

Peut-on dire que le garçon qui grimpe 5 mètres à la corde en 10" 7 — noté 14 au bac — est deux fois meilleur que celui qui les grimpe en 17" 7 — noté 7 au bac ? (Jacques GEORGE, E.N. des Battignoles).

Les discordances de la notation

De nombreux ouvrages de docimologie fournissent des exemples variés de discordances et de manques de fidélité (entre juges). Notre objectif n'est pas de les développer ici, nous nous contentons d'en synthétiser les conclusions.

Les raisons des divergences de notation sont doubles :

Facteurs internes d'abord

En l'absence de toute concertation sur les objectifs poursuivis, des juges, se basant sur leur expérience personnelle et les possibilités de leurs élèves, se donnent des critères différents et parfois peu stables. Même si un accord peut être atteint quant aux principaux critères à retenir dans l'évaluation, leur importance respective varie souvent de juge à juge. Ces problèmes prennent un relief d'autant plus accusé que la performance de l'élève est complexe.

Facteurs externes ensuite

Les facteurs que nous venons d'évoquer sont inhérents à la démarche d'évaluation menée de façon trop individuelle et trop empirique. En outre, la note peut être contaminée par l'intervention de facteurs totalement étrangers à la performance évaluée. G. De Landsheere (*Évaluation continue et examens, Précis de docimologie*, Bruxelles, Paris, Labor-Nathan, 1971) les présente comme suit :

— **Effet de stéréotypie** : le professeur maintient un jugement immuable sur la performance d'un élève, quelles que soient ses variations effectives.

— **Effet de halo** : le professeur, influencé par des caractéristiques de présentation de l'élève, sous-estime ou surestime la note. Plusieurs recherches ont mis en évidence une manifestation voisine de l'effet de halo : des caractéristiques de présentation de travail, telles que l'orthographe et l'écriture influencent significativement la note attribuée par le professeur, alors que ce dernier est averti que ces dimensions doivent être exclues de l'évaluation (voir D.-P. Svannel and J.-C. Marschall, *The effect of selected composition errors on grades assigned to essay examinations* ; C. Chase, *The impact of some obvious*

variables on essay test scores, in *Journal of educational measurement*, 1968, 5, 135-318 ; J. Marschall, *Composition errors and essay examination grades reexamined*).

— **Effet de contamination** : cet effet est de nature plus mécanique : les notes portées successivement pour différents aspects d'un même travail s'influencent mutuellement.

— **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.

— **Effet de trop grande indulgence et de trop grande sévérité** : certains juges sont systématiquement trop indulgents ou trop sévères dans toutes leurs évaluations.

— **Effet de l'ordre de correction** : devant un nouveau travail ou un nouveau candidat à évaluer, un juge se laisse influencer par la qualité du candidat précédent. Par exemple un travail moyen paraîtrait bon s'il suivait un travail médiocre.

— **Effet de relativation** : plutôt que de juger intrinsèque d'un travail, des professeurs jugent ce dernier en fonction de l'ensemble des travaux dans lequel il est inséré. Une expérience menée par Gjorgjevski a souligné ce phénomène. De 100 épreuves écrites jugées « bonnes » par 5 correcteurs, on a extrait 15 copies que l'on confie à 4 nouveaux correcteurs. On observe une nouvelle répartition gaussienne des notes attribuées à ces copies qui s'étalent de l'excellent au médiocre.

La présentation systématique adoptée ici ne doit pas laisser croire à l'indépendance de ces différents facteurs. Dans la réalité, leurs effets se conjuguent souvent.

Des techniques particulières existent pour pallier l'influence de certains défauts qui entachent la correction : corriger anonymement, corriger toutes les réponses à une question avant d'aborder la suivante, varier l'ordre de correction des copies, etc. Mais au-delà de ces techniques, une démarche plus fondamentale s'impose : définir en commun les objectifs de l'exercice proposé et mettre au point des critères d'évaluation relatifs à ces objectifs.