



Docimologie et biais de la notation chiffrée

Qu'est ce que la docimologie ?

- Terme inventé par **Henri Piéron** en 1923
- DEFINITION : Etude systématique des facteurs qui influencent la notation dans les examens ou lors des opérations de mesure.
- Elle étudie les écarts de notes entre correcteurs, l'application des barèmes, les échelles de notes, l'inter-corrélation entre examinateurs et la précision des correcteurs.
- Son objectif est d'atténuer dans toute la mesure du possible le rôle du hasard ou de la subjectivité dans l'attribution des notes.

Quelques exemples

1932: Expérience de multicorrection sur 100 copies de bac à Paris

- 6 groupes de 5 examinateurs
- Disciplines: le français, la philosophie, le latin, les mathématiques et la physique.

RESULTATS

- **Forte dispersion des notes** attribuées à chaque copie par les correcteurs.
- Aucune copie ne reçut deux fois la même note.
- Une même copie de **français** est notée **3 et 16** ;
- en **philo** et en **latin** l'écart maximum est de **12 points**.
- Les **mathématiques** et la **physique**: Ecart maximum de **9 points**

Quelques exemples

En 1975, l'Institut de recherche sur l'enseignement des mathématiques (IREM)

6 copies photocopées de mathématiques (niveau BEPC), soumis à 64 correcteurs, avec un barème, très précis, sur 40 points.

RESULTATS

- La dispersion des notes atteint près de 20 points !!!

Quelques exemples

En France, à la demande de **Laugier et Weinberg**, un professeur de physiologie de la **Faculté des sciences** accepta **37 copies** - dactylographiées et anonymes - qu'il avait corrigées **trois ans et demi auparavant**.

RESULTATS :

- Dans 7 cas seulement, il remit la même note au même devoir.
- Dans les 30 autres cas: écarts entre 1 et 10 points.
- L'admissibilité, avec ses nouvelles notations, aurait été modifiée: la moitié des précédents **admissibles** aurait été **refusée** et la moitié des **refusés** déclarée **admissible**

Quelques exemples

2018 : 6 copies du bac (épreuve de SES), soumises à la lecture de 30 correcteurs

Figure 9

Expérience de multicorrection de copies du bac
en sciences économiques et sociales (2006 et 2007)

	Copie 1	Copie 2	Copie 3	Copie 4	Copie 5	Copie 6
Note au bac	9	11	15	9	15	11
Note minimum	5	5	8	3	8	4
Note maximum	15	16	18	13	17	14
Écart maximum	10	11	10	10	9	10
Moyenne	8,8	9,0	13,0	8,8	12,9	8,0
Écart-type	2,4	2,7	2,5	2,1	2,1	1,8
Mode	7	8	14	9	13	8

Lecture : Pour la copie 1, notée par 30 correcteurs, la note minimum est de 5, la note maximum de 15.

Suchaut, 2008

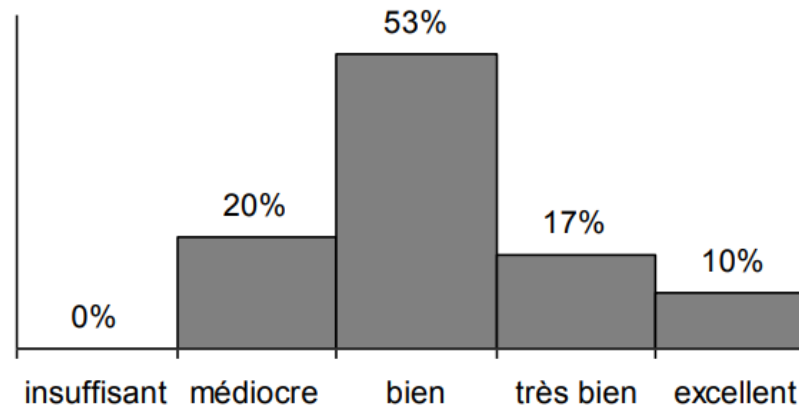
→ Ecart de note maxi de 10 pts.

→ Seules les 3 grandes catégories de copies : les faibles, moyennes, et bonnes ont été discriminées par la notation

Les biais de la notation chiffrée

- Loi de **Posthumus** (Enseignant Hollandais): L'enseignant se fonde sur la **courbe de Gauss** pour attribuer ses notes (70% de moyens ; 13% de bons; 13% de médiocres ; 2% d'excellents (génies) et 2% de mauvais (cancres))

Rot et Butas (1959) rapportent que Gjorgjevski a invité 5 professeurs d'une même branche de l'enseignement secondaire à noter indépendamment les uns des autres 100 copies de leur discipline sur une échelle à 5 degrés (1 = INSUFFISANT; 2= MEDIOCRE; 3 = BIEN, 4 = TRES BIEN ; 5 = EXCELLENT). Il a ensuite extrait 15 copies qui avaient toutes reçu la note « BIEN » par les 5 correcteurs. Elles ont été confiées, pour nouvelle correction, à 4 autres professeurs, qui ont à nouveau distribué les 15 copies à travers les 5 catégories de notes, comme l'indique la figure ci-dessous.



Les biais de la notation chiffrée

- **La constante macabre (André Antibé, 1988)**

Elle désigne le fait qu'il existerait un **pourcentage constant** de **mauvaises notes**, quel que soit le **niveau véritable** des étudiants par rapport aux connaissances réellement requises (résultats influencés par la courbe de Gauss).



Les biais de la notation chiffrée

- L'effet de **stéréotypie** : Le professeur maintient un jugement immuable sur la performance d'un élève quels que soient les efforts fournis ; il lui attribue presque toujours la même note.

Caverni, Fabre et Noizet (1975): A des professeurs de **sciences** de l'enseignement secondaire, ils ont demandé de noter (sur 20) chacun les 4 mêmes copies, accompagnées de « 5 notes censées avoir été obtenues précédemment par l'auteur de la copie ».

Copies :	a	b	c	d	Moyenne (sur les 4 copies)
Copies avec 5 notes précédentes bonnes, et des écarts faibles	12	8,5	15,25	3	9,69
Copies avec 5 notes précédentes faibles, et des écarts importants	9,75	6,5	11,75	2,75	7,69



Notes attribuées, bien plus favorables lorsque les 5 notes précédentes sont bonnes

Les biais de la notation chiffrée

- L'effet de **halo** : notation influencée par les caractéristiques de présentation de l'élève ou de sa copie. Exemple : Jugement influencé par la **tenue vestimentaire** lors d'un oral ou note d'écrit influencée par la **qualité d'écriture** ou la **présentation** d'une copie

Weiss (1969) : Deux rédactions dactylographiées ont été soumises à 2 groupes de 46 instituteurs de primaire.

Au **groupe 1**, il dit:

Le travail 1 est l'œuvre d'un élève moyen qui aime lire des BD ; son père et sa mère sont employés.

Le travail 2 a été fait par un enfant doué ; son père est rédacteur d'un quotidien connu

Pour le **groupe 2**, les commentaires ont été **inversés**.

Trois aspects (orthographe, Style, Fond) devaient être jugés indépendamment, en plus d'une « note Globale », chaque fois sur une échelle à 5 niveaux (1 = TB ; 5 = insuffisant).

RESULTATS :

-Les notes attribuées au travail pour lequel on a créé un **préjugé favorable** ont été significativement **supérieures** aux autres.

- A la copie 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 »

Les biais de la notation chiffrée

-L'effet de **contraste** : Résulte de l'interaction entre copies successives. Exemple : la copie qui suit une copie brillante risque d'être désavantagée et inversement.

De manière à mettre le phénomène en évidence, **Bonniol (1972)** a présenté une série de devoirs à corriger par deux groupes de 9 correcteurs.

Ce sont les mêmes devoirs, mais ils sont présentés dans l'ordre inverse dans les deux groupes.

RESULTAT :

Il observe que les différences (importantes) entre les deux groupes « sont plutôt imputables aux deux ordres de correction qu'aux différences de critères dont les examinateurs font état ».

Les biais de la notation chiffrée

- L'effet de **tendance centrale** : par crainte de surévaluer ou de sous évaluer un élève, le professeur groupe ses appréciations vers le centre de l'échelle.
- L'effet de **relativation** : Tendance à évaluer un travail en fonction du groupe plutôt qu'en fonction de sa valeur intrinsèque.