

## **L'« ASSESSMENT » À L'UNIVERSITÉ AUX ÉTATS-UNIS : COMPÉTENCES, ÉVALUATION, AMÉLIORATION**

Christiane Donahue  
Théodile-CIREL (ÉA 4354)  
Dartmouth College, États-Unis

### **INTRODUCTION**

Ce qu'on nomme l'« assessment » aux États-Unis domine le discours universitaire actuel, à la fois pour le rôle de responsabilisation qu'il joue dans les relations entre les diverses parties prenantes, et pour ce qu'il fournit comme modalités d'amélioration de l'enseignement. Dans ce courant de l'« assessment » aux États-Unis, on est à la recherche des signes du savoir et du savoir-faire construits par l'étudiant au cours de ses études. Depuis environ 20 ans, l'« assessment » est au cœur des intérêts et des soucis des institutions, des enseignants, des ministères de l'éducation, des parents, et par conséquent, des élèves et plus récemment des étudiants. L'objectif de cet article sera de présenter la perspective états-unienne concernant cette activité, afin de permettre un travail futur de mise en relation, de comparaison, avec des activités d'évaluation en France. L'intérêt de l'article devrait être de mettre en relief l'assessment, la diversité de façons de le comprendre, et ses enjeux. J'offrirai d'abord une définition de l'assessment et de son fonctionnement, suivi d'une exploration des différences entre « assessment » et recherches, et enfin une description de quelques projets d'assessment à l'université de Dartmouth.

## I. QU'EST-CE QUE L'ASSESSMENT ?

Dartmouth est une université reconnue aux États-Unis pour sa focalisation sur l'excellence de l'enseignement au premier cycle et en même temps pour les recherches. Il s'agit donc d'un établissement qui souhaite mener des enquêtes concernant l'étendue et la profondeur de l'apprentissage des étudiants et l'efficacité des enseignements – l'assessment – comme activité naturelle pour ses enseignants. Ce type d'enquête et d'analyse empirique des données est partie intégrante du processus d'apprentissage et d'instruction universitaire.

En voici une définition :

Un processus continu qui vise soit l'amélioration des programmes et des services, soit la description et l'amélioration de l'apprentissage des étudiants (le plus souvent, les deux en interaction). L'assessment est fondé sur l'idée que les enseignants ont besoin de meilleures informations sur l'apprentissage des étudiants, mais également que les programmes, voire les institutions, ont besoin d'assurer la qualité des enseignements par le biais du progrès des apprenants. (Betsinger 2014).

Autrement dit : l'assessment développe une approche systématique pour comprendre ce que les étudiants apprennent et comment, mais à un niveau local, ciblé, avec une applicabilité relativement immédiate, par le biais d'une collecte de données et avec l'objectif explicite d'améliorer le succès des étudiants, d'un curriculum, d'un programme, de l'efficacité d'une nouvelle initiative, voire même d'une institution dans son ensemble, en impliquant les enseignants dans l'analyse des données et la réflexion. L'assessment est souvent associé à une « défense » d'un enseignement, d'un curriculum, d'un établissement.

Dans ce processus, l'assessment n'est pas « la notation » mais l'exploration de questions spécifiques telles « qu'est-ce que les étudiants apprennent ? Y a-t-il des tendances partagées dans les difficultés qu'ils rencontrent ? Comment les étudiants apprennent-ils ? Quelles sont leurs expériences lors de la construction des savoirs dans nos cours ? Comment tissent-ils les liens entre l'apprentissage dans différents cours ? »

Voici des exemples-types d'objets de l'assessment :

– l'apprentissage des étudiants, à court et à long terme, dans un cours donné, dans une filière spécifique, ou dans « l'éducation générale » (des cours obligatoires en dehors de la filière d'études de l'étudiant, censés développer une culture générale qui nourrit la perspective interdisciplinaire de l'étudiant) ;

– l'apprentissage « équitable » : est-ce que tous les étudiants ont accès égal à l'apprentissage, en fonction de leurs styles d'apprentissage, de leurs ressources, de leurs « passés éducatifs » ? ;

– la question de la « persistance » : quels étudiants « persistent » (ou persévèrent) mieux suite à tel enseignement, à telle intervention ?

– la satisfaction des étudiants, et les attitudes des étudiants envers leurs profils d'apprentis ;

– l'efficacité des initiatives spécifiques ou des innovations (par exemple, un travail en TD qui cible un point faible, un cours pour étudiants qui montrent des lacunes, une réforme de contenu qui prend en compte de nouvelles découvertes scientifiques, etc.) ;

– les effets des outils technologiques dans l’enseignement ou l’apprentissage.

## **Les objectifs de l’assessment**

Au cours de l’assessment, on recueille des données de diverses sources afin de comprendre ce que les étudiants savent, comprennent, et peuvent faire avec leurs savoirs suite aux expériences éducatives ; ce processus s’achève quand les résultats servent pour étayer l’étape suivante d’apprentissage. On organise l’assessment à trois échelons, mais qui peuvent s’entretenir :

### *Niveau institutionnel*

À ce niveau, il s’agit d’un recueil de données afin de produire des informations utiles pour les décisions et le planning stratégique institutionnel au sens large. Ce type d’assessment peut être aussi simple qu’une observation de la fréquence d’utilisation des services de la bibliothèque afin de redistribuer les ressources, ou aussi compliqué qu’une analyse de l’université du point de vue des taux de réussite, de ce qu’on appelle « persistance », par type de population, afin de mettre en place des soutiens.

### *Niveau curriculum et programmes*

Ce niveau est lié aux analyses des cursus et réformes : est-ce que tous les étudiants atteignent les objectifs des programmes et des départements ?

Exemple : Un département de Neurosciences établit des compétences que doit acquérir tout étudiant au cours du programme. Les enseignants identifient quelles compétences sont liées à quels cours (à noter que plusieurs cours peuvent viser une même compétence). Ils notent s’il y a des compétences qui n’apparaissent pas. Ils questionnent les étudiants et les anciens concernant l’enseignement et les possibilités de développer les compétences en question, et ils comparent les résultats à ceux des établissements pairs. Suite à une présentation au département de ce qu’ils ont trouvé, ils modifient les maquettes et les cours afin de mieux offrir ce qui est nécessaire.

### *Niveau étudiants et cours*

À ce niveau, on se demande si l’étudiant est équipé pour s’autoévaluer (au-delà d’une tâche particulière – en continu – par rapport aux objectifs et compétences) ; si l’enseignant est mieux équipé pour comprendre l’apprentissage des étudiants et pour améliorer/ajuster son enseignement.

Exemple (un peu simple) : un enseignant souhaite redévelopper son cours. Il a constaté que les étudiants n’apprennent pas tout ce qu’il enseigne. Il identifie les compétences qu’il souhaite développer, utilise des examens qui s’alignent avec ces compétences (par exemple : « L’étudiant peut utiliser de façon stratégique les règles de... afin de résoudre des problèmes. »). Il analyse les examens afin de cerner les problèmes communs et révisé son enseignement en fonction de ces domaines de difficulté.

Récemment l’intérêt s’est tout particulièrement porté sur la question du « transfert » des savoirs et des connaissances – l’assessment de ce que les étudiants

arrivent à réutiliser de leurs savoirs dans un nouveau contexte. Mais l'assessment du transfert est très difficile : il est difficile à identifier si le savoir se transforme, surtout si une partie significative de ce qu'on cherche n'est pas directement observable.

## Les méthodes et les processus en jeu

Afin d'atteindre les objectifs de l'assessment, il faut :

- expliciter les objectifs des programmes et les compétences visées ;
- mettre en place des attentes exigeantes et des critères appropriés ; puis, de manière systématique, recueillir, analyser, et interpréter des données afin d'identifier le niveau de succès des programmes et des services qui visent à assurer la réalisation de ces objectifs et l'acquisition de ces compétences ;
- utiliser les résultats pour documenter, expliquer, et améliorer les programmes, les services, et les compétences d'étudiants.

Ceci est nommé « closing the loop » (« faire la boucle »).

La première étape d'un « assessment » est ainsi d'établir des « learning outcomes » (compétences d'apprentissage). Ces compétences sont ciblées par des enseignements censés les développer chez les étudiants. Pour un travail d'assessment, une compétence est une « définition de ce qui sera différent chez un étudiant (un groupe d'étudiants) après une expérience éducative », à savoir les « connaissances, capacités, stratégies, attitudes, habitudes, ou le savoir-faire, qu'un étudiant aura acquis ».

Au niveau institutionnel, ces compétences peuvent être très abstraites – « terminer ses études » ; « utiliser les ressources disponibles » ; « trouver un emploi dans les six mois après la fin des études ».

La notion de compétences est bien sûr connue en France ; « l'approche par compétences » aussi<sup>1</sup>. Mais l'approche par compétences est à différencier de la « pédagogie par objectifs », plus ancienne et très critiquée. Cette dernière découpe ce qui doit être enseigné en unités décontextualisées afin de les introduire et de vérifier leur acquisition ; « des contenus complexes sont progressivement réduits en unités de plus en plus simples de contenu d'enseignement » (Toumi, 2006, p. 1), et la transmission du savoir par l'enseignant est au centre. L'approche par compétences se centre davantage sur les actions de l'apprenant dans des situations d'apprentissage.

L'intérêt de l'assessment est de cerner la maîtrise des compétences chez les étudiants, mais, comme je l'ai dit plus tôt, pas uniquement au niveau de l'individu ; également au niveau du cours, du curriculum, du programme, de l'établissement. On ne peut pas faire l'assessment aux niveaux plus larges que celui de l'étudiant individuel sans commencer par lui mais dans l'assessment ce n'est pas l'individu en tant que tel qui « compte ». Pour les éducateurs aux États-Unis, c'est l'individu, son travail, qui est évalué, en fonction des normes et des critères utilisés pour noter.

---

1. La notion d'approche par compétences est très développée dans le domaine de la médecine, tant en France qu'aux États-Unis.

Mais pour un enseignant, ce n'est pas toujours évident de cerner la différence entre assessment de l'individu et assessment du curriculum, programme, établissement.

Les activités de l'assessment sont classées en « directes » et « indirectes » : une activité directe analyse un objet qui porte les traces de l'apprentissage alors qu'une activité indirecte essaie de cerner l'apprentissage en posant des questions, par exemple, à l'étudiant. Pour le recueil de données pour un assessment « direct », on peut viser des analyses de textes, des observations, des examens, des évaluations cliniques. Pour un assessment « indirect » les données seraient des entretiens, des questionnaires, des « focus groups », des textes de métaréflexion. La plupart des assessments se basent sur des formes indirectes pour sonder les expériences et les perceptions d'apprentissage des étudiants. L'assessment direct par l'analyse des productions écrites de ces étudiants est effectué moins souvent, et quand cela est fait, les formes des productions écrites et leurs contextes de production (sur table, à la maison, avec révisions ou non, etc.) sont objets de contestation. L'objectif pour les deux types est de confirmer si les compétences, stratégies, et savoirs désignés ont été maîtrisés dans le contexte en question. Il existe un tout autre débat concernant la possibilité de compétences « a-contextuelles », c'est-à-dire qui ne sont pas liées aux contenus spécifiques. Pour l'écrit universitaire, par exemple, on n'imagine pas de compétences a-contextuelles ; l'écrit est toujours produit dans un contexte et les savoirs nécessaires à sa production – conventions, règles, genres – sont contextuels. En revanche des stratégies rhétoriques et des savoir-faire sont éventuellement utilisables à travers de multiples contextes.

### **L'usage des résultats**

Les résultats sont utilisés afin de construire de nouvelles expériences pour les étudiants et d'améliorer leurs performances.

## **II. QUELLE RELATION/DIFFÉRENCE AVEC LES RECHERCHES ?**

Beaucoup de ce que je viens de décrire peut ressembler bien sûr aux recherches autour de l'enseignement et de l'apprentissage, en particulier au niveau des méthodes de recueil de données. Il est vrai qu'« assessment » et recherches partagent méthodologies et objets d'analyse. L'activité de recherches et l'activité d'assessment peuvent donc se ressembler. Ils ont des méthodes de recueil de données en commun : les « focus groups », les entretiens, les analyses de textes, par exemple. Ils partagent également les questions de validité d'outils, de fiabilité, de triangulation.

Cependant, bien qu'apparentées aux recherches, les questions de l'« assessment » n'en sont pas l'équivalent. L'assessment utilise des données empiriques (qualitatives et quantitatives) concernant l'apprentissage spécifiquement – et uniquement – afin de peaufiner les enseignements et d'améliorer les apprentissages, d'effectuer une « boucle » pédagogique et programmatique.

On voit bien les contrastes à partir de plusieurs points de repères :

– les profils des chercheurs par rapport aux « assessors » professionnels ne sont pas semblables ;

– le processus de « revue institutionnelle<sup>2</sup> » est nécessaire pour la recherche et non pour l'assessment ;

– les littératures que l'on doit lire (ou non) avant de procéder ne sont pas les mêmes (et en fait une « revue de la question » n'est pas attendue pour l'assessment) ;

– la nature et l'utilisation des résultats sont différentes, y compris comment et où les résultats sont publiés (ou non), rendus visibles au public (ou non) ; expertisés (ou non) ; repris par d'autres chercheurs (ou non).

Par exemple, un projet d'assessment ne s'attachera pas à une tradition existante de projets précédents, et même encore moins, à un cadre théorique particulier. Il est rare de choisir un paradigme pour un assessment, ou de revoir les méthodologies pertinentes avec un regard sur les utilisations précédentes. Pour l'assessment, donc, il n'y a pas de revue de la question, pas d'encadrement « scientifique », un travail sans théorisation (ou, plutôt, une théorie de l'apprentissage implicite), un objectif d'application. Surtout : il s'agit d'un travail local, entièrement pertinent pour un contexte particulier.

Dans la tradition éducative telle que je la connais, les recherches concernant l'enseignement et l'apprentissage universitaires cherchent à construire de façon cumulative un corps de savoirs, tout en résistant à l'applicationnisme, à l'application des résultats d'une seule analyse à une situation pédagogique générale. Cette résistance est une protection valable contre les soucis exprimés dans diverses disciplines concernant les recherches empiriques. En effet, des dommages pédagogiques peuvent se produire quand les résultats de recherches deviennent trop vite la fondation d'une nouvelle approche pédagogique. Les recherches, dans ce modèle des sciences sociales, sont empiriques, fondées sur des données. Elles utilisent de multiples formes de méthodologies quantitatives et qualitatives. Ce travail construit un corps de savoirs qui promeut la pensée, la théorie, et la construction des paradigmes d'une discipline.

Aux États-Unis, un appel spécifique à l'égard de l'évolution de ce type de recherches a été fait concernant l'écrit universitaire. Haswell (2005) spécifie ce besoin de produire des recherches « reproductibles, agrégeables, soutenues par des données » – un modèle bien connu en Europe, en sciences de l'éducation par exemple. Ce n'est qu'un modèle, mais ce qui intéresse les chercheurs en « writing studies » (champ des recherches sur l'écriture universitaire) aux États-Unis, c'est justement l'apport du modèle pour la création d'un corps de savoirs ouvert à la possibilité d'être reproduit, poussé plus loin, approfondi par de futures recherches, vérifié ou nuancé par de nouveaux travaux autour de la même question.

Il y a une valeur à faire des recherches qui ne sont pas destinées à l'application :

– la nature libératoire du fait qu'on sait que les résultats ne seront pas appliqués, et qu'une analyse très pointue a sa logique ;

---

2. IRB : Institutional Review Board est un processus d'expertise exigé aux États-Unis pour tout projet de recherches ; un comité analyse le projet afin de vérifier qu'il ne peut avoir d'effets négatifs pour les participants.

– l’opportunité de tester, de vérifier, de revérifier, de critiquer (de façon publique, disponible, parallèle) ;

– la profondeur permise par le fait qu’on n’a pas besoin de modifier quelque chose suite à un projet.

Dans ce cadre, on peut commencer à préciser nos discours. Les recherches dans les traditions des sciences sociales que je viens de décrire se tiennent à côté d’autres types de recherches : la recherche-action, par exemple, la recherche « humaniste, ou, ce qu’on appelle la « recherche institutionnelle » qui se rapproche justement de l’assessment.

Un exemple qui permet de bien voir la distinction recherches-assessment, est de regarder le modèle de « valeur ajoutée » soutenu par tant de responsables d’assessment. Bien sûr, tout établissement souhaiterait savoir comment il ajoute de la valeur aux vies, aux savoirs, aux expériences de ses membres. Mais, introduire un seul test en première et en dernière année des études universitaires afin de déterminer cette valeur ajoutée est plein de problèmes, en dehors des problèmes liés à tel ou tel test. Les problèmes en question sont traités amplement dans la littérature (cf. Ravitch 2015, Berliner 2014). Ils portent principalement sur l’impossibilité d’identifier ce qui aurait pu, spécifiquement, influencer le succès d’un étudiant en dernière année :

– pas de groupe contrôle ;  
– pas d’informations concernant les éventuelles sources des évolutions ;  
– pas de contrôle concernant la carrière de l’étudiant, par exemple, s’il a connu plus d’une université ;

– pas de prise en compte des profils des participants, de leurs motivations ;  
– et enfin, une relation assez identifiable entre le capital qui accompagne un étudiant à son entrée et sa performance dans le test.

Ainsi, l’on pourrait dire que le modèle « valeur ajoutée » est plutôt celui d’une « valeur accompagnée ».

### **III. EXEMPLES DE PROJETS D’« ASSESSMENT »**

Je vais maintenant développer quelques exemples d’assessment de l’apprentissage universitaire : trois projets avec leurs méthodes et leurs résultats, afin de mettre en relief les caractéristiques typiques de l’assessment. Les projets présentés illustrent, pour le premier le niveau pédagogique, pour les deux autres, plutôt le niveau curriculaire. Afin de situer les exemples, il faut savoir concernant l’éducation supérieure aux États-Unis qu’il s’agit de quatre années d’études en premier cycle, avec un appui sur une filière/discipline spécifique mais aussi un appui, souvent égal du point de vue des horaires de cours, sur une « éducation générale » – des matières en dehors du cursus de la discipline. Donc un étudiant en histoire va suivre des cours en biologie, en philosophie, en sciences politiques... L’assessment peut évoluer au sein des cours ou des cursus d’éducation générale ou disciplinaire. Les trois projets décrits ici, situés à Dartmouth College aux États-Unis, relèvent des contextes disciplinaires. Le premier concerne un cours d’initiation aux sciences de biologie et de chimie pour les étudiants qui souhaitent poursuivre une filière scientifique ; le deuxième, concernant un programme en sciences politiques ;

le troisième, un programme de première année focalisé sur l'initiation à l'écrit universitaire.

## **Niveau cours**

Tout étudiant qui souhaite poursuivre un diplôme en sciences naturelles doit passer par, soit le chemin traditionnel des cours de Biologie 11 et de Chimie 5, soit le chemin innovateur des cours de Biologie 8 et Chimie 9. Les professeurs ont analysé le taux de persistance en Biologie 11 et en Chimie 5 et ont trouvé qu'il était assez bas, notamment pour certaines populations : étudiants dits « première génération » (pour qui les parents et les grands-parents ne sont pas allés à l'université), étudiants minoritaires (noirs, hispaniques...). Les cours de Biologie 8 et de Chimie 9 sont enseignés en collaboration par une équipe de deux professeurs. Ceux-ci ont conçu, à base de recherches en pédagogie, un double cours biologie-chimie qui visait les intérêts et l'engagement des étudiants, avec centralement une introduction à la chimie illustrée par des exemples biologiques pertinents, et avec une partie laboratoire destinée à rendre plus concrets les travaux. L'équipe a suivi les taux de persistance en comparant le chemin traditionnel (11-5) au chemin innovateur (8-9), et ont pu établir une différence statistiquement significative entre les deux expériences. En revanche, cette différence ne tenait pas jusqu'au choix définitif de la filière scientifique, effectué par les étudiants en deuxième année. L'équipe est actuellement en train d'étudier cette deuxième question. Ce projet d'assessment part du principe que le cours en question est la pierre angulaire pour une préparation aux sciences. En identifiant le problème, un manque de persistance de la part des étudiants peu préparés, les professeurs mettent en marche un processus d'exploration des alternatives qui peuvent mener à de meilleurs résultats.

## **Niveau programme, dans un département**

Ce projet est en effet un projet de « pré- » assessment : les activités sont ce qui est nécessaire afin de monter un cycle d'assessment ensuite. Le département des sciences politiques à Dartmouth a été invité à analyser son programme afin de comprendre son fonctionnement. Aidé par l'Office de Recherches Institutionnelles et le Centre de Soutien pour l'Enseignement Universitaire de Dartmouth, la faculté a entrepris une analyse du contenu des programmes annoncés aux étudiants, y compris comment les consignes, les devoirs, les examens, les évaluations, et les critères d'évaluation ont été présentés. Cette première analyse a engendré une discussion approfondie concernant les compétences qu'on souhaitait développer chez les étudiants et l'adaptation des cours et des cursus afin de mieux soutenir ce développement. Les programmes et les structures des devoirs ont été ensuite modifiés, certains des professeurs ont adopté des « rubrics<sup>3</sup> », et, maintenant le département peut viser une analyse des effets de ces modifications sur l'apprentissage et les résultats.

---

3. Des grilles d'évaluation liées aux compétences.



## Niveau programme, éducation générale : e-portfolio

Le projet e-portfolio<sup>4</sup> qui sera décrit ici met en lien production textuelle et travail « méta » de réflexion critique autour de cette production et de l'expérience d'apprentissage-initiation plus globalement.

L'université demande aux étudiants de première année, de suivre obligatoirement deux cours d'initiation à l'écrit universitaire de suite (10 semaines chacun). Les compétences visées incluent par exemple la capacité d'énoncer une assertion, étayée, dans un texte bien structuré, clairement exprimé, et intégré dans une discussion plus large (par la lecture d'autres sources).

Nous avons pu constater au fil des années que les étudiants vivent ces deux cours, censés être liés/progressifs, comme des expériences autonomes, et ne voient pas comment « transférer » ou transformer leurs compétences d'un cours à l'autre. Dans le projet, 70 étudiants volontaires soumettent un « portfolio » de leurs écrits, avec des questionnaires qui demandent une réflexion, à la fin du premier et du deuxième trimestre.

Les questions du questionnaire sont du genre :

- Quel texte représente votre meilleur travail ? Pourquoi ?
- Qu'est-ce que vous avez appris qui va vous servir dans le deuxième cours ?
- Quels conseils pour les étudiants qui vont prendre ces cours l'année prochaine ?

Ces réflexions et les textes écrits dans les cours sont lus par un groupe d'enseignants-chercheurs qui analysent les énoncés dans les réflexions ainsi que les évolutions dans les écrits, en fonction des compétences que le cours est censé développer chez l'étudiant. Les observations servent à évaluer le programme afin d'améliorer l'enseignement et la réussite des étudiants.

## CONCLUSION

Ce petit exposé a souhaité dessiner les grandes lignes de ce que nous appelons aux États-Unis l'« assessment ». Cette forme d'assessment est spécifique à la culture d'éducation états-unienne. En Angleterre, par exemple, le mot désigne l'activité de notation. L'assessment opère à tous niveaux d'éducation, influence énormément toute institution aujourd'hui, et met en question une longue tradition universitaire qui consiste à faire confiance aux professeurs et à ne pas explorer de près les effets des enseignements dispensés. Les résistances à l'assessment sont ainsi souvent assez fortes, mais l'évolution de l'éducation supérieure aux États-Unis dans le 21<sup>e</sup> siècle, et notamment ses couts et les questions concernant sa valeur, suggèrent que nous ne sommes qu'au début d'un nouveau mouvement universitaire.

---

4. E-portfolio = un assemblage de travaux de l'étudiant, sauvegardé sur un site web, avec des questionnaires et une sélection par l'étudiant de ses meilleurs textes.

## RÉFÉRENCES

- Berliner, David. 2014. « Exogenous Variables and Value-Added Assessments : A Fatal Flaw ». *Teachers College Record* vol. 116 number 1.  
URL : <http://www.tcrecord.org/Content.asp?ContentID=17293>
- Betsinger, Alicia. 2014. « What Is Assessment ? ». Document interne, Université de Dartmouth.
- Haswell, Richard. 2005. « NCTE/CCCC's Recent War on Scholarship ». *Written Communication* vol. 22 n° 2, p. 198-223.
- Ravitch, Diane. 2015. « The Problems With Value-Added Assessment ». *Education Week*. URL : [http://blogs.edweek.org/edweek/Bridging-Differences/2010/10/dear\\_deborah\\_you\\_asked\\_what.html](http://blogs.edweek.org/edweek/Bridging-Differences/2010/10/dear_deborah_you_asked_what.html)
- Toumi, A. 2006. *L'approche par compétences*. Un nouveau regard sur l'élaboration des curricula et la planification des apprentissages. Oujda : Édition Hilal Impression.