HAL
archives-ouvertes.fr

L'évaluation par les pairs pourra-t-elle faire de l'examen une vraie activité pédagogique ?

L. Hadi Bouzidi, Alain Jaillet

To cite this version:

L. Hadi Bouzidi, Alain Jaillet. L'évaluation par les pairs pourra-t-elle faire de l'examen une vraie activité pédagogique ?. EIAH, Jun 2007, France. INRP, 2007. <hal-00161484>

HAL Id: hal-00161484

https://hal.archives-ouvertes.fr/hal-00161484

Submitted on 10 Jul 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

L'évaluation par les pairs pourra-t-elle faire de l'examen une vraie activité pédagogique ?

L'hadi Bouzidi*, Alain Jaillet**

** Cité des 20 logements*
BtB1 N° 66
06000 Béjaïa, Algérie
lhadi_bouzidi@yahoo.fr
*** ULP Multimédia*
16 rue Descartes
F 67000 Strasbourg, France
alain.jaillet@ulpmm.u-strasbg.fr

RÉSUMÉ. *L'examen, en tant qu'évaluation sommative, constitue un investissement important qui ne produit que la note de l'étudiant. Les copies d'un tel examen et les corrections des enseignants ne constituent pas, en général, une ressource pour l'apprentissage. Pour remédier à cette situation, nous proposons de faire de l'examen une vraie activité pédagogique. Nous suggérons notamment de le « prolonger » dans un campus virtuel et de l'exploiter pour l'apprentissage, en mettant en place un système d'évaluation formative par les pairs. Notre expérimentation qui a porté sur 68 étudiants a révélé, en plus d'un apport pédagogique, que l'évaluation par les pairs est équivalente à l'évaluation par l'enseignant et cela pour un examen composé de questions majoritairement fermées.*

MOTS-CLÉS : *évaluation par les pairs, évaluation formative, prolongement virtuel de l'examen en classe.*

ABSTRACT. *The exam, as a sommative assessment, constitutes a very important investment which produces only the notes of the student. The exam sheets and the teacher's feedbacks do not constitute, in general, a resource for training. For this reasons, we propose to make with the exam a true learning activity. We, in particular, suggest "to extend it" in a virtual campus and to exploit it for the training by setting up a system of formative peer assessment. Our experimentation, in which 68 students participated, reveals, in addition to one real training benefit, that the peer assessment is equivalent to the teacher assessment for exams where the questions are mainly closed.*

KEY WORDS: *Peer assessment, formative assessment, virtual prolongation of exam.*

1. Introduction

L'évaluation des apprentissages est l'une des tâches les plus délicates pour un enseignant. L'examen en classe exige un grand investissement de tous les acteurs gravitant autour du processus d'apprentissage (étudiants, enseignants et administration). Il sert, le plus souvent, à récolter des notes pour décider si l'étudiant peut ou non être accepté en regard du règlement des études, soit pour poursuivre dans le niveau supérieur, soit pour se voir décerner le diplôme. L'enseignant peut, par ailleurs, procéder à une évaluation dite formative, qui permet de prendre des mesures pour remédier aux lacunes éventuelles de son processus d'enseignement (pour le formateur) et d'apprentissage (pour l'étudiant). Dans ce contexte, l'évaluation par les pairs [TOPPING 1998] en classe peut constituer une technique très efficace pour l'apprentissage. Chacun peut constater que, malgré l'importance de l'investissement qui se fait pour les examens, ces derniers restent dans les archives et constituent plus une tâche administrative qu'une activité pédagogique. Nous proposons de remédier à cela en faisant de l'examen un processus au service de la pédagogie, grâce à l'utilisation de l'évaluation par les pairs dans un environnement virtuel d'apprentissage. Ainsi, nous présentons dans ce qui suit, l'état de l'art sur l'évaluation par les pairs via Internet, puis nous explicitons notre démarche d'implémentation d'un prolongement virtuel des examens effectués en classe impliquant une évaluation par les pairs. Nous présentons, ensuite le processus d'expérimentation que nous avons mis en place et nous livrons enfin, l'analyse et l'interprétation des données récoltées.

2. Evaluation par les pairs via Internet

L'utilisation de l'évaluation par les pairs via Internet est récente. Elle permet à des apprenants d'évaluer les travaux de leurs pairs et éventuellement d'évaluer leur propre travail. A la différence des techniques d'autoévaluation qui, le plus souvent, se limitent aux domaines cognitifs de base (apprendre, comprendre) [BLOMM, 1956] [ANDERSON & KRATHWOHL 2001], l'évaluation par les pairs permet de développer un apprentissage dans des domaines cognitifs élevés (application, analyse, évaluation). Il s'agit d'une approche qui implique les étudiants dans le processus de la révision, de l'évaluation et du feedback sur des travaux en ligne. J. A Gilles Doiron [DOIRON 2003] explique que certains auteurs critiquent l'utilisation des TIC pour l'évaluation par les pairs en argumentant qu'elle est moins rigoureuse que des formes d'évaluations traditionnelles, qu'elle exige trop des étudiants en leur mettant trop de pression, qu'elle n'est pas fiable, et qu'elle n'est pas nécessairement équitable. En réponse à cela, d'autres auteurs comme Bostock [BOSTOCK 2000] sont convaincus que l'évaluation (formative ou sommative) des travaux d'autres étudiants par les étudiants eux même a beaucoup d'avantages pour l'apprentissage, que ce soit pour l'évaluateur ou l'évalué. Bostock fait remarquer que l'évaluation par les pairs encourage l'autonomie des étudiants et développe des compétences dans des domaines cognitifs élevés. Il reconnaît certaines faiblesses de ce type

d'évaluation, mais explique qu'elles peuvent être évitées en mettant en place un système d'anonymat, de multiples évaluations et évaluateurs et une modération par l'enseignant. De plus, Bostock, argumente que l'Internet et les TIC permettent d'assister la gestion d'un grand nombre d'étudiants.

Il existe peu d'outils ou systèmes d'évaluations par les pairs en ligne. Nous citons les exemples suivants : Le programme CPR (UCLA-USA) [CHAPMAN 2001], le système OASYS [BHALERAO & WARD 2001] (université Warwick au Royaume-Uni), le programme « Peer Grader » [GEHRINGER, 2001] (université de l'Etat de la Caroline du Nord) et le système « NetPeas » [LIN et al. 2001]. L'outil d'évaluation par les pairs que nous avons utilisé pour notre expérimentation est le module « Atelier » de la plate forme Moodle.

3. Notre expérimentation

A la différence des systèmes existants, nous proposons d'étendre l'évaluation par les pairs en ligne aux travaux réalisés en classe (hors ligne), notamment pour les examens surveillés. Ainsi, le processus que nous proposons s'articule autour des phases suivantes : 1 - organiser un examen en classe, 2 - récupérer puis numériser les copies d'examen, 3 - traiter les copies numérisées pour assurer l'anonymat des étudiants, 4 - déposer les copies numérisées dans le campus virtuel pour chaque étudiant, 5 - préparer l'activité sur le campus virtuel en indiquant le barème, les critères d'évaluation et divers paramètres pour le calcul des notes et la distribution des copies aux étudiants, 6 - ouvrir les évaluations (pour l'enseignant et pour les pairs) et 7 - remettre les notes et les feedback aux étudiants et ouvrir les discussions.

Notre expérimentation a été réalisée, selon le processus indiqué ci-dessus (figure-1). Elle a porté sur un examen d'architecture d'ordinateur ayant impliqué 68 étudiants et comportant 17 questions majoritairement fermées.

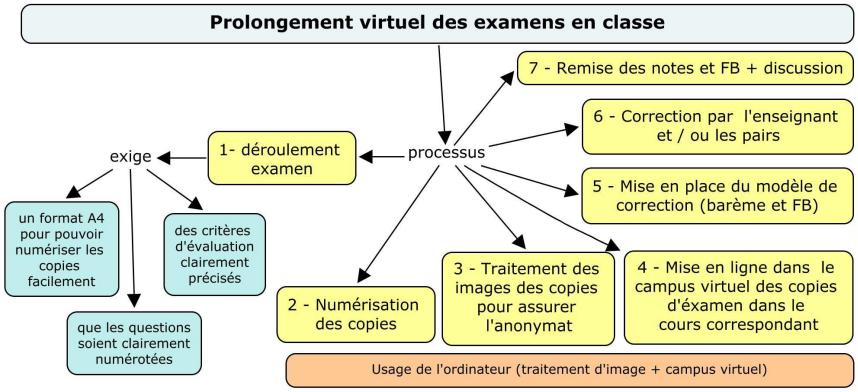

Figure 1. *Evaluation en ligne par les pairs des examens effectués en classe*

4. Récolte des données, analyse et interprétation

68 étudiants ont participé à l'expérience. 49 parmi ces 68 étudiants ont évalué leur propre travail. Nous avons récupéré, pour chaque copie : Une note d'autoévaluation, une à 4 notes attribuées par les pairs et une note attribuée par l'enseignant. Nous avons aussi récolté des données qualitatives à l'issue d'un questionnaire publié dans le campus virtuel et dont l'objectif est d'estimer l'apport pédagogique de notre expérimentation.

Nous avons voulu vérifier si l'autoévaluation et l'évaluation par les pairs donnent les « mêmes » résultats que l'évaluation par l'enseignant. Pour cela nous avons entrepris un test « *t* » de comparaison statistique appliquée à deux échantillons appariés (se rapportant aux mêmes individus, c'est-à-dire des notes). Il s'agit de calculer une valeur « *t* » et de la comparer à une valeur $t\alpha$ déterminée par un seuil d'erreur α (fixé à 5%) et un degré de libertés ν dans la table des dépassements de l'écart absolu (loi de Student). Nous avons défini un certain nombre de grandeurs résumées dans le tableau suivant:

n	Nombre d'individus de chaque échantillon (n = 69 notes par les pairs et 49 notes par autoévaluation)
Ne_i	Note attribuée par l'enseignant à la copie i
Na_i	Notes attribuées par l'auteur à sa propre copie i
Np_i	Moyenne des notes attribuées par les pairs à la copie i
$d_i = Ne_i - Np_i$	Différence entre la note attribuée par l'enseignant et la note attribuée par les pairs pour une copie i (cas d'une comparaison entre l'évaluation par l'enseignant et l'évaluation par les pairs).
$d_i = Ne_i - Na_i$	Différence entre la note attribuée par l'enseignant et celle attribuée par l'auteur de copie i à sa propre copie (cas d'une comparaison entre l'évaluation par l'enseignant et l'évaluation par les pairs).
$\bar{d} = \frac{\sum_{i=1}^{n} d_i}{n}$	Moyenne des différences d_i.
s_d	écart type de la série des différences d_i.
$s_{\bar{d}} = s_d / \sqrt{n}$	Erreur type.
$t = \bar{d} / s_{\bar{d}}$	La valeur t à rechercher

Tableau 2. *Grandeurs utilisées pour le test statistique de comparaison.*

Dans le cas d'une comparaison entre l'évaluation par l'enseignant et l'évaluation par les pairs, nous avons trouvé une valeur de $\boldsymbol{t}$ **= 1,54406.** Sachant que nous avons un degré de libertés $\nu = (n\text{-}1) = 67$, la table de dépassement de l'écart absolu (loi de Student) donne une valeur $\boldsymbol{t\alpha}$ **= 1.9944**. Le test stipule que nous ne devons accepter l'hypothèse selon laquelle les deux types d'évaluation soient équivalentes que si la valeur t calculée est inférieure ou égale (en valeur absolue) à $t\alpha$. Ceci étant le cas, on conclu que l'évaluation par les pairs est équivalente à l'évaluation par l'enseignant.

Dans le cas d'une comparaison entre l'évaluation par l'enseignant et l'autoévaluation, nous avons trouvé une valeur de $\boldsymbol{t}$ **= - 4,6121.** Sachant que nous avons un degré de libertés $\nu = 48$, la table de dépassement de l'écart absolu (loi de Student) donne une valeur $\boldsymbol{t\alpha}$ **= 2.0106**. Cette valeur n'étant pas inférieur à $\boldsymbol{t}$, on conclu que l'autoévaluation n'est pas équivalente à l'évaluation par l'enseignant. On pourra même affirmer, du fait que la moyenne des différences d_i est inférieure à 0, que l'autoévaluation surévalue les copies d'examen.

Les réponses au questionnaire soumis aux étudiants ayant participés à notre expérience révèlent des bénéfices pédagogiques. En effet, la participation à un processus d'évaluation (activité se situant au sommet de la taxonomie de Bloom) a fait que les étudiants acquièrent un apprentissage à la fois affectif (du fait qu'ils savent que les copies sont celles de leurs camarades) et efficace (du fait que cette activité les a obligé à relire leur cours, à voir leurs erreurs, à comprendre les fautes de leurs camarades et à construire des feedbacks justifiants leurs notes). Cependant, ces étudiants ont révélés quelques lacunes : des problèmes de connexion au campus virtuel, des difficultés d'appliquer les barèmes, des difficultés dans la remise des feedbacks et un manque d'interactions entre eux.

5. Conclusion

Les résultats obtenus démontrent la faisabilité et l'utilité de l'évaluation par les pairs sur des examens réalisés en classe. Le résultat stipulant que l'autoévaluation engendre des notes surévaluées est prévisible. L'autoévaluation constitue un très bon exercice pour l'évaluation formative. Le manque d'interactions entre les étudiants, pouvant être expliqué par l'anonymat des copies, suggère de dévoiler, après le processus d'évaluation par les pairs, l'identité des auteurs des copies en prévoyant une phase d'interaction. L'évaluation formative impliquant des interactions fréquentes entre l'enseignant et ses étudiants et l'élaboration de feedbacks, de consignes et d'orientations, alourdie la charge de travail de l'enseignant. La méthode proposée dans cet article, peu alléger cette charge en confiant une partie de cette évaluation aux étudiants. L'équivalence entre l'évaluation par les pairs et l'évaluation par l'enseignant pour des examens comportant des questions fermées ou semi fermées constitue un autre résultat important de notre recherche. Ce résultat pourra être exploité pour encourager à utiliser l'évaluation par les pairs comme évaluation sommative.

6. Bibliographie

[ANDERSON & KRATHWOHL 2001] Anderson, W., A., & Krathwohl, R.,, D., « A taxonomy for learning, teaching, and assessing; A revision of Bloom's taxonomy of educational objectives », *Addison Wesley Longman*, Inc., New York, 2001.

[BHALERAO & WARD 2001] Bhalerao, A., & Ward, A., « Toward Electronically Assisted Peer Assessment: A Case Study », *ALT-J*, Vol. 9, No. 1, pp. 26-37.

[BLOOM 1956] Bloom, B, S. « Taxonomy of Educational objectives: The Classification of Educationnal Goals », *Handbook I: Cognitive Domain*. New York: Longmans Green.

[GEHRINGER 2001] Gehringer, E., F., « Electronic peer review and peer grading in computer science course », *Proceeding of the thirty second SIGCSE technical symposium on Computer Science Education*, Vol. 33, No 1, ACM Press, New York, 2001, pp. 139-143.

[DOIRON 2003] Doiron, G., « The value of online student peer review, evaluation and feedback in higher education », *CDTL Brief*, Centre for Development of Teaching and Learning, National University of Singapore (NUS), vol.6 n°.9, September 2003.

[LIN et al. 2001] Lin, S. S., Liu, E. Z. and Yuan, S., « Web based peer assessment: attitude and achievement », *IEEE Transactions on Education*, vol.9, n°.4.

[TOPPING 1998] Topping, K., « Peer Assessment Between Students in Colleges and Universities », *Review of Educational Research*, 68(3), 249-276.

7. Références sur le WEB.

[BOSTOCK 2000] Bostock, S., « Student Peer Assessment, Learning Technology », Keele University, UK, (Electronic Citation 2006) http://www.keele.ac.uk/depts/aa/landt/lt/docs/bostock_peer_assessment.htm.

[CHAPMAN 2001] Chapman, O. L., « Calibrated Peer Review ™, The White Paper. A Description of CPR », (Electronic Citation 2006), http://cpr.molsci.ucla.edu.