



HAL
open science

Les pratiques alternatives de notation comme levier vers une évaluation plus équitable et qui accompagne mieux l'apprentissage

Caroline Cormier, Bruno Voisard, François Arseneault-Hubert, Véronique Turcotte

► To cite this version:

Caroline Cormier, Bruno Voisard, François Arseneault-Hubert, Véronique Turcotte. Les pratiques alternatives de notation comme levier vers une évaluation plus équitable et qui accompagne mieux l'apprentissage. Diversité, Réussite[s] dans l'Enseignement Supérieur (2024), Nantes Université [Nantes Univ], Apr 2024, Nantes, France. hal-04569339

HAL Id: hal-04569339

<https://hal.science/hal-04569339>

Submitted on 6 May 2024

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.

Les pratiques alternatives de notation comme levier vers une évaluation plus équitable et qui accompagne mieux l'apprentissage

Auteurs :

Voisard, Bruno ; Cégep André-Laurendeau ; bruno.voisard@clairendeau.qc.ca

Cormier, Caroline ; Cégep André-Laurendeau ; caroline.cormier@clairendeau.qc.ca

Arseneault-Hubert, François ; Cégep André-Laurendeau ; francois.arseneault-hubert@clairendeau.qc.ca

Turcotte, Véronique ; Cégep André-Laurendeau ; veronique.turcotte@clairendeau.qc.ca

Mots-clés : Évaluation des apprentissages ; Pratiques alternatives de notation ; Équité

Présentation de l'atelier

1. Objectifs de l'atelier et public visé

Cet atelier s'inscrit dans l'axe des **Design et pratiques pédagogiques inclusives** du colloque *Diversité et réussite[s]*. Dans le cadre de cet atelier, nous traiterons de pratiques alternatives de notation (PAN), qui sont des pratiques d'évaluation des apprentissages plus équitables que les pratiques traditionnelles. En effet, les pratiques de notation sont souvent un angle mort des réflexions relatives au soutien à la réussite de façon inclusive. C'est pourtant par la notation que des iniquités socioéconomiques peuvent être exacerbées par l'école : lorsque la notation est utilisée pour le classement et la sélection, et que les pratiques de notation sont biaisées en faveur de certains groupes, les pratiques de notation ont directement pour effet de contribuer à perpétuer ces iniquités. En proposant un atelier sur les pratiques alternatives de notation, notre intention est de renseigner les participantes et les participants sur ces iniquités et de leur présenter des alternatives.

Le public visé par cet atelier est le personnel pédagogique des établissements d'enseignement supérieur, principalement les personnes enseignantes et les conseillers et conseillères pédagogiques. Les étudiants et étudiantes en formation initiale des maîtres pourront tout autant bénéficier de l'atelier, qui leur permettra de réfléchir aux questions de l'équité, de la cohérence et de la transparence en évaluation des apprentissages.

L'objectif principal de l'atelier est de mieux connaître les fondements des PAN et leurs caractéristiques, et d'imaginer des modalités d'implantation de PAN qui s'appuient sur ces fondements, à travers des études de cas.

2. Ancrage théorique

Les visées de l'évaluation sont de soutenir les apprentissages et de témoigner des acquis (Conseil supérieur de l'éducation, 2018). Par exemple, l'évaluation formative a clairement comme objectif de soutenir l'apprentissage, alors que l'évaluation qui mènera à une note rapportée au bulletin a comme objectif de témoigner des acquis. L'utilisation des notes comme outil de sélection ou pour motiver les étudiants et étudiantes peut toutefois faire en sorte que l'évaluation ne puisse plus remplir adéquatement les visées de soutenir les

apprentissages et de témoigner des acquis. Ainsi, si l'évaluation et la notation sont deux processus distincts, les pratiques de notation ont un effet important sur l'évaluation.

La notation est utilisée comme instrument de motivation, par exemple, lorsque des points sont attribués à des devoirs ou des évaluations qui ont une visée formative, ou pour encourager ou décourager certains comportements tels que la participation ou les retards dans la remise de travaux. Or, les écrits de recherche montrent que lorsque la rétroaction est accompagnée d'une note, ceci nuit à la motivation intrinsèque (Koenka et al., 2021), en particulier pour les personnes en difficulté d'apprentissage. **L'utilisation de la note comme outil de contrôle du comportement des étudiants et des étudiantes nuit donc à la visée de l'évaluation de soutenir l'apprentissage.**

De plus, **lorsque la composition de la note dépend de cette manière d'autres facteurs que le degré d'atteinte des objectifs d'apprentissage, l'évaluation ne répond plus avec justesse à la visée de témoigner des acquis**, sans compter les risques d'introduire par de telles pratiques des **biais au détriment ou à l'avantage de certains groupes socioéconomiques**. Dans cette catégorie de pratiques, on pourrait aussi nommer la notation de production d'équipe, qui ne permet pas de conclure sur l'atteinte individuelle des objectifs d'apprentissage (Feldman, 2024 ; O'Connor, 2022).

Une dernière préoccupation majeure relative à certaines pratiques traditionnelles de notation est celle de l'utilisation de la moyenne pour composer la note finale. En calculant la note de cette façon, elle ne reflète pas nécessairement le degré d'atteinte des objectifs d'apprentissage et elle dépend beaucoup du profil de l'étudiante ou de l'étudiant. Par exemple, une personne mal préparée à un cours par son parcours antérieur pourrait mal performer au début du trimestre, mais tout de même le terminer en atteignant les cibles d'apprentissages. Une personne avec un tel profil sera pénalisée par les pratiques de notation qui sanctionne les erreurs en cours d'apprentissage : sa note dépendra beaucoup de son parcours, et pas seulement de son degré de maîtrise des objectifs terminaux.

Plusieurs systèmes de pratiques alternatives de notation (PAN) (Townesley & Schmid, 2020) visent à corriger ces biais dans les pratiques d'évaluation des apprentissages, comme la notation basée sur la maîtrise ou sur les standards et la notation par spécifications. La dénnotation est aussi une pratique qui tend notamment à remplacer des évaluations préparatoires qui entrent dans la composition de la note finale par différentes formes de rétroaction formative. Les caractéristiques propres à ces familles de pratiques de notation sont résumées dans le tableau ci-dessous.

Tableau 1 : Caractéristiques propres à trois familles de PAN

Notation basée sur la maîtrise ou les standards	Notation par spécifications ou par contrat	Dénnotation
Approches où c'est le degré d'atteinte des objectifs qui est noté plutôt que la performance à une évaluation et où chaque évaluation est clairement associée à un ou plusieurs objectifs	Approches où des ensembles de productions sont prédéfinis, chaque ensemble de productions de qualité satisfaisante correspondant à une note finale donnée	Approches où la rétroaction qualitative, l'autoévaluation et l'évaluation par les pairs sont favorisées et où le moins possible de travaux sont notés

Ces différentes familles de systèmes sont issues de préoccupations communes au sujet des effets délétères des pratiques de notation traditionnelles, et les personnes qui optent pour des PAN construisent souvent leur système en agrégeant des caractéristiques de différentes familles. Toutes les familles de PAN possèdent d'ailleurs plusieurs caractéristiques en commun (Townasley & Schmid, 2020) :

- des occasions de reprise ou de révision des évaluations ;
- la place laissée à l'erreur sans pénalité en cours d'apprentissage, en ne tenant compte que des traces les plus représentatives des apprentissages réalisés dans la composition de la note finale ;
- une notation de chaque production ou de chaque objectif à l'aide d'un nombre limité de niveaux (typiquement, de deux à cinq), donc qui ne fait pas appel à une notation sur 100 (ou sur 20), améliorant ainsi la fidélité intraévaluateur et interévaluateurs et diminuant les biais ;
- une amélioration de la transparence de l'évaluation à l'aide d'objectifs d'apprentissage clairement définis, présentés dans une forme qui est compréhensible pour les étudiantes et les étudiants ;
- une amélioration de la qualité de la rétroaction qualitative, qui vise à donner de l'information sur le trajet qu'il reste à parcourir pour atteindre les objectifs d'apprentissages (par opposition à donner de la rétroaction globale sur une production, qui porte un jugement sur la personne ou qui vise à une fonction essentiellement affective ou motivationnelle).

Il n'est donc pas étonnant que ces pratiques de notation alternatives soient vues comme une façon de corriger les iniquités inhérentes aux pratiques de notation habituelles, ou, du moins, de les atténuer.

3. Modalités et déroulement de l'atelier

L'atelier proposé est conçu pour une durée d'1 h 30. La **structure générale de l'atelier** est la suivante :

1. Présentation des animateur·rices et de l'objectif de l'atelier (5 min)
2. Première itération : étude de la thématique 1 (25 min)
3. Deuxième itération : étude de la thématique 2 (25 min)
4. Troisième itération : étude de la thématique 3 (25 min)
5. Conclusion : présentation du sommaire des principes étudiés (10 min)

L'atelier se déroule ainsi sous forme d'études de cas, groupées en trois thématiques. Voici la **séquence de chacune des itérations** :

a. Brève introduction théorique par les animateur·rices

Les fondements ou les recommandations tirées des écrits de recherche autour de la thématique sont présentés. Une diapo résume ces fondements ou recommandations et demeure affichée à l'écran pendant le temps de travail en équipe (étape suivante).

b. Distribution des cas et travail en petites équipes.

Les participants et participantes sont divisés en équipes de 3 à 5 personnes et reçoivent un cas. Deux cas différents sont distribués à chaque itération, donc la moitié des équipes travaille sur le cas A, l'autre moitié sur le cas B. Ces études de cas se présentent sous forme d'un texte décrivant une situation d'évaluation des apprentissages, rédigée au « je » et dont le narrateur est un enseignant ou une enseignante. Le contexte se situe en enseignement supérieur.

Les cas décrivent des réflexions, des contraintes et des enjeux pédagogiques, qui nécessitent l'application des principes théoriques présentés plus tôt. À chaque itération, même si le cas est différent (A ou B), il porte sur une thématique commune et nécessite l'application des mêmes concepts théoriques.

Le travail à réaliser en équipe consiste à répondre aux questions suivantes :

- i. Quelles sont les distorsions ou les iniquités que peuvent engendrer les pratiques d'évaluation des apprentissages dans la situation décrite ?
- ii. Quelles alternatives à ces pratiques pouvez-vous imaginer qui permettraient d'amoinrir ces distorsions et qui en feraient des pratiques plus équitables ou plus fidèles ?

c. Retour en plénière

Une équipe ayant travaillé le cas A et une équipe ayant travaillé le cas B résumant le cas respectif et présentent les alternatives imaginées. Les personnes responsables de l'animation de l'atelier présentent ensuite les « réponses » qu'ils et elles avaient imaginées.

Les trois thématiques abordées au fil de l'atelier sont les suivantes :

Thématique 1 : Comment laisser place à l'erreur sans pénaliser la note ?

Éléments de contenu : Articuler des modalités de reprises/de révisions, ou utiliser les traces les plus représentatives du degré d'atteinte des objectifs terminaux, plutôt qu'utiliser la moyenne des notes.

Thématique 2 : Comment s'assurer que la note soit fidèle ?

Éléments de contenu : Reconnaître les pratiques de notation qui entraînent des biais en pénalisant ou en cherchant à encourager certains comportements, et envisager des alternatives plus fidèles.

Thématique 3 : Comment concevoir un système de notation plus équitable ?

Éléments de contenu : Comprendre qu'un manque de transparence en évaluation favorise inégalement les personnes socialement mieux adaptées au système éducatif.

La conclusion générale de l'atelier vise à faire un sommaire des fondements et des recommandations, en lien avec les thématiques étudiées. L'atelier se conclut sur des trucs pour l'implantation progressive des PAN, tirés des écrits de Ken O'Connor (2022).

Références

Conseil supérieur de l'éducation. (2018). *Rapport sur l'état et les besoins de l'éducation 2016-2018 – Évaluer pour que ça compte vraiment*. Gouvernement du Québec. <https://www.cse.gouv.qc.ca/wp-content/uploads/2020/01/50-0508-RF-evaluer-compte-vraiment-REBE-16-18.pdf>

Feldman, J. (2024). *Grading for equity : What it is, why it matters, and how it can transform schools and classrooms* (Second edition). Corwin Press, Inc.

Koenka, A. C., Linnenbrink-Garcia, L., Moshontz, H., Atkinson, K. M., Sanchez, C. E., & Cooper, H. (2021). A meta-analysis on the impact of grades and comments on academic motivation and achievement : A case for written feedback. *Educational Psychology, 41*(7), 922-947.
<https://doi.org/10.1080/01443410.2019.1659939>

O'Connor, K. (2022). *A repair kit for grading : 15 fixes for broken grades* (Revised third edition). First Educational Resources.

Townsley, M., & Schmid, D. (2020). Alternative grading practices : An entry point for faculty in competency-based education. *The Journal of Competency-Based Education, 5*(3), e01219.
<https://doi.org/10.1002/cbe2.1219>

Colloque international Diversité et Réussite[s]
3 avril 2024, Nantes Université

Les pratiques alternatives de notation
comme levier vers une évaluation plus équitable et
qui accompagne mieux l'apprentissage




Caroline Cormier et Bruno Voisard
Cégep André-Laurendeau
Montréal, Canada



Cette présentation a bénéficié d'une contribution financière du ministère de l'Enseignement supérieur du Québec dans le cadre du programme d'aide à la diffusion des résultats de recherche (PADRR). Le contenu de la présentation n'engage que la responsabilité de l'établissement et des auteurs.

1

Quelles sont les finalités de l'évaluation?

-  Quelques secondes pour y penser
-  Partager avec une personne voisine, retenir une idée
-  Partager cette idée avec la salle

2

Les finalités de l'évaluation



Soutenir l'apprentissage

Renseigner les étudiant-es sur leur progression
Donner des pistes sur ce qu'il faut faire pour poursuivre les apprentissages



Témoigner des acquis

Informar avec justesse, au terme du parcours, des apprentissages réalisés

Les pratiques de notation habituelles répondent-elles bien à ces deux finalités?

Conseil supérieur de l'éducation (2018)

3

Thématique 1: comment laisser la place à l'erreur sans pénaliser la note?

- Lecture individuelle du cas
- Travail en équipe – répondre aux questions suivantes :
 - i. **Distorsions ou iniquités** que peuvent engendrer les pratiques d'évaluation des apprentissages dans la situation décrite ?
 - ii. Alternatives qui en feraient des pratiques **plus équitables, plus fiables ou qui soutiennent mieux l'apprentissage** ?
- Discussion en plénière

4

Nos pistes de réponses à Marie-Hélène (1a)

Ne pas accompagner son feedback d'une note	Baisse de motivation intrinsèque (Koenka et collab., 2021) Il faut changer le rapport des étudiants à la correction
Laisser place à l'erreur en cours d'apprentissage	Faire retravailler le même travail plutôt que de demander des travaux différents La rétroaction indique l'écart entre le niveau des apprentissages et l'atteinte des objectifs.
Seulement la version finale du travail sera notée	Toute évaluation est formative, jusqu'à ce qu'elle soit sommative. La note finale est sujette à être modifiée si les apprentissages sont réalisés ultérieurement

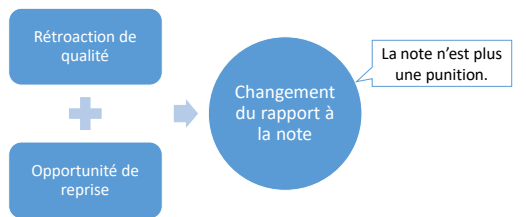
5

Votre avis sur les solutions de Clément (1b)?

- *Changer ma manière de corriger.*
- *Diminuer la pondération de cet examen.*
- *Ajouter des devoirs avec une faible pondération avant l'examen sur les mêmes sujets.*
- *Permettre à mes étudiants et étudiantes d'annuler la note d'un de leurs examens.*
- *Prévoir une reprise de l'examen récapitulatif accessible aux personnes qui n'ont pas obtenu la note de passage dans le cours, note finale max = la note de passage.*

6

Effet des PAN sur l'expérience éducative



Hattie & Timperley (2007), Koenka et coll. (2021)

7

Thématique 2: comment s'assurer de la cohérence de la note avec les objectifs du cours?

- Lecture individuelle du cas
- Travail en équipe – répondre aux questions suivantes :
 - i. **Distorsions ou iniquités** que peuvent engendrer les pratiques d'évaluation des apprentissages dans la situation décrite ?
 - ii. Alternatives qui en feraient des pratiques **plus équitables, plus fidèles ou qui soutiennent mieux l'apprentissage** ?
- Discussion en plénière

8

Nos pistes de réponses à Armelle (2a)

Rapporter la progression en donnant de l'info sur l'atteinte des objectifs d'apprentissage

- C'est l'objectif qui est noté, pas la question d'examen
- Améliore la qualité de la rétroaction

Calculer la note finale sur la base de l'atteinte des objectifs

- Si stoéchiologie est essentielle, alors un objectif « stoéchiologie » est maîtrisé seulement si l'étudiante démontre qu'elle a maîtrisé la stoéchiologie
- Idéalement : grille avec un nombre limité de niveaux
- Évite les problèmes d'Élodie (qui a eu la note de passage, mais n'a pas atteint les objectifs)

Fidélité : la note témoigne *uniquement* de l'atteinte des objectifs

- Si ce n'est pas un objectif, ça ne fait pas partie de la note
- Être en retard : ce n'est pas un objectif
- Ne pas contrôler les comportements en pénalisant la note.

Donner plus d'une occasion de démontrer la maîtrise des objectifs

- Sans pénaliser pour les reprises ou autres occasions

9

Nos pistes de réponse à Gabriel (2b)

À faire : clarifier les objectifs d'apprentissage et concevoir des activités d'évaluation qui se concentrent sur les cibles du cours

Rendre les cibles claires et fournir de la rétroaction pour accompagner leur développement

Éviter de confondre activité d'apprentissage et activité d'évaluation (risques de biais)

- Attention aux évaluations d'équipe: il faut s'assurer de l'atteinte individuelle des cibles d'apprentissage
- Attention aux devoirs qui entrent dans la composition de la note selon une logique de sommation: évaluation de la régularité du travail plutôt que des apprentissages

Feldman (2024)

10

Rapporter la progression en donnant de l'info sur l'atteinte des objectifs d'apprentissage

Pauvre en information		Riche en information	
Travaux et examens	Notes en cours de session	Énoncé abrégé des objectifs	Progression
Examen 1	13/20	Stoéchiométrie	En apprentissage
Examen 2	14/20	Structure des molécules	Ça y est presque
Examen final	/35	Polarité	Maîtrisé
Rapports de laboratoire	/20	Propriétés physiques	Maîtrisé
Projet affiche	4/5	Techniques de séparation	En apprentissage

Conseil supérieur de l'éducation (2018), O'Connor (2007)

11

Thématique 3: comment concevoir un système de notation plus équitable?

- Lecture individuelle du cas
- Travail en équipe – répondre aux questions suivantes :
 - i. **Distorsions ou iniquités** que peuvent engendrer les pratiques d'évaluation des apprentissages dans la situation décrite ?
 - ii. Alternatives qui en feraient des pratiques **plus équitables, plus fidèles ou qui soutiennent mieux l'apprentissage** ?
- Discussion en plénière

12

Nos pistes de réponses à Jean-Claude (3a)

<p>Transparence : les grilles doivent être fournies en même temps que les consignes</p> <ul style="list-style-type: none"> Manque de temps, nouveau prof, pas certains des standards? Se faire aider par ses collègues! Il faut bien connaître le contenu du cours pour décider des standards 	<p>Évaluation par les pairs : occasion de feedback</p> <ul style="list-style-type: none"> Pas nécessairement besoin de donner une note Si une note est assignée, un enseignement formel doit être prévu sur la façon de faire 	<p>Choisir les problèmes auxquels on répond : problème de fidélité de l'évaluation</p> <ul style="list-style-type: none"> Quelqu'un pourrait réussir le cours sans avoir jamais réussi tel type de problème. Peut-être que certains objectifs peuvent être considérés facultatifs. Mais l'essentiel ne doit pas pouvoir être écarté par les étudiants.
--	--	--

13

Nos pistes de réponses à Inès (3b)

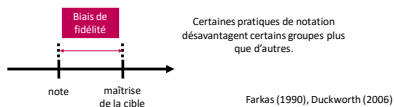
- Rappel: les deux finalités de l'évaluation sont de soutenir les apprentissages et de témoigner fidèlement des acquis
- Le classement des étudiant.es ne devrait pas être une finalité poursuivie par les profs
- La discrimination ne devrait donc pas être recherchée : si les cibles d'apprentissage sont clairement définies, qu'elles sont communiquées avec transparence, et que les étudiant.es sont accompagnés dans leur apprentissage de ces cibles : il y aura celles et ceux qui maîtrisent les cibles, et celles et ceux qui ne les maîtrisent pas.

Conseil supérieur de l'éducation (2018), O'Connor (2007)

14

Note qui évalue autre chose que les cibles: augmentation des biais de notation

- Exemples de **facteurs extérieurs** aux objectifs qui entrent parfois dans la composition de la note
 - Points de participation, pénalités pour absences
 - Pénalités de retard, pour matériel manquant
 - Points boni et productions supplémentaires
 - Points pour les devoirs et autres activités à visées formative
 - Plagiat
 - Productions d'équipes
 - Note de zéro pour un travail non remis (absence de trace ≠ absence d'apprentissage)
 - Trop grand nombre de niveaux de performance (introduction d'erreur aléatoire)



16

Les PAN: des pistes d'inspiration pour faire évoluer nos pratiques de notation

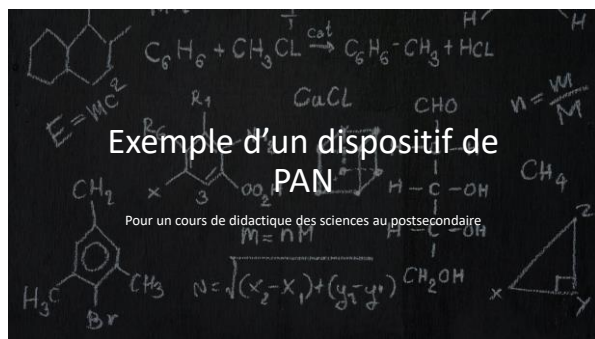
Différentes familles de PAN

Notation basée sur la maîtrise ou les standards	Notation par spécifications ou par contrat	Dénotation (comme pratique)
Notation basée directement sur l'évaluation du degré d'atteinte des objectifs du cours par opposition à une moyenne pondérée de la performance à des évaluations	Une note finale donnée correspond à un ensemble de productions prédéfini qui respecte les critères	La rétroaction qualitative , l' auto-évaluation et l' évaluation par les pairs sont favorisées. Le moins possible d'évaluations sont notées

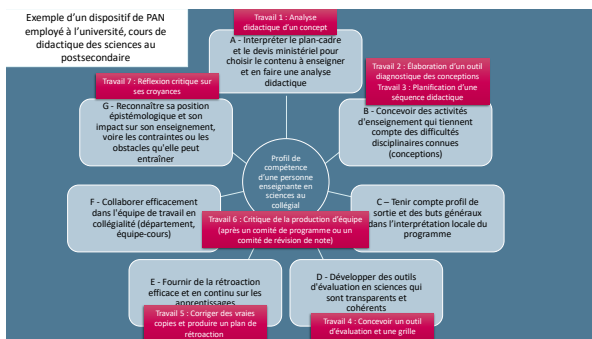
Si elle est présente, la « note » est un indicateur du progrès réalisé

Bloom (1968), Blum (2020), Clark & Talbert (2023), Townsley & Schmid (2020)

17



19



20

Objectifs du cours

- Réaliser l'analyse didactique d'un concept disciplinaire prescrit au devis ministériel (lié à l'élément A) ;
- Développer un instrument de détection des conceptions alternatives et démontrer sa pertinence (lié à l'élément B) ;
- Planifier une activité didactique pour l'enseignement d'un contenu reconnu pour être difficile ou être propice à des obstacles dus aux conceptions alternatives (lié à l'élément B) ;
- Concevoir un outil d'évaluation des apprentissages qui s'inscrit dans un système transparent, cohérent et équitable (lié à l'élément D) ;
- Fournir de la rétroaction individualisée, par écrit et portant sur la tâche à des étudiantes et des étudiants via la correction d'une évaluation (lié à l'élément E) ;
- Réfléchir de façon critique au déroulement d'un travail d'équipe, dans un contexte de comité de programme ou d'un comité de révision de note (lié aux éléments C et F) ;
- Évaluer l'impact sur sa pratique enseignante de ses croyances au sujet de la nature de la science (épistémologie), de l'enseignement et de l'évaluation (lié à l'élément G).

21

Consignes travail 4

Travail 4 Conception d'un outil d'évaluation des apprentissages Complémentaire

Description du travail

- Vous devez imaginer un système d'évaluation qui s'appuie sur des pratiques alternatives de notation dans un cours de sciences au collégial. Comme il s'agit d'un exercice dans le cadre de ce cours, vous n'avez pas à planifier l'ensemble du système d'évaluation pour un cours entier ; planifiez-le pour un module, un chapitre ou un élément de compétence seulement. Il doit tout de même y avoir plus d'un objectif d'apprentissage décrit dans votre système.
- Ce système doit prévoir un calendrier, des modalités d'évaluation, un outil d'évaluation complet (par exemple un minitest ou une description d'un travail), les consignes et la grille d'évaluation.
- Tout ceci doit être présenté dans un travail écrit au cours duquel vous répondez aux quatre questions listées par Townsley & Schmid (2020), en faisant la supposition que votre système serait étendu à toute une session.
- Vous y expliquez aussi en quoi votre outil d'évaluation et le système dans lequel il se situe respectent les valeurs de la cohérence, de la transparence et de l'équité. (En vous appuyant sur des sources pertinentes).
- Ce travail doit comporter trois pages environ, sans compter la bibliographie ni l'outil d'évaluation (qui doit être présenté en annexe) (Calibri 11 pts ou Arial 9 pts ou similaire, 1,5 interligne, marges habituelles).

Date de remise

- Lundi 4 mars 2024 à 23h59 sur Moodle

22

Consignes travail 4

Spécification pour le travail 4

- Le système d'évaluation présenté s'inscrit dans le cadre des pratiques alternatives de notation. Ce système inclut un calendrier, des modalités d'évaluation, un outil d'évaluation complet (par exemple un minitest ou une description d'un travail), les consignes et la grille d'évaluation.
- L'outil d'évaluation est cohérent avec les objectifs, les grilles permettent la transparence et les modalités d'évaluation respectent des principes d'équité, et le respect de ces valeurs est démontré en s'appuyant sur des sources pertinentes.
- Le système prévoit comment laisser l'étudiant.e.s faire des erreurs, sans que ça se transpose en pénalité sur la note finale. Le calendrier est conçu de façon à permettre de l'espace pour donner de la rétroaction aux étudiants.e.s.
- Les objectifs décrivent des observables clairs et qui correspondent bien au contenu disciplinaire choisi. L'outil d'évaluation présenté évalue avec cohérence les objectifs listés (ou un de ces objectifs).
- Le raisonnement autour du développement de ce système est cadré autour des quatre questions proposées par Townsley et Schmid (2020) et les réponses à ces questions sont formulées explicitement.
- Les sources sont citées au fil du travail, et ces sources sont pertinentes pour appuyer les propos.

23

23

Résumé des éléments essentiels

- Les occasions de reprises et de révision des évaluations permettent de mieux
 - soutenir les apprentissages en créant des boucles de rétroaction
 - témoigner de l'atteinte des objectifs en fin de session en ne pénalisant pas les erreurs réalisées en cours d'apprentissage
- Aligner l'ensemble du cours sur les objectifs d'apprentissage (incluant la note et la rétroaction), permet de
 - Mieux rétroagir (donner de l'information précise au sujet de l'écart entre les apprentissages réalisés et les cibles)
 - Donner une note finale en adéquation avec le degré d'atteinte des objectifs terminaux (degré d'atteinte des objectifs = sommation de notes à des évaluations)

24

L'équipe de recherche sur les pratiques alternatives de notation

François Arsenault-Hubert

Caroline Cormier

Veronique Turcotte

Bruno Voisard

Département de chimie, Cégep André-Laurendeau

25

Références

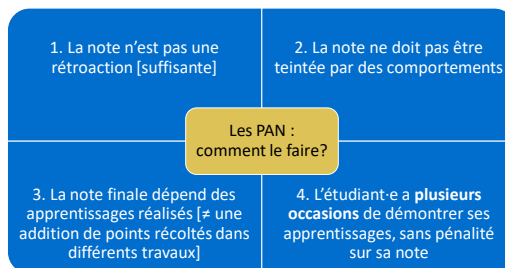
- Bloom, B. S. (1988). Learning for Mastery Instruction and Curriculum. Regional Education Laboratory for the Carolinas and Virginia, Topical Papers and Reports, Number 5. Data Evaluation Comment (Vol. 1, Numéro 3). Regional Education Laboratory for the Carolinas and Virginia, Mutual Plaza (Chapel Hill and Duke St. <https://eric.ed.gov/?id=ED253413>)
- Blum, S. D. (Ed.). (2020). *Ungrading : Why Rating Students Undermines Learning* (First Edition). West Virginia University Press.
- Clark, D., & Tabert, R. (2023). *Grading for growth : A guide to alternative grading practices that promote authentic learning and student engagement in higher education*. Routledge.
- Conseil supérieur de l'éducation. (2018). *Rapport sur l'état et les besoins de l'éducation 2016-2018 - Évaluer pour que ça compte vraiment*. REBE-16-18.pdf. Gouvernement du Québec. https://www.cse.gouv.qc.ca/wp-content/uploads/2020/01/20-0508_RE-évaluer-compte-vraiment-REBE-16-18.pdf
- Duckworth, A. L., & Seligman, M. E. P. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology*, 98(3), 198-208. <https://doi.org/10.1037/0022-0663.98.3.198>
- Farkas, G., Grobe, R. P., Sheehan, D., & Shuan, Y. (1990). Cultural Resources and School Success: Gender, Ethnicity, and Poverty Groups within an Urban School District. *American Sociological Review*, 55(1), 137-142. <https://doi.org/10.1086/208708>
- Feldman, J. (2024). *Grading for equity : What it is, why it matters, and how it can transform schools and classrooms* (Second edition). Corwin Press, Inc.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112. <https://doi.org/10.3102/003465430298487>
- Koehn, A. C., Linemerk-Garcia, L., Mohrhardt, H., Johnson, K. M., Sanchez, C. E., & Cooper, H. (2023). A meta-analysis on the impact of grades and comments on academic motivation and achievement : A case for written feedback. *Educational Psychology*, 41(7), 922-947. <https://doi.org/10.1080/00131302.2023.2209999>
- O'Connor, K. (2007). *A repair kit for grading : 25 fixes for broken grades* (Second edition). Educational Testing Service.
- Townsley, M., & Schmid, D. (2020). Alternative grading practices : An entry point for faculty in competency-based education. *The Journal of Competency-Based Education*, 9(3), 40-129. <https://doi.org/10.1002/cbe.1218>

26

Diapos en réserve!

En fonction des questions et des besoins

27



28

1. La note n'est pas une rétroaction [suffisante]

Est-ce nécessaire de donner des points à toutes les évaluations?

- Le petit examen qui prépare pour un autre examen...
- Le petit rapport de labo ou le petit texte préparatoire...
- Une courte entrevue qui prépare à la grosse entrevue...

Pas de points, mais rétroaction détaillée

- Individualisée
- Portant sur la tâche

Ne pas accompagner la rétroaction d'une note

- Sinon : diminution de la motivation intrinsèque (Koenka, 2019) et la rétroaction risqué de ne pas être lue

Hattie & Timperley (2007), Koenka et al. (2021)

29

2. La note ne doit pas témoigner des comportements

Retards, absence de matériel de sécurité (labo), voire absence non motivée à une évaluation

- Est-ce que ça fait partie du devis ministériel de notre cours? Des buts généraux de notre programme? Du profil de sortie?
- Pénalise d'abord les étudiant-es vulnérables → introduction de biais

La note doit témoigner du degré d'atteinte de la cible d'apprentissage du cours

- Et non pas être contaminée par des pénalités touchant d'autres aspects du comportement
- C'est une question de fidélité de l'évaluation (= la note correspond bien aux apprentissages réalisés)

Réduction du nombre de paliers

- Par ex. : note à 4 niveaux plutôt que 100
- = réduction des biais

Feldman (2024), O'Connor (2022)

30

3. La note finale dépend des apprentissages réalisés [≠ une addition de points récoltés dans différents travaux]

- Les notes partielles : montrent ce qui est acquis et ce qui ne l'est pas encore (donne plus d'information)

Travaux et examens	Pauvre en information		Riche en information	
	Notes en cours de session	Énoncé abrégé des objectifs	Progression	
Examen 1	13/20	Stoéchiométrie	En apprentissage	
Examen 2	14/20	Structure des molécules	Ça y est presque	
Examen final	/35	Polarité	Maîtrisé	
Rapports de laboratoire	/20	Propriétés physiques	Maîtrisé	
Projet affiche	4/5	Techniques de séparation	En apprentissage	

Conseil supérieur de l'éducation (2018), O'Connor (2022)

31

4. L'étudiant-e a plusieurs occasions de démontrer ses apprentissages, sans pénalité sur sa note

Stratégie 1: offrir des occasions de reprises

- Même mode d'évaluation
- Ou mode différent

Stratégie 2: prévoir des possibilités de révision

- Approprié pour les textes, les travaux de longs, les projets
- Amène les étudiant-es à profiter de la rétroaction

Stratégie 3: porter un regard holistique sur les productions

- Déterminer la note finale en fonction du degré d'atteinte des objectifs terminaux
- Choisir les traces les plus représentatives des apprentissages (ne pas pénaliser les erreurs en cours d'apprentissage)

Clark & Talbert (2023), O'Connor (2022)

32

Étude de cas #1a : Marie-Hélène, prof de psycho

Je m'appelle Marie-Hélène, enseignante en psychologie au cégep, et je me trouve confrontée à des défis majeurs dans le cadre de mes cours, notamment en ce qui concerne l'évaluation. Un problème récurrent qui suscite ma préoccupation est au sujet des travaux écrits. Les trois travaux au fil de la session consistent à analyser le comportement d'un personnage de film, de roman ou de série télévisée, en utilisant les concepts, les modèles et les théories vus en classe. Le problème est le fait que mes étudiants et mes étudiantes ne semblent pas lire attentivement les commentaires que je laisse sur leurs travaux, et ils semblent ignorer ces conseils lorsqu'ils abordent les travaux suivants. Au lieu de tirer profit de mes feedbacks, les étudiants se limitent à vérifier simplement leur note. Lors du travail suivant, je constate avec déception que les mêmes personnes réitèrent les mêmes erreurs, compromettant ainsi leur progression et leur compréhension des concepts clés de la psychologie.

Parallèlement, la structure d'évaluation de la session, qui prévoit trois examens (25 %, 25 %, et 30 %), a également révélé des préoccupations importantes. Le cas de Jean, un étudiant père monoparental, souligne les difficultés auxquelles certains font face. Incapable de se présenter à l'examen 1 en raison de la fermeture de la garderie de sa fille, Jean s'est vu attribuer un zéro, ce qui a eu un impact immédiat de 25 % sur sa note globale. Bien que j'aurais aimé offrir une solution adaptée à sa situation, les règles départementales strictes exigent un billet de médecin pour autoriser une reprise, ce qui me semble créer des barrières injustes pour les étudiants confrontés à des circonstances exceptionnelles. Toutefois, je reconnais que pour s'assurer d'une équité envers tous les étudiants, cette règle permet d'éviter les abus et ainsi tous les étudiants sont traités de la même façon.

En outre, je constate que des étudiantes ont considérablement progressé au cours de la session grâce à leur assiduité, à leur présence régulière et leur investissement dans les exercices.

Cependant, leur note finale ne le reflète pas en raison d'une performance décevante à l'examen 1. Il est déconcertant de constater que des étudiantes démontrant une réelle implication n'obtiennent pas toujours une note finale représentative de leurs apprentissages. Y a-t-il des solutions à toutes ces préoccupations?

Étude de cas #2a : Armelle, prof de chimie

Aujourd'hui, j'ai beaucoup de pain sur la planche : je dois corriger une évaluation que j'ai donnée la semaine dernière. Je m'appelle Armelle et je suis professeure de chimie au cégep. L'évaluation que je dois corriger est un examen qui porte sur la stœchiométrie, un sujet important dans mon cours.

Je commence par corriger la première copie que je prends au hasard dans la pile. Je lis attentivement les réponses et je vérifie si elles sont conformes au barème de correction que j'ai préparé à l'avance. Je souligne les erreurs, je fais des commentaires et j'enlève des points selon les critères que j'ai établis. Je passe à la deuxième copie et je fais la même chose. Puis à la troisième, à la quatrième, etc. Au bout d'un moment, je me rends compte que je dois revenir sur des copies que j'ai déjà corrigées pour m'assurer d'enlever le même nombre de points pour le même type d'erreurs. Je trouve ça fastidieux.

Je finis par arriver à la dernière copie. Je fais le total des points pour chaque étudiant et je calcule la moyenne du groupe. Elle est de 73 %. Je regarde de plus près les notes individuelles et je remarque qu'il y a une grande disparité entre les étudiants. Certains ont eu plus de 90 %, d'autres moins de 60 %.

Je prends une copie au hasard et je la relis avec attention. C'est celle d'Élodie, une étudiante qui a eu 62 %. C'est une note de passage, mais est-ce que ça veut dire qu'elle a maîtrisé les concepts essentiels? Je regarde ses réponses et je constate qu'elle a commis plusieurs erreurs graves. Elle a quand même réussi à avoir quelques points en répondant partiellement ou en appliquant des formules sans les comprendre. Je me dis qu'elle n'a pas atteint les objectifs les plus importants qui étaient évalués. Si elle ne complète pas ces apprentissages sur ces sujets, elle risque de rencontrer des difficultés dans le cours de chimie des solutions, qui est le cours de chimie suivant dans le programme de sciences.

En réfléchissant à ce problème, je me rappelle un événement qui s'est produit à la séance où a eu lieu l'examen : Julien est arrivé plus d'une heure en retard. À son arrivée, j'ai bien vu qu'il était en panique, mais il est allé s'asseoir à un bureau sans rouspéter, il a utilisé le temps qui lui restait et il est reparti en coup de vent à la fin du cours, comme il le fait souvent. En repensant à cet événement, je retourne à sa copie d'examen. Il n'a pas eu le temps de répondre à toutes les questions, mais ce qu'il a pu faire me démontre qu'il a réalisé l'essentiel des apprentissages. Malgré ça, sa note finale est de 53 %. Ça me laisse un malaise encore plus grand qu'après la consultation de la copie d'Élodie.

Je suis insatisfaite de mon évaluation. Je me pose alors des questions sur sa validité de mon évaluation et sur la valeur de la note. Est-ce qu'elle mesure vraiment ce que je voulais mesurer? Est-ce qu'elle reflète le degré d'atteinte des objectifs d'apprentissage que j'avais fixés? Est-ce qu'elle communique des informations assez précises sur les forces et les faiblesses de chaque personne?

Étude de cas #3a : Jean-Claude, prof de technologie de l'architecture

C'est ma passion pour l'architecture qui m'a poussé à réorienter ma carrière vers l'enseignement. J'avais envie de transmettre cette passion à mon tour. Je m'appelle Jean-Claude et je suis un nouveau professeur de technologie de l'architecture au cégep. Ma première session au collégial a été riche en enseignements, mais elle a également soulevé des défis en matière d'évaluation.

Pour évaluer mes étudiants, j'ai choisi de leur assigner des devis pour différents types d'édifices. Lors de la première présentation des consignes, je n'avais pas encore élaboré la grille de correction. À la réception des travaux, j'ai rapidement compris que mes attentes présentées en classe et les exigences formulées dans la grille n'étaient pas en adéquation. Plusieurs étudiants avaient omis des parties cruciales du devis, tandis que d'autres négligeaient des aspects importants que j'avais spécifiés dans la grille. Cela a entraîné des frustrations chez les étudiants et le climat de classe s'en est ressenti. J'avais l'impression d'avoir perdu la confiance de mes étudiants, et mes interactions avec eux ont été tendues le reste de la session.

Pour remédier à la situation, j'ai décidé d'adopter l'évaluation par les pairs, une méthode que j'avais apprise lors d'une formation en pédagogie. Pensant que mes étudiants, étant des adultes et en troisième année du programme, avaient une connaissance suffisante en architecture pour évaluer le travail de leurs pairs, j'ai mis en place ce processus. Certains des travaux de la session ont été notés par les étudiants eux-mêmes. Cependant, à ma grande surprise, les notes attribuées par les pairs étaient parfois très différentes de celles que j'aurais données. Certaines équipes se sont octroyé des notes élevées de manière uniforme, tandis que d'autres ont été plus sévères dans leurs évaluations. Cela a soulevé des questions sur la fiabilité de cette méthode et sur la capacité des étudiants à évaluer objectivement le travail de leurs camarades, et a ravivé les frustrations dans ma classe.

Face à la tendance de la moyenne des notes à être plus basse que prévu, j'ai pris une décision qui a suscité des réflexions. Dans les examens suivants, j'ai permis aux étudiants de choisir les problèmes auxquels ils répondaient, une décision motivée par le désir d'augmenter la moyenne globale de mon cours. Cette initiative visait à éviter que le département ne perçoive ma performance d'enseignant de manière négative.

En conclusion, ma première session d'enseignement en technologie de l'architecture m'a confronté à des défis dans le processus d'évaluation. Je ne sais pas quelles stratégies adopter, y a-t-il des solutions aux problèmes que j'ai vécus? Que pourrais-je mettre en place, à la prochaine session?

Étude de cas #1b : Clément, prof de mathématiques

Je suis Clément, prof de mathématiques depuis quatre ans maintenant. J'ai vécu une situation difficile à la dernière session : mon examen de mi-session n'a été réussi que par trois étudiants de mon groupe. J'avais pourtant l'impression que mon évaluation portait sur les sujets enseignés dans le cours, que les thèmes abordés avaient été clairement annoncés à l'avance et que le niveau de difficulté était représentatif de mes attentes. Comme je le fais avant chacune de mes évaluations, j'avais même pris la peine de discuter avec Annie, une prof d'expérience en qui j'ai confiance, de son point de vue sur la hauteur de mes attentes pour les sujets à cet examen. En plus de tout ça, beaucoup de temps de classe avait été consacré à la pratique de ces sujets, ce qui m'avait amené à croire que la moyenne du groupe allait être élevée.

Pour toutes ces raisons, je n'ai pas changé les résultats ou ma correction, mais j'ai l'impression que ça a eu un effet négatif sur la motivation de plusieurs étudiants et étudiantes du groupe, qui ont arrêté de venir au cours après la remise des notes de cet examen. Je crois aussi que cela a eu un effet négatif sur le climat de classe et sur l'attitude de plusieurs membres du groupe à mon égard. Pour être franc, ma fin de session a été un peu pénible.

Cette session, je donnerai à nouveau ce cours, et en repensant à tout ça alors que j'en suis à l'étape de la planification de ma session, j'ai décidé d'apporter des ajustements à mes pratiques d'évaluation. Pour m'aider à y réfléchir, je jette sur papier différentes idées qui me permettraient d'éviter que les problèmes que j'ai vécus à la dernière session ne se reproduisent. Avant d'écrire mon plan de cours, je reviens à ma liste d'idées pour arrêter mon choix sur une ou plusieurs de mes idées.

- Changer ma manière de corriger.
- Diminuer la pondération de cet examen.
- Avant l'examen, ajouter des devoirs portant sur les mêmes sujets. Ces devoirs compteraient dans la note finale, mais auraient une faible pondération.
- Permettre à mes étudiants et étudiantes d'annuler la note d'un de leurs examens.
- Prévoir une reprise de l'examen récapitulatif, après la fin de la session, qui serait accessible aux personnes qui n'ont pas obtenu la note de passage dans le cours. Pour éviter un sentiment d'injustice de la part des autres étudiant-es, la note maximale accessible aux personnes qui bénéficient de la reprise serait la note de passage.

C'est la première fois que rédiger un plan de cours me rend aussi anxieux. Je souhaite bien préparer mes étudiants et étudiantes aux autres cours du programme et faire en sorte que leurs notes soient représentatives de leurs apprentissages, mais d'un autre côté, je suis préoccupé de leur santé mentale et de la mienne. Je tiens à éviter d'avoir à vivre une autre fin de session aussi désagréable que la dernière. Laquelle ou lesquelles de ces idées devrais-je retenir? Est-ce qu'il y a des meilleures solutions que celles auxquelles j'ai pensé?

Étude de cas #2b : Gabriel, prof d'administration

Mes collègues le savent : une de mes préoccupations, comme prof, c'est d'amener mes élèves à développer des habitudes de travail qui favorisent la réussite professionnelle.

C'est pourquoi les devoirs, les tests de lecture et la participation aux discussions en classe comptent pour une partie de la note finale de mes cours. En effet, je sais que le travail régulier est synonyme de réussite et je sais aussi qu'en attribuant des points à des travaux, les élèves comprendront que j'y attribue de l'importance. Par cette pratique, j'espère aussi les amener à développer l'habitude de travailler et de faire leurs lectures régulièrement.

J'ai cependant remarqué que certains élèves qui performant bien lors des évaluations individuelles en classe perdent beaucoup de points dans leurs devoirs. Ça m'inquiète, puisque je souhaite que la note finale de mes étudiants soit représentative de leur performance. Je me vois comme un patron qui ne réussit pas à tirer le meilleur d'employés pleins de potentiel, quand je suis face à de telles situations.

J'intègre également des projets d'équipe qui contribuent significativement à la note de mes cours. Cette dimension collaborative est essentielle, car elle prépare les élèves à l'importance du travail d'équipe dans leur parcours académique et professionnel. C'est une compétence que je valorise particulièrement, car elle reflète la réalité du monde extérieur où les projets sont souvent le fruit d'un effort collectif. Je suis particulièrement fier de mes études de cas où mes élèves jouent le rôle d'une équipe de direction : elles contribuent vraiment à dynamiser mes cours de management.

Mon seul souci, par rapport aux travaux d'équipe, c'est qu'il m'est arrivé que des élèves faibles semblent obtenir la note de passage en raison de la note qu'ils obtiennent à ces travaux. Je songe à ajuster mes pratiques pour trouver une solution à ce problème. Comment s'assurer que les notes attribuées reflètent véritablement les compétences et la contribution de chaque élève, tant sur le plan individuel que collectif ? Je cherche à trouver un équilibre entre les évaluations individuelles, le travail d'équipe, la contribution aux discussions en classe et le développement de bonnes habitudes de travail, pour que la note finale soit le véritable reflet de l'engagement de mes élèves et de leurs compétences.

Étude de cas #3b : Inès, prof de physique

Je suis quelqu'un qui cherche toujours à s'améliorer, et après avoir corrigé un examen, je fais des statistiques avec les résultats à chacune des questions. Ça me permet de les ajuster – lesquelles sont trop faciles? Lesquelles sont trop difficiles? J'élimine les questions qui ont un facteur de discrimination ou un indice de difficulté trop faible, ou je les ajuste. Je m'assure aussi d'avoir une bonne corrélation entre chacune des questions et le résultat final à l'examen, puisque si cette corrélation est trop faible, ça signifie que la question mesure quelque chose de différent du reste de l'examen.

Je suis assez fière de mes examens de physique mécanique : c'est un cours que j'ai donné souvent, ce qui m'a permis d'arriver à de très bons examens. Pourtant, cette session, il s'est produit quelque chose de bizarre dans mon cours de mécanique. Quand je prépare un examen, je cherche à faire en sorte qu'il permette de bien distinguer le degré de compréhension de chaque personne. Pour y arriver, je m'assure d'avoir une belle distribution des résultats. Même en ayant sélectionné soigneusement les meilleures questions de ma banque, je n'obtiens pas les résultats que je souhaite. Les résultats sont plus élevés que lors des sessions antérieures, ce qui fait diminuer les indices de discrimination. La distribution des résultats finaux est beaucoup moins belle qu'à l'habitude, d'ailleurs. Je suis convaincue que cette anomalie n'est pas due à la tricherie. Après en avoir parlé avec des collègues, j'ai compris que les étudiants et étudiantes de cette cohorte sont plus forts que d'habitude : dans les cours de mathématiques et de chimie, les moyennes sont aussi plus élevées qu'aux sessions précédentes.

J'ai remarqué aussi que le climat de classe était particulièrement compétitif, cette session. Avec beaucoup d'étudiantes et d'étudiants qui cherchent à accéder à des programmes universitaires contingentés, c'est compréhensible. Je m'inquiète quand même des problèmes causés par l'anxiété de performance. De plus, je me rends compte qu'il suffit d'une distraction de temps en temps pour obtenir des notes qui peuvent remettre en question un choix de carrière. Je me demande entre autres, face à ça, si je devrais augmenter un peu davantage le degré de difficulté de mes questions. Comme ça, mon indice de discrimination devrait augmenter aussi. Le seul problème, avec une telle solution, c'est que des personnes plus faibles qui réussissaient le cours, avant, ne le réussiraient peut-être plus, alors qu'elles maîtrisent les concepts étudiés de façon suffisante. Est-ce que je dois concentrer mon évaluation sur la certification de l'atteinte des cibles d'apprentissage, ou ai-je raison de croire que mon rôle d'évaluatrice est de permettre de classer mes étudiants les uns par rapport aux autres?