

Variations sur le thème de l'évaluation de soutenances

Julien Pellé^{1,2}, Philippe Champagne^{1,3}, Dominique Bernier^{1,2}, Samuel Dupont^{1,2}

¹ Université Polytechnique Hauts-de-France

² ENSIAME – INSA Partenaire

³ ISTV

Résumé

Le déploiement de « la démarche compétences » au sein des écoles d'ingénieurs doit permettre d'améliorer la lisibilité des programmes par les professionnels et une bonne employabilité des diplômés, ce qui passe par l'alignement entre les formations et les référentiels métiers. La délivrance du diplôme ne doit donc pas être décorrélée des compétences professionnelles, au contraire, elle doit attester d'Acquis d'Apprentissage (AA) qui alimentent les compétences visées. L'école doit alors mettre en place des outils d'évaluation des AA permettant de juger les « apprenants » sur la base d'éléments observables, quantifiables et évaluables qui reflètent leurs capacités réelles. Dans cette communication, nous relatons les résultats d'une expérience réalisée à l'ENSIAME (futur INSA Hauts-de-France) concernant les modalités d'évaluation des soutenances de stage. Sur la base de soutenances réelles filmées et rediffusées à des enseignants volontaires, nous comparons 3 modalités d'évaluation. Pour chaque soutenance visualisée, nous avons analysé l'évolution de la moyenne des notes et des écarts-type en fonction de la méthode d'évaluation adoptée. L'utilisation d'une grille à critères formalisés tend à diminuer de façon significative la part de subjectivité de l'évaluation et permet de mieux jauger les niveaux atteints. Nous avons également comparé les résultats des évaluations critère par critère et nous mettons parfois en évidence de fortes différences d'interprétation entre les évaluateurs. Nous avons également recueilli le sentiment des participants sur la méthode d'évaluation et sur les résultats de l'expérience. Il s'avère que cette expérience a permis de les sensibiliser à un nouvel outil que certains souhaitent finalement tester dans leurs cours.

Mots-clés : Evaluation ; soutenances ; acquis d'apprentissage ; indicateurs

1. Contexte

1.1. Général

La mise en place du processus de Bologne a eu pour conséquence l'apparition de « la démarche compétences » qui se traduit notamment par une explicitation des Acquis d'Apprentissage (AA) visés par les formations et par chaque « grain » pédagogique : Unités d'Enseignement (UE) et Eléments Constitutifs d'Unités d'Enseignements (ECUE). Cette formalisation doit permettre

d'améliorer l'alignement entre les formations et les référentiels métiers, de permettre aux étudiants de comprendre l'intérêt des enseignements, d'améliorer la lisibilité des programmes par les professionnels. Enfin, elle doit idéalement permettre de juger de manière plus juste les « apprenants ». Pour cela l'appréciation se fait sur la base d'éléments observables, quantifiables et évaluables qui reflètent leurs capacités réelles en lien avec les AA visés. L'adoption d'une démarche compétences suscite donc un questionnement quant à l'adéquation des pratiques d'évaluations. Nous présentons dans cet article une expérience pédagogique portant sur l'évaluation des AA lors des soutenances de stage.

1.2. Evaluation des soutenances de stage

Le stage en entreprise est une expérience qui doit permettre à nos élèves de découvrir le monde professionnel, le métier de l'ingénieur et le travail en équipe. Il permet de développer et de mettre en pratique les savoirs, savoir-faire et savoir-être, ainsi que l'esprit d'analyse, l'autonomie, le sens de l'organisation et l'esprit d'initiative. Il s'agit pour l'élève ingénieur de cogérer un projet avec le tuteur d'accueil : il doit acquérir la réelle dimension de l'ingénieur à travers la recherche de méthodes, la prise de décision, la mise en œuvre d'outils, l'approche de la dimension économique, etc.

Dès lors, le stage en entreprise est une situation authentique où l'exercice est en lien direct avec le métier d'ingénieur et qui se prête bien à l'évaluation des AA au plus près des compétences attendues par les industriels. Dans le cadre de la réflexion menée au sein de l'école sur l'évaluation des AA résultant de la démarche compétences qui se déploie dans les formations d'ingénieurs sous l'impulsion de la CTI, nous avons choisi de tester la méthode des grilles critériées lors de l'évaluation des soutenances de stage.

1.3. Choix de la grille critériée

Une grille critériée est un tableau explicitant les critères utilisés pour évaluer l'apprentissage des élèves et à la fois les niveaux d'atteinte pour chaque critère. La grille critériée est réputée participer à la diminution de la subjectivité et permettre la clarification de critères d'évaluation et des indicateurs de réussite pour les étudiants [1]. Elle est plus adaptée pour l'évaluation des rapports écrits et des soutenances orales. Enfin, nous avons également fait ce choix pour montrer aux collègues un nouvel outil d'évaluation, la grille critériée étant pour le moment assez peu utilisée dans l'école.

2. Méthodologie

Nous avons sélectionné des soutenances de stage de deuxième de cycle ingénieur (stage de 18 semaines minimum entre septembre et janvier), non confidentielles, pour lesquelles les étudiants représentant les trois spécialités de l'ENSIAME acceptaient d'être filmés par un système de captation le moins intrusif possible.

Les vidéos ont ensuite été visionnées pour sélectionner celles qui allaient servir à l'expérience : élimination des vidéos présentant des problèmes de qualité d'image et sélection d'un échantillon représentatif des différents niveaux de qualité de la soutenance proprement dite.

Treize enseignants participant traditionnellement à l'évaluation des soutenances de stages se sont déclarés volontaires pour participer à cette expérience qui s'est déroulée sur une demi-journée selon le schéma suivant : après une présentation de l'expérimentation (contexte, enjeu et méthode), les enseignants ont visionné ensemble chacune des 3 vidéos sélectionnées et ont procédé de manière individuelle et synchronisée à l'évaluation suivant trois méthodologies différentes (chaque vidéo n'a été visionnée qu'une seule fois) :

- E1 : l'enseignant attribue une note en réponse à la question : « Note /20 que vous donneriez sans le support d'une grille d'évaluation »,

- E2 : la deuxième évaluation est basée sur l'utilisation de la grille de notation utilisée à l'ENSIAME (grille à laquelle tous les participants sont rompus). Cette grille comporte 10 critères portant sur la forme et sur le fond de la présentation. L'échelle de notation comporte 7 niveaux 0, 1, 3, 5, 6, 7, 8 et 10 illustrés chacun par un qualificatif (3=faible, 8=très bien, 10=exceptionnel). Chaque critère est explicité et connu au préalable des étudiants mais les attendus de chaque niveau ne sont pas décrits. Une note sur 20 est calculée ensuite,
- E3 : la troisième évaluation est basée sur l'utilisation d'une nouvelle grille (présentée pour la première fois lors de cette expérimentation). Les acquis d'apprentissage visés lors du stage sont divisés en critères pour lesquels quatre niveaux d'atteinte (A = très bien, B = correct, C = le minimum acceptable, D = défaillant) sont explicitement détaillés. Chaque critère étant affecté d'une note, une note globale sur 20 peut être calculée.

L'ensemble des résultats est synthétisé à l'issue de l'expérience pour une première analyse présentée aux participants qui sont ensuite invités à partager leur perception de la formalisation des attendus et leur inclination à mettre en œuvre cette méthode pour l'évaluation de leurs enseignements.

3. Résultats et discussions

La table I montre pour chaque soutenance et chaque type d'évaluation, la moyenne et l'écart-type des 13 notes obtenues. Les notes obtenues avec la méthode E3 sont calculées en utilisant 2 barèmes. Les tables II et III mettent en évidence les différences parfois observées entre les 13 évaluateurs.

TABLE I. MOYENNE (/20) ET ECART-TYPES (ITALIQUE) DES 13 EVALUATIONS

Soutenance	Méthode E2 (référence)		Méthode E1		Méthode E3 : (A=8 ; B=7 ; C=6 ; D=3)		Méthode E3 : (A=9 ; B=7 ; C=5 ; D=0)	
	1	11.62	<i>2.14</i>	11.42	<i>2.05</i>	11.65	<i>1.48</i>	10.01
2	13.90	<i>1.34</i>	13.67	<i>1.30</i>	13.68	<i>1.36</i>	13.89	<i>2.22</i>
3	15.61	<i>0.91</i>	15.88	<i>0.91</i>	15.33	<i>0.43</i>	16.5	<i>0.86</i>

TABLE II. ANALYSE DETAILLEE DES 13 NOTES (/10) SUR UN CRITERE DE LA GRILLE DE REFERENCE (E2) POUR LA SOUTENANCE 2

Critère	Note mini	Note maxi	Moyenne	Ecart-type
« ... facteur humain... »	1	10	6.4	2.5

TABLE III. ANALYSE DETAILLEE DES 13 EVALUATIONS SUR UN CRITERE DE LA GRILLE CRITERIEE (E3) POUR LA SOUTENANCE 1

Critère	Nbre de A	Nbre de B	Nbre de C	Nbre de D
« ... graphiques/tableaux ... »	3	4	3	3

3.1. Analyse des résultats

Dans ce paragraphe, la grille avec formalisation des indicateurs et des critères compile une note finale en utilisant la conversion suivante : A=8/10 ; B=7/10 ; C=6/10 ; D=3/10, chaque critère ayant ensuite le même poids. Ce choix de pondération a été fait pour être cohérent avec les notes les plus utilisées dans la grille de références afin de permettre une comparaison pertinente des outils. D'après les résultats de la table 1, les 3 évaluations donnent alors des notes très proches (écarts compris entre 0.3 points et 0.6 points en fonction de la soutenance). En revanche, l'analyse de l'évolution des écart-types nous permet de mettre en évidence que la formalisation

des critères et des indicateurs semble suggérer une baisse de la subjectivité sur le résultat de l'évaluation pour 2 soutenances sur les 3 de l'expérience. L'outil avec une formalisation complète des critères et des indicateurs est donc tout à fait pertinent. On constate cependant qu'il n'y a presque pas d'écart entre la formalisation partielle (E2) et aucune formalisation (E1) aussi bien sur les moyennes que sur les écart-types. Du fait que les 13 évaluateurs ont l'expérience de la grille de référence, il semblerait qu'ils appliquent intuitivement ses critères, même en son absence.

Dans une volonté d'être plus discriminants vis-à-vis de l'atteinte d'un seuil minimal de compétences liées aux soutenances, nous avons également compilé des notes en utilisant la conversion suivante : A=9/10 ; B=7/10 ; C=5/10 ; D=0/10. L'objectif est atteint dans la mesure où la soutenance 1, qui avait la note la plus faible, voit sa note diminuer de 13.9% tandis que la soutenance 3, qui avait la note la plus élevée, voit sa note augmenter de 5.7%. La note de la soutenance 2, quant à elle, n'évolue quasiment pas.

L'analyse critères par critères des réponses permet de mettre en évidence ceux qui ont été interprétés différemment. La table 2 montre un exemple pour l'évaluation de référence (E2) tandis que la table 3 montre un exemple pour l'évaluation avec critères et indicateurs formalisés (E3). Les auteurs n'auraient pas pensé a priori que de tels écarts pouvaient être observés du fait de la formalisation des attentes. On identifie ainsi des critères qui sont soit mal formalisés, soit non pertinents dans le cadre de l'évaluation des soutenances, mais en tout cas sur lesquels une réflexion et un travail supplémentaires sont nécessaires.

3.2. Retour des participants

Il ressort de l'échange qui a eu lieu avec les évaluateurs à l'issue de l'expérience que :

- Ce type de grille peut allonger la durée des soutenances du fait d'un temps de lecture plus important (NB : les évaluateurs découvraient la nouvelle grille critériée)
- Le format des grilles est un facteur pouvant avoir une influence sur l'évaluation finale.
- Les collègues ont trouvé l'expérience intéressante et certains sont prêts à tester une grille critériée dans leurs propres enseignements afin d'évaluer son intérêt pédagogique dans d'autres contextes (apprentissage par projet).

4. Conclusions et perspectives

Dans cette communication, nous relatons une expérience pédagogique concernant l'évaluation de soutenances de stage en école d'ingénieurs suivant 3 méthodes, dont une utilisant une grille critériée avec une formalisation des critères d'évaluations et des niveaux d'atteinte. Les résultats semblent montrer la pertinence de cette grille pour diminuer la subjectivité d'une évaluation. Les collègues restent réticents face à l'allongement de la durée de l'évaluation même si certains pensent tester le dispositif au sein de leurs propres enseignements. L'expérience permet aussi de mettre en évidence les critères ou niveaux d'atteinte mal formulés sur lesquels nous allons retravailler. Lorsque la grille sera finalisée, elle sera aussi un outil d'auto-évaluation efficace pour nos étudiants qui connaîtront alors clairement les attentes, que nous comptons différencier en fonction du niveau d'étude.

Références

- [1] ERTHIAUME, D., DAVID J., DAVID T. Réduire la subjectivité lors de l'évaluation des apprentissages à l'aide d'une grille critériée : repères théoriques et applications à un enseignement interdisciplinaire. *Revue internationale de pédagogie de l'enseignement supérieur*, 2011. Vol. 27 ; n° 27-2.