

L'évaluation

Problèmes liés à la correction des évaluations

2006 - n°3

S'il existe de nombreux problèmes liés à la correction des évaluations, je souhaite, en guise de préambule, attirer votre attention sur les inégalités engendrées par la formulation des questions.

Etant donné que, dans notre système d'enseignement, l'enseignant à toute latitude pour composer ses propres questions, l'étudiant a plus ou moins de chances de réussite selon qu'il est interrogé dans une école A ou une école B pour un même cours. Cette « injustice » est d'autant plus marquée si d'autres paramètres sont par ailleurs identiques : par exemple, au sein d'une même école, plusieurs enseignants sont titulaires du même cours, annoncent les mêmes objectifs de départ et enseignent à partir du même support de cours (syllabus). Si, dans ce cas, les questions d'évaluation divergent selon les enseignants, cela signifie qu'un étudiant peut très bien échouer alors qu'il aurait réussi chez un autre professeur et vice-versa.

Par ailleurs, un autre facteur entre en ligne de compte dans la conception des questions. Celles-ci doivent être en adéquation avec ce qui a été enseigné. Peut-on imaginer un enseignant en hôtellerie qui passerait son temps de cours à expliquer comment cuisiner (de manière théorique, sans exercice pratique) et qui demanderait à ses étudiants lors de l'évaluation de « réaliser » un mets ? Cette méthode semble inconcevable. Nous nous accorderons probablement à penser que le cours doit permettre une mise en pratique si on évalue la pratique. Pourtant, bien souvent, les enseignants utilisent les temps de cours pour transmettre des informations aux étudiants alors qu'ils leur posent des questions, à l'examen, qui nécessitent d'autres processus jamais exercés durant les périodes de cours : analyser, comparer, critiquer, etc.

La méthode d'enseignement doit donc être en adéquation avec l'évaluation (voir REPI n°1).

Voici à présent une liste non exhaustive des problèmes liés à la correction des évaluations. En prendre conscience, c'est pouvoir agir afin de les éviter ou, du moins, les amenuiser.

1. Divergences entre correcteurs

Pierre, 18 ans, rentre dépité à la maison : il a demandé à son frère, prof de français, de rédiger pour lui un résumé de texte. Aujourd'hui, il a reçu sa note : Faible.

Ce genre d'anecdote est fréquente... De manière plus rigoureuse, de nombreuses expériences ont démontré qu'un étudiant classé « médiocre » par un correcteur pouvait être classé « excellent » par un autre et vice-versa. Chaque correcteur disposant de critères, parfois inconscients, qui lui sont propres et qui diffèrent probablement de ceux des autres correcteurs.

Eviter ce biais passe par une définition des critères de correction la plus précise possible et une transmission de ces critères aux étudiants (en donnant si nécessaire des exemples de réponses considérées comme correctes ou incorrectes). Une grille de correction peut être utilisée lorsque plusieurs correcteurs sont concernés, elle permet également de limiter les problèmes développés aux points suivants.

2. Fidélité dans le temps chez un même correcteur

G. De Landsheere rapporte le constat suivant : 14 historiens ont été invités à réexaminer des copies un an à un an et demi après les avoir évaluées une première fois (toute trace de correction ayant été effacée). Dans 92 cas sur 210, le verdict de réussite ou d'échec a différé entre les 2 corrections !

Manifestement, un même correcteur n'est pas égal à lui-même... ceci s'explique notamment par les biais développés aux points 3, 4, 5 et 6.

3. Effets d'ordre de correction

Stéphanie s'effondre en découvrant l'ordre de passage de l'examen oral : elle passera directement après Layla, la meilleure de la classe...

Il apparaît en effet que les correcteurs notent par contraste avec la copie précédente. Par ailleurs, les dernières copies d'une série sont notées plus sévèrement que les premières (selon une expérience de Bonniol en 1972). Par conséquent, Stéphanie aurait probablement de meilleurs résultats si elle se trouvait dans les premiers étudiants et après un élève plutôt médiocre...

4. Le biais de clémence ou sévérité

Voici quelques commentaires d'enseignants :

- « *Personne n'est nul au point d'obtenir un zéro. Il faut au moins accorder un demi-point pour la feuille et l'encre...* »
- « *Un 20/20, ça n'existe pas. Personne n'est parfait !* »
- « *Si l'étudiant répond correctement à la question, c'est-à-dire en donnant tous les éléments vus au cours, il a 12/20. Pour avoir plus, il faut qu'il m'épate.* »

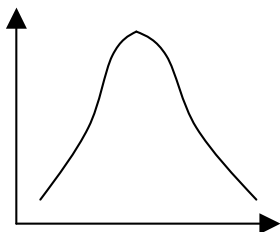
Certains correcteurs sous-évaluent ou surévaluent systématiquement une performance.

5. Effet de tendance centrale

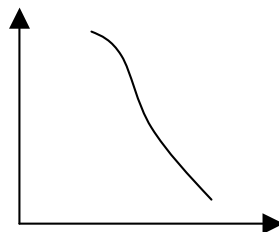
A contrario, de peur de sur ou sous-évaluer, d'autres enseignants rassemblent leurs cotes vers le centre de l'échelle.

6. La loi de POSTHUMUS et la courbe de GAUSS

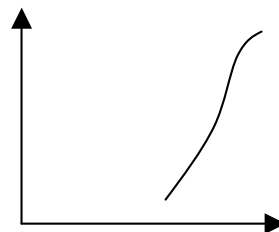
Formulée dès 1947, la loi de POSTHUMUS peut s'exprimer, selon Crahay, de cette manière : « *Un enseignant tend à ajuster le niveau de son enseignement et ses appréciations des performances des élèves de façon à conserver d'année en année, approximativement la même distribution (gaussienne) des notes* ».



Gr.1 : Courbe de Gauss



Gr. 2 : Courbe en i



Gr.3 : Courbe en j

L'enseignant aurait donc tendance à coter de manière à ce que la majorité des étudiants se regroupent autour de la moyenne, quelque-uns s'en écartant positivement ou négativement (graphique 1). Or, la situation « idéale » en formation serait la suivante : les individus en début de parcours se classeraient selon une courbe en i (graphique 2) à un « prétest » (ce qui signifie que la plupart maîtrisent peu les objectifs) et selon une courbe en j (graphique 3) à un « post-test », ce qui voudrait dire que la plupart ont acquis les objectifs poursuivis.

7. La prise en compte de critères non académiques

- « *Judith travaille tellement, elle en veut. Elle est motivée, elle fait beaucoup d'efforts pour essayer d'atteindre le niveau demandé* »
- En délibération de juin : « *C'est vrai que Jérôme commet énormément d'erreurs orthographiques mais il a déjà décroché un boulot où l'écrit a peu de place... Il doit commencer le 1^{er} juillet. On ne va quand-même pas lui refuser son diplôme !* »

Des critères tels que l'effort ou les ambitions futures ne peuvent être pris en compte dans l'attribution d'une note (à moins qu'ils ne soient justement l'objet même de l'évaluation).

Une piste de solution : la correction anonyme. Par cette méthode, l'enseignant ne sait pas de qui il corrige la copie. Cette démarche n'est pas toujours évidente, notamment en cas de questions ouvertes pour de petits groupes d'étudiants, cas dans lequel l'enseignant peut facilement reconnaître une écriture par exemple.

8. L'effet de halo

Il s'agit de l'influence de caractéristiques externes sur la note. Par exemple, l'écriture, le soin, l'attitude d'une personne. Le type d'examen joue ici un rôle déterminant. Cet effet se produira plus souvent dans le cas d'examens oraux ou d'écrits à questions ouvertes que dans le cas de QCM.

9. L' Effet de contamination

Il s'agit de l'influence mutuelle des cotes respectives attribuées aux différents aspects d'un travail. Par exemple, l'enseignant commence par noter l'orthographe d'un travail. Si les erreurs sont nombreuses, voire empêchent la compréhension, l'enseignant est enclin à une cote négative également pour la syntaxe, pour le fond, etc. Par contre, si l'orthographe est impeccable (ou presque), l'enseignant se trouve dans de « bonnes dispositions » pour les autres aspects de son évaluation. Dans tous les cas, il va chercher une confirmation de sa première impression.

Ici aussi, l'utilisation d'une grille de correction est un outil utile.

Par ailleurs, cet effet se marque également lors de la correction d'évaluations à questions ouvertes. Lorsque la réponse à la première, voire à la deuxième question est considérée comme « bonne » ou « mauvaise » par l'enseignant, celui-ci a tendance à aborder la correction de la troisième question avec ce même a priori.

Une piste de solution consiste alors à corriger toutes les copies question par question.

10. L'effet de stéréotypie

Il s'agit de la tendance pour un enseignant à fonder son jugement sur l'étudiant (par exemple après la correction d'un premier travail) et non sur sa performance réelle du moment. L'étudiant est ainsi « catalogué », on lui appose une « étiquette » de laquelle il a bien des difficultés à se débarrasser.

11. Effet pygmalion ou effet oedipien

Si la prédiction de l'oracle à propos d'Œdipe s'est réalisée, c'est parce qu'elle a été révélée... Sinon, Œdipe aurait connu son père (et ne l'aurait pas tué) et sa mère (et ne serait pas devenu son amant).

Rosenthal et Jacobson ont prédit à des enseignants le fabuleux potentiel de 5 élèves (à la suite d'un test qu'ils ont fait passer dans leur classe) et ce, de manière tout à fait aléatoire ! Au terme de l'année scolaire, ces 5 élèves ont réalisés d'étonnants progrès !

D'autres chercheurs ont tenté l'expérience sans toutefois obtenir des résultats aussi impressionnants. Pourtant, apparemment, les attentes des enseignants par rapport à la réussite de leurs étudiants se manifestent par des comportements qui influencent les résultats de manière non négligeable.

Sources :

Le trait d'Union Express – Université de Sherbrooke – « Evaluation des apprentissages. Comment éviter les pièges de la correction ? » ; http://www.usherbrooke.ca/ssf/tu/vol_4/no_2/correct.html (consulté le 31/1/2006)

Crahay M, Cours de Pédagogie, « Chapitre 2. L'effet POSTHUMUS ou les mirages de l'évaluation normative » : <http://www.ulg.ac.be/pedaexpe/cours/dos5b/d5b-p2.htm> (consulté le 31/1/2006)

Pour toute question, pour tout commentaire, pour une collaboration, pour poursuivre la réflexion dans le contexte de votre cours, le Service Pédagogique d'Hemes reste à votre disposition.

Laurence Oger - 04/344.62.25 - Courriel: service.pedagogique@hemes.be - Pages web: www.hemes.be/spp

Exemple de grille de correction

Voici un exemple de grille de correction pour un résumé de texte, accompagné d'une pondération en fonction des critères.

Critères	Commentaire	Points - total sur 20
1. Respect de la consigne chiffrée (longueur). Exemple pour un texte de 418 mots → résumé attendu : 100 mots.	Si l'étudiant dépasse la limite de <ul style="list-style-type: none"> ➤ +15 % ou -15 % (<i>plus de 115 mots ou moins de 85 mots dans notre exemple</i>): - 2 points ➤ entre + ou - [15-30%] : - 4 points ➤ entre + ou - [30-50%] : - 6 points ➤ au-delà de 50% : - 10 points (ce n'est plus un résumé, l'étudiant est en échec automatiquement) 	Retrait de points sur le score obtenu
2. Fond : <u>1) Fidélité informationnelle (/6)</u> <ul style="list-style-type: none"> ➤ Exhaustivité ➤ Parmi les idées reprises, exactitude de l'information ➤ Insertion d'idées extérieures (jugement, sentiment personnel, info externe, etc.) <u>2) Réécriture personnelle (/4)</u> <ul style="list-style-type: none"> ➤ Vocabulaire <ul style="list-style-type: none"> ➤ Syntaxe 	→ L'enseignant cible les informations qu'il juge essentielles, la pondération s'établit en fonction de la proportion d'infos reprises → Maintien/perte de sens → Retrait : un point par idée extérieure → Utilisation de <u>synonymes</u> et maintien des <u>termes spécifiques</u> → retrait de points en cas de confusion de termes phonétiquement proches, inadéquation de mots dans le contexte → Reconstruction personnelle des phrases	<u>Total sur 10</u> → sur 4 → sur 2 → --- → sur 3 → --- → sur 1
3. Forme <u>1) Cohérence de la phrase (/4)</u> <ul style="list-style-type: none"> ➤ Sens global de la phrase ➤ Syntaxe de base maîtrisée ➤ Phrases complexes <u>2) Cohérence du texte (/2)</u> <u>3) Orthographe (/4)</u> <ul style="list-style-type: none"> ➤ Nbre d'erreurs d'usage : ____ ➤ Nbre d'erreurs grammaticales : ____ 	→ Si phrases incompréhensibles, retrait de points → sujet-verbe-compléments → Coordination et subordination à bon escient et correctement construites → Si erreur de ponctuation, retrait de points → Enchaînements logiques des idées → de 0 à 2 erreurs : 4/4 ; de 3 à 4 erreurs : 3/4 ; de 5 à 6 erreurs : 2/4 ; de 7 à 8 erreurs : 1/4 ; plus de 8 erreurs : 0/4	<u>Total sur 10</u> → --- → sur 2 → sur 2 → --- → sur 2 → sur 4