montre qu'il sait que les altérations principales de l'hypoglycémie chronique atteignent les neurones du cerveau, que l'hypoglycémie aiguë provoque une altération de la substance de Nissl des neurones et que l'hypoglycémie chronique provoque la destruction de cellules nerveuses.

- 4° On peut déterminer si le candidat comprend le mécanisme des troubles, dans la question suivante :

INSTRUCTIONS. Dans la liste alphabétique sont indiqués cinq mécanismes différents pouvant conduire à l'hypoglycémie.

Inscrire la lettre appropriée après chacun des exposés numérotés en associant au trouble le mécanisme responsable :

- A. Augmentation de l'activité métabolique.
 - B. Hyperinsulinisme.
 - C. Mise en réserve d'un glycogène anormal dans le foie.
 - D. Absence de mise en réserve de glycogène dans le foie.
 - E. Hypofonctionnement de l'hypophyse ou des surrénales.
- (B) 1. Adénomes des îlots de Langerhans.
 - (A) 2. Exercice physique violent.
 - (A) 3. Hyperthyroïdisme.

- (E) 4. Maladie de Simmonds.
- (C) 5. Maladie de Von Gierke.
- (D) 6. Hépatite épidémique.
- (C) 7. Maladie d'Addison.

Si le candidat répond correctement à cette série, cela montre qu'il comprend les bases du métabolisme des glucides et connaît les facteurs influençant ce métabolisme.

Récapitulation.

Ainsi, au moyen de ces dix-sept « questions objectives », nous avons mis en évidence toutes les connaissances requises pour pouvoir donner une note supérieure à 90/100, à savoir :

1. Liste des quatre causes principales.
2. Reconnaître la différence entre les effets de l'hypoglycémie chronique et aiguë.
3. Décrire les altérations au niveau du cerveau.
4. Compréhension des mécanismes.
5. Liste des causes mineures.
6. Différencier les lésions de l'hypoglycémie aiguë et chronique.

AN

COMPOSITION DE
ANGLETERRE —
OF EDU
FIN DU SECONDAIR

EPREUVE I. Deux questions -

1. Choisissez un des sujets suivants :
 - a. Une nuit brumeuse.
 - b. Un marchand achète un objet d'occasion. Décrivez le.
 - c. Quelle serait votre politique des radios ou de télévisions ?
 - d. Une grande foule se rassemble.
 - e. Plaisir de la photographie ou du cyclisme.
 - f. Pensez-vous que garçons et filles ont une idée de carrière ?
 - g. Quelles sont vos réalisations de l'exploration ?
2. Choisissez un des sujets suivants :
 - a. Décrivez un entraînement sportif et les performances dans un sport.
 - b. Après avoir visité une usine, écrivez un article sur les perspectives d'avenir qu'elle offre.
 - c. Décrivez clairement un appareil électroménager. Expliquez comment il fonctionne. (sèche-cheveux, un aspirateur, un mixer).
 - d. Dans votre ville, on veut améliorer la circulation. Écrivez une lettre à un journal local sur votre projet.

COMPOSITION DE LANGUE MATERNELLE
ANGLETERRE — GENERAL CERTIFICATE
OF EDUCATION 1967
FIN DU SECONDAIRE — NIVEAU ORDINAIRE.

EPREUVE I. Deux questions — 1½ h.

1. Choisissez un des sujets suivants. Consacrez-y environ 1 h.
 - a. Une nuit brumeuse.
 - b. Un marchand achète et revend, le même jour, un objet d'occasion. Décrivez les deux scènes.
 - c. Quelle serait votre politique si vous dirigiez les programmes de radio ou de télévision ?
 - d. Une grande foule se disperse. Décrivez la scène.
 - e. Plaisir de la photographie *ou* du dessin *ou* de la danse *ou* du cyclisme.
 - f. Pensez-vous que garçons et filles ont les mêmes chances de carrière ?
 - g. Quelles sont vos réactions devant les progrès et les réalisations de l'exploration spatiale ?

2. Choisissez un des sujets suivants. Consacrez-y environ ½ h.
 - a. Décrivez un entraînement destiné à améliorer vos performances dans un sport de votre choix.
 - b. Après avoir visité une entreprise, faites un rapport à vos compagnons sur les conditions de travail et les perspectives d'avenir qu'elle offre.
 - c. Décrivez clairement *un* des appareils suivants et expliquez comment il fonctionne : un « walkie talkie », un sèche-cheveux, un aspirateur, un mélangeur-batteur (mixer).
 - d. Dans votre ville, on veut créer un centre commercial où la circulation sera interdite aux véhicules. Ecrivez une lettre à un journal local pour exposer vos vues sur le projet.

EPREUVE II. Quatre questions — 1³/₄ h.

1. Résumez le passage suivant en bonne prose continue et en 110 mots maximum. A la fin, indiquez combien de mots vous avez utilisés. Le passage compte 314 mots.

Suit un texte sur la confiance exagérée dans la science et la technologie.

2. Lisez le passage suivant ; ensuite, répondez aux questions.

(Le texte décrit deux grands types de promeneurs — ceux qui suivent un guide et ceux qui partent à l'aventure — et souligne l'intérêt des excursions géologiques.)

- a. Expliquez avec vos propres mots la différence entre les deux types de promeneurs.
 - b. Quel conseil l'auteur donne-t-il à propos des grottes ? Formulez-le avec vos propres mots.
 - c. Expliquez brièvement pourquoi l'auteur croit que la géologie est : un *hobby* amusant — un *hobby* instructif.
 - d. L'auteur écrit qu'il utilise l'expression « creuser un fossé » au sens littéral. Expliquez pourquoi le sens est ici littéral.
 - e. Expliquez les expressions suivantes : précautions prescrites ; être parfaitement conscient de la nature de ses actes.
 - f. Choisissez quatre des mots suivants. Remplacez-les par des synonymes ou des périphrases qui pourraient être utilisées dans le texte sans en changer le sens (...).
3. Répondez, au choix, à une des deux questions suivantes :
- a. Choisissez trois mots parmi les suivants. Construisez des phrases (6 en tout) montrant que ces mots peuvent être employés dans deux sens différents (...).
 - b. Définissez en une phrase trois des mots suivants : monopole - interlude - préface - microscope - antidote.
4. Répondez au choix à une des deux questions suivantes :
- a. Expliquez clairement, mais brièvement, la différence de sens entre chaque paire de phrases (porte sur *could-should* ; *can-may* ; *will-shall* ; *might-must*).
 - b. Récrivez correctement le passage suivant, en respectant toutes les idées. Vous pouvez changer l'expression, l'ordre des mots et des idées, l'orthographe et la ponctuation.
(Suit un texte défectueux d'une centaine de mots).

« Par exemple si on
à travers une ville
dans son chemin
gés, entraînant
contourner le bâti
et la rendant auss
déviation de cette
causés mais le p
spécialement si la
la distance ajouté
camion portant ce
camion consomm
plus de temps po
enlevés de sa vie.

h.
bonne prose continue et en
uez combien de mots vous
14 mots.

exagérée dans la science

, répondez aux questions.

*types de promeneurs —
ceux qui partent à l'aven-
s excursions géologiques.)*

ots la différence entre les

il à propos des grottes ?
ots.

l'auteur croit que la géo-
— un *hobby* instructif.

pression « creuser un fos-
z pourquoi le sens est ici

vantes : précautions pres-
cient de la nature de ses

ivants. Remplacez-les par
rases qui pourraient être
changer le sens (...).

deux questions suivantes :

suivants. Construisez des
ue ces mots peuvent être
rents (...).

des mots suivants : mono-
roscope - antidote.

x questions suivantes :

èvement, la différence de
phrases (porte sur *could-
night-must*).

sage suivant, en respec-
vez changer l'expression,
l'orthographe et la ponc-

centaine de mots).

« Par exemple si on projetait de faire passer une route à travers une ville mais qu'une maison historique était dans son chemin alors les plans devraient être changés, entraînant des dépenses considérables, pour contourner le bâtiment créant un virage dans la route et la rendant aussi dangereuse pour les autos. Par la déviation de cette route non seulement des frais sont causés mais le prix de différentes choses augmente spécialement si la déviation est grande parce que si la distance ajoutée est disons dix kilomètres et qu'un camion portant certains articles parcourt la route, le camion consommerait alors plus d'essence prendrait plus de temps pour arriver et dix kilomètres seraient enlevés de sa vie. »

CONSIGNES POUR LA CORRECTION.

EPREUVE I. — Maximum 50 points.

Le schéma de notation qui figure ci-dessous ne constitue qu'un guide préliminaire. Des additions et des amendements pourront être apportés lors de la réunion des examinateurs qui sera convoquée après une première lecture des travaux.

Question 1 (maximum 35 points).

On attend un minimum de 400 mots, mais les compositions ne doivent pas être principalement notées en fonction de la longueur. Tenir compte du sujet choisi et de la façon dont il est traité. Même si elle est courte, une composition où l'argumentation est serrée et où l'expression est bonne doit obtenir plus de points qu'une longue narration informe.

L'examinateur devrait avoir une idée claire de ce qu'est une composition recevant tout juste la note de passage (16 points). Pareille composition doit contenir des idées raisonnables, mais non très originales. L'expression doit être claire, mais sans distinction particulière. On ne devrait rencontrer, dans le travail, que quelques erreurs mécaniques. Les candidats qui ont dépassé ce niveau général doivent être récompensés et ceux qui ont travaillé en dessous de ce niveau doivent être pénalisés.

On trouvera dans le document annexé des notes détaillées sur les qualités à observer dans les compositions. On compte toutefois que les examinateurs noteront en fonction de leur impression générale, n'alloueront donc pas une proportion fixe des points pour les différents aspects. Si le sujet se prête à la controverse, les idées et leur enchaînement peuvent être plus importants que dans une composition descriptive où le vocabulaire pourrait prendre une plus grande place. Nous nous fions au jugement des examinateurs.

En réservant 35 points à cette première question, on a voulu marquer son importance majeure dans l'ensemble des examens. Beaucoup de candidats sont médiocres. Toutefois, si l'examinateur ne disperse pas largement ses notes, cette première question ne pèsera pas d'un poids suffisant dans l'ensemble des résultats.

Groupe A (29-35 points).

La composition est de qualité exceptionnelle.

Groupe B (22-28 points).

La qualité du travail est au-dessus de la moyenne.

Groupe C (14-21 points).
Le travail est de qualité m

Groupe D (7-13 points).
Le travail est d'un niveau
comme satisfaisant.

Groupe E (0-6 points).
Le candidat est incapable
cohérence.

Cette épreuve est un exa
l'épreuve 1, nous examinons la
ses opinions, ses expériences
et ses intérêts. Il ne s'agit ni
rales, ni d'une évaluation de
jeune fille décrit clairement
obtenir une note favorable si
nents, même si l'examinateur
tion aborde le sujet de façon
si, derrière la masse de pro
l'impression que se trouvent
attitudes morales élevées, il n
aspect : seul, ce qui est écrit
la clarté de l'expression et la

Question 1 (Maximum 35

- (a) Une narration ou une de
une narration, la nuit brun
- (b) En gros, les deux scènes
penser la vivacité de narra
et le contraste.
- (c) On attend une définition cl
points doivent être déve
choisi la radio ou la télév
deux.
- (d) Le sujet du travail est la
une courte introduction.
- (e) Une réponse cohérente, d
plus que de longs errement
- (f) Ce sujet n'est pas facile,
bonne ordonnance des a
choisis.
- (g) Un traitement purement na
mais quelques exemples d
nécessaires pour expliquer

CORRECTION.

ci-dessous ne constitue
ons et des amendements
ion des examinateurs qui
lecture des travaux.

ts, mais les compositions
notées en fonction de la
isi et de la façon dont il
ne composition où l'argu-
on est bonne doit obtenir
informe.

idée claire de ce qu'est
la note de passage (16
tenir des idées raisonna-
ssion doit être claire, mais
evrait rencontrer, dans le
ues. Les candidats qui ont
être récompensés et ceux
eau doivent être pénalisés.

annexé des notes détaillées
compositions. On compte
ront en fonction de leur
c pas une proportion fixe
. Si le sujet se prête à la
nement peuvent être plus
n descriptive où le voca-
de place. Nous nous fions

mière question, on a voulu
s l'ensemble des examens.
es. Toutefois, si l'examina-
es notes, cette première
suffisant dans l'ensemble

ceptionnelle.

us de la moyenne.

Groupe C (14-21 points).

Le travail est de qualité moyenne.

Groupe D (7-13 points).

Le travail est d'un niveau inférieur à ce que l'on considère
comme satisfaisant.

Groupe E (0-6 points).

Le candidat est incapable de présenter ses idées avec
cohérence.

Cette épreuve est un examen de langue maternelle. Dans
l'épreuve 1, nous examinons la capacité du candidat à exprimer
ses opinions, ses expériences, ses impressions, ses sentiments
et ses intérêts. Il ne s'agit ni d'un test de connaissances géné-
rales, ni d'une évaluation des aptitudes du candidat. Si une
jeune fille décrit clairement une scène en anglais, elle doit
obtenir une note favorable si les matériaux utilisés sont perti-
nents, même si l'examineur pense que la jeune fille en ques-
tion aborde le sujet de façon trop sentimentale. Par ailleurs,
si, derrière la masse de prose incohérente, l'examineur a
l'impression que se trouvent des émotions profondes et des
attitudes morales élevées, il n'a pas à s'occuper de ce dernier
aspect : seul, ce qui est écrit compte. Pour réussir l'épreuve,
la clarté de l'expression et la précision du style sont essentiels.

Question 1 (Maximum 35 points)

- (a) Une narration ou une description sont acceptables. Dans
une narration, la nuit brumeuse doit jouer un rôle essentiel.
- (b) En gros, les deux scènes doivent être équilibrées. Récom-
penser la vivacité de narration, de dialogue, de description
et le contraste.
- (c) On attend une définition claire de la politique dont plusieurs
points doivent être développés. Le candidat doit avoir
choisi la radio ou la télévision ; il ne peut avoir réuni les
deux.
- (d) Le sujet du travail est la dispersion ; autoriser néanmoins
une courte introduction.
- (e) Une réponse cohérente, claire, de longueur modérée fait
plus que de longs errements et des répétitions.
- (f) Ce sujet n'est pas facile, récompenser généreusement la
bonne ordonnance des arguments et les exemples bien
choisis.
- (g) Un traitement purement narratif ne peut pas être accepté,
mais quelques exemples de progrès réalisés peuvent être
nécessaires pour expliquer les réactions.

Question 2 (Maximum 15 points).

- (a) On attend une description claire et logique.
- (b) Le rapport doit traiter les trois aspects de la question. On insiste surtout sur l'information rapportée ; on n'exigera donc pas une forme de rapport particulier.
- (c) On exige à la fois une description de l'objet et une explication de son fonctionnement. Ne punissez pas sévèrement les erreurs matérielles ; notez simplement la clarté de l'expression.
- (d) On peut envisager de nombreux aspects. Aussi, pénalisez toute idée inadéquate.

Décomptez :

- 1 point pour une mauvaise rédaction de l'adresse.
- 1 point pour un manque de cohérence entre la vedette et les salutations.
- 1 point si l'élève a signé « M. John Smith » ou « Mademoiselle Jeannette Smith ».
- 1/2 point pour d'autres erreurs de disposition ou de ponctuation, ou d'orthographe dans des mots essentiels.

Tous ces sujets fournissent un matériau suffisant pour une demi-heure de travail. On attend un minimum de 200 mots. Le choix est riche.

EPREUVE II — Maximum 50 points.

Question 1 (maximum 16 points).

Les points attribués à cette question sont habituellement beaucoup plus bas que pour les autres. Les examinateurs sont priés de ne pas considérer 11 comme le maximum attribuable au résumé.

A. — Attribuer un maximum de 2 points pour chacun des aspects suivants. Pour obtenir 2 points, le candidat doit avoir clairement compris l'idée et l'avoir exprimée correctement. Nuancez vos notes par 1 1/2, 1 ou 1/2 point. A la fin du travail, comptez une nouvelle fois l'ensemble en fonction de la fluidité et de la cohérence du résumé complet. Si un passage est incohérent, supprimez au maximum 1/4 des points attribués ; s'il est plutôt rocailleux, soustrayez 1/8. Cette correction devrait apparaître sur la composition en écrivant par exemple 10 — 1 = 9. Toute soustraction de points destinée à pénaliser un texte plus long que la limite fixée ou une proposition non seulement inexacte mais absurde doit apparaître séparément. Le total final doit être entouré d'un cercle dans la marge.

- 1° L'homme de la rue ad...
scientifiques,
- 2° sans douter de leur or...
effets.
- 3° et 4°. La demande de r...
le niveau de vie ne se ra...
- 5° Les hommes ont confian...
son travail,
- 6° et croient qu'on ne peut
- 7° Bien qu'ils reconnaisse...
sont pas toujours d'acc...
les nouvelles découvertes
- 8° le public est convaincu...
par trouver l'unanimité

B. — La limite de 120...
Supprimez un point pour c...
le maximum. Ne comprenez...
productifs tels que « Dans c...
Recomptez les mots. N'a...
inscrit par le candidat.

Question 2 (maximum 2...

En notant cette questio...
que section au candidat qu...

- (a) (1) Ceux qui aiment qu...
(2) et ceux qui préfèr...
plan général perme...
moment (esprit d'e...
- (b) Visiter des grottes.
A moins que vous ne s...
connaît très bien le t...
Respectez toutes les r...
- (c) (1) Il ne faut guère d'...
(2) Le géologue fait...
découvertes
à petite échelle.
- (d) Quand le géologue cas...
un acte unique, car p...
casser à nouveau cette...
- (e) (1) Mesures de sécurit...
imposées ou recom...
(2) Pour comprendre e...
les conséquences c...

logique.
pects de la question. On
rapportée ; on n'exigera
ticulier.

de l'objet et une explica-
hissez pas sévèrement les
ment la clarté de l'expres-

aspects. Aussi, pénalisez

édaction de l'adresse.
phérence entre la vedette

John Smith » ou « Made-
».

rs de disposition ou de
graphe dans des mots

matériau suffisant pour une
minimum de 200 mots. Le

stion sont habituellement
s. Les examinateurs sont
e le maximum attribuable

points pour chacun des
ts, le candidat doit avoir
exprimée correctement.
point. A la fin du travail,
en fonction de la fluidité
et. Si un passage est in-
des points attribués ; s'il
Cette correction devrait
vant par exemple 10 — 1
tinée à pénaliser un texte
proposition non seulement
séparément. Le total final
large.

- 1° L'homme de la rue accepte aujourd'hui les découvertes scientifiques,
- 2° sans douter de leur origine, de leur validité, ou de leurs effets.
- 3° et 4°. La demande de nouveaux progrès destinés à élever le niveau de vie ne se ralentit jamais.
- 5° Les hommes ont confiance dans l'homme de science et dans son travail,
- 6° et croient qu'on ne peut arrêter le progrès.
- 7° Bien qu'ils reconnaissent que les hommes de science ne sont pas toujours d'accord sur les sécurités apportées par les nouvelles découvertes,
- 8° le public est convaincu que les hommes de science finiront par trouver l'unanimité ou au moins un large accord.

B. — La limite de 120 mots donne une marge généreuse. Supprimez un point pour chaque tranche de 5 mots dépassant le maximum. Ne comptez pas dans le comptage des mots introductifs tels que « Dans ce passage, l'auteur explique que... ». Recomptez les mots. N'acceptez pas simplement le chiffre inscrit par le candidat.

Question 2 (maximum 20 points).

En notant cette question, accordez le maximum dans chaque section au candidat qui expose nettement le point.

- (a) (1) Ceux qui aiment que l'on prévoie tout à leur place (1)
(2) et ceux qui préfèrent disposer simplement d'un plan général permettant de suivre l'inspiration du moment (esprit d'exploration) (2)
- (b) Visiter des grottes. (1)
A moins que vous ne soyez guidé par quelqu'un qui connaît très bien le terrain. (1)
Respectez toutes les règles de sécurité. (2)
- (c) (1) Il ne faut guère d'équipement spécial. (2)
(2) Le géologue fait continuellement de nouvelles découvertes (1½)
à petite échelle. (½)
- (d) Quand le géologue casse une pierre, il pose, en fait, un acte unique, car personne après lui ne pourra casser à nouveau cette même pierre. (2)
- (e) (1) Mesures de sécurité imposées ou recommandées. (1)
(2) Pour comprendre exactement les conséquences de ce qu'on fait. (1)

(f) Souterrain :

qui est sous le sol, sous la surface de la terre ; (1)
sous terre. (1)

Spéléologues :

personnes qui étudient les grottes scientifique- (1)
ment ; explorateurs de grottes.

Consciencieusement :

de bonne foi, sans se laisser distraire, avec (1)
beaucoup d'application
honnêtement (1/2)
fidèlement (1/2)

Etc.

Pour les sous-questions (a), (b) et (e), n'attribuez aucun point aux élèves qui se bornent à recopier une partie du texte.

Question 3 (maximum 6 points).

(a) Attribuez un point pour toute phrase construite correctement.

N'attribuez aucun point en cas de construction incorrecte.

(b) Accordez un point pour chaque définition exacte exprimée dans une phrase correcte (maximum 3 points).

Accordez un point pour chaque phrase où le sens du mot commençant par le même préfixe ressort clairement (maximum 3 points).

1/2 point pour une définition exprimée en une phrase incorrecte.

1/2 point pour une définition fournie en phrase incomplète.

Aucun point si le mot commençant par le même préfixe n'est pas présenté dans une phrase.

Un *monopole* est une propriété exclusive détenue par une firme (ou)

Un *monopole* est le nom donné à une firme qui détient des droits commerciaux exclusifs.

Un *interlude* est un intervalle ménagé au cours de la représentation d'une pièce (ou un intervalle dans le déroulement d'un événement).

Etc.

Question 2 (maximum

(a) Accordez un point par

Vérifiez la présence de

N'accordez aucun point

quelle phrase le candidat

(b) Les candidats répondent

Déduisez 1 point par

toute construction b

incorrecte, 1/2 point

d'une virgule essentielle

Revoyez la note totale

du passage.

Pour les questions 1

des points, arrondissez

Si les deux questions

point, arrondissez l'u

Mêmes remarques p

face de la terre ; (1)
(1)

ottes scientifique- (1)
s.

er distraire, avec (1)
(1/2)
(1/2)

et (e), n'attribuez aucun
pier une partie du texte.

ase construite correcte-

construction incorrecte.
finition exacte exprimée
n 3 points).

phrase où le sens du mot
ressort clairement (maxi-

ée en une phrase incor-

e en phrase incomplète.
nt par le même préfixe

clusive détenue par une

ne firme qui détient des

gé au cours de la repré-
alle dans le déroulement

Question 2 (maximum 8 points).

(a) Accordez un point par phrase.

Vérifiez la présence des idées suivantes...

N'accordez aucun point si l'on ne voit pas clairement à
quelle phrase le candidat se réfère.

(b) Les candidats répondront de façons différentes.

Déduisez 1 point par faute d'orthographe, 2 points pour
toute construction boiteuse, 1 point pour une expression
incorrecte, 1/2 point pour l'omission ou l'emploi erroné
d'une virgule *essentielle*, 1 point pour chaque idée omise.

Revoyez la note totale en fonction de l'impression générale
du passage.

Pour les questions 1 et 2, si le candidat obtient la moitié
des points, arrondissez à l'unité supérieure (ex. : $6 \frac{1}{2} = 7$).

Si les deux questions donnent une note comprenant 1/2
point, arrondissez l'un par excès et l'autre par défaut.

Mêmes remarques pour les questions 3 et 4.

EXEMPLE D'ENSEIGNEMENT SEMI-INDIVIDUALISE. ¹

La New Trier Township High School, Winnetka.

La *New Trier Township High School* est une grande école du degré secondaire supérieur où l'enseignement est semi-individualisé et où l'étudiant peut corriger son orientation jusqu'au terme de l'adolescence.

Elle accueille les élèves à partir de 14 ans ² et jouit d'une grande réputation tant aux Etats-Unis qu'à l'étranger.

3 740 étudiants suivent régulièrement les cours ³. Le corps professoral compte 255 membres dont :

40	professeurs d'anglais
31	» de mathématiques
30	» d'éducation physique
29	» de langues étrangères
27	» de « Social Studies » (Histoire, géographie, civisme, sociologie et économie)
17	» de sciences
9	» de musique
7	» de cours techniques
7	» de commerce
6	» d'art dramatique
6	» de peinture-dessin
6	» d'automobile
4	» d'économie domestique
2	» d'hygiène
1	» d'enseignement spécial (retardés mentaux édu-cables).

On compte un enseignant pour 15-16 étudiants et les classes réunissent généralement 25 élèves environ.

¹ Extrait d'un article que nous avons publié dans *Education*, en mai 1961.

² On sait que l'organisation de l'enseignement des Etats-Unis n'est pas uniforme. La *N.T. High School* relève du système « NK 8-4 » : un an de *Nursery school*, un an de jardin d'enfants, 8 ans de primaire et 4 ans de secondaire. Les deux autres systèmes les plus fréquents sont : NK 6-3-3 et NK 6-6. On rencontre aussi NK 7-5, NK 6-2-4 et NK 6-4-4.

³ Ces chiffres se réfèrent à l'année scolaire 1959-1960. Ils sont extraits de : *Information for College Admission Officers*, N.T. Township High School, nov. 1959, ou ont été recueillis sur place.

Les prestations d'un p
sabilités spéciales (telles
ex.) comportent 24 période
de 5 périodes et 4 période

Le niveau intellectuel
tests classiques, est élevé
leurs études à New Trier
la moyenne nationale, dan
sances (SCAT, STEP, N
continuent des études sup

L'individualisation des pr

Le principe fondamen
Trier High School est dé
programme de cours : « P
hommes sans éprouver un
variété infinie. » ¹

Restant donc ainsi im
netka, les études sont org
possibilité de se dévelop
capacités et ses penchar
fantaisie stérile ou que les

Selon le système rép
peut obtenir un diplôme
celles-ci, il a gagné un r
fixé par le conseil d'admin
compte de normes géné
unité conventionnelle l'in
des différents cours, pour
par exemple, à un semes
cinq périodes par semain
l'étudiant atteint une no
fixée.

Pour obtenir le cert
School, il faut y avoir ga

Le jeu des *credits* es
sécurité, car il permet d
liberté de l'élève.

En principe, celui-ci é
on compose le menu d'u
ce qui attire et convient le

MI-INDIVIDUALISE. ¹

Winnetka.

ol est une grande école
enseignement est semi-
r son orientation jus-

e 14 ans ² et jouit d'une
qu'à l'étranger.

ent les cours ³. Le corps

Histoire, géographie, ci-
nomie)

(retardés mentaux édu-

15-16 étudiants et les
es environ.

Education, en mai 1961.

Etats-Unis n'est pas uniforme. La
an de Nursery school, un an de
ndaire. Les deux autres systèmes
contre aussi NK 7-5, NK 6-2-4 et

0. Ils sont extraits de : *Information*
ool, nov. 1959, ou ont été recueil-

Les prestations d'un professeur qui n'assure pas de respon-
sabilités spéciales (telles que président d'un département, par
ex.) comportent 24 périodes de 40 minutes par semaine : 4 cours
de 5 périodes et 4 périodes de *counseling*.

Le niveau intellectuel des étudiants, jugés sur la base des
tests classiques, est élevé. Environ 80 % de ceux qui terminent
leurs études à New Trier obtiennent des résultats supérieurs à
la moyenne nationale, dans les tests d'aptitudes et de connais-
sances (SCAT, STEP, *National Merit*). 92 % des diplômés
continuent des études supérieures.

L'individualisation des programmes.

Le principe fondamental de l'action pédagogique de la New
Trier High School est défini dans la première phrase de son
programme de cours : « Personne ne peut croire à la dignité des
hommes sans éprouver un plaisir profond au spectacle de leur
variété infinie. » ¹

Restant donc ainsi imprégnées de l'esprit du Plan de Win-
netka, les études sont organisées de façon à fournir à chacun la
possibilité de se développer à son rythme propre, selon ses
capacités et ses penchants. Ceci ne signifie nullement que la
fantaisie stérile ou que les solutions de facilité soient permises.

Selon le système répandu dans tout le pays, l'étudiant ne
peut obtenir un diplôme de fin d'études que si, au cours de
celles-ci, il a gagné un nombre total de points ou de *credits*
fixé par le conseil d'administration de l'école qui tient lui-même
compte de normes générales. Ces *credits* expriment en une
unité conventionnelle l'importance qualitative et quantitative
des différents cours, pour un semestre (5 *credits* correspondent,
par exemple, à un semestre de cours « mineur », à raison de
cinq périodes par semaine). Les *credits* ne sont acquis que si
l'étudiant atteint une note supérieure à une limite minimum
fixée.

Pour obtenir le certificat d'études de la New Trier High
School, il faut y avoir gagné 350 *credits* en 4 ans.

Le jeu des *credits* est à la fois source de souplesse et de
sécurité, car il permet de délimiter exactement le champ de
liberté de l'élève.

En principe, celui-ci établit son programme d'études comme
on compose le menu d'un repas en choisissant, sur la carte,
ce qui attire et convient le mieux.

¹ NEW TRIER TOWNSHIP HIGH SCHOOL, *Curriculum Guide*, déc. 1959, p. 1.

Voici, par exemple, la liste des cours offerts pour la première année, avec mention des *credits* qu'ils rapportent s'ils sont suivis avec succès pendant un semestre¹ :

Cours « majeurs »

(Rapportant 10 *credits* par semestre ; sauf indication contraire, ils doivent être suivis pendant 2 semestres consécutifs)

Anglais	
Algèbre	Mathématiques générales
Alimentation (1 semestre)	Peinture-dessin (2 périodes par jour)
Civisme	Photographie
Commerce	Radio amateur
Dessin industriel	Travail du bois
Electricité (1 semestre)	Travaux manuels
Histoire universelle	Vêtement (1 semestre)
Langues étrangères	

Cours « mineurs »

(Nombre de *credits* entre parenthèses)

Art dramatique (6)	Dessin industriel (1 période par jour) (5)
Chant choral (3)	Diction (6)
Dactylographie (5)	Harmonie (5)
Cours d'harmonie (6)	Orchestre symphonique (5)
	Travaux manuels (1 période par jour) (5)

Néanmoins, plusieurs restrictions influencent le choix de l'élève. En premier lieu, un certain nombre de cours, jugés indispensables à la culture de base de tous les membres de la nation, sont obligatoires :

<i>Cours</i>	<i>Durée obligatoire</i>	<i>Credits</i>
Langue maternelle	4 ans	80
Mathématiques	2 ans	40
« Social Studies »	2 ans ²	40
Sciences	1 an	20
Education physique	4 ans	16
Automobile : théorie et pilotage	1 semestre	3

Total 199

¹ Cf. NEW TRIER TOWNSHIP HIGH SCHOOL, *Registration Bulletin for Freshmen 1959-1960. Planning a Course of Study*, p. 2.

² Dont un an obligatoirement consacré à l'histoire des Etats-Unis.

Environ 2/3 du total des cours imposés. Toutefois, être satisfait à ces exigences, nous avons relevé une dizaine de d'accomplir valablement l'a

Un second facteur important de son programme : les choses auxquelles il aspire. Etamine ses conditions d'admission bien de *credits* l'étudiant du secondaire, pour des branches

Enfin, les parents, les mentors et les orienteurs s'efforcent d'intérêts, veillant à ce qu'il tait ses potentialités.

La semaine scolaire compte 30 minutes, — réunion quotidienne (20') et durée du lunch toutes les semaines, — on suggère généralement un programme moyennement d'étude personnelle. Voici d'ailleurs la première année¹ :

I		II	
<i>Cours</i>	<i>pér. sem.</i>	<i>Cours</i>	<i>pér. sem.</i>
Anglais	5	Anglais	5
Algèbre	5	Algèbre	5
Latin	5	Sciences	5
Peint./dess.	10	Commerce	5
Gymn. (filles)	4	Gymn. (garç.)	4
	—	Musique instr.	4
	29		
Etude personnelle	11		
	—		
	40		

On remarquera combien, avec l'émiettement de l'effort, tous les cours académiques sont terminés à la matinée.

¹ NEW TRIER TOWNSHIP HIGH SCHOOL

Cours offerts pour la première année qu'ils rapportent s'ils sont suivis pendant le premier semestre¹ :

»
sauf indication contraire, les cours sont offerts pendant deux semestres consécutifs)

Mathématiques générales
Arts et dessin (2 périodes par semaine)

Graphie
Carpenterie
Menuiserie
Travaux manuels
Arts et dessin (1 semestre)

»
Cours de parenthèses)

Arts et dessin (1 période par jour) (5)
(6)
Arts et dessin (5)
Arts et dessin (5)
Arts et dessin (1 période par jour) (5)

Les facteurs suivants influencent le choix de l'ensemble des cours, jugés importants par tous les membres de la commission :

Cours obligatoires	Credits
1 ^{er} semestre	80
2 ^e semestre	40
3 ^e semestre	40
4 ^e semestre	20
5 ^e semestre	16
6 ^e semestre	3
Total	199

Registration Bulletin for Freshmen
des Etats-Unis.

Environ 2/3 du total des *credits* exigés sont donc fournis par des cours imposés. Toutefois, comme nous le verrons, il peut être satisfait à ces exigences de façon fort libre ; ainsi, nous avons relevé une dizaine de possibilités différentes permettant d'accomplir valablement l'année de sciences réclamée.

Un second facteur important guide l'étudiant dans l'élaboration de son programme : la profession ou les études supérieures auxquelles il aspire. En particulier, chaque université détermine ses conditions d'admission et spécifie notamment combien de *credits* l'étudiant doit avoir acquis dans l'enseignement secondaire, pour des branches déterminées.

Enfin, les parents, les maîtres, les conseillers pédagogiques et les orienteurs s'efforcent de guider l'étudiant au mieux de ses intérêts, veillant à ce qu'il tire le meilleur profit possible de ses potentialités.

La semaine scolaire comptant 5 jours de 8 périodes de 40 minutes, — réunion quotidienne avec le conseiller pédagogique (20') et durée du lunch toujours pris à l'école (25') non comprises, — on suggère généralement à l'élève de se constituer un programme moyennement chargé, ménageant du temps pour l'étude personnelle. Voici quatre exemples types pour la première année¹ :

I		II		III		IV	
Cours	pér. sem.	Cours	pér. sem.	Cours	pér. sem.	Cours	pér. sem.
Anglais	5	Anglais	5	Anglais	5	Anglais	5
Algèbre	5	Algèbre	5	Algèbre	5	Algèbre	5
Latin	5	Sciences	5	Histoire	5	Histoire	5
Peint./dess.	10	Commerce	5	Aliment.	7	Sciences	5
Gymn. (filles)	4	Gymn. (garç.)	5	Chant	3	Lang. étrang.	5
	—	Musique instr.	5	Gymn. (filles)	4	Gymn. (garç.)	5
	29		—		—		—
Etude personnelle	11		30		29		30
	—		10		11		10
	40		40		40		40

On remarquera combien ces plans de travail contrastent avec l'émiettement de l'effort si fréquent chez nous. Pratiquement, tous les cours académiques peuvent être concentrés dans la matinée.

¹ NEW TRIER TOWNSHIP HIGH SCHOOL, *Courses for Freshmen*, avril 1959, p. 4.

Environ un quart du temps passé à l'école est réservé aux études personnelles. Les heures ainsi laissées libres seront souvent consacrées à des travaux de recherche dans la bibliothèque scolaire qui, aux Etats-Unis, joue un rôle incomparablement plus important qu'en Belgique.

L'individualisation de l'enseignement.

Non seulement l'élève choisit les branches qui lui conviennent le mieux, mais encore l'enseignement de chacune de celles-ci sera adapté à ses possibilités.

En effet, toutes les branches importantes peuvent être étudiées à cinq niveaux d'aptitudes différents : inférieur, moyen faible, normal, accéléré, avancé. De cette façon, l'effort réclamé à l'étudiant, qu'il possède une intelligence supérieure ou soit peu doué, qu'il soit fort dans un domaine et en retard dans un autre, est toujours en proportion avec ses possibilités.

Comme l'élève suit rarement tous ses cours à un même niveau, le danger d'une ségrégation générale, selon les aptitudes, semble minime : l'école reste d'ailleurs très attentive à ce problème, regroupant systématiquement tous les étudiants à l'occasion de certaines activités. On met aussi tout en œuvre pour inculquer une véritable tolérance vis-à-vis du plus ou moins grand talent des compagnons d'études et pour encourager chacun à se dépasser : « En sport, tout le monde n'est pas capable de jouer en excellence ; c'est aussi le cas dans les études. Mais on attend de chacun le meilleur de lui-même.. »¹

Le système d'enseignement à différents niveaux est pratiqué à New Trier, depuis la fin de la première guerre mondiale, à la plus grande satisfaction de tous. En moyenne, les étudiants se répartissent selon les pourcentages suivants :

— supérieurs	15—20 %
— normaux (moyens forts)	40—55 %
— moyens faibles	36—40 %
— limités	5—8 %
— avancés (seniors)	± 10 %

Quand les élèves entrent à la New Trier High School, les groupes sont provisoirement déterminés sur la base des résultats scolaires antérieurs et d'autres renseignements réunis selon un système que nous étudierons plus loin. Par après, les résultats obtenus dans l'école même corrigeront et guideront les affectations, celles-ci n'étant jamais définitives, quelle que soit la branche.

¹ NEW TRIER HIGH SCHOOL, *Guide Book to New Trier*, Winnetka, 1959, p. 39.

Pareille souplesse complexe, mais il ne faut cependant pas se laisser aller. Dans le tableau suivant, on aura donc 1, 2, 3 et 4 possibilités offertes pendant l'année scolaire 1959-1960. Quelques exemples de possibilités des études¹.

- Chaque cours est désigné par un chiffre. On aura donc 1, 2, 3 et 4 possibilités offertes pendant l'année scolaire 1959-1960. Quelques exemples de possibilités des études¹.
- a) Le chiffre des centaines désigne le semestre :
 1 = 1^{er} semestre
 2 = 2^e semestre
 0 = peut être soit le 1^{er} ou le 2^e semestre
 3 = cours de fin de cycle
- b) Le chiffre des dizaines désigne le niveau :
 1 = niveau inférieur
 2 = niveau moyen faible
 3 = niveau normal
 4 = niveau accéléré
 5 = niveau avancé
 6 = séminaire
 9 = tous les niveaux

Branches et niveaux de cours	
Langue maternelle	125
	124
	123
	122
	122 R
	121
	224 - 223 - 222 - 221
	324 - 323 - 322 - 321
	424 - 424 (+) - 423 - 422 -

¹ Les deux documents de base sont : NEW TRIER HIGH SCHOOL, *Curriculum Guide*, 1959.

à l'école est réservé aux
 i laissées libres seront
 echerche dans la biblio-
 e un rôle incomparable-

branches qui lui convien-
 ent de chacune de cel-

portantes peuvent être
 érents : inférieur, moyen
 e façon, l'effort réclamé
 ence supérieure ou soit
 ne et en retard dans un
 es possibilités.

ses cours à un même
 générale, selon les aptitu-
 leurs très attentive à ce
 nt tous les étudiants à
 met aussi tout en œuvre
 e vis-à-vis du plus ou
 études et pour encoura-
 tout le monde n'est pas
 t aussi le cas dans les
 meilleur de lui-même.. »¹

érents niveaux est prati-
 emière guerre mondiale,
 s. En moyenne, les étu-
 entages suivants :

- 15—20 %
- s) 40—55 %
- 36—40 %
- 5— 8 %
- ± 10 %

w Trier High School, les
 es sur la base des résul-
 eignements réunis selon
 oin. Par après, les résul-
 igeront et guideront les
 éfinitives, quelle que soit

w Trier, Winnetka, 1959, p. 39.

Pareille souplesse conduit certes à une organisation com-
 plexe, mais il ne faut cependant pas exagérer la difficulté.

Dans le tableau suivant, nous faisons apparaître toutes les
 possibilités offertes pendant le second semestre de l'année
 scolaire 1959-1960. Quelques notes marginales précisent l'esprit
 des études¹.

Chaque cours est désigné par un nombre de trois chiffres :

- a) Le chiffre des centaines indique en quelle année il se donne ;
 on aura donc 1, 2, 3 et 4.
- b) Le chiffre des dizaines indique le semestre :
 1 = 1^{er} semestre
 2 = 2^e semestre
 0 = peut être suivi au 1^{er} ou au 2^e semestre
 3 = cours de vacances d'été.
- c) Le chiffre des unités indique le niveau d'aptitude :
 1 = niveau inférieur
 2 = moyen faible
 3 = normal
 4 = enseignement accéléré
 5 = avancé
 6 = séminaires
 9 = tous niveaux réunis.

Branches et niveaux de cours	Observations
Langue maternelle	
125	— Pour étudiants supérieurement intel- ligents, excellents en langue mater- nelle (intègre la langue maternelle, la biologie et l'histoire)
124	— Cours enrichi, fondé sur les œuvres littéraires de niveau universel
123	— Cours moyen
122	-- Cours moyen inférieur ; s'adresse plus particulièrement aux élèves qui n'ont pas encore acquis une métho- de de travail rationnelle
122 R	— R = « remedial ». Destiné aux élè- ves identifiés par le bureau de testing comme présentant une défi- cience marquée en lecture (com- préhension et rapidité) et en ortho- graphe.
121	— Accueille les élèves présentant une faiblesse générale dans toutes les branches de la langue maternelle.
224 - 223 - 222 - 221	Les différences de niveaux en 2 ^e , 3 ^e et 4 ^e années sont <i>mutatis mutandis</i> parallèles à celles que nous indiquons ci-dessus pour la 1 ^{re} . 323 = journalisme. 424 (+) = « Great Books ».
324 - 323 - 322 - 321	
424 - 424 (+) - 423 - 422 - 421	

¹ Les deux documents de base suivants ont été utilisés : NEW TRIER TOWNSHIP
 HIGH SCHOOL, *Curriculum Guide*, 1959 et *Program of Classes*, 1959-1960.

Branches et niveaux de cours	Observations
Mathématiques	La différence entre les niveaux 2, 3 et 4 réside moins dans l'accélération que dans la méthode d'enseignement et l'approfondissement.
Algèbre A (accélééré) 124	
» — 124	
» E (expérimental) 124	
» — 123	
» E 123	<i>Au niveau 2</i> , les explications sont détaillées ; beaucoup d'applications ; pas d'incursions dans les domaines voisins.
» — 122	
» E 122	<i>Au niveau 3</i> , l'étudiant doit plus travailler par lui-même : théorie plus rigoureuse, enrichissement.
Mathématiques de base 121	<i>Au niveau 4</i> , les concepts sont traités rapidement ; étude plus approfondie ; nombreuses incursions dans les domaines voisins.
Mathématiques approfondies (20 credits) 225	Le même manuel est employé pour ces 3 niveaux et tous les élèves étudient le même chapitre en même temps, ce qui permet, à tout moment, le passage d'un niveau à l'autre.
Géométrie 224	
» 223	
» 222	
Mathématiques 222	
Mathématiques de base 221	
Mathématiques approfondies (20 credits) 315	
Mathématiques 324	
Algèbre 323	
» 322	
Géométrie 303	
Mathématiques approfondies 425	4 semestres de mathématiques sont obligatoires ; toutefois, 2 de ceux-ci peuvent être consacrés à la comptabilité ou à l'arithmétique commerciale.
Mathématiques 424	
Algèbre (niveau universitaire) 403	
Trigonométrie 403	
Usage de la règle à calculs 323 (2 credits ; va avec algèbre 322)	
« Social studies »	4 semestres de « Social Studies » sont obligatoires ; 2 de ceux-ci doivent être consacrés à l'histoire des Etats-Unis.
Civisme 123, 122, 121	
Histoire universelle 125, 124, 123	
Histoire : temps modernes 224	
Histoire : antiquité 204	
Histoire : moyen âge 202, 203, 204	
Histoire : universelle 221	
Géographie 222 - 223	
Histoire : temps modernes 323	
Histoire : Grande-Bretagne 303	
Histoire : U.S.A. 324 - 323 - 322 - 321	
Histoire : U.S.A. (approfondie) 425	
Histoire : contemporaine (Europe) 425	
Histoire : Grande-Bretagne (approf.) 404	
Histoire : Amérique latine 403	
Histoire : Extrême-Orient 403	
Histoire : Universelle, XX ^e S. 402	
Histoire : Etats-Unis, XX ^e S. 402	
Civisme 403	
Sociologie 403	
Sciences économiques 403	

Branches et niveaux de cours	
Sciences	
Biologie 123 - 124 - 1	
Radio amateur 1	
Biologie 224 - 223 - 2	
Chimie 2	
Biologie 323 - 322 - 3	
Electronique 3	
Physique 3	
Chimie 324 - 3	
Chimie (séminaire) 4	
Physique 425 - 424 - 4	
Sciences 4	
Langues	
Latin 124 - 123 - 1	
224 - 223 - 2	
324 - 3	
424 - 4	
Allemand 124 - 1	
224 223 223	
324 - 3	
42	
Russe 12	
22	
32	
Français 125 124 12	
224C 224 223C 22	
324C 324 323C 323 324	
424C 423	
Espagnol 125 124 123 12	
224 224C 223 223C 22	
324 324= 323 32	
424 42	
Arts	
Dessin-peinture 12	
Travail manuel 12	
Dessin-peinture 229C 22	
Travail manuel 22	
Dessin-peinture 32	
Céramique 32	
Joaillerie 32	
Histoire de l'art 32	
Céramique 42	
Joaillerie 42	
Peinture-dessin 42	

Observations

ence entre les niveaux 2, 3 de moins dans l'accélération; la méthode d'enseignement approfondissement.

iveau 2, les explications sont; beaucoup d'applications; incursions dans les domaines

iveau 3, l'étudiant doit plus travailler lui-même: théorie plus riche, enrichissement.

iveau 4, les concepts sont traités; étude plus approfondie; incursions dans les domaines.

me manuel est employé pour niveaux et tous les élèves étudient le même chapitre en même temps, permet, à tout moment, le passage d'un niveau à l'autre.

estres de mathématiques sont; toutefois, 2 de ceux-ci doivent être consacrés à la comptabilité et l'arithmétique commerciale.

estres de « Social Studies » obligatoires; 2 de ceux-ci doivent être consacrés à l'histoire des Etats-

Branches et niveaux de cours

Observations

Sciences	
Biologie	123 - 124 - 125
Radio amateur	123
Biologie	224 - 223 - 222
Chimie	225
Biologie	323 - 322 - 321
Electronique	323
Physique	325
Chimie	324 - 323
Chimie (séminaire)	426
Physique	425 - 424 - 423
Sciences	409

Les élèves qui désirent continuer des études dans les Facultés de Sciences des Universités sont invités à suivre 6 semestres de sciences et 8 de mathématiques.

Langues	
Latin	124 - 123 - 122 224 - 223 - 222 324 - 323 424 - 423
Allemand	124 - 123 224 223 223C 324 - 323 426
Russe	124 224 324
Français	125 124 123 224C 224 223C 223 324C 324 323C 323 324= 424C 423C
Espagnol	125 124 123 122 224 224C 223 223C 222 324 324= 323 322 424 423

Aucun cours de langue n'est obligatoire. Les étudiants sont cependant fort encouragés à les suivre s'ils se destinent aux études supérieures. On pense que, très prochainement, les universités exigeront à l'entrée que l'étudiant ait étudié au moins une langue étrangère pendant 3 ans.

C indique qu'il s'agit de la continuation d'un cours précédent.
= cours de conversation.

Arts	
Dessin-peinture	129
Travail manuel	121
Dessin-peinture	229C 229
Travail manuel	221
Dessin-peinture	329
Céramique	329
Joaillerie	329
Histoire de l'art	329
Céramique	429
Joaillerie	429
Peinture-dessin	429

Tous ces cours peuvent être suivis à raison de 2 périodes par jour. (= cours « majeur » : 10 credits) ou de 1 période par jour (= cours « mineur » : 5 credits).

Toutefois, Histoire de l'art 329 est toujours un « majeur ».

Branches et niveaux de cours		Observations
Commerce		Objectifs poursuivis par ce cours : 1. Préparation aux études supérieures de sciences commerciales ; 2. Formation générale ; 3. Préparation pour les étudiants qui veulent travailler part-time dans le commerce pendant leurs études universitaires ; 4. Préparation professionnelle pour les élèves qui ne feront pas d'études supérieures ; 5. Possibilité pour les étudiants qui cherchent leur voie, de voir si la carrière commerciale pourrait les intéresser. Le 422 = une réunion de 30' tous les matins à l'école ; le reste consiste en stages pratiques (20 h/semaine).
Commerce général	122 123	
Commerce général	222	
Comptabilité	223	
Le point de vue du consommateur		
Vente	302	
Publicité	303	
Organisation commerciale	303	
Sténographie	323	
Dactylographie	303	
Pratique du bureau	322	
Secrétariat	423	
Sténographie	423	
Droit commercial	423 403	
Pratique commerciale	422	
Automobile		Une loi, passée en 1955, rend ce cours obligatoire dans toutes les High Schools de l'Illinois. (3 « credits »).
Théorie et pratique (pilotage)	209	
Economie domestique		Destiné principalement aux jeunes filles ; toutefois : — les garçons s'intéressant à la restauration ou à l'hôtellerie peuvent suivre les cours d'alimentation ; — les garçons s'intéressant à l'architecture peuvent suivre les cours de décoration intérieure.
Alimentation	109	
Vêtement	109	
Alimentation	209	
Vêtement	209	
Décoration intérieure	329	
Cours techniques		— Le but poursuivi n'est pas de donner une formation professionnelle, mais de faire de l'étudiant « un consommateur intelligent des produits de l'industrie » (p. 64). — Chacun de ces cours rapporte 5 ou 10 « credits » selon le temps qu'on leur consacre. — Les universités reconnaissent ces « credits » ; de plus en plus, elles souhaitent que les étudiants acquièrent une formation technique.
Dessin industriel	123 122	
Bois	123 121	
Electricité	103	
Dessin (architecture)	223	
Bois	222	
Métallurgie générale	223	
Dessin (mécanique)	323	
Métallurgie générale	323	
Notions d'architecture	323	
Lecture de plans	302	

Branches et niveaux de cours
Musique
Chant
Orchestre
Cours d'harmonie
Appréciation musicale
Chant
Petit groupe vocal
Appréciation musicale
Composition
Madrigaux (groupe vocal)
Composition
Chorale (sélection de 32 étud.)
Opéra
Grande chorale
Harmonie (cadets - groupe de sorties - groupe d'honneur)
Orchestre (seniors)
Ensemble instrumental
Piano
Education physique - Sports -
Art oratoire - Art dramatique
Photographie
Croix-Rouge
L'étude de chaque élève a
Nous avons dit que l'école s'occupe immédiatement une place, à côté de la formation intellectuelle que nous venons de mentionner.
La majorité des étudiants viennent de six écoles publiques des alentours.
Dès le mois de janvier les <i>Schools</i> fournissent aux six écoles publiques des alentours, des renseignements, d'aptitudes et de caractéristiques des enfants qui ont manifesté l'intérêt pour les études. Les élèves des écoles de New Trier, les élèves des écoles de Winnetka, en avril, pour y s'inscrire.

Observations

poursuivis par ce cours :
 ration aux études supérieures
 sciences commerciales ;
 ation générale ;
 ration pour les étudiants qui
 nt travailler part-time dans le
 erce pendant leurs études uni-
 aires ;
 ration professionnelle pour les
 s qui ne feront pas d'études
 ieures ;
 ilité pour les étudiants qui
 ent leur voie, de voir si la
 re commerciale pourrait les
 sser.
 = une réunion de 30' tous les
 l'école ; le reste consiste en
 atiques (20 h/semaine).

oi, passée en 1955, rend ce
 ligatoire dans toutes les High
 de l'Illinois. (3 « credits »).

principalement aux jeunes
 utefois :
 arçons s'intéressant à la res-
 ion ou à l'hôtellerie peuvent
 les cours d'alimentation ;
 arçons s'intéressant à l'archi-
 e peuvent suivre les cours de
 ation intérieure.

t poursuivi n'est pas de don-
 ne formation professionnelle,
 de faire de l'étudiant « un con-
 ateur intelligent des produits
 ndustrie » (p. 64).
 n de ces cours rapporte 5 ou
 redits » selon le temps qu'on
 onacre.
 niversités reconnaissent ces
 its » ; de plus en plus, elles
 itent que les étudiants acquiè-
 ne formation technique.

<i>Branches et niveaux de cours</i>	<i>Observations</i>	
<i>Musique</i>	Il s'agit soit d'une initiation pour l'amateur, soit d'études approfondies pouvant préparer aux carrières musicales.	
Chant		129
Orchestre		129
Cours d'harmonie		129
Appréciation musicale		129
Chant		229
Petit groupe vocal		229
Appréciation musicale		229
Composition		329
Madrigaux (groupe vocal)		429
Composition		429
Chorale (sélection de 32 étud.)		029
Opéra		029
Grande chorale		029
Harmonie (cadets - groupe de sorties - groupe d'honneur)		029
Orchestre (seniors)		029
Ensemble instrumental		029
Piano		029

Education physique - Sports - Danse

Art oratoire - Art dramatique

Photographie

Croix-Rouge

L'étude de chaque élève avant son entrée à New Trier.

Nous avons dit que l'école recueille des informations détaillées sur ses futurs élèves afin que ceux-ci puissent trouver immédiatement une place, au moins provisoire, dans le système complexe que nous venons de décrire. Voici comment on procède.

La majorité des étudiants qui fréquentent New Trier proviennent de six écoles publiques et de sept écoles confessionnelles des alentours.

Dès le mois de janvier, les services de testing de la *High School* fournissent aux six écoles publiques des tests d'intelligence, d'aptitudes et de connaissances à administrer aux enfants qui ont manifesté l'intention de s'inscrire aux cours de New Trier. Les élèves des écoles confessionnelles viendront à Winnetka, en avril, pour y subir des épreuves similaires.

Chaque école doit en outre remplir une fiche individuelle qui porte, outre les résultats des tests mentionnés, un tableau résumant l'appréciation générale des professeurs¹ :

APPRECIATIONS GLOBALES

Intelligence	1	2	3	4	5
Application	1	2	3	4	5
Sens des responsabilités	1	2	3	4	5
Conduite à l'école	1	2	3	4	5
Qualités de chef	1	2	3	4	5

APPRECIATIONS SPECIALES

Anglais	1	2	3	4	5
Mathématiques :					
Raisonnement	1	2	3	4	5
Connaissances de base	1	2	3	4	5
« Social Studies »	1	2	3	4	5
Sciences	1	2	3	4	5

Code : 1 = supérieur ; 2 = au-dessus de la moyenne ; 3 = moyen ; 4 = en dessous de la moyenne ; 5 = pauvre. Le chiffre retenu est encadré.

Après l'administration des tests, le directeur du service de *counseling* de New Trier rencontre les conseillers des différentes écoles inférieures afin de recueillir des renseignements complémentaires concernant non seulement l'instruction, mais aussi l'histoire, la famille, la santé des futurs étudiants.

Ensuite, chacun de ceux-ci est interviewé dans son école. L'entretien porte d'abord sur les talents (arts, sports, bricolage) et sur les goûts (branches préférées), puis sur la famille, l'enquêteur tâchant d'identifier les problèmes éventuels. Une question choc termine l'entrevue.

Au mois de mai, parents et élèves — qui ont déjà pu étudier le programme des cours — sont invités à assister à une réunion d'information à la High School. Les inscriptions aux différents cours sont recueillies.

La direction de New Trier possède alors assez d'éléments pour constituer des groupes d'environ 30 étudiants qui, pendant toutes leurs études, seront confiés au même conseiller pédagogique et le rencontreront chaque matin. Dans ces groupes, l'hétérogénéité est systématiquement recherchée, tant en ce qui concerne les aptitudes que les origines socio-économiques. Toutefois, les sexes sont séparés, le conseiller étant toujours du même sexe que ses étudiants. De plus, on s'efforce de donner

¹ Cf. NEW TRIER HIGH SCHOOL, *Test and Personal Data Card*.

une personnalité équilibrée ment que trop d'adolescents réunis.

Le jour de la rentrée, n... res sont organisés, mais c... qu'il doit faire : il sait, par... doit se joindre au groupe... groupe supérieur ; il sait qu... déjà quels livres, quels ca... acheter et on lui a fait conn... cun de ces articles.

Terminons ces indicati... en signalant qu'environ 15... les parents sont invités à u... treront le conseiller de leurs... les raisons du classement a... cation des manifestations ex...

La semaine suivante, l... l'école avec leur fils et suiv... à ses côtés (la journée est, à... midi et la soirée). On devin... rience facilite la compréh...

Ayant commencé ses é... ces conditions excellentes... jour, par son conseiller.

Les conseillers pédagogique

Grâce à ses conseiller... School réussit ce véritable t... ment à guider efficacement... le dédale des programmes... plan psychologique que su...

Le *counseling* est assur... 126 conseillers pédagogiqu... 1/5 de leur temps à cett... 10 conseillers full-time dir... 6 psychologues full-time.

Au cours de la réunion... pe de 30 étudiants dont... rencontre son conseiller po... ne, de carrière et, en géné... majorité. L'atmosphère dé... une excellente transition... commencer.

...tir une fiche individuelle
...s mentionnés, un tableau
... professeurs¹ :

OBBALES

1	2	3	4	5
1	2	3	4	5
1	2	3	4	5
1	2	3	4	5
1	2	3	4	5

ECIALES

1	2	3	4	5
1	2	3	4	5
1	2	3	4	5
1	2	3	4	5
1	2	3	4	5

...sus de la moyenne ; 3 =
...e ; 5 = pauvre. Le chiffre

...e directeur du service de
...conseillers des différen-
...illir des renseignements
...ement l'instruction, mais
...futurs étudiants.

...interviewé dans son école.
...s (arts, sports, bricolage)
...puis sur la famille, l'en-
...mes éventuels. Une ques-

... qui ont déjà pu étudier
...s à assister à une réunion
...criptions aux différents

...e alors assez d'éléments
...30 étudiants qui, pendant
...même conseiller pédago-
...atin. Dans ces groupes,
...e cherchée, tant en ce qui
...ines socio-économiques.
...conseiller étant toujours du
...s, on s'efforce de donner

...l Data Card.

une personnalité équilibrée à chaque groupe en évitant notam-
ment que trop d'adolescents « à problèmes » ne s'y trouvent
réunis.

Le jour de la rentrée, non seulement les cours et les horai-
res sont organisés, mais chaque étudiant sait exactement ce
qu'il doit faire : il sait, par exemple, que pour telle branche, il
doit se joindre au groupe des moyens et pour telle autre, au
groupe supérieur ; il sait qu'un conseiller l'attend ; il sait même
déjà quels livres, quels cahiers, quel équipement sportif il doit
acheter et on lui a fait connaître d'avance le prix exact de cha-
cun de ces articles.

Terminons ces indications concernant les débuts scolaires
en signalant qu'environ 15 jours après la rentrée des classes,
les parents sont invités à une première réunion où ils rencon-
treront le conseiller de leurs enfants et où ils seront éclairés sur
les raisons du classement aux différents niveaux, sur la signifi-
cation des manifestations extra-scolaires, etc.

La semaine suivante, les pères et les mères se rendent à
l'école avec leur fils et suivent une journée de classe complète
à ses côtés (la journée est, à cette occasion, répartie sur l'après-
midi et la soirée). On devine aisément combien une telle expé-
rience facilite la compréhension entre l'école et la famille.

Ayant commencé ses études secondaires supérieures dans
ces conditions excellentes, l'étudiant sera suivi, jour après
jour, par son conseiller.

Les conseillers pédagogiques.

Grâce à ses conseillers pédagogiques, la New Trier High
School réussit ce véritable tour de force qui consiste, non seule-
ment à guider efficacement chacun de ses 3 740 étudiants dans
le dédale des programmes, mais aussi à les aider tant sur le
plan psychologique que sur le plan social et médical.

Le *counseling* est assuré par :

- 126 conseillers pédagogiques part-time, professeurs consacrant
1/5 de leur temps à cette mission.
- 10 conseillers full-time dirigeant les précédents.
- 6 psychologues full-time.

Au cours de la réunion quotidienne du matin, chaque grou-
pe de 30 étudiants dont nous avons signalé la constitution
rencontre son conseiller pour discuter de problèmes de discipli-
ne, de carrière et, en général, de tout ce qui peut intéresser la
majorité. L'atmosphère détendue de ces rencontres constitue
une excellente transition vers le travail ardu qui va bientôt
commencer.

L'étudiant peut aussi consulter individuellement son conseiller, chaque fois qu'il le désire.

De plus, pendant l'année, le conseiller rendra au moins une visite personnelle à la famille de chaque élève appartenant à son groupe.

Il va sans dire qu'une mission aussi délicate ne s'improvise pas : elle réclame des hommes ouverts aux problèmes de la jeunesse et spécialement formés pour l'aider.

Lors du recrutement des professeurs, le conseil d'administration de l'école attache une importance toute particulière aux qualités intellectuelles et morales indispensables au parrainage des étudiants et donne la priorité à ceux qui les possèdent.

Pendant sa première année de fonction, le conseiller doit assister à 50 réunions de formation d'environ 1/2 heure chacune et, après chaque cycle de 4 ans, il est soumis à un nouvel entraînement.

La préparation de base se déroule de la façon suivante :

I. Avant la rentrée des classes :

Un certain nombre de séances sont d'abord consacrées à l'étude de l'administration de l'école. Ensuite, le conseiller reçoit en communication toutes les informations recueillies sur ses futurs élèves.

Enfin, quelques jours avant la fin des vacances, le directeur des études de première année réunit les « freshman helpers », c'est-à-dire des étudiants de dernière année qui vont aider le conseiller dans sa tâche. Pendant les 9 premières semaines, l'assistant participera à toutes les réunions matinales. Pendant les 9 semaines suivantes, il ne fournira plus que deux prestations par semaine ; par après, il n'interviendra plus que sur demande du conseiller.

II. Pendant l'année scolaire.

Sous la présidence d'un directeur spécialisé, la formation se poursuit lors de causeries familières faites par les directeurs des départements (anglais, mathématiques, sciences, etc.), par les chefs des différents services (directeurs du testing, psychologue, bibliothécaire, médecin-assistant social, délégué à l'association des parents) et par les adultes responsables des clubs étudiantins.

Le directeur des conseillers convoque aussi certaines réunions spéciales pour envisager des problèmes nouveaux concernant l'admission d'élèves, la révision des programmes, etc.

Enfin, en avril, une conférence de clôture permet de dresser le bilan d'activité de l'année qui touche à sa fin et d'établir les premières prévisions pour l'année suivante.

AGAZZI, A., *Les aspects pédagogiques de l'Europe*, C.C.C., 1967.

BACHER, F., *La normalisation*, numéro spécial de la revue 1969, 131-156.

BACHER, F., *L'évaluation des résultats*, in *Le Travail humain*, 1968.

BLOCK, J., *Mastery Learning*, New York, 1968.

BLOOM, B., HASTING and MADSEN, *Student Learning*, New York, 1968.

BONNARDEL, R., *Application de la méthode à l'étude de la notation des résultats*, 1946, 130-139.

BRITTON, J., *Experimental Marking of Year-Olds*, in *Educational Research*, 1968.

BRITTON, J., MARTIN, N. et ROBERTS, *An Account of an Experiment*, 1968.

BRUNELLE, L., *Pourquoi des examens*, 1968.

CHOPPIN, B. et PURVES, A., *Choice items dealing with the teaching of English*, 3, 1, 1968.

DEMANGEON, M. et LARCEBEAUX, *BINOP*, 1958, 14, 131-156.

Docimologie et Education, numéros 2-3, 1969, 166 p.

EDGEWORTH, F.V., *The Statistician*, *Stat. Society*, 1888, 51, 599-603.

Examen des examens, N° spécial, 1970.

Examinations Bulletins, Londres, 1970.

N° 1. *The Certificate of Competence for teachers and examiners*, 1964.

N° 2. *The C.S.E. : Examinations*, 1964.

N° 3. *The C.S.E. : An introduction*, 1964.

N° 4. *The C.S.E. : An introduction*, 1964.

N° 5. *The C.S.E. : School-leaving*, 1964.

N° 6. *The C.S.E. : Examinations*, 1964.

N° 7. *The C.S.E. : Examinations*, 1964.

BIBLIOGRAPHIE

- AGAZZI, A., *Les aspects pédagogiques des examens*, Strasbourg, Conseil de l'Europe, C.C.C., 1967.
- BACHER, F., La normalisation de la notation, in *Docimologie et Education*, numéro spécial de la revue « Les Sciences de l'Education », n° 2-3, 1969, 131-156.
- BACHER, F., L'évaluation des résultats scolaires au niveau de l'école moyenne, in *Le Travail humain*, 1965, 28, 219-230.
- BLOCK, J., *Mastery Learning*, New York, Holt, Rinehart and Winston, 1970.
- BLOOM, B., HASTING and MADAUS, *Formative and Summative Evaluation of Student Learning*, New York, Mc Graw-Hill, 1970.
- BONNARDEL, R., Application de la méthode d'analyse factorielle de Thurstone à l'étude de la notation des copies d'examens, in *Le Travail humain*, VIII, 1946, 130-139.
- BRITTON, J., Experimental Marking of English Composition Written by Fifteen-Year-Olds, in *Educational Review*, Birmingham, Vol. 16.1, 1963, 17-23.
- BRITTON, J., MARTIN, N. et ROSEN, H., *Multiple Marking of English Composition, An Account of an Experiment*, London, H.M.S.O., 1966.
- BRUNELLE, L., *Pourquoi des examens ?* Paris, Société des Editions rationalistes, 1968.
- CHOPPIN, B. et PURVES, A., A comparison of open-ended and multiple choice items dealing with literary understanding, in *Research in the Teaching of English*, 3, 1, 1969, 15-24.
- DEMANGEON, M. et LARCEBEAU, S., Une expérience de correction multiple, in *BINOP*, 1958, 14, 131-156.
- Docimologie et Education*, numéro spécial de la revue *Les sciences de l'Education*, 2-3, 1969, 166 p.
- EDGEWORTH, F.V., The Statistics of Examinations, in *Journal of the Royal Stat. Society*, 1888, 51, 599-635.
- Examen des examens*, N° spécial des *Cahiers de pédagogie*, 92, septembre 1970.
- Examinations Bulletins*, Londres, H.M.S.O.
- N° 1. *The Certificate of secondary education : some suggestions for teachers and examiners*, 1963.
- N° 2. *The C.S.E. : Experimental examinations - Mathematics*, 1964.
- N° 3. *The C.S.E. : An introduction to some techniques of examining*, 1964.
- N° 4. *The C.S.E. : An introduction to objective-types examinations*, 1964.
- N° 5. *The C.S.E. : School-based examinations*, 1965.
- N° 6. *The C.S.E. : Experimental examinations : Technical drawing*, 1965.
- N° 7. *The C.S.E. : Experimental examinations - Mathematics 2*, 1965.

- N° 8. *The C.S.E. : Experimental examinations : Science*, 1965.
- N° 9. *The C.S.E. : Trial examinations : Home economics*, 1966.
- N° 10. *The C.S.E. : Experimental examinations : Music*, 1966.
- N° 11. *The C.S.E. : Trial examinations - Oral English*, 1966.
- N° 12. *Multiple marking of English compositions*, 1966.
- N° 13. *The C.S.E. : Trial examinations : Handicraft*, 1966.
- N° 14. *The C.S.E. : Trial Examinations - Geography*, 1966.
- N° 15. *Teachers' experience of school based examining (English and Physics)*, 1967.
- N° 16. *The C.S.E. : Trial examinations - Written English*, 1967.
- N° 17. *The C.S.E. : Trial examinations - Religious knowledge*, 1967.
- N° 18. *The C.S.E. : The place of the personal topic - History*, 1968.
- N° 19. *C.S.E. : Practical work in science*, 1969.
- HARTOG, P. and RHODES, E.C., *An Examination of Examinations*, London, McMillan, 1936.
- HARTOG, P., *The Marking of English Essays*, London, McMillan, 1941.
- HINTON, E.M., *An Analytical Study of the Qualities of Style and Rhetoric Found in English Compositions*, New York, 1940.
- HOTYAT, F., *Les examens*, Paris, Bourrelier, 1962.
- LAUGIER, H. et SCHREIDER, E., Recherche docimologique sur un examen de l'enseignement supérieur, in *Biotypologie*, 1958, 19, n° 2, 61-72.
- LAUWERYS, J.A. et SCANLON, D.G., Ed., *Examinations*, The World Year Book of Education, 1969, London, Evans, 1969.
- MATHER, D., FRANCE, N. et SARE, G., *The C.S.E., A Handbook for Moderators*, London, Collins, 1965.
- McINTOSH, D., WALKER, D. and McKAY, D., *The Scaling of Teachers' Marks and Estimates*, Edinburgh, Oliver and Boyd, 1962, 2^e éd.
- JOINT MATRICULATION BOARD, *The Marking of Scripts in Advanced Level History*, Universities of Manchester, Liverpool, Leeds, Sheffield and Birmingham, 1964.
- OTTER, H.S., *A Functional Language Examination*, Oxford Univ. Press, 1968.
- PASSERON, J.C., Sociologie des examens, in *Education et Gestion*, 1970, 2, 6-16.
- PIDGEON, D. et YATES, A., *An Introduction to Educational Measurement*, Londres, Routledge et Kegan Paul, 1968.
- PIERON, H., *Examens et docimologie*, Paris, P.U.F., 1963.
- PIOBETTA, J.B., *Examens et Concours*, Paris, P.U.F., 1943.
- REMONDINO, C., Recherche sur les systèmes numériques d'évaluation scolaire, in *Le Travail humain*, 1965, 18, 3-4, 263-265.

- Reports of the Secondary School Examinations*, 1947, (1st) - 1964 (8st).
- REUCHLIN, M., *L'orientation de l'enseignement de l'Europe*, C.C.C., 1966.
- REUCHLIN, M. et BACHER, F., Les examens, in *Revue française de psychologie*, 1966, 15, 1, 1-10.
- ROT, N. et BUJAS, Z., Les distributions des résultats des examens, in *Le Travail humain*, 1959, 22, 1, 1-10.
- THORNDIKE, R.L., *Marks and Grades*, *Journal of Educational Research*, 1914, 27, 1, 1-10.
- Symposium sur la docimologie de Psychologie (Rome, 1959), juin 1959.
- VALENTINE, C.W., *The Reliability of Examinations*, 1932.
- VERNON, P., *Secondary School Examinations*, 1966.
- WALKER, A.S., *Pupils' School Examinations*, 1966.
- WISEMAN, S., The Marking of Examinations, in *British Journal of Educational Psychology*, 1961, 31, 1, 1-10.
- WISEMAN, S., *Examinations and the Teacher*, 1961.
- WRIGLEY, J., The Relative Efficiency of Examinations, in *Journal of Educational Psychology*, 25, 1955, 1, 1-10.
- YATES, A. and PIDGEON, D., *An Introduction to Educational Measurement*, N.F.E.R., 1957.

ions : Science, 1965.
me economics, 1966.
ons : Music, 1966.
English, 1966.
ions, 1966.
dicraft, 1966.
graphy, 1966.
used examining (English and
ritten English, 1967.
religious knowledge, 1967.
nal topic - History, 1968.
89.
on of Examinations, London,
ondon, McMillan, 1941.
alities of Style and Rhetoric
1940.
mologique sur un examen de
1958, 19, n° 2, 61-72.
ations, The World Year Book
E., A Handbook for Modera-
e Scaling of Teachers' Marks
1962, 2^e éd.
of Scripts ni Advanced Level
pool, Leeds, Sheffield and
on, Oxford Univ. Press, 1968.
ducation et Gestion, 1970, 2,
o Educational Measurement,
F., 1963.
F., 1943.
numériques d'évaluation sco-
3-265.

- Reports of the Secondary School Examinations Council*, Londres, H.M.S.O., 1947, (1st) - 1964 (8st).
- REUCHLIN, M., *L'orientation pendant la période scolaire*, Strasbourg, Conseil de l'Europe, C.C.C., 1964.
- REUCHLIN, M. et BACHER, F., L'appréciation des élèves par leurs professeurs, in *Revue française de Pédagogie*, 1968, 2, 19-25.
- ROT, N. et BUJAS, Z., Les distributions de notes scolaires comparées aux distributions des résultats obtenus dans les tests de connaissances, in *Le Travail humain*, 1959, 22, 19-26.
- THORNDIKE, R.L., Marks and Marking Systems, in R.L. EBEL, *Encyclopedia of Educational Research*, Londres, McMillan, 1969, pp. 759-766.
- Symposium sur la docimologie, XIII^e congrès de l'Association Internationale de Psychologie (Rome, 1958), in *Le Travail humain*, XXII, 1-2, janvier-juin 1959.
- VALENTINE, C.W., *The Reliability of Examinations*, University of London Press, 1932.
- VERNON, P., *Secondary School Selection*, Londres, Methuen, 1957.
- WALKER, A.S., *Pupils' School Records*, Newnes, Educ. Publ., 1955.
- WISEMAN, S., The Marking of English Composition in Grammar School Selection, in *British Journal of Educational Psychology*, XIX, 1949, 200-209.
- WISEMAN, S., *Examinations and English Education*, Manchester, University Press, 1961.
- WRIGLEY, J., The Relative Efficiency of Intelligence and Attainment Tests as Predictors of Success in Grammar Schools, in *British Journal of Educational Psychology*, 25, 1955, 107-116.
- YATES, A. and PIDGEON, D., *Admission to Grammar Schools*, London, N.F.E.R., 1957.

TABLE ANALYTIQUE

INTRODUCTION	7
------------------------	---

PREMIERE PARTIE

DEFINITIONS

I. Docimologie, docimastique et doxologie	13
II. Examens et concours :	
Observation et évaluation continues	14
Examens internes et examens externes	15
III. Notes et scores	16

DEUXIEME PARTIE

L'ACCUSATION ET LA DEFENSE

CHAPITRE 1. — Critique des examens	19
1. Corps étrangers dans l'éducation, au service d'une pédagogie dépassée	19
2. Anxiété et <i>stress</i>	20
3. Inégalité - injustice	21
4. L'échec, générateur d'échecs	25
5. Rupture entre enseignement et examens	26
6. Désaccord entre correcteurs	27
a) Composition française	
b) Mathématiques	
c) Médecine	
d) Divers	
e) Combien de correcteurs pour stabiliser la note ?	
7. Infidélité d'un même correcteur	
Un schéma pour continuer la recherche	32
8. Stéréotypes et effets de halo	33

9. Manque de validité	
10. Un instrument d'immobilisation	
a) Effets irréversibles de	
b) Les examens ne sont	
11. Faiblesse de beaucoup d'e	
12. Autres critiques	

CHAPITRE 2. — Défense de l'

1. La mesure rigoureuse est p	
2. Les maîtres jugent bien le	
3. Validité limitée mais réelle	
4. S'endurcir pour la vie	
5. Se situer par rapport aux	
6. Large synthèse et intégrat	
7. L'examen externe contrôlé	
8. L'examen externe, <i>feedba</i>	

CONSTR

Les grandes phases - Vue d'e

CHAPITRE 1. — L'objet et les

I. L'objet	
A. Le pronostic	
1. Tests de maturité sp	
2. Vérification des con	
3. Essai	
B. L'inventaire	
C. Le diagnostic	
II. Les objectifs	
A. Les objectifs généraux	
1. Les objectifs cogniti	

UE

7
13
14
15
16

DEFENSE

19
d'une pédagogie dé- 19
20
21
25
26
27
la note ?
32
33

9. Manque de validité	36
10. Un instrument d'immobilisme social	37
a) Effets irréversibles de la certification scolaire	37
b) Les examens ne sont pas neutres	38
11. Faiblesse de beaucoup d'expériences docimologiques	40
12. Autres critiques	40
CHAPITRE 2. — Défense de la note subjective et de l'examen	42
1. La mesure rigoureuse est peut-être impossible	42
2. Les maîtres jugent bien les élèves	44
3. Validité limitée mais réelle des examens traditionnels	46
4. S'endurcir pour la vie	47
5. Se situer par rapport aux autres	47
6. Large synthèse et intégration des connaissances	47
7. L'examen externe contrôle le professeur	47
8. L'examen externe, <i>feedback</i> pour le professeur	48

TROISIEME PARTIE

CONSTRUCTION DE L'EXAMEN

Les grandes phases - Vue d'ensemble	51
CHAPITRE 1. — L'objet et les objectifs	52
I. L'objet	52
A. Le pronostic	52
1. Tests de maturité spécifique (<i>readiness</i>)	53
2. Vérification des connaissances-clés ou notions critiques	53
3. Essai	54
B. L'inventaire	54
C. Le diagnostic	54
II. Les objectifs	56
A. Les objectifs généraux	57
1. Les objectifs cognitifs	57

a) La taxonomie de Bloom	58
b) Le modèle de J.-P. Guilford	60
2. Domaine affectif	64
B. Les objectifs spéciaux	66
Illustration	68
CHAPITRE 2. — Rédaction des questions	73
I. Observations générales	73
A. Des questions compréhensibles	74
B. Tenir compte du niveau d'information	74
C. Essayer ou prétester les questions	74
D. Calcul de la facilité des questions	75
E. Calcul de l'efficacité. Pouvoir discriminatif	75
1. Méthode simple	75
2. Méthode plus fine	76
II. Réponses ouvertes ou fermées ?	79
A. Réponses ouvertes	79
B. Réponses fermées - Questions à choix multiple	81
1. Utilité	81
2. Constituer une provision de questions	81
3. Exploiter la gamme des possibilités logiques	82
a) Question à complément simple	82
b) Association simple	83
c) Association composée	83
d) Association à terme exclu	84
e) Analyse de relations de cause à effet	84
f) Analyse d'observations	84
g) Comparaisons quantitatives	86
h) Compléments groupés	87
4. Calcul de l'efficacité des distracteurs	87
5. Critiques et réfutation partielle	88
a) Une objectivité trompeuse	88

b) Choix « corrects »	
c) Un jeu de hasard	
d) Acrobatie mentale	
e) Inconvénients incertains	
C. En guise de conclusion	
CHAPITRE 3. — La notation	
I. Un préambule indispensable	
A. La courbe de Gauss, imagerie	
B. La courbe de Gauss, image	
C. L'écart type ou sigma, image	
1. Signification	
2. Estimation rapide	
D. La concentration des résultats	
E. Courbe de Gauss voulue	
F. Comment savoir si une courbe est normale	
II. La notation subjective : l'évaluation	
A. Introduction	
B. Nature et faiblesses des notations	
C. Utilité	
D. Construction	
1. Combien de degrés de liberté	
2. Définir l'objet de l'évaluation	
E. Utilisation	
1. Combien d'élèves par question	
— Elève comparé à lui-même	
— Elèves comparés entre eux	
2. Lutter contre la contamination	
F. Comment synthétiser les résultats	

.	58
.	60
.	64
.	66
.	68
.	73
.	73
.	74
.	74
.	74
.	75
f	75
.	75
.	76
.	79
.	79
multiple	81
.	81
.	81
iques	82
.	82
.	83
.	83
.	84
t	84
.	84
.	86
.	87
.	87
.	88
.	88

b) Choix « corrects » contestables	89
c) Un jeu de hasard	89
d) Acrobatie mentale	90
e) Inconvénients incertains	91
C. En guise de conclusion : un compromis	91
CHAPITRE 3. — La notation	93
I. Un préambule indispensable : la courbe de Gauss	93
A. La courbe de Gauss, image de la probabilité	93
B. La courbe de Gauss, image de l'enseignement non individualisé	95
C. L'écart type ou sigma, indice précieux	96
1. Signification	96
2. Estimation rapide	97
D. La concentration des résultats autour de la moyenne	100
E. Courbe de Gauss voulue par les maîtres	101
F. Comment savoir si une distribution est normale ?	101
II. La notation subjective : l'échelle d'évaluation	105
A. Introduction	105
B. Nature et faiblesses des échelles d'évaluation	106
C. Utilité	107
D. Construction	107
1. Combien de degrés ?	108
2. Définir l'objet de l'évaluation	108
E. Utilisation	111
1. Combien d'élèves par échelon ?	111
— Elève comparé à lui-même	111
— Elèves comparés entre eux	113
2. Lutter contre la contamination et la tendance centrale	114
F. Comment synthétiser les évaluations ?	115

G. Un cas particulier : la notation de la composition française	118
1. Quatre méthodes d'évaluation	118
a) La méthode de l'impression générale	118
b) Les échelles de spécimens	120
c) La méthode analytique	121
Quelles qualités observer ?	122
Un exemple	124
d) La méthode des comptages de fréquences	127
2. Plusieurs sujets au choix ?	128
3. Conclusion	129
III. La notation objective	130
IV. L'étalonnage ou mesure de la position relative	130
A. Le centilage	131
B. Les notes standard ou notes Z	134
C. L'échelle à cinq classes	136
D. L'échelle à neuf classes (stanines)	137
CHAPITRE 4. — Contrôle de la fidélité de l'examen	138
1. Eviter toute ambiguïté dans les questions	138
2. Des questions en nombre suffisant	139
3. Un contrôle mathématique	139
a) La méthode pairs-impairs	139
b) Deux formes parallèles	140
4. Répétition de la notation	140
CHAPITRE 5. — Contrôle de la validité	141
I. La validité du contenu	141
II. La validité prédictive	143

QUATRIEME PARTIE

LES PROCEDURES DE MODERATION

CHAPITRE 1. — Position du problème	147
1. Définition	147

2. Modérer n'est pas cap	
3. Modération volontaire c	
4. La modération commen	
5. Pas de comparabilité	
6. Peut-on se fier aux tes	
CHAPITRE 2. — Quelques sys	
I. La formule la plus libérale	
II. Système imposé de modér	
connaissances	
III. Modération volontaire par a	
IV. Un système simple au servi	
V. En Angleterre, un système	
A. Préliminaires	
B. Les professeurs notent	
C. Correction par les mod	
1° Contrôle de la sévéri	
2° Contrôle de la discri	
3° Contrôle de la con	
D. Correction des échantill	
E. Comment ajuster des no	
1° Ajustement du média	
2° Ajustement du média	
VI. La note de fin d'année	
Travail de l'année + trava	

LE MYTHE D

CHAPITRE 1. — Evolution de la	
I. La courbe des aptitudes	
II. La courbe des connaissances	

Composition française	118
.	118
.	118
.	120
.	121
.	122
.	124
équences	127
.	128
.	129
.	130
ative	130
.	131
.	134
.	136
.	137
men	138
.	138
.	139
.	139
.	139
.	140
.	140
.	141
.	141
.	143

E

MODERATION

.	147
.	147

2. Modérer n'est pas caporaliser	148
3. Modération volontaire ou imposée ?	149
4. La modération commence au début de l'année scolaire	149
5. Pas de comparabilité sans fidélité élevée	150
6. Peut-on se fier aux tests ?	151
CHAPITRE 2. — Quelques systèmes de modération	153
I. La formule la plus libérale : le système suédois	153
II. Système imposé de modération par branche à partir d'un test de connaissances	156
III. Modération volontaire par appel à une banque de questions	156
IV. Un système simple au service de la sélection	157
V. En Angleterre, un système de modération complet	160
A. Préliminaires	160
B. Les professeurs notent les examens	161
C. Correction par les modérateurs	161
1° Contrôle de la sévérité	163
2° Contrôle de la discrimination	164
3° Contrôle de la conformité	165
D. Correction des échantillons restants et contrôles	165
E. Comment ajuster des notes discordantes ?	168
1° Ajustement du médian	168
2° Ajustement du médian et de l'écart type	168
VI. La note de fin d'année.	
Travail de l'année + travaux pratiques + test	171

CINQUIEME PARTIE

LE MYTHE DE LA COURBE DE GAUSS

CHAPITRE 1. — Evolution de la courbe des connaissances	180
I. La courbe des aptitudes	180
II. La courbe des connaissances	181

CHAPITRE 2. — Une pédagogie de la courbe en j	184
CHAPITRE 3. — La théorie de l'évaluation formative	186
I. Jalonner l'ascension du savoir	187
II. Guider l'élève	188
III. Limites pratiques du système	189
CONCLUSIONS ET RECOMMANDATIONS	191

ANNEXES

I. Etude comparée d'une question d'examen présentée selon la méthode traditionnelle et selon la méthode par questions à choix multiple	199
— Méthode traditionnelle	199
— Méthode par choix multiple	200
III. Composition de langue maternelle - Angleterre, G.C.O., 1967 - Fin du secondaire	205
— Questions	205
Epreuve I	205
Epreuve II	206
— Consignes pour la correction	208
Epreuve I	208
Epreuve II	210
III. Exemple d'enseignement semi-individualisé.	
— La New Trier Township High School	214
— L'individualisation des programmes	215
— L'individualisation de l'enseignement	218
— L'étude de chaque élève avant son entrée à New Trier	223
— Les conseillers pédagogiques	225

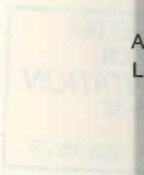
INTRODUCTION	
PREMIERE PARTIE — <i>Défini</i>	
DEUXIEME PARTIE — <i>L'accus</i>	
Chapitre 1: Critique de	
Chapitre 2: Défense de	
TROISIEME PARTIE — <i>Constr</i>	
Chapitre 1: L'objet et	
I. L'objet	
II. Les objectifs	
Chapitre 2: Rédaction d	
I. Observations g	
II. Réponses ouv	
Chapitre 3: La notation	
I. Un préambule	
II. La notation s	
III. La notation c	
IV. L'étalonnage	
Chapitre 4: Contrôle de	
Chapitre 5: Contrôle de	
QUATRIEME PARTIE — <i>Les</i>	
Chapitre 1: Position du	
Chapitre 2: Quelques s	
I. La formule s	
modération p	
sances	
II. Système impo	
III. Modération vo	
IV. Un système s	
de la sélecti	

en j	184
native	186
.	187
.	188
.	189
.	191
h présentée selon la par questions à choix	199
.	199
.	200
erre, G.C.O., 1967 - Fin	205
.	205
.	205
.	206
.	208
.	208
.	210
.	214
.	215
.	218
se à New Trier	223
.	225

TABLE DES MATIERES

INTRODUCTION	7
PREMIERE PARTIE — <i>Définitions</i>	11
DEUXIEME PARTIE — <i>L'accusation et la défense</i>	17
Chapitre 1 : Critique des examens	19
Chapitre 2 : Défense de la note subjective et des examens	42
TROISIEME PARTIE — <i>Construction de l'examen</i>	49
Chapitre 1 : L'objet et les objectifs	52
I. L'objet	52
II. Les objectifs	56
Chapitre 2 : Rédaction des questions	73
I. Observations générales	73
II. Réponses ouvertes ou fermées ?	79
Chapitre 3 : La notation	93
I. Un préambule indispensable. La courbe de Gauss	93
II. La notation subjective. Echelle d'évaluation	105
III. La notation objective	130
IV. L'étalonnage ou mesure de la position relative	130
Chapitre 4 : Contrôle de la fidélité	138
Chapitre 5 : Contrôle de la validité	141
QUATRIEME PARTIE — <i>Les procédures de modération</i>	145
Chapitre 1 : Position du problème	147
Chapitre 2 : Quelques systèmes de modération	153
I. La formule suédoise, la plus libérale - Système de modération par branche à partir d'un test de connais- sances	153
II. Système imposé sur la même base.	156
III. Modération volontaire, par appel à une banque d' <i>items</i>	156
IV. Un système simple de modération globale au service de la sélection	157
	237

V. En Angleterre. Un système de modération complet . . .	160
VI. Travail de l'année + travaux pratiques + test. Quel grade attribuer en fin d'année ?	171
CINQUIEME PARTIE — <i>Le mythe de la courbe de Gauss</i> . . .	177
Chapitre 1: Evolution de la courbe des connaissances . . .	180
Chapitre 2: Une pédagogie de la courbe en <i>J</i>	184
Chapitre 3: La théorie de l'évaluation formative	186
CONCLUSION ET RECOMMANDATIONS	191
ANNEXES	197
I. Etude comparée d'une question d'examen	199
II. Exemple de questions pour une composition en langue maternelle	205
III. Exemple d'enseignement semi-individualisé	214
BIBLIOGRAPHIE SUCCINCTE	227
TABLE ANALYTIQUE	230



modération complet . . .	160
pratiques + test. Quel e ?	171
be de Gauss	177
connaissances	180
be en J	184
formative	186
.	191
.	197
n d'examen	199
e composition en langue	205
ni-individualisé	214
.	227
.	230

ACHEVE D'IMPRIMER
LE 26 JANVIER 1971.