



## ÉVALUATION DES EXAMENS

### CONTEXTE

J'ai commencé à donner des cours à l'université en 1987, ce qui ne fait pas nécessairement de moi une sommité en pédagogie, loin de là.

Mais cette expérience trentenaire me permet un certain recul, une certaine perspective sur l'enseignement et son corollaire l'évaluation.

J'ai eu la responsabilité des évaluations de mes cours au moment où je suis devenu professeur en 1995 et, contrairement à plusieurs collègues, je me suis intéressé à la pédagogie universitaire dont l'évaluation est une partie indissociable. J'ai travaillé depuis 1997 avec des spécialistes du domaine et expérimenté plusieurs outils d'évaluation des connaissances et des compétences dans des cours divers où la taille oscillait entre 5 et 450 étudiants, ici mais en Europe, en Afrique et en Amérique du Sud, aux premiers, 2<sup>e</sup> et 3<sup>e</sup> cycles universitaires, dans des Écoles de génie mais aussi en sciences de l'éducation.

### MEILLEURE FORME D'ÉVALUATION

Arrivé au terme de ma carrière, je suis convaincu que la meilleure forme d'évaluation est le compagnonnage tel qu'il se pratiquait il y a des siècles. Dans un tel contexte, si on transpose aux notes que nous avons aujourd'hui, un étudiant maintient la note échec (E ou F ou 0/20, selon les systèmes) tant qu'il n'obtient pas la cote la plus élevée (A ou A+ ou encore 20/20) : il est déclaré compétent, ou pas. Pour ce faire, transposant cette forme d'apprentissage du passé à aujourd'hui, vous seriez appelé(e) à travailler avec moi directement sur des contrats de recherche (pas sur une cathédrale!). Je vous assisterais pour préparer les demandes de subventions et pour réaliser mesures, simulations ou calculs théoriques puis je réviserais tous vos calculs, toutes vos simulations et toutes vos calibrations d'instruments de mesure. Je travaillerais AVEC vous toute la journée. Cela fonctionne optimalement avec un rapport superviseur/supervisé de 1:1. Dans un contexte d'ingénierie, je vous ferais préparer les réponses aux appels de projets, la supervision de chantier et/ou de ressources humaines en ayant là aussi un contact continu et quotidien AVEC vous.

Ce qui se rapproche le plus de cette formule à l'université est le projet de recherche de doctorat ou de maîtrise. Mais, puisque l'université s'est mercantilisée à partir des années 1990 (au Québec en 1994 avec la réforme Legault, alors ministre de l'Éducation, qui a créé les contrats de performance – je simplifie un peu l'histoire ici), de nos jours les chercheurs universitaires supervisent jusqu'à 30 étudiants, voire davantage, ce qui est absolument impossible à faire rigoureusement (basé sur la taille maximale de mon groupe de recherche en 2013 qui comportait plus de 20 personnes dont 3 chercheurs postdoctoraux). Ce mentorat, se retrouve aussi en entreprise lorsqu'un(e) ingénieur(e) sénior vous est assigné(e) afin de vous aider à vous intégrer à la boîte qui vous a embauché(e).

Mais, ces formes d'évaluation continue et personnalisée ne sont pas possibles dans les institutions du 21<sup>e</sup> siècle. Enfin dans la majorité d'entre-elles, surtout lorsqu'elles sont publiques. Le ratio en classe dépasse 30:1, en moyenne dans les établissements de l'Université du Québec (et tel je le mentionnais plus haut, je



donnais un cours à 450 personnes à l'Université Laval) qui affiche les moyennes les plus basses au Québec.

## **PIRE FORME D'ÉVALUATION**

A l'opposé du spectre des évaluations, on retrouve l'examen. A mon avis la pire forme d'évaluation des compétences et connaissances. Pire que les exercices dirigés, les travaux pratiques, etc.

La performance à un examen dépend de plusieurs facteurs qui n'ont pas de lien avec les compétences et connaissances. L'état d'esprit, la santé, etc.. Pire, des gens ayant de très bonnes notes à un examen pourraient s'avérer incapables de porter un jugement professionnel alors qu'à l'inverse, on peut se planter royalement à un examen et avoir néanmoins une fructueuse carrière devant soi.

Encore en cette première demie du 21<sup>e</sup> siècle, il existe des universités (en Espagne où j'ai enseigné en 2000) où un examen ANNUEL CUMMULATIF sanctionne toute une année de cours (qui là-bas ne se termine pas en mai...). Pour illustrer toutes les variations possibles.

Mais pour comparer avec les examens d'ici, mon second exemple sera tiré des examens des cours des programmes de médecine. À l'Université Laval, pendant les trois années de formation générale (ce qualitatif – générale – est important ici) les étudiants ont 90% de leurs évaluations sous la forme des quiz de ENR. À la différence près qu'ils n'ont PAS droit à leurs notes, leur ordinateur ou leur téléphone. Crayon (avant COVID) et c'est tout.

Si vous formulez le commentaire suivant dans votre esprit : « bourrage de crâne », je retorquerai oui et non. Cet exercice permet de faire pénétrer des notions dans la mémoire à court terme. Il faut évidemment des heures de pratique pour que ça pénètre plus loin. Il faut répéter, utiliser, reformuler, et travailler souvent. Si l'apprentissage d'un cours no.1 se produit SEULEMENT la veille de l'examen et que le soir suivant vous recommencez à étudier un cours no.2 (sans avoir ouvert un livre ou fait un exercice AVANT), je vous l'accorde, c'est du bourrage de crâne. Car le cours no. 2 viendra remplacer le cours no. 1 en 24 ou 36 heures. Le cours no.1 sera évacué (*flushé* diront certains) de la mémoire à court terme au profit du second. Si par contre, vous devez vous astreindre à préparer/étudier/lire/synthétiser une fois, deux fois, trois fois les mêmes concepts. Ils resteront plus profondément dans votre esprit. Vous les apprenez alors par cœur, dans le sens véritable du terme.

Bien entendu, il faut une pratique considérable pour que vous sachiez par cœur la résistance ohmique d'un chauffe-eau ou la conductivité thermique de l'aluminium. Il faut de l'expérience pour déclarer sans calcul, à 4 km de distance, que 400 tonnes de climatisation sont en fonction dans un hôpital un matin de février et sans avoir jamais consulté un plan du bâtiment. Et il est impossible de faire de vous de tels experts en 13 semaines.

Mais, en proposant les auto-évaluations de 5 questions, le cours force à travailler souvent, sur une longue période (non restreinte à la veille de l'examen), et il reste aux étudiants qui le font depuis 2020, davantage qu'à ceux qui l'ont fait précédemment. Ce n'est pas la panacée, c'est simplement mieux.



Pourquoi est-ce que je permets les notes de cours?

Il y a un lien à faire tout d'abord avec les trois premières années de médecine : les cours ENR sous ma responsabilité sont très généraux, ils ratissent très large. Il faut donc des questions simples et variées qui seront les plus claires possible et c'est pourquoi je vous demande de me signaler, outre les erreurs, les ambiguïtés de certaines questions. Je permets les notes de cours en sachant que celles et ceux qui auront mieux lu/étudié/parcouru/complété les notes, n'auront pas à chercher sans arrêt dans un PDF pour retrouver la page qui correspond au concept. Ils termineront de répondre aux 20 questions sans systématiquement aller avec « ctrl-F » dans les PDFs. Ou, ils n'auront pas à faire de calcul (Quelle est la déclinaison solaire à Tokyo le jour de l'équinoxe à 14h43?). S'ils prennent 15 minutes pour répondre aux 20 questions, ils disposent de 25 de plus (on budgete de 30 à 40 minutes pour répondre aux questions) pour faire les problèmes (ce qui ne garantit pas la performance, mais permet de réviser les calculs).

En ce qui concerne les problèmes à calcul, il n'y a pas de techniques complexes (hormis le calcul de l'angle d'incidence du soleil sur une surface inclinée dans le module 11) à appliquer. Les problèmes proposés dans le cours et donc aux examens sont tous d'une désarmante simplicité.

## **CORRECTION DES EXAMENS**

Ce qui importe dans la correction est l'équité envers tous et toutes. Pour être équitable, il faut d'abord corriger d'un seul trait le même problème pour toute la classe et le même jour. Ensuite, il faut un barème pour attribuer des notes à différents éléments. Puis, il faut pénaliser les mêmes erreurs (ou équivalentes) de la même manière. Puis, une fois un problème évalué, on passe au problème suivant de la même manière.

Si, par exemple, on décide de rétribuer les unités de 2/5, on doit retirer 2/5 à toutes les personnes qui n'ont pas mis les unités appropriées ou se sont trompées dans la conversion. Mais, si on ferme les yeux sur ça, il faut les fermer pour tout le monde. Et il existe une troisième variante : une personne a mis les unités MAIS a fait une erreur de conversion. Il est alors possible de mettre 0 ou 2 ou 1 voire même autre chose 0,5 ou 1,5 : l'important est d'être conséquent pour toutes et tous.

Exemple, dans la dernière question d'un INTRA 2 récent de ENR811, on demandait le coût annuel de chauffage au bois. Il fallait multiplier le coût unitaire par la quantité calculée antérieurement. Si les gens avaient calculé la mauvaise quantité de bois, ils avaient nécessairement la réponse numérique fautive. Ainsi, ai-je accordé tous les points de la solution à toutes celles et ceux qui ont écrit « Pour obtenir le coût annuel de chauffage au bois il faut multiplier le coût unitaire par la quantité calculée ». J'ai même octroyé les points à une personne qui, ayant fait une erreur d'ordre de grandeur (facteur 1000), a indiqué : « J'ai dû faire une erreur de calcul, cette réponse est impossible ». Mais, je l'ai fait pour tous et tous. Ce qui bien entendu a augmenté la moyenne par rapport à ce qu'elle aurait été si j'avais sanctionné toutes celles et ceux qui n'avaient PAS présenté de calcul classique dans cette case.



Dans les EXAMENS, au départ il s'agit, à peu de choses près, de problèmes simples.

Des problèmes simples du type  $F=ma$  : On donne  $F=50N$ ,  $m = 10kg$ , que vaut l'accélération  $a$ ? (10 points). Parfois je donne aussi des données inutiles.

Si je ne demande QUE la réponse numérique, j'espère obtenir 100% de 5. Je ne puis octroyer des points si on m'affirme avoir fait une erreur en multipliant 50 par 10. Le résultat est 0/10 ou 10/10.

Si je demande aussi une solution, il y aura 5 (ou 3 ou 4) points pour la réponse et 5 (ou 7 ou 6) points pour le calcul. Même si le 5 est donné, j'accorde 0/10 au total s'il n'y a pas de solution (plagiat potentiel).

Si je décide que j'octroie 2/5 pour les unités dans la solution, que la réponse est correcte, que le « calcul » présenté est correct mais qu'il n'y a pas d'unités, la personne obtient 8/10.

Plus un problème est simple, et plus la pondération de la réponse seule est prépondérante. Généralement 5+5 ou 4+6 mais j'ai aussi implanté 10/0.

Pourquoi pas des problèmes à développement?

C'est un choix très COVID19, les problèmes sont désormais trop simples, les évaluations dans le cours sont nombreuses. Les examens dits à développement souffrent des mêmes problèmes que les autres et induisent souvent encore davantage de biais possibles puisqu'ils proposent souvent des problèmes complexes pouvant donner lieu à une foule de réponses correctes mais différentes de celle qui est dans le corrigé proposé par le prof. C'est pourquoi, pour ce cours, les questions sont devenues très simples mais l'erreur de calcul est sévèrement sanctionnée.

## NOTES ET COTES

Les notes ne vous serviront que pendant votre passage universitaire. Qui se souciera de vos résultats académiques? Ils ne servent qu'à octroyer des bourses. Votre entretien d'embauche sera beaucoup plus déterminant.

Aller consulter ce document qu'il est obligatoire de préparer et que personne ne lit : le **Plan de cours**. Pour établir les cotes, les seuils sont aménagés de manière à obtenir une moyenne de supérieure à 3,0/4,3. C'est une particularité qui n'est propre qu'à mon analyse de la situation à l'ÉTS où l'institution exige B de moyenne pour obtenir un diplôme de second cycle. Donc, si en raison d'une correction « sévère » personne ne pouvait obtenir A+, le seuil du A+ pourrait être modifié à la et conséquemment pour les cotes inférieures sans altérer le seuil de passage à 50%. Qui plus est, lorsque la meilleure note d'une épreuve est trop éloignée de 100%, il arrive que toutes les notes soient majorées d'un supplément fixe  $[100 - \text{max}]$  ou que toutes les notes soient multipliées par  $[100/\text{max}]$  (ce qui répartit davantage les notes).

Ainsi, les notes ne traduisent pas nécessairement la performance réelle aux épreuves et plus important, elles peuvent avoir un rapport éloigné avec la compréhension réelle ou effective des notions du cours.