

ÉVALUATION PAR LES PAIRS : UN TRAVAIL D'ÉTUDIANT ?

Faire participer les étudiants à leur notation

J.-P. Jacquot

LORIA — Université Henri Poincaré – Nancy 1

Résumé

Cette contribution présente deux expériences d'intégration de l'évaluation par les pairs dans l'évaluation finale d'unités d'enseignement. Évaluer est un savoir-faire important. Les expériences relatées ici montrent comment il est possible d'aborder ce thème d'enseignement comme un des objectifs d'une activité de type « projet ».

Mots-clés

Evaluation, pédagogie active

I INTRODUCTION

L'évaluation des étudiants est une de nos activités les plus sensibles, pour 5 raisons :

- bonne : nous sommes garants des connaissances et des compétences de ceux à qui nous décernons un diplôme grâce à la rigueur de la procédure d'évaluation,
- idéale : l'évaluation aide l'étudiant à situer ses connaissances et ses compétences et à se focaliser sur celles qu'il doit acquérir
- douteuse : l'évaluation est ressentie comme opaque, subjective et donc injuste,
- mauvaise : l'évaluation est trop souvent assimilée à une simple notation qui réduit la mesure d'une réalité complexe à un simple exercice mono-dimensionnel,
- terrible : les étudiants définissent des stratégies qui visent à optimiser le rapport effort sur note plutôt que la qualité de l'apprentissage.

Ces raisons ont un point en commun : l'évaluation est perçue comme un pouvoir extérieur, hors de portée des étudiants. Ainsi, l'évaluation devient un sujet qui n'est pas considéré comme « enseignable » et donc ne l'est pas.

Nous pensons que la formation à l'évaluation a sa place à l'université. C'est évident en formation des maîtres : noter est une activité routinière du métier. De même pour les ingénieurs qui devront organiser et construire des équipes, donc évaluer d'une

certaine façon. Dans les deux cas, une bonne compréhension de ce qu'est une évaluation et de son fonctionnement est essentielle.

Nous présentons deux expériences, une en licence, l'autre en master, dans lesquelles les étudiants participent à la notation de leurs camarades. Nous décrivons les unités d'enseignements (UE) et la procédure d'évaluation, puis nous discutons les observations que nous avons faites au cours des années.

II SPIL

L'UE intitulée « Simulateur de Projet en Ingénierie du Logiciel » est le cours de production industrielle de logiciels en deuxième année de master informatique. Il est réalisé sous la forme d'un projet collectif (Jacquot, 90). Tous les étudiants de la promotion, de 12 à 25 personnes, doivent construire une équipe-projet efficace pour réaliser un logiciel complexe. Les étudiants ont la responsabilité totale du projet ; ce point est crucial. De ce fait, ils découvrent ce que des notions « simples » comme la gestion de projet, la dynamique de groupe, la planification, la qualité, la documentation, etc. recouvrent réellement lorsqu'elles doivent être mises en œuvre.

II.1 Objectif de l'évaluation

La première motivation pour introduire l'évaluation par les pairs a été notre insatisfaction avec la procédure traditionnelle de notation. L'objectif de SPIL est plus de construire un projet, c'est-à-dire des structures abstraites, des comportements et de les mettre en œuvre, que de produire un logiciel. Ainsi il serait contraire à notre objectif de « noter » chaque étudiant sur la base des artefacts qu'il a produit. Par ailleurs, nous n'avons pas suffisamment d'information sur l'intérieur du projet puisque nous en donnons la responsabilité totale aux étudiants. Donc, nous étions contraints de donner à chaque participant la même note, attribuée de fait collectivement au projet. La clause d'individualisation utilisable en cas de comportement fortement déviant par rapport à la moyenne n'a jamais été utilisée.

La seconde motivation est une réflexion sur le métier auquel nous préparons : chef de projet. Cette fonction implique deux missions : construire des équipes performantes et affecter des tâches au mieux des compétences des membres de l'équipe. Dans les deux cas, une évaluation juste des personnes est indispensable. Dès lors, introduire formellement un travail sur l'évaluation dans l'UE est légitime.

II.2 Procédure

Le principe de l'évaluation est décrit dans le document qui définit le projet. La nature et l'usage de cette auto-évaluation y sont précisés. Chaque membre du groupe sait qu'il devra observer et évaluer ses pairs.

Rapidement, nous demandons au groupe de définir, de discuter et d'obtenir un consensus sur une feuille d'évaluation et sur une procédure. Nous insistons sur la nécessité que le groupe définisse précisément les critères observés et évalués afin qu'ils soient compris de la même façon par tous les participants. Les feuilles

d'évaluation nous sont retournées remplies juste après la fin du projet. Le groupe choisit le niveau d'anonymat qu'il souhaite.

Les feuilles permettent de calculer le profil individuel de chaque étudiant, c'est-à-dire la perception qu'en a le groupe, et la variation individuelle de la note. Celle-ci est déterminée comme une variation du comportement par rapport à la moyenne du comportement du groupe. Elle est limitée à ± 1 et ajoutée à la note collective pour obtenir la note individuelle sur 20 que notre système impose d'utiliser.

III INITIATION À LA FORMATION DES MAÎTRES

L'initiation à la formation des maîtres est une UE optionnelle offerte en deuxième année de licence (Jacquot, 07). Elle vise à aider les étudiants à s'orienter vers le professorat des écoles. L'unité comporte trois parts : des cours sur la pédagogie et la didactique, un projet en sciences avec des classes d'écoles élémentaires et un module sur l'exposé scientifique. La notation par les pairs concerne cette dernière partie.

III.1 Objectif de l'évaluation

Un enseignement sur la formation des maîtres doit traiter de ce sujet si présent dans le métier ! Notre vœux est de démythifier ce sujet, en particulier les a priori d'arbitraire et de subjectivité. Plutôt qu'un cours abstrait et ennuyeux, nous faisons réfléchir les étudiants en les mettant en situation réelle. Cette pratique est conforme à l'esprit général de cet enseignement où nous privilégions l'action.

III.2 Procédure

La grille d'évaluation a été extraite de (Brown, 98). Nous l'avons simplifiée pour la ramener à 5 critères notés sur une échelle de 1 à 6.

Tout d'abord, nous présentons et discutons la problématique de l'évaluation et la grille avec la classe. Puis nous effectuons une séance d'évaluation « à blanc » sur des présentations d'articles techniques que les étudiants ont préparées en petit groupe. Chaque groupe présente son exposé qui est évalué immédiatement. Nous demandons que la grille soit remplie collectivement afin de susciter une discussion au sein de chaque groupe pour aboutir à un consensus argumenté. Les évaluations, dont la nôtre, sont recueillies, affichées au tableau et discutées. La discussion porte sur la justification, par chaque groupe, de leur « note » pour chaque critère.

L'évaluation en vraie grandeur a lieu lors de la présentation des exposés qui forment l'épreuve finale du module. Une grille d'évaluation est remplie par chaque groupe à l'issue de chaque exposé. La note finale est la moyenne de la note que nous attribuons et de celle attribuée par les pairs. Tous les étudiants d'un groupe d'exposé ont la même note. Celle-ci représente environ 25% de la note finale de l'UE.

IV CONSTATS ET DISCUSSION

La notation par les pairs a été introduite 2000 pour SPIL et en 2003 pour la formation des maîtres. L'idée d'introduire cette forme d'évaluation est plus ancienne mais plusieurs questions nous ont fait hésiter quelques années: « Est-ce une évaluation juste ? », « Est-ce acceptable par les étudiants ? », « Ne risque-t-on pas de casser les groupes ? », « Est-ce que les étudiants joueront le jeu ? », « Est-ce un exercice intéressant ? », « La note obtenue est-elle académiquement admissible ? ».

IV.1 Évaluation juste

Les étudiants sont très sensibles à la notion d'évaluation « juste », sans toujours sentir l'ambiguïté du terme : « juste » comme précis (le juste poids) ou « juste » comme impartial (juste décision). Notre objectif est de dépasser cette ambiguïté.

En apprenant qu'ils vont participer à l'évaluation, les étudiants sont le plus souvent sceptiques : on ne saura pas parce qu'on n'a pas appris et ce ne sera pas juste.

L'absence de savoir faire est un argument facile à contrer : en cours d'apprentissage, il est « normal » de ne pas savoir. Le vrai problème est de définir un enjeu suffisamment important pour susciter l'apprentissage mais sans excès pour ne pas induire des stratégies de rejet ou d'évitement.

L'argument sur le juste est plus délicat à traiter. Nous insistons sur le fait qu'il ne s'agit pas d'évaluer une personne mais une compétence. Une grille d'évaluation avec des critères précis est rassurante. Lors de l'évaluation « à blanc », les étudiants sont toujours surpris de constater l'homogénéité des évaluations de chaque exposé. Ils sont donc rassurés pour l'exercice en vraie grandeur.

IV.2 Acceptabilité par les étudiants

Avant d'imposer l'évaluation par les pairs, nous l'avons proposée plusieurs fois. Le même scénario s'est alors répété. Les étudiants refusaient au double motif « rationnel » que le projet étant collectif la note doit l'être également, et qu'un étudiant ne peut pas noter un autre. Le projet faisait apparaître des différences importantes entre les participants. Après l'attribution d'une note unique, nous entendions des remarques sur l'injustice de la situation. Une majorité considérait tout aussi anormal que les meneurs du projet ou les « boulets » n'aient pas été distingués. L'opposition de principe à l'évaluation par les pairs n'est donc pas très solide. Pour la rendre acceptable, nous avons joué sur deux paramètres : l'importance relative de l'évaluation dans la note finale et la transparence de la procédure.

Dans nos deux UE, l'évaluation par les pairs influe suffisamment sur la note pour être visible, mais est suffisamment cadrée pour que l'échec d'un étudiant à l'UE ne puisse pas être attribué à ses seuls camarades. La réticence initiale se transforme en une certaine satisfaction du fait que chacun ait pu donner un avis pris en compte.

La transparence de la procédure repose sur le principe d'obtention d'un consensus. En licence, le consensus est obtenu lors du remplissage de la grille, après un travail

explicite sur la grille et la problématique de l'évaluation. En master, le consensus est utilisé pour construire la grille et définir la procédure.

IV.3 Impact sur le fonctionnement du groupe

La forme pédagogique des deux UE repose sur des projets en groupe. Nous devons être vigilants à ne pas en perturber la dynamique. L'évaluation par les pairs a deux risques : privilégier des activités payantes pour l'évaluation au détriment du projet et faire exploser le groupe en exacerbant les affinités et antagonismes personnels.

Le principe d'une évaluation multi-critères prémunit contre le premier risque. Les grilles permettent de « récompenser » des investissements de nature différente. Les stratégies utilitaristes néfastes se situent à un autre niveau : des étudiants optimisent leur effort en considérant les coefficients de chaque UE dans le calcul de moyenne.

Nous n'avons pas constaté que l'évaluation par les pairs amenait le groupe à éclater ou à se coaliser. Tout au plus permet-elle une dernière expression de la réalité préexistante du fonctionnement du groupe. La seule fois où la remise des résultats de l'évaluation s'est mal passée pour SPIL, la promotion n'avait jamais réussi à former un groupe stable et les conflits interindividuels étaient déjà perceptibles.

IV.4 Evaluation et honnêteté

La qualité d'une évaluation par les pairs dépend de l'honnêteté des participants. Même en tant qu'exercice, l'évaluation perd tout intérêt si la malice s'introduit. Nous souhaitons aussi que l'évaluation par les pairs soit pertinente : qu'elle fournisse un résultat exploitable. L'honnêteté de la procédure doit être admise et reconnue.

La plaisanterie usuelle (« Ils vont tous se donner 20 ») est infondée ! Nous n'avons jamais relevé de cas flagrant de tricherie. Au contraire, nous sommes toujours surpris, année après année, de la proximité de nos propres évaluations et celle des étudiants dans l'UE de formation des maîtres. Les étudiants ont un esprit critique vis-à-vis d'eux-mêmes et de leurs camarades et ils savent l'exercer à bon escient.

Avec SPIL, l'évaluation porte clairement sur des compétences ; même si elle n'a pas pour objectif de juger l'individu, elle pourrait être ressentie comme telle. Hors cette promotion atypique citée précédemment qui a échoué à former un groupe, nous n'avons jamais noté de réactions violentes à la remise des résultats. La majorité des étudiants admet facilement les résultats. Les quelques exceptions correspondent à des étudiants qui découvrent alors que le groupe ne les perçoit pas comme ils le croyaient. D'une certaine façon, l'évaluation leur ouvre les yeux.

La feuille d'évaluation des enseignements, remplie à l'issue de la formation, comporte une rubrique sur la méthode d'évaluation de chaque UE. Si nous y relevons quelques critiques sur le principe même de l'évaluation par les pairs, la tonalité générale est plutôt une « neutralité bienveillante ».

IV.5 Intérêt pour l'évaluation

Les étudiants sont intéressés par l'activité d'évaluation. Deux observations : tous les étudiants participent au remplissage des grilles, il y a très peu de grille « neutre ».

Les étudiants participent malheureusement peu aux activités qui n'ont pas un intérêt immédiat. Les procédures d'évaluation par les pairs que nous avons mises en place ne « rapportent » rien aux participants ; l'absence de participation ne peut pas avoir d'effet négatif sur la note individuelle. Or nous récupérons toutes les grilles.

Une façon d'échapper à la corvée du remplissage d'une grille est de cocher la même colonne pour toutes les lignes. De telles réponses sont de fait très rares. Dans le cadre de la formation des maîtres, les grilles donnent toujours un avis sur l'exposé qu'elles évaluent ; dans le cadre de SPIL, on trouve au moins une ou deux personnes sur lesquelles une évaluation est portée dans les feuilles les plus neutres.

IV.6 Acceptabilité académique

La question de l'acceptabilité académique de l'évaluation par les pairs ne nous a pas été posée dans le cadre de ces deux expériences. Bizarrement, notre système repose sur des notes mais ne s'interroge pas sur la façon dont elles sont élaborées. Notre démarche pourrait, ou devrait, être interpellée ; le silence est significatif.

V CONCLUSION

Dans un monde où les indicateurs, notations, évaluations, indices, de préférence chiffrés, sont devenus omniprésents au nom de la rationalisation des décisions, nous devons intégrer à nos enseignements une réflexion sur l'élaboration de tous ces nombres. Très modestement, nous essayons de montrer comment un indicateur particulier, la note, se construit. La forme que nous avons choisie repose sur deux idées. Un processus se comprend surtout lorsqu'on le pratique. L'implication des étudiants dépend de l'enjeu. On peut faire confiance à la maturité et au professionnalisme des étudiants. Face à un objectif pédagogique clair, ils savent dépasser les préjugés et les comportements « bêtement scolaires ».

Dans ces expériences, la notion d'évaluation formative est poussée à la limite puisque les étudiants sont doublement acteurs : évalués et évaluateurs. Elles ne sont probablement pas transposables en l'état dans d'autres cours. Néanmoins, l'enseignement universitaire ne se limite pas à la transmission de connaissances, mais doit inclure l'apprentissage du raisonnement, du doute, de la critique, de l'observation. Évaluer met en œuvre ces dernières compétences et il est permis de penser que l'évaluation par les pairs peut permettre de renforcer leur acquisition.

Nous souffrons de la dictature des notes. En partie parce qu'elles sont vécues comme l'expression d'un pouvoir absolu et mystérieux. En montrant comment s'élabore une note, en impliquant tous les acteurs, nous espérons contribuer à assainir cette notion.

RÉFÉRENCES

- Brown, G. (1998). Bien faire un cours, un exposé, une conférence. Paris. Les Éditions d'organisation
- Jacquot, J.-P., Guyard, J. & Boidot, L. (1990). Modeling Teamwork in an Academic Environment. SEI'90 Conference, Pittsburgh (PA). LNCS N°423 – Springer.
- Jacquot, J.-P. & Devaux S. (2007). Initiation à l'enseignement des sciences à l'école. Actes du 4e colloque Questions de pédagogie dans l'enseignement supérieur 24-26 janvier 2007. Louvain : Presses Universitaires de Louvain