

"FEUILLETON" de l'EVALUATION, par Jean-Paul GUICHARD, de PARTHENAY

Episode 8 : DOCIMOLOGIE (1). La note : évaluation objective ?

La docimologie est un terme et une discipline créés par PIERON (1881-1964), psychologue français. Elle s'est définie comme une "*étude systématique des examens*". Pour apprécier le champ d'étude de la docimologie, voici les questions essentielles que s'est posé et se pose la docimologie, et qui sont d'une actualité brûlante :

1. Les notes, classements, scores attribués à une même prestation d'élève :

- par différents correcteurs sont-ils identiques ou non ?
Si non quels sont les écarts observables ?
- par un même correcteur à des moments différents sont-ils identiques ou non ?
Si non quels sont les écarts observables ?
- En cas d'écarts, comment les interpréter ?

2. Les examens :

- sont-ils justes ? (justesse)
- sont-ils réussis par les candidats les plus compétents ?
- sont-ils justes ? (justice)
- N'y a-t-il pas risque d'éliminer du système des ressources humaines ? de les gaspiller ?
- Quelle est la place du hasard dans la distribution des diplômes ?

Un des premiers problèmes auxquels s'est attaqué PIERON a été le problème de la notation et de ses fluctuations. Tout le monde est plus ou moins convaincu que deux correcteurs différents ne mettront pas la même note à une même copie. Mais peut-on quantifier les écarts ? Peut-on mettre en évidence un certain nombre de facteurs et quantifier leur importance ? La première idée est de faire une expérience de multicorrection de copies ; si vous n'avez jamais tenté l'expérience, faites-la dans votre établissement avant de lire la suite : vous serez plus que surpris. Voici les résultats obtenus par PIERON pour des copies de Baccalauréat :

Matières	Ecart moyen	Ecart le plus fréquent	Ecart max.	Moyennes		Dispersion (σ)	
				Générale	Etalement des M. des 5 juges	Générale	Etalement des σ . des 5 juges
Composition franç.	3,29	6 et 7	13	8,64	6,32 - 10,00	3,42	2,91 - 3,85
Verson latine.....	2,97	5	12	7,62	5,15 - 9,37	4,22	3,70 - 4,66
Anglais.....	2,24	4	9	8,57	7,17 - 9,60	3,39	2,86 - 3,92
Mathématiques	2,05	4	9	8,06	7,01 - 9,16	4,24	4,00 - 4,30
Philosophie.....	3,36	5 et 7	12	9,26	7,65 - 11,23	3,04	2,22 - 4,25
Physique.....	1,88	4	8	7,87	7,11 - 9,48	3,45	3,31 - 3,63
Moyenne générale.....	7,98					3,62	

A. BONBOIR. LA DOCIMOLOGIE. 1972, p. 108

Une étude attentive du tableau montre que les mathématiques, contrairement à un préjugé répandu ne se distinguent guère des autres matières, et l'écart maximum de notes sur 20 est de 9 points pour une même copie! L'auriez-vous pensé ? Ces résultats incroyables, en particulier pour les mathématiques, ont été maintes fois renouvelés. Voici les résultats d'une expérience menée par le C.R.D.P. de Lyon auprès de 150 prof. de maths, exerçant dans une 3^{ème}, qui avaient à noter les copies de 3 élèves ayant eu à résoudre le même problème.

Des résultats impressionnants :

	Max. 20	Moyenne	Marge de variation
Elève I		5,70	0,5 - 11,5
Elève II	16		11,5 - 20
Elève III	08		3,5 - 11,5

(G.de LANDSHEERE. EVALUATION CONTINUE ET EXAMENS. PRECIS DE DOCIMOLOGIE. 1974, p 32)

Quelles conséquences pour les examens ? Les désaccords entre correcteurs font entre autres qu'un candidat peut être reçu ou collé suivant par qui il est corrigé. Combien cela peut-il concerner de candidats à votre avis ? 5% ? 10% ? 20% ? Eh bien, hélas, c'est pire que ça! Cela varie de 36% à 81% dans l'enquête précédente de PIERON faite sur le Baccalauréat. Voici le tableau des résultats :

	Refusés par les six correcteurs	Admis par les correcteurs	Admis par les uns et refusés par les autres
Version latine	40%	10%	50%
Composition française	21%	09%	70%
Anglais	37%	16%	47%
Mathématiques	44%	20%	36%
Philosophie	09%	10%	81%
Physique	37%	13%	50%

(G.de LANDSHEERE. EVALUATION CONTINUE ET EXAMENS. PRECIS DE DOCIMOLOGIE. 1974, p 33.)

Pour atténuer ces effets désastreux on pourrait, me direz-vous, généraliser la double correction, comme cela se fait pour certains **concours**, le CAPES par exemple. Hélas encore! PIERON a essayé de calculer le nombre de correcteurs qu'il faudrait pour que la moyenne des notes se stabilise. Même si c'est en maths où il en faudrait le moins, il faudrait néanmoins 13 correcteurs environ pour une même copie, et 78 dans une étude de BONNIOL! Evidemment cela n'est envisageable ni au point de vue matériel, ni au point de vue financier, d'autant que cela pourrait aller jusqu'à 762 corrections!

Composition française.....	78
Version latine.....	19
Anglais.....	28
Mathématiques.....	13
Dissertation philosophique.....	127
Physique.....	16
G.de LANDSHEERE. EVALUATION CONTINUE ET EXAMENS. PRECIS DE DOCIMOLOGIE. 1974, p. 34.	

Vraie note = moyenne de :		
114	évaluations en version latine	
468	" en français	
168	" en anglais	
078	" en mathématiques	
762	" en philosophie	
095	" en physique	
J.J. BONNIOL. CONNEXIONS N°19. 1976, p.71		

La vraie note, la note objective n'existe donc pas et ne peut exister. C'est un **mythe**, une notion que J.-J. BONNIOL qualifie de "*fallacieuse sur le plan théorique, abusive sur le plan méthodologique et absurde sur le plan pratique*".

Que faire, me direz-vous ? Nous en reparlerons dans notre prochain épisode, mais avant nous irons encore d'étonnement en étonnement à propos des fluctuations de la note.