

Zeitschrift: Études pédagogiques : annuaire de l'instruction publique en Suisse
Band: 60/1969 (1969)

Artikel: La valeur objective des examens
Autor: Panchaud, Georges / Hensler, Hélène / Fall, Séga
DOI: <https://doi.org/10.5169/seals-115670>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. [Mehr erfahren](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. [En savoir plus](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. [Find out more](#)

Download PDF: 09.02.2026

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

La valeur objective des examens

ÉTAT DES RECHERCHES

L'institution des examens est depuis longtemps l'objet de critiques. On reproche à ce mode d'évaluation son caractère subjectif, voire arbitraire; on insiste sur l'inégalité des conditions dans lesquelles se trouveraient les candidats, soit à cause de leur plus ou moins grande émotivité, soit par le hasard qui leur attribue tel sujet plutôt que tel autre.

Aujourd'hui, on réclame leur suppression. Ils représentent, aux yeux des contestataires, le symbole de la pédagogie traditionnelle, sous la forme d'un examinateur tout-puissant, vérifiant la somme des connaissances apprises par cœur par des candidats en condition d'infériorité.

L'opinion que l'on se fait des examens est elle-même très subjective. Chacun puise ses arguments dans son expérience personnelle de candidat. De son côté, le corps enseignant est divisé. Si les uns reconnaissent la relativité des appréciations qu'ils donnent, d'autres pensent qu'à tout prendre le système est le moins mauvais des systèmes possibles.

L'une des questions qui revient le plus souvent est de savoir si les examens écrits offrent davantage de garanties que les examens oraux; chaque examinateur a sa petite théorie à ce sujet.

En définitive, la forme et l'importance des examens d'une institution scolaire découlent dans la plupart des cas d'une décision empirique prise par l'administration ou par le corps enseignant lui-même.

Ce problème n'a pourtant pas échappé à des recherches scientifiques. Quelles sont ces recherches et quelles ont été leurs conclusions, c'est ce que des circonstances particulières nous ont amenés à étudier.

Depuis 1964, un organisme dont le siège est à Genève¹, s'efforce de mettre sur pied un baccalauréat international auquel pourraient se présenter des candidats dans les diverses parties du monde. Une expérience va débuter l'année prochaine avec l'appui d'un grand nombre d'universités et de gouvernements.

L'un des problèmes majeurs auxquels les organisateurs se trouvent confrontés est celui d'assurer un niveau d'exigence égal à des sessions d'examens qui se tiendraient simultanément à Santiago du Chili, à Téhéran, New-York, Oxford, Genève, Paris, etc.

¹ Il s'agit d'une Fondation qui porte le nom d'« Office du Baccalauréat international ».

Les examens écrits peuvent se prêter à un contrôle préalable et postérieur à la passation des épreuves, mais comment assurer des conditions égales aux examens oraux? Faut-il les supprimer et ne garder que des écrits? Nous avons essayé de trouver une réponse à ces questions dans les recherches entreprises jusqu'à maintenant.

Si le problème du baccalauréat est dans cette étude au premier plan, les constatations que nous présentons s'appliquent toutefois à la plupart des situations d'examen à d'autres niveaux.

1. Objectifs du baccalauréat

Dans le passé, du moins pour l'examen que subissaient en Sorbonne les candidats originaires de divers pays d'Europe, le baccalauréat, simple épreuve orale, devait constater que l'enseignement avait atteint son but essentiel, à savoir « former des hommes capables de parler, d'argumenter et de discuter ». On exigeait des candidats se présentant à cet examen un certificat prouvant qu'ils avaient effectué des études de rhétorique. Ces certificats de scolarité ne présentant plus la garantie d'études réellement faites, à partir de 1830 on introduisit petit à petit des épreuves écrites pour augmenter l'efficacité du baccalauréat.

C'est dans cet esprit qu'il faut considérer les multiples réformes qu'a connues cet examen.

Aujourd'hui comme à l'origine, le baccalauréat vise, non seulement à sanctionner les études secondaires, mais aussi à sélectionner la fraction de jeunes gens susceptibles de poursuivre des études supérieures. En plus de ce double rôle, le baccalauréat essaie de contrôler le rendement et le niveau de l'enseignement secondaire.

C'est dans cette pluralité de fonctions qu'il faut chercher les raisons qui font douter actuellement de la valeur de ce diplôme. Ce caractère hybride entraîne l'ambiguïté du rôle de l'examineur. Et les divergences qu'on signale dans la cotation des épreuves traditionnelles résultent bien souvent de ce que chaque correcteur ou interrogateur se place sur un plan conforme à la fonction qu'il attribue au baccalauréat.

2. Les méthodes d'évaluation traditionnelles

A. L'EXAMEN ÉCRIT

L'usage de ce mode d'évaluation étant l'un des plus répandus, c'est dans l'analyse du rôle attribué à l'écrit qu'on peut trouver la justification de la préférence pour ce type d'épreuves.

a) Rôle de l'écrit

On reconnaît à l'écrit un certain nombre d'avantages :

- Il semble en effet que la perspective des épreuves écrites constitue pour l'élève une sorte d'aiguillon, de stimulant, dans l'acquisition et la maîtrise des connaissances contenues dans un programme défini.
- Il exige du candidat une réflexion profonde qui lui permette de rassembler, clarifier, ordonner et communiquer ses idées dans une rédaction satisfaisante. De plus, la difficulté de l'expression écrite est en elle-même un élément éducatif important.
- L'absence de contact direct entre examinateur et examiné diminue le poids des facteurs subjectifs dans la notation.

Dans certains pays, l'examen écrit ne constitue que la première partie de la vérification des connaissances et des aptitudes. Le sort du candidat n'est décidé qu'au terme d'autres types d'épreuves : pratiques, orales, tests, etc. Ailleurs, l'écrit a un rôle décisif puisqu'il a une fonction d'élimination, de pré-sélection : seuls les candidats ayant franchi le barrage qu'il constitue peuvent subir les autres genres d'épreuves. Parfois même une très forte moyenne à l'écrit dispense de la suite de l'examen.

La prédilection accordée à l'examen écrit est-elle justifiée ? Il s'agit de voir comment se présente concrètement la situation de l'examen écrit et dans quelle mesure elle permet la réalisation du rôle qu'on lui assigne.

b) Situation de l'écrit

La situation de l'écrit présente les caractéristiques spécifiques suivantes :

- Du fait de l'absence de contact direct avec le correcteur de sa copie, le candidat est laissé à lui-même pour exprimer ce dont il est capable. Cette situation d'isolement — en dépit de la présence d'autres personnes dans la salle d'examens — peut provoquer un état de tension, voire de panique, chez le candidat.
- L'incompréhension initiale de la question d'écrit par le candidat entraîne souvent l'attribution d'une note très basse à son travail.
- Le temps consacré à la correction d'une épreuve écrite peut être très variable et influencer la note, alors que la durée limitée de l'oral exige une appréciation immédiate.
- La présentation (écriture, style, etc.) influence le correcteur dans la notation, alors que c'est le contenu de la copie du candidat qui constitue l'indice de ses connaissances.
- L'examen écrit sous forme d'essai (dissertation) exagère l'importance de savoir exprimer une idée et sous-estime l'idée elle-même.

De ce fait, elle risque de pénaliser les candidats qui, pour une raison ou pour une autre, ne connaissent pas toutes les subtilités de la langue.

Il semble donc que ces différents facteurs de la situation d'examen écrit ne sont pas de nature à révéler de manière très objective la valeur du candidat.

c) *Objectivité de l'écrit*

Plusieurs recherches ont porté sur l'objectivité des examens écrits, fondées essentiellement sur l'analyse du degré d'accord entre examinateurs pour la note attribuée à une même copie.

L'enquête réalisée en 1930 par la commission Carnegie a révélé des écarts dans la notation d'une même copie pouvant aller jusqu'à 13 points sur 20 en français, 12 points en philosophie et version latine, 9 points en anglais et mathématiques et 8 en physique¹.

Depuis lors, de nombreuses études ont été consacrées à la correction des épreuves écrites. Il en est ressorti un certain nombre de conclusions:

1. La personnalité du correcteur intervient largement dans la réussite ou l'échec des candidats. Une typologie des examinateurs les classe grossièrement en « sévères » tendant à recourir à toute l'échelle de notes (0 à 10 par exemple) et en « indulgents » utilisant seulement une partie de cette échelle.

L'étude faite par Mialaret² sur la cotation de 17 copies d'arithmétique par 17 élèves-inspecteurs est assez significative à ce point de vue: on y voit un correcteur mettre la même note à deux copies dont la moyenne des notes attribuées par les 17 correcteurs était respectivement de 5,44 et 10,61.

2. Il existe incontestablement des matières et des sujets plus faciles à corriger que d'autres, selon que l'accord sur un barème et des critères communs peut être réalisé ou non.

3. Une copie est cotée non d'une façon absolue, d'après le savoir qui s'y trouve exprimé, mais par rapport au groupe de copies où elle se trouve. N. Rot et Z. Bujas³ ont fait corriger 100 copies par 5 professeurs, puis en ont prélevé 15, toutes cotées « bien » qu'ils font noter une seconde fois par 4 autres correcteurs. On remarque alors une nouvelle hiérarchisation de ces 15 copies:

¹ *La correction des épreuves écrites dans les examens*. Paris, La Maison du Livre, 1936.

² G. Mialaret, *Comment augmenter l'objectivité des examens et des concours*. Revue Internationale de Pédagogie, N° 3, 1958.

³ N. Rot et Z. Bujas, *Les distributions de notes scolaires comparées aux distributions des résultats obtenus dans les tests de connaissance*. Le Travail Humain, N° 22, 1959.

- 20 % sont classées « Médiocre »
- 53 % sont classées « Bien »
- 17 % sont classées « Très bien »
- 10 % sont classées « Excellent »

D'où un manque de fidélité de la correction des épreuves écrites. Un élève passable voit ses chances de réussite augmentées ou diminuées selon qu'il appartient à un lot de candidats médiocres ou brillants, selon qu'on juge sa copie après une série de copies très faibles ou au contraire excellentes. C'est ce qu'on appelle l'effet de contraste, qui se produit également à l'oral.

4. L'entente entre correcteurs se réalise plus facilement sur les bonnes ou mauvaises copies que sur les copies moyennes. Or, l'appréciation de ces copies est la plus délicate puisque le seuil d'admissibilité se trouve dans cette zone.

5. Les coefficients affectés à certaines matières en vue d'augmenter leur poids dans les résultats des examens réalisent rarement cet objectif, comme l'indiquent les recherches de Valin¹ et de Delaporte².

6. Le manque de fidélité de l'écrit se révèle également dans l'inconstance d'un même correcteur. C'est ce qu'illustrent les travaux de Laugier et Weiberg³: sur 37 compositions de sciences naturelles notées par le même professeur à trois ans et demi d'intervalle, seules 7 ont obtenu la même note les deux fois.

Ce rappel, nullement exhaustif, des conclusions formulées au terme de multiples études docimologiques, n'a pour but que de nous rendre attentifs aux difficultés des sélections par l'écrit traditionnel.

Ces études mentionnent parfois des réformes susceptibles d'améliorer la notation et d'augmenter l'objectivité de ce genre d'examen. Mais avant d'aborder ces aménagements, voyons si l'examen oral présente plus de garantie.

B. L'EXAMEN ORAL

L'examen oral traditionnel occupe une place importante dans un grand nombre de systèmes scolaires, tant à l'admission qu'au terme des études. Son maintien s'explique par le rôle spécifique qu'on lui attribue en tant que moyen d'évaluation. Cependant seuls l'analyse de la situation concrète de l'oral et l'examen de son degré d'objectivité permettent d'établir dans quelle mesure il rend possible l'expression et l'appréciation de la valeur réelle d'un candidat.

¹ E. Valin, *La valeur des examens. Etude docimologique réalisée au Liban*. Unesco. Etudes et documents d'éducation, 1961.

² Cf. H. Piéron, *Examens et docimologie*. P.U.F., Paris 1963, p. 98.

³ H. Laugier et D. Weinberg, *Le facteur subjectif dans les notes d'examen*. Année Psychologique, 1927, 28 et 1930, 31.

Nous examinerons successivement :

- a) le rôle de l'oral;
- b) la situation de l'oral;
- c) l'objectivité de l'oral.

a) *Rôle de l'oral*

Méthode d'évaluation unique ou complémentaire, moyen de repêchage ou d'élimination, l'oral est censé exercer diverses fonctions spécifiques :

- Grâce au contact direct qui s'établit entre l'interrogateur et le candidat, l'examen n'est plus dépersonnalisé : l'oral devient un moyen d'évaluation d'une grande souplesse, permettant l'adaptation de l'examineur à l'examiné, et l'inverse également. Une question peut être reformulée plusieurs fois de façon différente, alors qu'à l'écrit une mauvaise compréhension de la question par le candidat peut déterminer son échec.
- L'oral permet une observation plus complète dans un climat apparemment plus favorable : on peut apprécier la vivacité d'esprit du candidat, son habileté à se tirer d'embarras, sa façon de s'exprimer, la solidité de ses connaissances, son émotivité, etc. L'examineur, en un mot, peut se faire une idée de la valeur « globale » du candidat ¹.

b) *Situation de l'oral*

L'examen oral peut être défini comme un processus d'interaction se déroulant dans un *cadre* donné, entre deux personnes dont l'une — l'*examineur* — cherche à tirer des renseignements de l'autre — l'*examiné* — renseignements qui lui permettront d'évaluer ses aptitudes et ses connaissances relatives à un *sujet* déterminé. Le cadre, le sujet, l'examineur et l'examiné constituent les quatre dimensions principales de l'analyse de la situation de l'examen oral.

Le cadre de l'examen oral dépend de son organisation. Si l'examen est interne, administré par le professeur qui enseigne la matière de l'interrogation, ses conditions ressemblent à celles d'un simple contrôle effectué en cours d'année scolaire. Dans ce cas, il peut constituer également pour le professeur un moyen de contrôle de l'efficacité de son enseignement, ce qui influence son attitude face au candidat. S'il est externe, il se déroule dans des conditions qui ne

¹ Or, nous le verrons par la suite, c'est précisément le caractère global du jugement de l'interrogateur, où interviennent un certain nombre de facteurs étrangers au but de l'examen, qui compromet son objectivité.

sont pas familières au candidat. La présence d'autres personnes que l'examineur, anonymes ou non, a des répercussions et sur le comportement de l'examineur et sur celui de l'examiné.

Le déroulement de l'interrogation diffère selon que l'oral est indépendant ou complémentaire d'un écrit dont l'examineur connaît déjà les résultats. Si l'interrogation débute dès que la question a été posée, le résultat n'est pas le même que lorsque le candidat dispose d'un temps de préparation. Les conditions de cette préparation — matériel à disposition, durée, lieu, etc. — jouent également un rôle. Enfin, le nombre de candidats à interroger par les examinateurs et le temps qui leur est imparti peuvent déterminer l'atmosphère de l'interrogation.

Le sujet de l'examen oral constitue à proprement parler le stimulus du processus d'interaction entre l'examineur et l'examiné. C'est un des principaux déterminants du contenu et de la forme de la communication.

La question à traiter peut être imposée directement par l'examineur ou tirée au sort par le candidat. Elle constitue le plus souvent un facteur de hasard, dans la mesure où elle ne représente qu'une parcelle de la matière sur laquelle a porté la préparation du candidat. La clarté de l'énoncé de la question, sa formulation écrite ou orale, est un autre élément à prendre en considération. L'interrogation diffère également selon que la question sert à vérifier des connaissances ou met en jeu la créativité du candidat. Enfin la discipline à laquelle est rattaché le sujet à traiter est souvent déterminante pour la forme de l'interrogation :

- Pour les sciences humaines,
l'examen consiste généralement en un exposé fait par le candidat. Dans ce cas, comme pour la dissertation, une grande importance est attachée à l'aisance dans l'expression d'une part et au coefficient personnel du candidat d'autre part.
- Pour les langues,
l'examen peut prendre la forme d'un exposé ou d'un dialogue entre examinateur et examiné. Pour les langues étrangères, il sert souvent à apprécier conjointement la faculté de s'exprimer et les connaissances littéraires du candidat.
- Pour les sciences exactes,
il s'agit le plus souvent d'un problème à résoudre devant l'examineur. C'est alors la rigueur du raisonnement qui joue un grand rôle.

L'examineur. — Le rôle de la personnalité de l'examineur dans l'examen oral est renforcé du fait de l'absence quasi générale de tout accord théorique, de toute norme précise quant à la procédure

qu'il convient d'adopter dans la conduite de l'interrogation et dans son évaluation. L'attitude et le comportement de l'examineur sont infiniment variables en fonction d'une série d'éléments parmi lesquels on peut citer : la conception que l'examineur a de son rôle et de son statut, son âge, son sexe, sa formation, ses motivations, l'idée ou l'impression qu'il a du candidat avant même que celui-ci ait pris la parole.

C'est ainsi que l'examineur peut être ouvert, approuvant fréquemment son candidat et intervenant rapidement s'il est embarrassé ; parfois, il ira même jusqu'à amorcer les réponses aux questions posées. S'il est très loquace, il lui arrive d'oublier la situation d'examen et de donner un cours au candidat au lieu de l'interroger. A l'opposé, on trouve l'examineur froid et distant, se cantonnant dans une attitude passive ou intervenant uniquement pour rectifier une affirmation du candidat.

L'examineur estime parfois que la fonction de l'examen est de vérifier également les connaissances du candidat en dehors du point à traiter. Dans ce cas, il pose au candidat des questions qui l'écartent de son sujet et peuvent l'amener à perdre pied.

Ces différents types de comportement de l'examineur ainsi que sa physionomie conditionnent l'atmosphère de l'examen et impressionnent très diversement les examinés selon leur propre tempérament. La conception de l'examineur quant à la matière sur laquelle porte l'examen est un autre facteur déterminant dans l'évaluation : il établit implicitement une hiérarchie des questions posées, laquelle varie d'un examineur à l'autre pour la même matière, d'où son caractère souvent arbitraire.

Relevons encore le fait de la division de l'attention de l'interrogateur entre la conduite de l'entretien et l'appréciation des réponses. Dans ces conditions, il peut difficilement garder un souvenir assez précis du déroulement des opérations pour porter un jugement objectif au terme de l'examen. D'où l'intérêt de la présence, aux côtés de l'examineur, d'une ou plusieurs personnes fonctionnant comme experts. Ceux-ci peuvent d'ailleurs entraîner l'examineur à être plus ou moins exigeant vis-à-vis des candidats, parce qu'il se sent lui-même jugé.

L'examiné. — Le comportement de l'examiné dépend pour une bonne part de sa compétence dans le sujet à traiter. Cependant, il ne faut pas négliger l'influence de facteurs liés à sa personnalité (âge, sexe, présentation, voix, façon de s'exprimer, etc.) sur le déroulement et l'évaluation de l'examen¹.

¹ Des facteurs tels que la nationalité ou l'appartenance ethnique peuvent également jouer un rôle si le milieu socio-culturel du candidat et de l'examineur n'est pas homogène, comme cela peut se produire dans les examens internationaux.

Si le candidat est émotif, des réactions d'inhibition peuvent intervenir, réactions plus ou moins accentuées selon l'examineur. Le candidat peut au contraire être stimulé par la situation d'examen et donner l'impression d'une grande assurance malgré ses connaissances superficielles. L'idée que le candidat se fait des attentes de son interlocuteur détermine son comportement dans la mesure où il essaie de s'y conformer, puisque l'issue de l'examen aura des conséquences directes pour son avenir. Enfin, les opinions personnelles du candidat jouent parfois un rôle dans l'appréciation.

L'analyse nous a conduit à distinguer dans la situation de l'oral un certain nombre de facteurs dont l'influence respective sur l'évaluation finale n'est pas facile à mesurer. Et cela d'autant plus que dans la réalité ces différents éléments, en interaction continue, se renforcent ou se neutralisent mutuellement.

Ce caractère essentiellement dynamique de l'examen oral peut évidemment compromettre l'objectivité de l'appréciation.

c) *L'objectivité de l'oral*

A notre connaissance, l'étude de l'incidence des différents facteurs de la situation de l'oral sur l'appréciation n'a fait l'objet d'aucune investigation systématique.

La plupart des recherches sur l'oral — elles ne sont pas nombreuses — concernent son objectivité. Elles ont souvent été réalisées à partir des données statistiques que constituent les notes attribuées aux candidats.

Pour illustrer notre propos, nous allons examiner quatre de ces recherches.

I. Une des premières recherches expérimentales sur le problème des examens oraux a été celle de O. Trimble, publiée en 1934 aux Etats-Unis¹. Il s'agissait de déterminer le degré d'objectivité de l'examen oral au moyen des critères de validité et de fidélité.

La validité d'un instrument de mesure est sa rectitude matérielle. Dans le cas de l'examen oral, sa validité est parfaite s'il permet de rendre compte exactement de l'aptitude ou du savoir que l'on entend mesurer.

La fidélité peut être définie comme la concordance des résultats d'évaluations diverses réalisées avec un même instrument et dans des conditions identiques. C'est la rectitude formelle. Dans le cas de l'examen, il peut s'agir de la constance d'un même examinateur après un certain laps de temps (fidélité de l'examineur) ou de la concordance entre les évaluations de plusieurs examinateurs notant un même candidat (fidélité de l'épreuve).

¹ O.C. Trimble, *The oral examination: its validity and reliability*. School and Society (New York), N° 1009, vol. 39, pp. 550-552.

L'investigation de Trimble a porté sur 25 étudiants en psychologie appliquée se présentant à un examen final. Plusieurs mois avant l'interrogation, les candidats furent invités à préparer un sujet de leur choix, sujet qu'ils considéraient comme important en psychologie appliquée. L'auteur (A) réalisa lui-même l'interrogation, assisté de deux de ses collègues (B et C), tous deux professeurs de pédagogie et de psychologie.

Pour l'attribution de la note, chaque professeur disposait d'une échelle d'appréciation standardisée comportant les 9 points suivants:

- | | |
|--|---|
| 1. Discernement dans le choix du sujet | } Traits spécifiques
(aptitudes particulières) |
| 2. Précision dans l'énoncé des faits | |
| 3. Aptitude à traiter un sujet | |
| 4. Méthode de travail | |
| 5. « Bagage » de connaissances | |
| 6. Connaissances théoriques et pratiques | |
| 7. Aptitude à réfuter des arguments | |
| 8. Qualité du vocabulaire | |
| 9. Appréciation globale du niveau du candidat par rapport aux autres étudiants | } Valeur globale du candidat |

Pour chacun des 9 éléments à apprécier, la valeur maximale était de 100 points. Les résultats furent les suivants:

Fidélité des appréciations

a) Pour chacun des 8 traits spécifiques

On calcula la corrélation entre les appréciations des examinateurs A et B, A et C, B et C, puis la moyenne des 3 coefficients de corrélation obtenus.

Dans l'ensemble, les coefficients de corrélation sont faibles, sauf pour le point 5 (« bagage » de connaissances) où l'on obtient un coefficient de corrélation moyen de 0,644.

b) Pour la moyenne des 8 traits spécifiques (estimation synthétique)

La comparaison des moyennes entre les 8 appréciations des 3 examinateurs a donné des coefficients de corrélation relativement élevés:

- 0,489 entre l'examineur A et l'examineur B
- 0,611 entre l'examineur A et l'examineur C
- 0,712 entre l'examineur B et l'examineur C

Il se produit donc une certaine compensation des différences d'appréciation de chacun des 8 traits spécifiques (constatées sous a).

c) *Pour l'appréciation globale du niveau du candidat (point 9)*

Les coefficients de corrélation calculés sont faibles :

0,404 entre A et B

0,225 entre A et C

0,333 entre B et C

Or, ce jugement global est généralement le seul à déterminer la note du candidat dans un examen oral traditionnel. Il est très rare que les examinateurs disposent de critères d'évaluation précis permettant une notation analytique.

Validité des appréciations

Pour déterminer la validité des appréciations résultant de l'examen oral, on a comparé pour chaque candidat, l'estimation synthétique (moyenne des 8 traits spécifiques) aux résultats d'un test administré quelque temps auparavant.

La corrélation entre les deux types de résultats s'est révélée faible : 0,494 pour l'ensemble des examinateurs. L'auteur en déduit que l'épreuve « objective » (test) et l'examen oral n'ont pas les mêmes fonctions.

Bien que cette recherche ait porté sur une population restreinte, on peut conclure avec O. Trimble que l'examen oral, soigneusement préparé et administré de façon systématique à l'aide de critères standardisés peut présenter une fidélité aussi grande que les tests¹.

Par contre, la recherche a montré que l'examen oral administré sans l'aide de critères d'évaluation précis est un instrument de mesure peu fidèle et ne saurait être maintenu tel quel.

II. Dans une étude réalisée au Liban, E. Valin aborde un autre aspect du problème des examens oraux, à savoir quelle est la liaison statistique entre les résultats obtenus aux épreuves écrites et ceux obtenus aux épreuves orales². Il s'interroge sur l'homogénéité de l'examen de la deuxième partie du baccalauréat libanais. Selon l'auteur, l'examen peut être considéré comme homogène dans la mesure où l'écrit et l'oral classent les candidats de la même manière.

Les résultats obtenus se présentent comme suit :

1. Le coefficient de corrélation calculé entre les notes globales, à l'écrit et à l'oral, de chaque candidat ayant subi ces deux parties, s'est révélé assez faible : 0,4.

¹ A notre avis, la recherche a porté sur une population trop restreinte pour permettre de vérifier la validité de l'oral par rapport au test.

² E. Valin, *La valeur des examens. Etude docimologique réalisée au Liban*. Unesco — Etudes et documents d'éducation.

2. On serait tenté d'attribuer le caractère peu élevé de cette corrélation à la différence de contenu entre la partie écrite et la partie orale de l'examen, vu que trois matières seulement sont communes aux deux parties.

3. Cette hypothèse doit cependant être écartée car en comparant les notes d'écrit et d'oral pour une même matière, on obtient des coefficients de corrélation encore plus faibles :

0,14 en philosophie arabe
0,20 en philosophie française
0,27 en mathématiques et en sciences

4. Le diagnostic par l'oral est donc sensiblement différent du diagnostic par l'écrit.

5. La corrélation écrit-oral est légèrement plus élevée pour les matières scientifiques, ce qui peut s'expliquer par le fait qu'en mathématiques et en sciences, la notation réalisée à l'aide d'un barème atteint une plus grande objectivité.

Notons que H. Piéron et ses collaborateurs ont également analysé la liaison statistique écrit-oral au cours d'une recherche effectuée en 1955 sur le baccalauréat français et portant sur plusieurs milliers de candidats¹.

Pour l'ensemble des candidats admis, les corrélations entre les épreuves écrites et les épreuves orales pour les mêmes matières ont été les suivantes :

<i>Section philosophie</i>		<i>Section mathématiques</i>	
Philosophie	0,22	Mathématiques	0,38
Physique	0,31	Physique	0,26
Sciences naturelles	0,32	Philosophie	0,20

De ces données statistiques, il ressort que si la corrélation écrit-oral est faible indépendamment des matières d'examen, elle est légèrement plus élevée pour les épreuves portant sur les sciences exactes, comme dans la recherche effectuée par Valin.

La variation des conditions d'examen qui ont fait l'objet de ces deux recherches n'a eu aucun effet notable sur le parallélisme des résultats, ce qui semble confirmer la thèse d'une plus grande objectivité pour les épreuves en sciences exactes.

A cet égard, la recherche effectuée par H. Piéron, M. Reuchlin et F. Bacher sur les épreuves orales de physique est intéressante.

III. Il s'agit d'une étude expérimentale concernant le baccalauréat de mathématiques². Dix-neuf professeurs de physique de

¹ H. Piéron, M. Reuchlin, F. Bacher et M. Demangeon, *Analyse des corrélations entre notations à une session de baccalauréat*. Biotypologie, mars-juin 1962.

² H. Piéron, M. Reuchlin et F. Bacher, *Une recherche expérimentale de docimologie sur les examens oraux de physique au niveau du baccalauréat de mathématiques*. Biotypologie, mars-juin 1962.

lycées de Paris et de province ont participé à l'expérience. Les variations dues à la différence de sexe ont été éliminées: il n'a été fait appel qu'à des professeurs et des candidats de sexe masculin.

Trois de ces professeurs ont interrogé quarante élèves de mathématiques élémentaires, dans les mêmes conditions qu'au baccalauréat, mais avec enregistrement de l'interrogation. Les enregistrements réalisés par les trois professeurs ont été divisés en deux séries de vingt interrogations entremêlées.

L'expérience proprement dite s'est déroulée en plusieurs phases dont nous ne rapporterons que les plus marquantes:

a) L'enregistrement direct des interrogations des 20 élèves de la première série a été soumis à l'écoute des 16 professeurs qui n'avaient pas participé à l'examen et qui furent chargés de noter les candidats avec les mêmes exigences qu'au baccalauréat.

b) L'autre série des 20 interrogations a été rapportée par un parleur unique, répétition également enregistrée sur magnétophone: le contact vocal avec les examinateurs et les examinés était donc supprimé. Les 16 professeurs procédèrent à la notation de cette deuxième série de candidats.

c) Ces mêmes professeurs notèrent en outre une troisième série de candidats qui avaient rédigé leurs réponses à de brèves questions écrites, du même type que celles qui auraient pu être formulées oralement.

Les principales données recueillies sont les suivantes:

1. La durée d'intervention de l'examineur pendant l'interrogation varie dans d'assez fortes proportions d'un candidat à l'autre et d'un examinateur à l'autre. Il serait évidemment intéressant de savoir dans quelle mesure un tel facteur détermine la performance du candidat.

2. Pour les trois séries de notations concernant chacune un même groupe d'examinés par tous les examinateurs, on a obtenu les corrélations moyennes suivantes entre examinateurs:

Présentation écrite	0,813
Enregistrements directs	0,727
Enregistrements répétés	0,568

3. Si l'on examine les valeurs extrêmes des corrélations des 120 couplages d'examineurs, on constate une importante marge de variation des accords suivant le couple:

	<i>Valeurs extrêmes</i>	
	Minima	Maxima
Présentation écrite	0,335	0,973
Enregistrements directs	0,424	0,947
Enregistrements répétés	0,127	0,891

4. Des différences considérables dans les échelles utilisées sont mises en évidence par la comparaison des moyennes des notes attribuées par chaque examinateur. Pour les trois séries de notation, la moyenne des notes attribuées est de 8,66 sur 20 pour l'examineur le plus sévère et de 12,8 sur 20 pour l'examineur le plus indulgent. L'écart maximum entre les moyennes pour chacune des trois séries a été de

5,37 pour les enregistrements directs
3,67 pour les enregistrements répétés
4,02 pour la présentation écrite

Toutes ces données indiquent que l'oral mérite une moindre confiance que l'écrit.

IV. Nous examinerons une dernière recherche réalisée en Italie au niveau universitaire et publiée en français en 1964. Cette recherche¹ comportant trois enquêtes avait pour objectif de déterminer la relation existant entre la note attribuée par le professeur et celle que s'est donnée l'étudiant sans connaître le jugement du professeur.

Nous ne relaterons que la première de ces enquêtes, dite rétrospective, dans laquelle on a comparé les notes données par le professeur et celles que les étudiants se sont attribuées au terme de l'interrogation.

Cette enquête a porté sur quatre matières:

Statistique I
Statistique II
Démographie
Economie et Politique agraire

Notons que pour chaque examen, l'interrogateur était le professeur enseignant la matière en question. Les principaux résultats se présentent comme suit:

1. La corrélation entre les notes des étudiants et celles attribuées par les professeurs se révéla relativement élevée: entre 0,64 et 0,79 selon les matières.

2. Les écarts entre chaque note singulière que l'étudiant s'était attribuée et la note correspondante donnée par le professeur étaient très faibles.

Comment interpréter de tels résultats? Peut-on voir là une garantie d'objectivité de la notation?

Arrivés à ce stade, les chercheurs ont tenté d'isoler un certain nombre de facteurs pour voir dans quelle mesure ils influençaient la capacité de l'étudiant de se rapprocher de la note du professeur.

¹ D. de Castro, *Recherches docimologiques sur quelques matières inscrites au programme d'examens universitaires*. Le Travail humain, N° 3-4, 1965.

Les étudiants qui ont passé plus d'un examen de statistique réalisent-ils des écarts moindres que ceux qui n'ont passé qu'un seul examen? L'école d'origine (institut technique ou lycée), la différence de sexe, peuvent-elles favoriser un tel rapprochement? On a également tenu compte des caractéristiques des étudiants telles que la note moyenne réalisée dans toutes les matières pour les examens déjà passés au moment de l'enquête, le nombre d'ajournements antérieurs en d'autres matières ou pour les matières statistiques, etc. On n'a pu constater aucun lien significatif entre ces différentes caractéristiques et la capacité des étudiants de se rapprocher de la note du professeur.

Au vu de ces observations, l'auteur quitte le plan docimologique pour se placer dans une perspective psychologique. Il analyse alors les processus psychologiques du professeur et de l'étudiant qui déterminent leur jugement: tous deux formulent leur jugement de manière empirique, c'est-à-dire sans se référer à un modèle rigoureux. En suivant cette démarche, l'auteur aboutit aux *conclusions* suivantes:

1. Etant donné la quantité d'éléments qui interviennent dans l'élaboration d'un jugement complexe au cours d'un examen et la différence de stimulations auxquelles sont exposés les deux sujets de l'expérience, la concordance des jugements devrait être impossible.

2. Celle qui a été constatée s'explique par un processus d'empathie qui permet à l'étudiant de conjecturer le jugement du professeur.

3. C'est d'une part parce qu'il connaît les exigences du professeur et d'autre part par une série d'indices qu'il peut relever pendant le déroulement de l'interrogation (signe approbateur ou désapprobateur, intonation, exclamations, etc.) que l'examiné s'attribue finalement une note très proche de celle de l'examineur.

3. Conclusions

Les méthodes d'évaluation traditionnelles, telles que nous les avons décrites, constituent-elles des instruments suffisamment sûrs pour garantir la valeur prédictive assignée à l'examen du baccalauréat?

L'analyse qui précède a mis en évidence l'intervention d'un certain nombre de facteurs dans l'évaluation et nous a amenés à constater que les méthodes traditionnelles n'assurent pas le maximum d'objectivité.

Il faut toutefois constater que les études conduites jusqu'ici ont surtout prouvé l'absence de fidélité des méthodes d'examens et celle des examinateurs.

La corrélation entre le savoir et le pouvoir des candidats d'une part et les notes obtenues, d'autre part, n'apparaît pas d'une façon rigoureuse. Il en est de même du facteur affectif et émotif dont on

parle beaucoup pour expliquer l'incohérence de certains résultats. La plupart des recherches entreprises portent sur le contrôle de candidats que les examinateurs ne connaissent pas.

Il est important de souligner ici les deux situations fort différentes entre des examens internes et des examens externes. En Suisse, mise à part la maturité fédérale, les examens de baccalauréat se passent à l'intérieur des établissements et sont dirigés par les maîtres qui ont eux-mêmes préparé les candidats. La situation est tout autre en France, par exemple, où des jurys d'examens sont appelés à juger des candidats anonymes. Beaucoup des défauts relevés, mais pas tous, disparaissent dans le système des examens internes, d'autres surgissent au contraire (parti pris du professeur, passé scolaire trop connu, etc.).

Il nous reste à voir dans quelle mesure les moyens d'évaluation peuvent être améliorés à partir des données existantes et en fonction de celles que pourraient nous fournir certaines recherches dans ce domaine.

A. D'ores et déjà, il devrait être possible d'agir sur un certain nombre d'éléments de la situation d'examen qui constituent des facteurs de distorsion dans l'appréciation.

Le sujet

Celui-ci doit être formulé clairement, de manière à ce qu'il puisse être immédiatement compris, non seulement par le candidat, mais également par tous les examinateurs d'une même matière. Ce point est particulièrement important pour les épreuves écrites, où le candidat ne bénéficie pas de la présence de l'examineur qui pourrait préciser le sens de la question. A l'intérieur d'une même épreuve, les questions doivent être distinctes, c'est-à-dire en rapport étroit avec un aspect bien défini des connaissances que l'on entend mesurer¹, de façon à susciter des réponses précises, susceptibles d'être interprétées de la même façon par tous les examinateurs d'une matière.

La notation

Le recours à un barème de notation standardisé semble indispensable. On utilisera un système de notation analytique plutôt que synthétique: à chaque élément à évaluer correspondra un nombre de points maximum, et la valeur du candidat par rapport à cet élément sera précisée sur une échelle d'appréciation. La note finale sera fournie par le total des points obtenus pour chaque réponse (ou item).

¹ Pour les langues vivantes, par exemple, on évitera qu'une même question ait pour but d'évaluer simultanément les connaissances littéraires du candidat et sa faculté de s'exprimer.

Les échelles utilisées seront suffisamment sensibles pour rendre compte autant que possible des différences dans la qualité des réponses fournies par les candidats.

La standardisation de l'« outil » utilisé par les examinateurs limite le rôle des facteurs subjectifs et permet d'éviter les divergences entre examinateurs dues à des différences dans la dispersion des notes et la moyenne par rapport à laquelle se fait cette dispersion.

Rôle de l'examineur

Celui-ci est particulièrement important pour l'examen oral. Dans toute la mesure du possible, le comportement de l'examineur devra être standardisé, afin d'éviter l'influence de facteurs personnels sur les résultats de l'examen. C'est pourquoi les rapports consolidés entre examinateur et examiné sont à éviter: l'interrogateur sera donc différent du professeur enseignant la matière de l'examen.

Le rôle de l'interrogateur d'oral comporte trois aspects principaux:

1. Créer une atmosphère cordiale afin que l'examiné puisse s'exprimer dans les meilleures conditions.
2. Diriger l'attention de l'examiné sur l'objet précis de chaque question.
3. Enregistrer de manière objective les réponses aux questions posées. Dans cette tâche, l'examineur pourra être secondé par un expert ou un membre du jury. On peut également envisager l'emploi d'un magnétophone.

L'attitude de l'examineur sera aussi neutre que possible, susceptible cependant de s'adapter aux réactions particulières du candidat.

A cet égard, la formation de l'examineur est importante: il doit être au courant des différentes situations qui peuvent se présenter et connaître parfaitement les facteurs qui risquent de perturber l'appréciation, afin de pouvoir les contrôler.

B. En dépit des mesures que l'on peut prendre pour accroître l'objectivité de l'écrit et de l'oral, les progrès susceptibles d'être obtenus seront forcément limités tant qu'on n'a recours qu'à ces deux méthodes d'évaluation. C'est pourquoi il serait souhaitable de leur adjoindre d'autres moyens d'appréciation (tests, questionnaires, examen-bilan, épreuves pratiques, etc.) afin d'assurer le maximum de validité.

On peut aussi se demander si dans la perspective de réformes plus profondes, il ne conviendrait pas de déterminer d'une façon beaucoup plus rigoureuse les objectifs des disciplines qui composent un programme d'études et de n'interroger que dans le cadre de ces objectifs. Les différenciations entre examinateurs les plus frappantes viennent

le plus souvent du fait qu'ils ne jugent pas la même chose. Pour les branches littéraires, l'un apprécie le style, l'autre la grammaire, un troisième le goût et l'imagination, un quatrième voudra tenir compte de tout. Les recherches de taxonomie de l'éducation lancées par l'Américain Bloom peuvent être d'une grande utilité¹.

Enfin, il est de plus en plus évident que l'ordinateur est appelé à jouer un rôle considérable dans la passation des examens. Il est le seul à pouvoir offrir une objectivité presque totale entre candidats en même temps qu'une disponibilité à l'égard des réponses offertes par eux que l'examineur n'a souvent pas.

Le Professeur Cardinet, dans un exposé qu'il faisait récemment sur l'emploi des ordinateurs pour les examens, terminait par cette conclusion :

« La souplesse de l'ordinateur rend possibles des contrôles inconcevables autrement. L'autoévaluation que peut effectuer l'étudiant est favorable à sa motivation et développe son autonomie. Surtout, l'aspect inhumain de l'évaluation d'autrui est laissé à la machine, de sorte que l'enseignant peut plus facilement jouer son rôle d'appui, aussi bien affectif qu'intellectuel. Etudiants et professeurs se trouvent alors en situation de collaboration plutôt que de conflit ».

Si l'emploi de l'ordinateur n'est pas pour aujourd'hui, il le sera certainement pour demain.

GEORGES PANCHAUD
HÉLÈNE HENSLER
SÉGA FALL

¹ Benj. Bloom, *Taxonomy of educational objectives*. The classification of educational goals. David Mc Kay, New York, 1956.