La Docimologie

- En 1930, le professeur Laugier sème un malaise pernicieux dans les milieux universitaires, en effectuant une expérience de multi-correction de copies d'agrégation d'histoire puisées dans les archives. 166 copies ont été corrigées par 2 professeurs travaillant séparément, sans connaître leurs appréciations respectives. Tous les deux avaient une longue expérience et corrigeaient méticuleusement. Les résultats furent surprenants.

La moyenne des notes du premier correcteur dépassait de près de deux points celle du second.

Le candidat classé avant dernier par l'un était classé second par l'autre.

Les écarts de notes allaient jusqu'à 9 points.

Le premier correcteur donnait un 5 à 21 copies cotées entre 2 et 14 par le second ;

Le second donnait un 7 à 20 copies cotées entre 2 et 11,5 par le premier.

La moitié des candidats reçus par un correcteur était refusée par l'autre.

- En 1932, *La Commission Carnégie* effectua une expérience de multi-correction en prélevant, au hasard, cent copies dans les archives du baccalauréat à Paris. Ces copies furent distribuées à 6 groupes de 5 examinateurs. Les disciplines concernées étaient : le français, la philosophie, le latin, les mathématiques et la physique. On demanda aux examinateurs de noter les copies,

Les résultats montrèrent une forte dispersion des notes attribuées à chaque copie par les correcteurs. Aucune copie ne reçut deux fois la même note. Une copie de français est notée 3 et 16; en philosophie et en latin l'écart maximum est de 12 points.

Les mathématiques et la physique, réputées pour des sciences exactes, ne sont pas épargnées : l'écart maximum est respectivement de 9 et 8 points.

- En 1975, l'Institut de Recherche sur l'Enseignement des Mathématiques (IREM) de Grenoble entreprend une expérience analogue de multi-correction.

Un échantillon de 6 copies photocopiées de mathématiques (niveau 3ème) est soumis à 64 correcteurs, avec un barème, très précis, sur 40 points. La dispersion des notes est près de 20 points.

Pas d'accord entre eux, les examinateurs ne se montrent pas davantage fidèles à eux-mêmes

Les enquêteurs anglais en avaient été frappés. Laugier et Weinberg en France, avaient fait la même constatation. Un professeur de physiologie de la Faculté des Sciences accepta 37 copies -dactylographiées et anonymes - qu'il avait corrigées trois ans et demi auparavant.

Dans 7 cas seulement, il remit la même note au même devoir.
Dans les 30 autres cas, il y eut des divergences comprises 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.

Le degré d'accord de ce professeur avec lui-même ne fut pas plus élevé qu'avec deux de ses collègues chargés de la même tâche : les coefficients de corrélation atteignant respectivement 0,58, 0,59 et 0,56.

Les expériences existent depuis bien longtemps
mais tout se passe comme si elles n'avaient jamais eu lieu.
Soit on ne les connaît pas, soit on n'y "croit pas" !