



INSA

6^{ème}

Colloque Pédagogie & Formation

16 et 17 mai 2019 – INSA Centre Val de Loire

Variation sur le thème de l'évaluation de soutenances

Julien Pellé, Philippe Champagne,
Dominique Bernier, Samuel Dupont

julien.pelle@uphf.fr



Commission des Titres d'Ingénieur



Le processus de Bologne – La démarche compétences:



Objectifs :

- **Améliorer :**
 - l'alignement entre la formation et les référentiels métiers
 - la compréhension de l'intérêt des enseignements par les étudiants
 - la lisibilité des programmes par les professionnels
- **Evaluer** les capacités des étudiants au cours et à la fin de la formation

Construction :

Une démarche formalisée autour des Acquis d'Apprentissages :

- Terminaux (AAT) : déclinaison à l'échelle de la formation
- Intermédiaires (AAI) : déclinaison à l'échelle de chaque UE
- Spécifiques (AAS) : déclinaison à l'échelle de chaque ECUE

Une démarche structurante :

- Alignement des objectifs inter ECUE; inter-UE
- Aide au « chaînage » chronologique

Une démarche collégiale

**Evaluer des AA? oui !
mais comment ?**

Expérience



Evaluation des capacités :



Dans l'idéal, une évaluation doit être basée sur :

- une **situation d'évaluation** proche du métier, de préférence transversale
- l'**observation** d'un ensemble d'éléments :
 - **observables, quantifiables, mesurables,**
 - **en lien avec les Acquis** d'Apprentissage visés,
 - **connus et compris** de tous (évalués et évaluateurs)
- l'**interprétation de l'observation** permettant de porter un jugement fiable, reproductible, juste, indépendant de l'évaluateur... sur l'atteinte ou pas des attendus
 - par exemple, des **niveaux d'atteinte** explicites

Situation: la soutenance de stage

- communiquer avec efficacité à l'oral dans un temps limité l'expérience acquise sur:
 - la découverte du monde professionnel, le travail en équipe,
 - le développement et la mise en pratique des savoirs, savoir-faire et savoir-être,
 - l'esprit d'analyse, l'autonomie, le sens de l'organisation et l'esprit d'initiative,
 - la gestion d'un projet avec le tuteur d'accueil et la prise des décisions

Expérience collégiale d'évaluation des soutenances de stage :

Contraintes/objectifs :

- Comparer plusieurs méthodes d'évaluation sur plusieurs soutenances
- Faire participer un maximum de collègues volontaires
- Ne pas perturber les étudiants pendant leurs « vraies » soutenances
- Evaluer la pertinence de l'utilisation d'une grille critériée¹

Méthodologie :

- Enregistrements d'environ 10 soutenances non confidentielles avec accord (Janvier 2019)
- Visionnage et choix
 - Élimination des vidéos de mauvaise qualité,
 - Sélection d'un échantillon représentatif de différents niveaux (3 vidéos)
- Appel à participation : 13 collègues volontaires et habitués à l'évaluation des soutenances
- Expérience sur ½ journée (Mars 2019) :
 - Présentation du contexte, des enjeux et de la méthode pendant 15min
 - Visionnage unique et collectif de chaque vidéo
 - Evaluation individuelle et synchronisée de chaque soutenance selon 3 dispositifs E1, E2 et E3.
 - Analyse des premiers résultats en temps réel
 - Retour et discussion avec les participants (30 minutes)

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.

3 dispositifs d'évaluation:

- E1 : à l'expérience

Note /20 que vous donneriez sans le support d'une grille d'évaluation	10.00
---	-------

- E2 : grille « habituelle » de référence avec 10 critères explicités. Echelle de notation comporte 8 niveaux (entre 0 et 10) pour chaque critère, les attendus de chaque niveau ne sont pas explicités

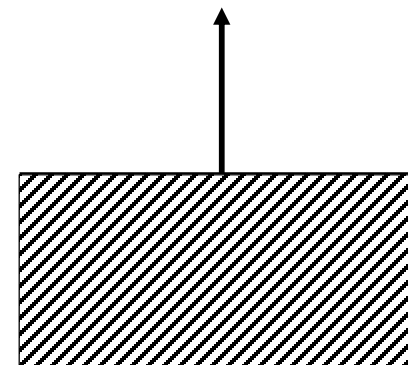
	INACCEPTABLE	INSUFFISANT	FAIBLE		MOYEN	ASSEZ BIEN	BIEN	TRES BIEN		EXCEPTIONNEL
	0	1	3		5	6	7	8		10
Critère 1										
Critère 2										

3 dispositifs d'évaluation:

- E3 : « nouvelle » grille avec 15 critères et 4 niveaux d'atteintes explicités pour chaque critère

Le temps imparti est utilisé judicieusement	
La durée allouée est respectée et le découpage temporel est pertinent ;	A
La durée allouée est respectée et le découpage temporel est pertinent mais le débit souffre d'accélération ou de ralentissement ;	B
La durée allouée n'est pas respectée (> 10% et < 20 %) ou le découpage temporel n'est pas pertinent ;	C
La durée allouée n'est pas respectée (> 10% et < 20%) et le découpage temporel n'est pas pertinent ou la durée allouée n'est pas respectée (> 20%) ou le découpage temporel n'est pas pertinent au point de nuire à la compréhension du travail effectué.	D

Seuil d'acceptabilité



Une note sur 20 est ensuite compilée en attribuant les valeurs suivantes (sur 10) :

- A=8; B=7; C=6; D=3 (pour permettre la comparaison avec la méthode E2)
- A=9; B=7; C=5; D=0 (pour être plus discriminants)

Subjectivité de l'évaluation (sur la base des 13 évaluations) :

Soutenance	Méthode E2 (référence)		Méthode E1		Méthode E3 : (A=8 ; B=7 ; C=6 ; D=3)	
	Moyenne	Ecart-type	Moyenne	Ecart-type	Moyenne	Ecart-type
1	11.62	2.14	11.42	2.05	11.65	1.48
2	13.90	1.34	13.67	1.30	13.68	1.36
3	15.61	0.91	15.88	0.91	15.33	0.43

- 3 soutenances de niveaux différents
- Plus de disparités quand la qualité de la soutenance diminue
 - volonté de certains évaluateurs de « sauver » l'étudiant ?
 - plus difficile d'évaluer précisément quand le niveau est moyen : 5/10, 6/10, 7/10 ?
- L'évaluation sans grille (E1) donne des résultats très similaires à l'évaluation habituelle (E2)
 - Application intuitive de la grille même en son absence ?
- L'évaluation avec grille critériée (E3) donne des résultats très similaires en terme de moyenne (fait volontaire)
- Chute de l'écart-type pour 2 soutenances sur 3.
- les 3 méthodes donnent des évaluations corrélées

Choix du barême :

Soutenance	Méthode E3 : (A=8 ; B=7 ; C=6 ; D=3)		Méthode E3 : (A=9 ; B=7 ; C=5 ; D=0)	
	Moyenne	Ecart-type	Moyenne	Ecart-type
1	11.65	1.48	10.01	2.54
2	13.68	1.36	13.89	2.22
3	15.33	0.43	16.5	0.86

- La moins bonne note chute de 13.9 %
- La meilleure augmente de 5.7%
- La note intermédiaire n'évolue quasiment pas
- En contre-partie, l'écart-type augmente avec ce nouveau barême

Identification de critères qui posent problème:

L'analyse détaillée des 13 évaluations pour chaque soutenance permet de mettre en évidence des critères où les disparités sont les plus fortes.

Exemple 1 : Evaluation E2 (Note / 10) pour la soutenance 2

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

Exemple 2 : Evaluation E3 pour la soutenance 1

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

- Formulations maladroités du critère et/ou du niveau attendu
- Critère non-pertinent pour l'évaluation d'une soutenance (redondance?)
- Critère difficile à évaluer dans le cadre de cette soutenance (ajout d'une case à cocher « non-évaluable »? Comment la prendre en compte dans la note?)

Retour « à chaud » des collègues ayant participé:

- Temps de lecture important donc allongement de la durée des soutenances
- Expérience collégiale intéressante qui a permis de mettre en évidence les écarts d'appréciation entre collègues en l'absence de formulation claire
- Certains sont prêts à tester la grille critériée dans leurs propres enseignements afin d'évaluer son intérêt pédagogique dans d'autres contextes (APP, ...)
- Avec ce type de grilles, les étudiants peuvent s'auto-évaluer plus facilement et connaissent plus précisément nos attentes.
- La présentation (forme) de la grille peut avoir une influence sur l'évaluation finale

	INACCEPTABLE	INSUFFISANT	FAIBLE		MOYEN	ASSEZ BIEN	BIEN	TRES BIEN		EXCEPTIONNEL
	0	1	3		5	6	7	8		10
Critère 1										
Critère 2										

Le temps imparti est utilisé judicieusement	
La durée allouée est respectée et le découpage temporel est pertinent ;	A
La durée allouée est respectée et le découpage temporel est pertinent mais le débit souffre d'accélération ou de ralentissement ;	B
La durée allouée n'est pas respectée (> 10% et < 20 %) ou le découpage temporel n'est pas pertinent ;	C
La durée allouée n'est pas respectée (> 10% et < 20%) et le découpage temporel n'est pas pertinent ou la durée allouée n'est pas respectée (> 20%) ou le découpage temporel n'est pas pertinent au point de nuire à la compréhension du travail effectué.	D

Conclusions

Expérience pédagogique collégiale en école d'ingénieurs

- Evaluation des Acquis d'Apprentissage lors des soutenances de stage
- 3 dispositifs comparés: expérience, grille actuelle, grille critériée formalisée

La grille critériée :

- permet une baisse de la subjectivité dans la note obtenue
- peut apparaitre comme difficile à utiliser de par sa longueur
- doit faire l'objet d'itérations pour s'assurer de la bonne compréhension et pertinence des critères et des niveaux
- peut devenir un outil efficace d'auto-évaluation

Cette expérience a permis de :

- de sensibiliser les collègues aux disparités existantes entre évaluateurs
- de sensibiliser les collègues aux « bonnes » pratiques
- de donner l'envie aux collègues de tester différentes évaluations dans leurs enseignements

Questions

Julien Pellé, Philippe Champagne,
Dominique Bernier, Samuel Dupont

julien.pelle@uphf.fr