



MINISTÉRIO DA EDUCAÇÃO E CULTURA

Psicologia Educacional

DISTRIBUIÇÃO

*"La Docimologie dans les examens
et les concours"*

Henri Piéron

*Arg. L.P. 1
Jan. 3*

LA DOCIMOLOGIE DANS LES EXAMENS ET LES CONCOURS

Par M. HENRI PIÉRON
Professeur honoraire à la Sorbonne et
au Collège de France

Les examens jouent un rôle considérable dans l'enseignement, tant au point de vue des entrées que des sorties. Dans le premier degré, ils ne comptent guère à l'entrée, bien que certains enfants ne soient pas susceptibles de recevoir l'enseignement normal. Ils peuvent présenter des anomalies sensorielles ou autres, des états de débilité mentale, etc., mais, à cet âge, le problème n'est pas aigu. L'entrée principale est dans le second degré. Le problème de l'examen d'entrée en sixième est un de ceux qui provoquent le plus de discussions. Enfin, il y a l'entrée dans l'enseignement supérieur.

Jusqu'à ces dernières années, l'examen de sortie constituait en même temps l'examen d'entrée, et cela a été une cause de confusion très fâcheuse. En effet, les examens de sortie d'un enseignement doivent porter essentiellement sur un certain contrôle de la formation et de l'acquisition des connaissances au cours d'une certaine période. Il faut s'assurer que les mécanismes du langage, du calcul, ont été bien acquis, et puis une formation de pensée, en particulier la formation logique. Enfin, il faut contrôler les connaissances indispensables, selon les différences de niveau correspondant à l'âge.

Mais quand il s'agit des examens d'entrée, le problème essentiel n'est pas le passé, c'est l'avenir. Il faut savoir quelles seront les capacités d'acquisition ultérieure. Or, le baccalauréat était à la fois l'examen d'entrée dans l'enseignement supérieur et aussi un examen de contrôle de la formation éducative au cours du second degré. C'est la confusion que nous dénonçons et sur laquelle nous reviendrons. En tout cas, il faut savoir la valeur des méthodes. L'examen

fait appel à une balance pour peser, mais que valent les poids, que valent les mécanismes même de la balance? Ce problème de validité est celui que j'ai été amené à me poser d'assez bonne heure, en vue d'aboutir à une technique scientifique des examens.

LES ORIGINES DE LA DOCIMOLOGIE

Dès la fin de mes études secondaires, je fis la connaissance d'Alfred Binet, qui avait créé un laboratoire de psychologie à la Sorbonne; puis j'entrai en rapport avec Edouard Toulouse, un médecin psychiatre qui recherchait aussi des méthodes de mesure intellectuelle. En 1904, avec ce dernier, nous avons mis au point une technique de psychologie expérimentale reposant sur l'examen différentiel, en vue de déterminer les caractéristiques psychologiques différentes des individus. A cette époque, la tradition considérait la psychologie comme un des domaines de la philosophie. Ayant passé d'assez bonne heure une licence de philosophie, je préparais l'agrégation de philosophie en étudiant la technique de la psychologie expérimentale; mais certains de mes juges n'admettaient pas qu'on eût besoin de la médecine, de la physiologie et du laboratoire pour faire un professeur de philosophie. A l'une de mes leçons, un des examinateurs proposait de me donner une note très proche du maximum, un 18, tandis qu'un autre disait que cela ne méritait pas plus de 2. Il y eut alors, paraît-il, des discussions très vives dans le jury. Finalement, je réussis tout de même, mais grâce au hasard, parce que je sus expliquer un texte grec de Plotin en suivant l'interprétation de Bergson. J'en tirai des conclusions assez ironiques sur l'incertitude des examens.

LES RECHERCHES AVEC HENRI LAUGIER

Plus tard, en 1922, je m'associâi avec Henri Laugier pour mettre sur le plan expérimental la technique des examens. Nous commençâmes par le degré inférieur du certificat d'études. Nous constatâmes que, lorsqu'on classait les groupes d'épreuves, certaines concernaient uniquement des acqui

sitions de mémoire, d'autres des aptitudes mentales, d'autres étaient très secondaires, comme la gymnastique, le chant, le dessin chez les garçons, et la couture chez les filles. Or, il y avait les mêmes corrélations entre toutes ces épreuves, comme si elles étaient extrêmement parentes. Le bon élève d'examen a certaines qualités qui apparaissent, quelle que soit la nature des épreuves.

En revanche, quand nous avons appliqué des tests spéciaux aux capacités intellectuelles, les corrélations ont été extraordinairement faibles, presque nulles, même avec des épreuves qui, cependant, portaient sur la capacité mentale. Il y avait là un désaccord complet entre la notion d'aptitude psychologique et la réussite du bon élève au certificat d'études.

Alors, Laugier passa au niveau le plus élevé. En examinant la double correction faite pour le concours d'entrée à l'Ecole normale supérieure, section des lettres, il s'aperçut que la corrélation entre les corrections était de 0,82 (1 représente l'identité parfaite) pour l'histoire et la géographie, mais de 0,42 seulement pour la philosophie. Cela voulait dire pratiquement que les deux correcteurs s'accordaient pour recevoir 8 des candidats mais n'étaient pas du tout d'accord pour les 12 autres. Pour l'histoire, il y avait accord pour 15 sur 20.

L'INCERTITUDE DES NOTATIONS

Laugier étudia ensuite les examens des certificats d'études supérieures de sciences. On confia les textes des copies à trois professeurs. Entre eux, la corrélation varia de 0,36 à 0,81, et elle passa même de 0,58 à 0,81 quand les mêmes copies étaient corrigées à quelque temps d'intervalle. Un des correcteurs était plus stable que l'autre.

Selon la cotation des trois examinateurs, il y aurait eu 41 ou 73% des candidats admissibles, c'est-à-dire ayant plus de la moyenne. Donc, avec un examinateur il y en aurait

41 reçus et avec un autre 73 27% des candidats auraient l'accord pour être admis, c'est-à-dire pour avoir plus que la moyenne.

Chose curieuse, on fit faire les corrections par une jeune bachelière qui venait de sortir de philosophie et qui n'avait pas du tout de connaissances scientifiques. Ses notes étaient en corrélation de 0,51 avec celles des correcteurs qualifiés, c'est-à-dire que, rien qu'à la vue de la composition, il n'y aurait pas eu plus d'inconvénients à prendre au hasard un correcteur ou la jeune bachelière. C'étaient là des indications que la validité des examens n'était pas parfaitement satisfaisante. C'est ce qui m'a poussé à poser les principes d'une docimologie, fondée sur les méthodes psycho-techniques, c'est-à-dire ayant une base scientifique.

L'ENQUÊTE SUR LE BACCALAURÉAT

A ce moment, la Fondation Carnegie décida de subventionner une enquête docimologique fondamentale qui porterait sur le baccalauréat. Ses résultats ont été publiés en 1936. Avouons qu'ils n'ont pas été entièrement satisfaisants au point de vue des opinions exprimées. Sur trois mille questionnaires envoyés, il en revint cent seulement. D'une manière générale, on avait grande confiance dans les examens, et il n'y avait pas de raison de changer. Beaucoup d'examineurs repoussaient toute idée de méthodes plus techniques de contrôle. On voyait répéter le droit des éducateurs à examiner. C'est évidemment l'esprit, plutôt artistique que scientifique, qui dominait dans ces cas-là. On arrive à juger un candidat comme on juge un tableau.

Ce qui fut important, dans l'enquête Carnegie, ce sont les expériences. Sur six matières, on choisit cent copies dans une session de baccalauréat (1930), et elles furent confiées à six groupes de cinq correcteurs compétents. Ces trente correcteurs, après avoir donné indépendamment leurs notes,

furent appelés aussi à expliquer leur manière de les donner. On trouva une extraordinaire diversité dans les points de vue. Par exemple, pour les mathématiques, deux notes étaient données dont on faisait la moyenne: une question de cours, un problème. De même pour la physique. Pour les langues, il y avait une version et un thème.

Les corrélations ont été, pour le problème de mathématiques, de 0,838; pour le cours, de 0,807. En physique: problème, 0,861; cours, 0,798. En version latine: 0,839. En anglais: thème, 0,748; version, 0,701. En dissertation française: 0,568. En dissertation philosophique: 0,489.

L'accord au-dessus de la moyenne se rencontrait à raison de 20% pour les mathématiques, de 13% pour la physique, de 10% pour la version latine, de 16% pour l'anglais, de 9% pour la dissertation française et de 10% pour la dissertation philosophique.

Au-dessous de la moyenne, il y a davantage d'accord: 40% en mathématique, 37% en physique, 40% en version latine, 37% en anglais, 21% en dissertation française, 9% en dissertation philosophique. Cela veut dire que le hasard jouait avec certitude dans 36% des cas pour les mathématiques, dans 50% des cas en physique et en version latine, dans 47% des cas en anglais, dans 70% des cas en dissertation française et dans 81% des cas pour la dissertation philosophique. Pour les écarts de notation entre correcteurs, la moyenne était de 2,05 en mathématiques, 1,88 en physique, 2,97 en version latine, 2,24 en anglais, 3,29 en dissertation française, et 3,36 en dissertation philosophique. L'écart maximum obtenu entre les cinq correcteurs était de 9 points en mathématiques sur 20,8 points en physique, 12 points en version latine, 9 points en anglais, 13 points en dissertation française, 12 points en dissertation philosophique.

LE NOMBRE DES CORRECTEURS

D'après les corrélations obtenues, on peut calculer

le nombre de correcteurs qui serait nécessaire pour obtenir la note en laquelle on puisse avoir pleine confiance. Ce nombre varie beaucoup. Pour la dissertation française, il est de 78. On choisit 78 correcteurs compétents et on leur donne à corriger seulement trois copies. Les notes moyennes pour ces trois copies étaient 6,5, 8,7 et 10. Pour la première, les notes extrêmes étaient de 1 et 13 points, pour la deuxième de 4 et 14 points, pour la troisième de 3 et 16 points. Il faut dire que c'étaient des correcteurs professionnels. Avec un seul correcteur, selon qu'on tombe sur l'un ou sur l'autre, une même copie peut avoir 1 point ou 13 points alors que sa valeur moyenne est de 6,5! Sans doute, les accords sont plus faciles pour les notes extrêmes, mais on a une idée des écarts de jugement.

Dans les examens que j'ai eu à faire passer, notamment pour le diplôme des conseillers d'orientation professionnelle, on a institué, à ma demande, la double correction pour l'écrit, et ainsi j'ai pu être renseigné sur les différences entre les correcteurs, avec des précisions numériques sur les différences entre les deux correcteurs très compétents pour chacune des compositions: psychologie, pathologie, sciences économiques, étude du travail technique des métiers. Le résultat, c'est qu'on trouve toujours à peu près, quels que soient les examens, la même variabilité. La corrélation moyenne est de 0,50. Les écarts de 5 points et plus, sur 20, vont de 4 à 45%, c'est-à-dire que, dans certains cas, il y a 45% des candidats pour lesquels la note donnée diffère de cinq points ou plus. Les différences entre examinateurs sont notables, ainsi que les marges de notation. A l'écrit, certains cotent de 7 à 17, à l'oral un peu moins de 5 à 12. Les notes moyennes sont 8 pour l'un, 15 pour l'autre. Un examinateur note en moyenne autour de 15, un autre autour de 8. Il est difficile d'arriver à une grande précision dans ces conditions, et les faits donnent à réfléchir.

LE PROBLÈME DE L'ORAL

Ces résultats se sont limités au domaine de l'écrit et d'un écrit anonyme, c'est-à-dire dénué de toute influence possible. Au point de vue de l'oral, il n'y a pas de données précises, et c'est pourquoi le Conseil supérieur de la Recherche scientifique m'a demandé de reprendre l'étude, avec mon excellent collaborateur M. Reclin, de l'Institut d'orientation. Nous ferons, grâce à l'emploi du magnétophone, des recherches expérimentales parallèles aux recherches sur l'écrit.

On voudrait supprimer l'oral, mais il y a beaucoup de résistances. On peut dire que c'est la meilleure et la pire des choses. On prétend que l'oral peut rectifier les erreurs de l'écrit. Dans un seul sens, puisque ne passent l'oral que les admissibles à l'écrit, et cela pour des raisons de facilité, étant donné le nombre d'examineurs et le temps qui serait nécessaire. Que rectifie-t-on? Uniquement trop d'indulgence en éliminant quelques candidats supplémentaires. D'après les enquêtes qui sont faites avec grande compétence par M. Chattelun, je crois que, sur 1.500 candidats en 1936, 52,5% étaient éliminés à l'écrit et 8% seulement à l'oral, c'est-à-dire qu'en somme l'écrit est tout de même la chose dominante.

Pour que l'oral permette vraiment de bien connaître un candidat, il faut un examinateur excellent, et, dans le cas contraire, l'oral peut être extrêmement dangereux. En effet, il y a des examinateurs qui ont des lubies, des sympathies et des antipathies, comme dans le cas personnel que cite le regretté Leriche. Avec un excellent examinateur, l'oral permet d'approfondir les connaissances et, surtout, de rechercher les capacités intellectuelles d'un candidat, de distinguer le perroquet de l'esprit réfléchi, ce qui n'est pas toujours aussi facile à l'écrit. L'oral peut jouer un rôle utile, mais il exige des conditions très particulières.

Aussi, le problème du perfectionnement des examens

est complexe. Il exige que l'on forme des examinateurs, et cela d'emblée. Est-il admissible que, sans aucune préparation, un professeur possède la technique et les qualités que l'on peut exiger d'un examinateur?

LA MOYENNE STATISTIQUE

Lorsqu'on considère des séries d'individus, ils se répartissent autour d'un certain état moyen; par exemple pour la taille, on a beaucoup d'individus de taille moyenne, puis on a des tailles de plus en plus rares qui s'écartent de cette moyenne, en plus ou en moins. Pour les aptitudes intellectuelles, les examens psycho-techniques nous montrent qu'il en est de même. On obtient les mêmes "courbes en cloche" de la fréquence des états différents. Par conséquent, il est utile au correcteur de savoir que, si on lui soumet un très grand nombre de copies, il a des chances d'avoir une courbe en cloche. La courbe en cloche n'est valable que dans la mesure où l'on considère des épreuves correspondant à des niveaux différents. Avec des épreuves trop faciles, à peu près tout le monde réussit. La courbe est dite en G, c'est-à-dire que le sommet des fréquences atteint les hautes valeurs. Au contraire, avec des épreuves trop difficiles, tout le monde a zéro, et il n'y a pas moyen de coter. Quand on fait de la psycho-technique avec des tests, le test est satisfaisant pour un niveau dans la mesure où l'on obtient la courbe en cloche. Quand on ne l'obtient pas, c'est que le niveau d'épreuve n'est pas adapté au milieu que nous voulons examiner.

Dans le cas le plus satisfaisant, l'habitude de prononcer des admissions ou des admissibilités d'après la moyenne est une erreur fondamentale. La zone moyenne autour de 10 est en effet trop large pour qu'il soit indifférent de classer au-dessus ou au-dessous de la moyenne. Alors, ce n'est pas la peine d'examiner: il n'y a qu'à tirer au sort. Si l'on veut faire une coupure dans ces conditions il faut prendre un tiers ou deux tiers des candidats, mais jamais la

moitié; quand on en prend la moitié, on est sûr de se tromper dans un nombre de cas considérable. On ne s'accorde bien que pour les tout premiers ou les tout derniers. Voilà les faits que les données expérimentales nous ont apportés.

DES ÉPREUVES PLUS OBJECTIVES

Dans le contrôle des connaissances comme dans celui des mécanismes, on peut imaginer des épreuves plus objectives, par exemple les sondages, analogues à ceux qu'on emploie en statistique démographique. Il faut sonder les connaissances non pas au hasard d'une ou deux questions, mais sur une certaine étendue. Pour les mécanismes, on peut opérer de même. Si l'on veut connaître la capacité en orthographe, le meilleur test est la dictée, la dictée rationalisée, la même pour tout le monde, avec les mêmes difficultés et les mêmes possibilités de notation. On dit depuis longtemps que la dictée est un procédé commode. Autrefois, dans tous les examens, on imposait la dictée. Pour les téléphonistes, la dictée jouait un rôle essentiel. En matière d'orientation, on a fini par protester. Il n'importe pas de savoir si la téléphoniste connaît l'orthographe; pour une dactylo, c'est autre chose. Il faut donc se demander à quoi vise l'examen quand on veut faire une sélection.

Enfin, il ne faut pas juger les candidats sur une préparation de dernière heure concernant des connaissances qui sont rapidement oubliées. Il est beaucoup plus important d'apprendre à se servir des instruments de travail que d'accumuler des notions dans la mémoire. Il n'est d'ailleurs pas nécessaire d'apprendre à se servir de tous les instruments. On m'a appris à extraire des racines carrées et je n'ai jamais eu à extraire une racine carrée: il y a des tables de logarithmes. Pourquoi imposer un effort spécial pour quelque chose de difficile qu'on ne fera jamais? Il faut adapter l'enseignement aux besoins et à l'avenir de l'individu.

Seulement, il faut considérer l'aspect sélectif. C'est

là le problème le plus important et le plus difficile, car il décide de l'entrée dans le second degré et dans l'enseignement supérieur. Aux Etats-Unis, l'entrée à la faculté de médecine est rendue particulièrement difficile. On donne la première place aux épreuves de capacités intellectuelles. Il s'agit de savoir ce qu'on pourra faire, et non ce qu'on a fait jusque-là. Il y a de bons élèves du premier degré qui échouent au second, et il y a de bons élèves du second degré qui échouent dans l'enseignement supérieur, parce que, chaque fois, on demande quelque chose de plus. Même dans le second degré, on voit s'écrouler au baccalauréat des élèves qui tenaient la tête dans les premières classes. C'est cette évolution qu'il faut prévoir, et c'est là le problème essentiel de cette orientation professionnelle à laquelle je me suis consacré depuis bien des années.

PAS DE RECETTES DANS LES CONCOURS

Un autre point est l'influence qu'exercent les professeurs dans les concours. A l'heure actuelle, c'est autant les professeurs que les candidats qu'on reçoit dans les grandes écoles. Je sais bien que la formation par le professeur est quelque chose, mais il y a aussi chez certains une tendance à enseigner des "colles", à indiquer les réponses qui conviennent à tel ou tel examinateur et qui facilitent beaucoup l'admission.

En résumé, on fausse la balance, on fausse les poids par la préparation des concours. Je soutiens qu'il ne doit pas y avoir de concours sur des épreuves pour lesquelles la préparation a pu être faite. Certes, on a besoin de s'assurer d'un certain bagage de connaissances. Cela, c'est l'affaire d'un examen préalable, mais ne doit pas intervenir dans le classement final: c'est un élément de sélection. On ne peut admettre, évidemment, que des gens avant un certain niveau de connaissances; mais l'essentiel consiste dans les capacités à résoudre les problèmes qui n'ont pas de recette.

Le vice de notre enseignement est d'indiquer des recettes pour résoudre les problèmes. A des enfants moyens il faut apprendre la règle de trois car ils ne la trouveraient pas tout seuls. Mais les concours sont faits pour discerner les intelligences capables de résoudre les problèmes sans ce genre de mécanisme. Je ne crains pas d'être révolutionnaire en disant que le concours de Polytechnique est une catastrophe pour le pays. On épuise des jeunes gens dans une préparation qui ressemble à un entraînement sportif pour battre un record. Quand ils sont arrivés, souvent ils sont vidés par cet effort intensif et, dans leur carrière, ils risquent de commettre des erreurs graves.

L'ORIENTATION NÉCESSAIRE

En ce qui concerne les examens sélectifs proprement dits, il faut envisager des réformes, et on y songe depuis longtemps à la Commission de la réforme de l'enseignement. Il faut instaurer un cycle d'orientation dans lequel on pourra suivre un enfant pendant assez longtemps au moyen d'épreuves qui permettront de mieux le connaître et de l'orienter. Il serait faux d'employer des méthodes brutales de sélection qui laissent de côté ceux qui ne sont pas élus. On doit tirer parti de tout le monde et, par conséquent, explorer chacun pour savoir de quel côté il pourra réussir. C'est l'orientation qui est essentielle dans la réorganisation de l'enseignement, et non pas seulement au niveau du second degré, mais au niveau même des écoles spéciales, à l'heure actuelle, les écoles d'ingénieurs. Je sais que l'on s'en préoccupe en haut lieu. C'est de cette façon qu'on évitera les erreurs d'examen et qu'on renoncera aux néfastes préparations de concours.

/nb.