

Le « Classroom Assessment » dans l'Enseignement Supérieur : Diagnostiquer des Faiblesses dans son Enseignement

Marianne Poumay, juillet 2005

A l'origine de la professionnalisation de l'enseignement et de l'apprentissage, le courant de *l'évaluation de classe* (en anglais « classroom assessment ») est une composante d'un autre courant, celui de la *recherche en classe*, né aux Etats-Unis au début des années 1990. Il a pour but d'aider les enseignants du supérieur à mieux comprendre ce que les étudiants apprennent à leur cours et avec quelle qualité ils apprennent.

La *recherche en classe* pousse les enseignants à tenter de mieux comprendre les processus d'apprentissage en œuvre chez leurs étudiants et à isoler les facteurs sur lesquels leur enseignement peut avoir une influence.

Une prise de conscience au quotidien

L'*évaluation de classe* fait un large usage de questionnaires, d'observations, de tests et de feedback multiples pour sans cesse améliorer l'apprentissage des étudiants. Ce courant est centré sur l'apprenant mais est conduit par l'enseignant. Il bénéficie autant à l'un qu'aux autres. Il est formatif et très dépendant du contexte (différent dans chaque cours). Il est permanent et se base sur la connaissance de ce qu'est un enseignement efficace. Il tente de rendre les enseignants plus conscients de ce qui se passe bien et moins bien dans leur groupe d'étudiants, de façon à agir sur les points faibles et renforcer ainsi l'apprentissage des étudiants. Il incite l'enseignant à prendre des mesures de façon continue dans son groupe classe, à émettre des hypothèses quant à des facteurs d'amélioration de l'apprentissage, à tester ces hypothèses par des interventions sur son propre groupe et à en mesurer les effets en termes de satisfaction des étudiants et, surtout, en termes d'impact sur leur apprentissage.

Pour aller (un peu) plus loin, nous conseillons vivement la lecture du bref article en ligne

(<http://honolulu.hawaii.edu/intranet/committees/FacDevCom/guidebk/teachtip/assessment-1.htm>) de Thomas A. Angelo et K. Patricia Cross issu de l'ouvrage *Classroom Assessment Techniques, A Handbook for College Teachers*, (1993, 2nd Ed), décrivant les principes et caractéristiques de cette approche, ses présupposés et des suggestions pour sa mise en œuvre (environ 4p). Des mêmes auteurs, un autre lien (environ 2p, voir <http://honolulu.hawaii.edu/intranet/committees/FacDevCom/guidebk/teachtip/assessment-2.htm>), mène vers 5 exemples de techniques très faciles et rapides pour l'*évaluation de classe*.

Autres liens intéressants :

- Descriptions et exemples de techniques d'*évaluation de classe* proposées par l'U. d'Oregon sur <http://tep.uoregon.edu/resources/assessment/cats/cats.html>
- Développé par un large groupe d'enseignants, le « Field-testing Learning Assessment Guide » (FLAG) propose une série de techniques et les décrit de façon systématique et opérationnelle. <http://www.flaguide.org/cat/cat.php>. Il mêle cependant les techniques simples et complexes, celles qui rendent compte de façon assez brute de l'apprentissage d'un groupe et celles qui renseignent finement sur l'apprentissage de chaque étudiant. Intéressant pour donner des idées d'action.

Des mesures de l'apprentissage individuel

Au-delà de ces techniques très simples à pratiquer au quotidien avec son groupe classe, d'autres techniques, plus poussées et visant chaque étudiant en tant qu'individu, servent à évaluer l'apprentissage de chacun et à lui donner des cotes. Il s'agit par exemple du traditionnel testing destiné à mesurer les apprentissages à des moments-clé. En plus de donner une idée de ce que les étudiants ont retenu, compris, appliqué, synthétisé, etc. d'une partie du cours, ces jugements sur les performances individuelle des étudiants aideront à se faire une idée plus fine de leur niveau de compétence... et donc aussi de l'impact d'une action de formation.

En effet, pour évaluer l'**impact** d'un cours en ligne ou, plus généralement, d'une action de formation, il est important de se donner les moyens de mesurer finement l'apprentissage des étudiants. Cet apprentissage se mesure généralement par une multitude de notes attribuées à l'étudiant au fur et à mesure des différentes activités jalonnant le cours.

La mesure doit être objectivée au maximum, de façon à approcher de façon valide l'impact de la formation. Mais donner des cotes aux étudiants n'est pas chose facile. L'enseignant doit objectiver la cotation et la rendre fiable, mais aussi multiplier les cotes et les rendre suffisamment diagnostiques pour donner aux étudiants l'occasion de prendre conscience de leurs manques et de s'améliorer. Des cotes les plus objectives possibles, nombreuses et diagnostiques aideront l'enseignant à mesurer l'impact de son action de formation.... et faciliteront en même temps la réussite de ses étudiants.

Nous tenterons ci-dessous de présenter quelques outils qui facilitent la cotation de travaux par l'enseignant et la compréhension de ces cotes par les étudiants.

Les échelles de jugement

Les échelles de jugement (en anglais « rubrics ») sont des outils qui aideront l'enseignant à objectiver sa cotation. Nous proposons ici un aperçu de cette technique, à adapter par chacun en fonction des activités développées dans son cours. Les paragraphes qui suivent sont fortement inspiré de Allen 2003 et de Mullinix 2004, librement traduits, complémentés et adaptés.

Une échelle de jugements indique comment seront évaluées les productions des étudiants. Cette échelle est un ensemble de catégories ordonnées (et parfois illustrées) de façon à pouvoir y comparer un travail ou un comportement. Nous en proposons des exemples en fin de document. Cette échelle précise les qualités ou les processus dont la performance devra faire preuve pour se voir attribuer telle ou telle cote. Ces qualités sont placées sur un continuum, correspondant au continuum des cotes que l'étudiant obtiendra si sa performance correspond aux qualités listées.

Ces échelles peuvent théoriquement être utilisées pour classifier tout produit ou comportement, comme par exemple les rapports de recherche, portfolios, travaux artistiques, présentations orales, activités de groupe, etc. Elles peuvent servir l'auto-évaluation (par les étudiants aux-mêmes) ou des évaluations par les pairs, par des enseignants ou par des tuteurs. Elles peuvent être utilisées pour donner aux étudiants des feed-back formatifs, pour certifier leurs performances ou pour juger de la qualité d'un programme de cours.

Quand et pourquoi constituer une échelle de jugements ?

L'enseignant construit une telle échelle, avec ou sans l'aide de ses étudiants, lorsque l'activité dont il souhaite évaluer le résultat est complexe (constitue en fait un ensemble de performances

plutôt qu'une simple performance isolable) et difficilement objectivable sans création de catégories. Une échelle est nécessaire lorsque juger la performance des étudiants serait non fiable sans outil. Par exemple, il est difficile, voire impossible, de juger objectivement et de façon fiable les rapports écrits d'étudiants qui avaient à résoudre un problème complexe.

Voici quelques-uns des **avantages** de ce type de technique :

- permet de juger efficacement les produits ou comportements complexes
- aide les enseignants à préciser leurs propres attentes
- permet à des juges bien entraînés d'appliquer les mêmes critères et standards (fiabilité), donc utile en cas de multiples correcteurs
- permet de mettre à jour les forces et faiblesses d'un groupe d'étudiants, pour mieux y remédier
- se réfère à des critères et non à une norme. Les juges se demandent par exemple « L'étudiant a-t-il rencontré le critère qui permet de lui mettre une note de 5 ? » plutôt que « l'étudiant Dupont a-t-il un meilleur travail que l'étudiant Meulemans ? ». Cette approche est donc plus compatible avec des environnements d'apprentissage collaboratif et coopératif que les modèles basés sur la comparaison et la compétition entre étudiants. Cela permet aussi l'évaluation du cours ou du programme, puisque l'on connaît la proportion d'étudiants qui ont rencontré les critères fixés comme minimaux pour la réussite de chaque activité.
- Peut être appliqué par les étudiants eux-mêmes, par des enseignants, par des tuteurs, individuellement ou collectivement, de façon formative ou certificative, bref, offre une large flexibilité d'application.

Quelques **exemples** de telles échelles de jugements, conçues à l'université de Monmouth (New Jersey, US)

- pour l'auto-évaluation de présentations orales de travaux de groupes :
<http://its.monmouth.edu/facultyresourcecenter/Rubrics/Rubric-Self-Eval-Presentation-LessonCritique.htm>
- pour l'évaluation par les pairs :
<http://its.monmouth.edu/facultyresourcecenter/Rubrics/Rubric-PeerEval-Presentation-LessonCritique.htm>
- pour l'évaluation de portfolios :
<http://its.monmouth.edu/facultyresourcecenter/Rubrics/PracticumAssessmentRubric.htm>
- pour l'évaluation de compte-rendus de recherches (échelle beaucoup plus simple) :
<http://its.monmouth.edu/facultyresourcecenter/Rubrics/ReactionPaperAssSht.htm>

Pour aller plus loin, quelques liens intéressants :

- L'un des sites de l'université d'état de Californie (<http://www.calstate.edu/acadaff/sloa/links/rubrics.shtml>) propose une série de liens vers d'autres sites proposant des listes de critères pour l'évaluation des performances des étudiants.
- Mary Allen propose un « pas à pas » pour le développement d'échelles de jugements. Très précieux, elle donne aussi 8 idées pour l'utilisation pratique de telles échelles dans l'enseignement supérieur, que ce soit pour évaluer son propre cours ou pour évaluer les performances de ses étudiants (http://www.calstate.edu/acadaff/sloa/links/using_rubrics.shtml)
- Le site de Bollie Mullinix (Monmouth) explique l'intérêt de ces échelles de jugements (<http://its.monmouth.edu/facultyresourcecenter/rubrics.htm>), fournit un exemple d'encodage et de traitement des données issues de ces échelles puis propose une « échelle des échelles » pour aider les enseignants à évaluer la qualité de leur propre grille (<http://its.monmouth.edu/facultyresourcecenter/Rubrics/A%20Rubric%20for%20Rubrics.htm>)
- Sur ce même site, des exemples d'échelles de jugements pour différents types de compétences (<http://its.monmouth.edu/facultyresourcecenter/rubrics.htm#SampleRubrics>)

Les copies-types

En complément aux échelles de jugement, fournir aux étudiants des travaux corrigés (« copies-type » pour les Canadiens, « exemplars » ou « diagnostic statements » en anglais) leur permet de se rendre compte de ce qui est exactement attendu d'eux, ce qui sera considéré comme une erreur et la façon dont leurs erreurs seront sanctionnées.

Le principe est que les étudiants peuvent consulter des travaux existants, comparables à la performance qui leur est demandée. Ces travaux sont classés en fonction de la cote qu'ils ont obtenue en regard de différents critères de correction. Les étudiants voient donc la réponse de l'étudiant (anonymisée, de façon à pouvoir aussi bien montrer des travaux insuffisants que des travaux excellents) et les commentaires de l'enseignant, justifiant ses cotes dans différentes rubriques.

Les étudiants sont souvent rassurés par l'aspect très pratique de ces copies-types. Ils peuvent situer leur propre travail lorsqu'ils se préparent à le soumettre pour évaluation et se demander quel commentaire d'enseignant il recevrait s'il soumettait son travail sans le revoir.

L'idéal est de fournir un commentaire par critère et par cote, de façon à ce que les étudiants comprennent bien à la fois les critères et l'échelle de jugement attachée à chacun d'entre eux. Pour ce faire, l'enseignant doit avoir conçu une échelle de jugement (cfr ci-dessus).

Exemple :

A l'université de Californie, plusieurs enseignants fournissent aux étudiants des « diagnostic statements » des années précédentes, petits paragraphes rédigés par les enseignants en guise de feed-back sur la performance de leurs anciens étudiants.

http://www.essayeval.org:82/uc_diagnostic/ (choisir un sujet, puis cliquer sur « show diagnostic statements »)

Références :

- Allen, M., (2003). Using Scoring Rubrics, référencé le 10 juillet 2005 depuis le site http://www.calstate.edu/acadaff/sloa/links/using_rubrics.shtml
- Mullinix, B., B. (2004). Rubrics, référencé le 10 juillet 2005 depuis le site <http://its.monmouth.edu/facultyresourcecenter/rubrics.htm>