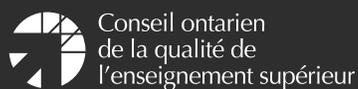




Évaluation des résultats d'apprentissage : MANUEL DU PRATICIEN



Auteurs : Lori Goff (McMaster University), Michael K. Potter (University of Windsor), Eleanor Pierre (EJPCommunications & Mohawk College), Thomas Carey (Transforming Learning Together/San Diego State University/Kwantlen Polytechnic University), Amy Gullage (McMaster University), Erika Kustra (University of Windsor), Rebecca Lee (McMaster University), Valerie Lopes (Seneca College), Leslie Marshall (Mohawk College), Lynn Martin (McMaster University), Jessica Raffoul (University of Windsor), Abeer Siddiqui (McMaster University), Greg Van Gastel (McMaster University)

Remerciements :

Nous tenons à souligner les contributions essentielles de Peter Marval, de l'Université de Windsor, et des conseillers internationaux Gordon Joughin, Arshad Ahmad et Lesley-Jane Eales-Reynolds. La création de ce manuel a bénéficié du soutien du Conseil ontarien de la qualité de l'enseignement supérieur (COQES).

TABLE DES MATIÈRES

PRÉFACE	4
----------------	----------

PARTIE 1 : VUE D'ENSEMBLE ET CADRE	7
-------------------------------------------	----------

1.1 Évaluation des résultats d'apprentissage au niveau d'un programme	8
1.2 Concepts et principes	13
1.3 Création d'un plan pour l'évaluation des résultats d'apprentissage au niveau d'un programme	20

PARTIE 2 : PRATIQUES D'ÉVALUATION	29
------------------------------------------	-----------

2.1 Pensée critique, résolution de problèmes, jugement et perspicacité	31
2.2 Travaux de recherche et d'érudition	34
2.3 Communication	36
2.4 Créativité et conception	39
2.5 Autonomie et compétence professionnelle	40

PARTIE 3 : DÉVELOPPER LA CAPACITÉ DE L'ÉTABLISSEMENT	43
-------------------------------------------------------------	-----------

3.1 Modifier la culture de l'établissement et accroître le recours aux pratiques d'évaluation authentique	44
3.2 Évaluation authentique	53
3.3 Tendances futures	57

RÉSUMÉ	60
---------------	-----------

TRAVAUX CITÉS	61
----------------------	-----------

AUTRES RÉFÉRENCES	64
--------------------------	-----------

PRÉFACE

Ce manuel se veut un outil de référence à l'usage du corps professoral, des membres du personnel, des hauts dirigeants universitaires et des pédagogues qui prennent part à la conception ou à l'examen des programmes et des cours, ainsi qu'à l'évaluation des résultats d'apprentissage au niveau d'un programme aux fins de l'amélioration des programmes. En évaluant les résultats d'apprentissage au niveau d'un programme, on peut plus facilement apporter des améliorations aux plans de cours, d'enseignement et d'évaluation.

COMMENT UTILISER CE MANUEL

Ce manuel est conçu de telle manière que vous pouvez soit passer à une section en particulier soit lire les parties dans l'ordre.

Les définitions, exemples, cas et recommandations que renferme ce document visent à vous faciliter l'élaboration d'outils d'évaluation efficaces des résultats d'apprentissage au niveau d'un programme, mais ces outils devront être évalués et adaptés au contexte propre à votre établissement.

Dans la Partie 1, nous explorons la théorie, les principes, les raisons d'être et les méthodes qui touchent la production de résultats d'apprentissage au niveau d'un programme.

La Partie 2 aborde une variété de pratiques et nouveautés en évaluation des résultats d'apprentissage.

Dans la Partie 3, nous fournissons des conseils pratiques et des techniques pour accroître la capacité de l'établissement par l'instauration d'une culture d'établissement, par l'accroissement de la participation du corps professoral à l'évaluation des résultats d'apprentissage au niveau d'un programme et par l'examen des méthodes d'évaluation qui font partie intrinsèque des programmes d'enseignement.

En vous servant de ce guide pour élaborer un plan d'évaluation des résultats d'apprentissage au niveau d'un programme, vous serez mieux à même d'orienter la collecte systématique de données et de vous en servir pour améliorer constamment l'efficacité de vos programmes et démontrer aux autres le degré d'apprentissage de vos élèves.



Définitions et exemples :
définition des principaux termes.



Études de cas :
exemples d'applications et d'idées.



Recommandations :
exemples de questions et de suggestions.

À l'international, l'évaluation des résultats d'apprentissage au niveau d'un programme d'enseignement a été étudiée comme méthode d'évaluation de la qualité et d'amélioration continue de la qualité des programmes. D'après un rapport de l'UNESCO (Altbach, Reisberg et Rumbley, 2009), l'accroissement de l'intégration et de l'échange tant d'élèves que d'enseignants au niveau mondial a été un important objectif international du milieu de l'enseignement supérieur ces dernières années. Cette tendance requiert des établissements qu'ils définissent des normes de qualité, ce qui se traduit par un accent mis davantage tant sur les résultats d'apprentissage que sur les preuves de cet apprentissage tirées des évaluations de cours, pour démontrer que les élèves ont fait l'apprentissage escompté. Si l'on s'entend généralement pour dire que l'autonomie des établissements est importante, les chercheurs s'entendent aussi sur le fait qu'une plus grande clarté au sujet des résultats d'apprentissage serait profitable aux élèves (Harris, 2009; Tremblay, Lalancette et Roseveare, 2012). Le processus de « Tuning » (établissement de points de référence, de convergence et de compréhension commune entre les programmes d'enseignement) de Bologne¹, auquel participent plus de 40 pays, a grandement contribué aux discussions internationales sur la qualité, les résultats d'apprentissage et les autres processus en vue de favoriser la transparence, la mobilité et l'employabilité des étudiants (Altbach et al., 2009; Barrie, Hughes, Crisp et Bennison, 2011). D'autres mouvements se sont employés à définir et évaluer les résultats d'apprentissage à des niveaux globaux et entre les disciplines; par exemple, les Dublin Descriptors définissent les résultats d'apprentissage à l'échelle du secteur européen de l'enseignement supérieur (Harris, 2009).

Récemment, l'attention s'est aussi tournée vers les résultats d'apprentissage au niveau d'une discipline. Par exemple, la United Kingdom Quality Assurance Agency (agence d'assurance de la qualité du Royaume-Uni) a amorcé un processus de définition des résultats d'apprentissage davantage axée sur une discipline (ce processus est appelé Subject Benchmark Statements) dans lequel ces résultats sont évalués par des examinateurs externes. Le processus de « Tuning » de Bologne (1999), les projets latino-américains de « Tuning » (2004) et Tuning USA (2009) ont travaillé à la définition des résultats d'apprentissage au niveau de la discipline (Harris, 2009; Barrie et al., 2011). De la même façon, l'initiative Valid Assessment of Learning in Undergraduate Education (VALUE – évaluation valide de l'apprentissage dans les études

de premier cycle) était centrée sur l'évaluation par la production de grilles de notation aux É. U. (AACU, 2009). Le gouvernement australien joue un rôle actif dans l'assurance de la qualité depuis les années 1980 (Chalmers, 2007), et l'Australian Learning and Teaching Council a facilité la définition des résultats d'apprentissage au niveau des disciplines dans le cadre du projet Learning and Teaching Academic Standards (normes universitaires d'apprentissage et d'enseignement) en 2010. Bien que l'on ait consacré beaucoup d'attention à l'élaboration des résultats d'apprentissage, l'harmonisation des résultats de cours au moyen d'une évaluation des cours et des programmes par les établissements pose des difficultés à l'échelle internationale (Barrie, Hughes et Smith, 2009; Barrie, Hughes, Crisp et Bennison, 2012). L'Organisation de coopération et de développement économiques (OCDE) a lancé un projet international visant à lever cette difficulté. L'étude de faisabilité Assessment of Higher Education Learning Outcomes (AHELO – évaluation des résultats d'apprentissage de l'enseignement supérieur) visait à examiner les résultats obtenus par les élèves à des examens uniformisés que l'on utilisait pour mesurer les résultats d'apprentissage au niveau d'une discipline et entre les disciplines, et ce, dans le but de produire des données qui pourraient guider les établissements dans leurs mesures d'amélioration (Harris, 2009; Tremblay, Lalancette et Roseveare, 2012; Lennon et Jonker, 2014).

Le Canada a également participé à des projets internationaux visant à faire l'essai de tests normalisés/uniformisés afin d'évaluer le degré d'atteinte des résultats d'apprentissage des élèves au niveau d'une discipline. Par exemple, le Conseil ontarien de la qualité de l'enseignement supérieur (COQES) a fait l'essai de la méthode « Collegiate Learning Assessment » avec des étudiants en génie civil (Lennon, 2014). Des préoccupations ont été soulevées quant à la taille de l'échantillon et au biais d'autosélection, notamment les craintes que l'on puisse se servir des données ainsi obtenues pour établir un palmarès et réaffecter des ressources publiques au détriment de certains établissements et que l'information limitée tirée de tests normalisés soit trop simpliste pour qu'on l'utilise dans les contextes complexes des établissements (Tremblay et al., 2012; Lennon, 2014). Bien que des projets internationaux commencent à chercher des méthodes pour évaluer les résultats d'apprentissage au niveau des programmes ou des établissements, l'évaluation demeure une tâche difficile et complexe.

¹ Pour en savoir plus sur le processus de Bologne, consulter la page <http://www.ehea.info/article-details.aspx?ArticleId=5>

À l'instar des tendances observées à l'international, le réseau ontarien de l'enseignement postsecondaire participe depuis de nombreuses années à la définition des résultats d'apprentissage au niveau des programmes, avec de solides chefs de file, en particulier au niveau collégial. En 1990, le gouvernement provincial a publié un document intitulé *Vision 2000: Quality and Opportunity*, dans lequel il recommandait la création d'un conseil ayant pour mandat d'élaborer des normes de programmes pour le réseau collégial (Conseil ontarien des affaires collégiales, 1990; ministère de la Formation et des Collèges et Universités (MFCU), 2006a et 2006b). Les programmes tant collégiaux qu'universitaires utilisent maintenant une série de résultats d'apprentissage au niveau des programmes pour déterminer ce que les étudiants sont censés acquérir jusqu'à l'obtention de leur diplôme. De nombreux établissements en Ontario ont aussi défini des attentes au niveau de l'établissement souvent appelées « caractéristiques des diplômés » (Barrie et al., 2009).

Les programmes collégiaux comportent 1) une norme de formation professionnelle propre au programme, 2) des normes de compétences essentielles à l'employabilité qui sont déterminantes pour réussir au travail, dans la vie quotidienne et au chapitre de l'apprentissage la vie durant et 3) des exigences en matière de formation générale aux fins de l'étendue des connaissances et du développement des citoyens (MFCU, 2006b), autant d'exigences auxquelles les étudiants doivent satisfaire avant de décrocher leur diplôme. Une équipe de programme, à la lumière des conseils fournis par un spécialiste des programmes d'enseignement, définira un ensemble de résultats d'apprentissage au niveau du programme qui suivent la description de programmes du ministère de la Formation et des Collèges et Universités (MFCU) et feront partie du dossier historique du programme. La Commission d'évaluation de la qualité de l'éducation postsecondaire (CEQEP) de l'Ontario s'attend à ce que tous les programmes collégiaux débouchant sur un diplôme respectent des normes au niveau du diplôme. Le Credential Validation Service (CVS – service de validation des attestations/titres de compétences) est un service qui assiste le réseau collégial dans le processus de définition des résultats et des évaluations d'apprentissage en sorte qu'il se développe en conformité avec

la directive exécutoire du ministre, soit le Cadre d'élaboration des programmes d'enseignement. Ce service est une filiale de Collèges Ontario et, avec les programmes d'études approuvés par le bureau des gouverneurs (conseil d'administration) des collèges locaux, il constitue une composante du mécanisme d'autorégulation mis en place pour le réseau collégial (Service de l'assurance de la qualité des collèges de l'Ontario, 2014).

Avec l'instauration du Cadre d'assurance de la qualité (Conseil d'assurance de la qualité des universités de l'Ontario ((CAQUO), 2012) et les attentes au niveau du diplôme (Ontario Council of Academic Vice-Presidents (OCAV – conseil des vice-recteurs aux études de l'Ontario), 2007; Conseil des universités de l'Ontario (CUO), non daté), les universités ontariennes s'attachent désormais à évaluer les résultats d'apprentissage au niveau du programme et à améliorer continuellement l'apprentissage universitaire. Tant les collèges que les universités suivent des processus d'autorégulation. Les universités présentent leurs résultats d'apprentissage au niveau des programmes par l'intermédiaire du CAQUO, et ce, à la suite d'un processus suivi par les établissements qui est conforme au Cadre d'assurance de la qualité (CAQUO, 2012).

L'évaluation des résultats au niveau du programme ou du diplôme fait partie intégrante de l'éducation centrée sur l'apprentissage. Elle représente un mécanisme permanent pour ce qui est de remettre en question les présuppositions au sujet de l'efficacité des programmes, de cerner les éléments contradictoires d'un programme et de garantir l'atteinte des objectifs d'apprentissage des étudiants. Elle permet aussi l'amélioration continue des buts et objectifs des programmes. Le COQES, de concert avec le CUO et le SAQCO (Service de l'assurance de la qualité des collèges de l'Ontario), a vanté l'évaluation des résultats d'apprentissage au niveau des cours et des programmes tant dans les collèges que dans les universités. Si la production de résultats d'apprentissage a été entièrement adoptée dans la plupart des établissements d'enseignement postsecondaire de l'Ontario, il demeure que l'évaluation efficace des résultats au niveau des programmes reste difficile pour nombre d'établissements. Les stratégies que renferme ce manuel sont susceptibles de faciliter la résolution de cette difficulté.

PARTIE 1

VUE D'ENSEMBLE ET CADRE

Résumé

Dans cette partie, nous énonçons les raisons et avantages d'évaluer les résultats d'apprentissage au niveau d'un programme comme moyen d'améliorer et d'enrichir les programmes. Nous décrivons également le fondement théorique de ce manuel, qui souligne l'importance de l'authenticité, de la validité et de la fiabilité des pratiques et principes d'évaluation et décrit l'harmonisation des objectifs des cours aux résultats d'apprentissage plus généraux au niveau d'un programme. En dernier lieu, nous décrivons brièvement la façon dont vous pouvez créer un plan d'évaluation des résultats d'apprentissage au niveau d'un programme. La Partie 2 explore les moyens pratiques d'appliquer ces concepts théoriques.

1.1 ÉVALUATION DES RÉSULTATS D'APPRENTISSAGE AU NIVEAU D'UN PROGRAMME

Les résultats d'apprentissage au niveau d'un programme précisent ce que les étudiants devraient connaître, valoriser ou être capables d'accomplir après avoir terminé leur programme avec succès. Ces résultats sont souvent atteints au moyen d'activités particulières d'apprentissage, lesquelles sont intégrées au cours et visent à un apprentissage général au niveau des programmes.

POURQUOI ÉVALUER LES RÉSULTATS D'APPRENTISSAGE AU NIVEAU D'UN PROGRAMME?

Il peut être de la plus grande utilité d'évaluer intentionnellement vos propres **résultats d'apprentissage au niveau d'un programme**, de même qu'une évaluation efficace, bien planifiée, pour :

- veiller à ce que les étudiants acquièrent les plus importantes compétences, idées, attitudes et valeurs de la discipline ou profession;
- documenter les éléments probants de l'apprentissage des étudiants en fonction des résultats réels qu'ils ont obtenus, aux fins d'attestation et de responsabilisation;
- s'assurer que les attentes sont communiquées clairement aux étudiants et qu'ils les comprennent (y compris ceux qui sont intéressés à s'inscrire à un programme);
- vous permettre d'accroître l'efficacité de votre programme en fonction des résultats réellement obtenus par les étudiants;
- faire valoir la qualité de votre programme; rendre vos diplômés attrayants pour les employeurs et votre programme attrayant pour les étudiants et donateurs éventuels;
- mettre l'accent sur les priorités courantes de l'établissement en matière d'enseignement et d'apprentissage.

Le suivi des données tant quantitatives que qualitatives peut permettre aux établissements de prendre des décisions menant à une instruction améliorée, de plus solides programmes d'enseignement ainsi que des politiques plus efficaces et efficientes en matière d'évaluation des résultats d'apprentissage, avec l'objectif général d'améliorer l'enseignement et l'apprentissage. Cela peut renforcer l'engagement continu chez les membres du corps professoral, les étudiants, le personnel et les administrateurs, qui travaillent à la fois individuellement et en équipes pour concevoir et évaluer des programmes ainsi qu'améliorer l'apprentissage chez les étudiants.

Les établissements peuvent aussi se servir de ces analyses pour faciliter l'exécution de leur mission et renforcer les arguments à l'appui d'un financement accru et/ou de la répartition des ressources. L'information tirée de l'évaluation des résultats d'apprentissage au niveau d'un programme peut aider à satisfaire aux exigences des organismes d'accréditation et de financement, éclairera les diverses décisions axées sur la responsabilisation et témoigne du temps et des ressources consacrés à des stratégies de participation des apprenants qui cadrent avec les résultats d'apprentissage et l'évaluation. Utilisée avec application, l'évaluation des bons résultats d'apprentissage peut guider les établissements dans l'atteinte de leurs buts et justifier leurs déclarations que les diplômés sont bien préparés pour réussir dans leurs entreprises futures.

UN CADRE AUTHENTIQUE ET HARMONISÉ AXÉ SUR LES RÉSULTATS

L'approche fondée sur les résultats est utilisée pour évaluer les résultats d'apprentissage au niveau d'un programme depuis un certain nombre d'années à l'échelle internationale, largement dans le système collégial ontarien et plus récemment dans le système universitaire. Elle met l'accent sur la capacité des étudiants de démontrer ce qu'ils savent, ce à quoi ils accordent de la valeur et ce qu'ils peuvent faire à la fin de leur cours et de leur programme. Ces attentes sont exprimées sous la forme de résultats d'apprentissage, qui sous-tendent programmes et cours (Angelo et Cross, 1993). Les cours s'articulent intentionnellement autour de l'exécution d'un programme cohérent. Il est important que toutes les évaluations soient situées dans le contexte de la discipline ou profession voulue. Si les programmes sont bien planifiés, les nombreuses tâches d'évaluation au niveau du programme seront probablement intégrées et incorporées aux cours. Ainsi, les évaluations de cours peuvent être conçues de manière à démontrer l'atteinte des résultats d'apprentissage au niveau du programme en plus des résultats au niveau d'un cours. Souvent, la conception des cours et programmes est la plus efficace et la plus réussie lorsqu'elle est accomplie par des équipes formées de chargés de cours, de pédagogues et d'administrateurs.

L'une des approches courantes en conception intentionnelle au niveau du programme et au niveau du cours est l'**harmonisation constructive** (Biggs et Tang, 2011). Dans un cours harmonisé de façon constructive, les résultats d'apprentissage, les expériences d'apprentissage et les tâches d'évaluation concourent à un renforcement mutuel de l'atteinte des résultats d'apprentissage au niveau du cours (figure 1). Étant par nature des descriptions du degré de réussite, les résultats d'apprentissage sont liés à l'évaluation, laquelle est la plus efficace lorsqu'elle est effectuée dans le contexte d'une discipline, d'une profession, d'un programme et/ou d'un cours. L'apprentissage axé sur les résultats, dans le cadre d'harmonisation constructive (Biggs et Tang, 2007), aide les concepteurs de cours à observer, mesurer et évaluer les résultats d'apprentissage. Une fois les résultats d'apprentissage et les tâches d'évaluation déterminés, on peut planifier les activités d'apprentissage, les méthodes d'enseignement ainsi que les ressources nécessaires pour obtenir ces résultats et évaluations; à son tour, chaque élément d'un cours soutient l'apprentissage voulu.



Résultats d'apprentissage au niveau d'un programme

Énoncés précisant ce que les étudiants qui réussissent devraient savoir, valoriser ou être capables d'accomplir d'ici la fin d'un programme.

De quelle façon rédige-t-on des résultats d'apprentissage?

Pour décrire un résultat d'apprentissage, commencez par employer un verbe d'action observable (représentant des connaissances, compétences et valeurs), puis enchaînez avec un énoncé précisant l'apprentissage à démontrer (contenu).

Dans le document « The ABCD of Learning Outcomes »¹, il est recommandé de se poser les questions suivantes :

Auditoire : Qui sont les apprenants?

Comportement : Que seront-ils capables de connaître, valoriser ou accomplir?

Condition : Dans quelles conditions ou quel contexte se produira l'apprentissage?

Degré : Qu'est-ce qui sera accompli et à quel niveau?

Exemples :

<http://www.uoguelph.ca/tss/pdfs/What%20is%20a%20Learning%20Outcome%20Handout.pdf>
<http://www1.uwindsor.ca/ctl/system/files/PRIMER-on-Learning-Outcomes.pdf>

¹ Adapté de : <https://www.iusb.edu/weave/Goals%20and%20Objectives%20Powerpoint>



Harmonisation constructive

Création de cours dont les résultats d'apprentissage, les expériences d'apprentissage et les tâches d'évaluation cadrent de façon interne et tendent vers les résultats du programme.

Harmonisation constructive au niveau du cours



FIGURE 1

Dans un **programme** harmonisé de façon constructive, les **cours** sont soigneusement coordonnés pour assurer le développement ou le déroulement continu, depuis l'introduction jusqu'à la maîtrise des résultats d'apprentissage, le tout aboutissant à l'obtention des résultats escomptés au niveau du programme (figure 2).

Harmonisation constructive au niveau du programme

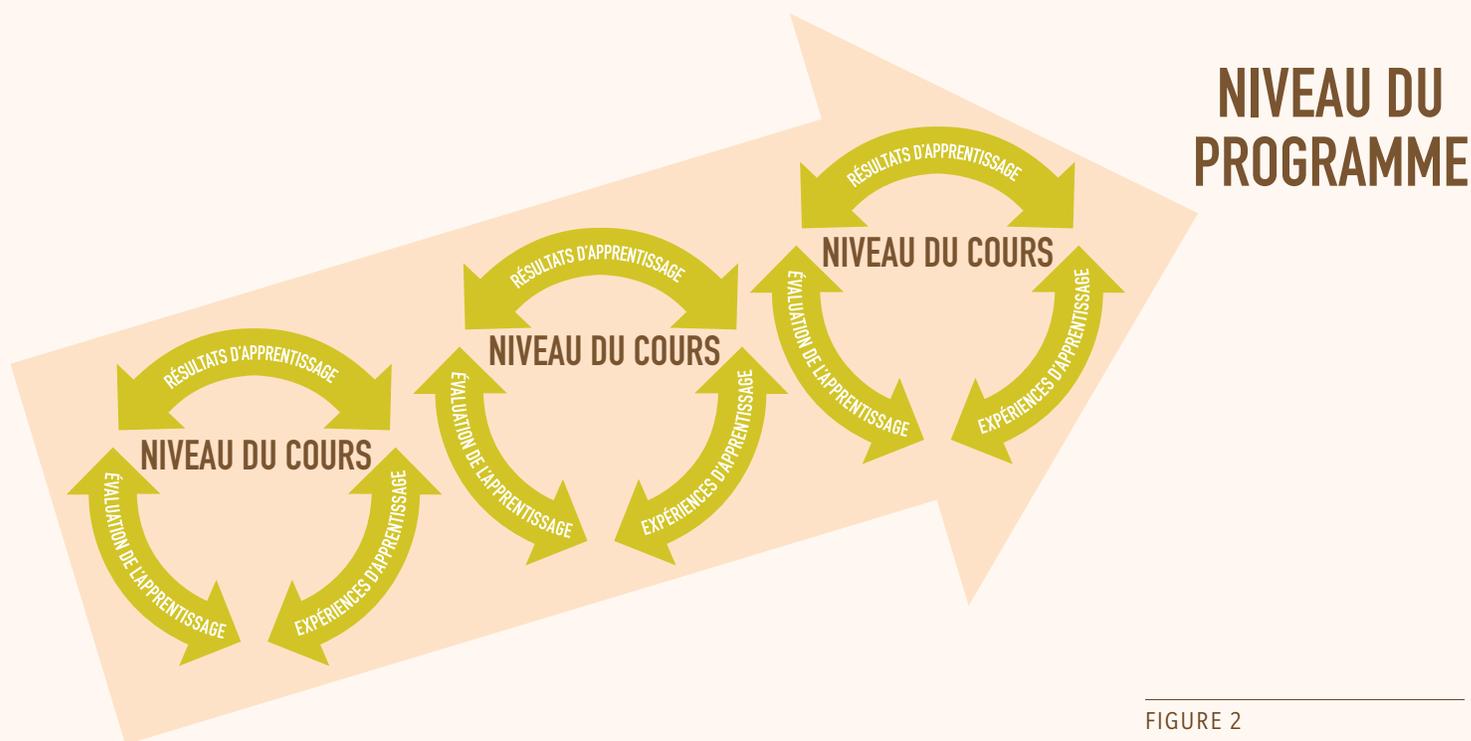


FIGURE 2

Selon Biggs et Tang (2007), on devrait harmoniser tous les aspects de l'enseignement et de l'apprentissage avec les résultats d'apprentissage correspondants au niveau du programme afin de clarifier le point de mire de l'attention. Dans les activités d'évaluation, les étudiants démontrent leur degré d'atteinte des résultats d'apprentissage du cours. Les stratégies d'enseignement, les activités d'apprentissage, les évaluations et les ressources devraient toutes être conçues et organisées de manière à aider les étudiants à atteindre les résultats d'apprentissage au niveau du cours. Wiggins et McTighe (2005) proposent une approche de conception inversée qui facilite cette harmonisation à l'étape de la conception du cours.

Les programmes qui se composent de cours harmonisés de façon constructive contribuent à l'atteinte des résultats d'apprentissage au niveau d'un programme. Cette harmonisation des programmes devrait être appréhendée du point de vue du développement pour permettre aux étudiants de répondre aux attentes exprimées dans les résultats au niveau du programme de la première session à la dernière. Comme le font valoir Biggs et Tang (2007), lorsque cette évaluation est réalisée de cette façon, les résultats d'apprentissage au niveau d'un programme deviennent plus stratégiques et efficaces; ils deviennent une composante organique des stratégies globales d'apprentissage. Les résultats d'apprentissage au niveau d'un programme sont aussi conçus pour refléter le contexte des caractéristiques des diplômés au niveau de l'établissement ainsi que les attentes en matière d'accréditation ou d'agrément au niveau provincial ou de la discipline (figure 3).

Harmonisation constructive aux niveaux de l'établissement et de la province



FIGURE 3

1.2 CONCEPTS ET PRINCIPES

Du fait que les contextes varient beaucoup d'un établissement à l'autre ou selon la faculté ou le département, il est très difficile de formuler des approches uniques qui permettent d'évaluer efficacement les résultats d'apprentissage au niveau d'un programme. Nous croyons qu'en faisant valoir d'importants concepts et en mentionnant une série de principes pouvant être appliqués à des environnements divers, nous fournissons la possibilité d'adopter une approche cohérente qui reste néanmoins souple.

Notre approche en matière d'évaluation des résultats d'apprentissage au niveau d'un programme repose sur quatre concepts sous-jacents, à savoir l'amélioration de la qualité, l'harmonisation constructive, l'évaluation authentique et la taxonomie de la structure du résultat d'apprentissage observé (SRAO). Ces concepts peuvent être considérés comme un fondement pour passer du niveau macro au niveau micro. Le présent manuel repose sur ces concepts et ceux-ci fondent les trois principes qui guident l'évaluation des résultats d'apprentissage au niveau d'un programme dont nous faisons état à la fin de cette partie.

FONDEMENT THÉORIQUE

La notion d'**amélioration de la qualité (au niveau de l'établissement ou niveau macro)** renvoie à des cultures d'établissement qui encouragent une honnête réflexion sur la concentration des politiques et pratiques sur l'apprentissage et leur efficacité pour assurer la réussite des étudiants, cette réflexion devant servir à éclairer les politiques et les pratiques en sorte de favoriser une amélioration de l'apprentissage des étudiants bien raisonnée et reposant sur des éléments probants (Biggs, 2001). Cette approche d'« assurance de la qualité prospective » encourage une amélioration continue qui va au-delà de l'assurance de la qualité et de son « évaluation aux fins de responsabilisation » connexe. L'assurance de la qualité aux fins de responsabilisation a tendance à viser les universités en tant qu'entités commerciales, mais, souvent, fait fi d'une réforme pédagogique significative ou entrave pareille réforme et peut se détourner de la qualité de l'éducation (Biggs, 2001). Nous recommandons une évaluation aux fins de l'amélioration continue de la qualité. L'utilisation intentionnelle de l'information tirée des évaluations au niveau du programme pour améliorer continuellement les programmes en est un exemple. À la fin de la présente section est présenté un exemple précis de la façon dont on peut se servir des évaluations pour améliorer les programmes, dans l'encadré « Étude de cas : Carnegie Mellon University – Évaluation aux fins d'une amélioration intégrée aux processus d'examen des programmes ». Les facteurs qui influent sur la culture et la capacité d'un établissement au chapitre de l'amélioration de la qualité sont étudiés plus avant à la Partie 3.



Amélioration de la qualité

Une culture d'établissement qui encourage une honnête réflexion sur la concentration sur l'apprentissage et qui utilise cette réflexion pour éclairer les politiques et les pratiques en vue d'une amélioration continue.

L'**harmonisation constructive (au niveau du programme ou niveau méso)** assure que les résultats d'apprentissage, les expériences d'apprentissage et les évaluations cadrent, à l'interne, avec un cours et que, de la même façon, les cours soient cohérents entre eux afin de renforcer l'obtention des résultats d'apprentissage au niveau d'un programme ainsi que des caractéristiques des étudiants au niveau de l'établissement ou d'une discipline (figures 1, 2 et 3; Biggs et Tang, 2011). Nous considérons que l'évaluation des résultats au niveau d'un programme est la plus efficace lorsque tout le programme est harmonisé de façon constructive.

L'**évaluation authentique (niveau micro)** est une évaluation qui procure la preuve directe d'une application efficace de l'apprentissage (Angelo, 1999; Maki, 2010). Nous croyons que cette évaluation ne devrait pas être vue comme une fin en soi, mais plutôt considérée comme un moyen de guider les stratégies d'enseignement et d'améliorer l'apprentissage des étudiants. Dans les tâches d'évaluation authentique, il est demandé aux étudiants de mettre à profit les compétences, connaissances, valeurs et attitudes acquises dans des situations reproduisant le mieux possible le contexte de rendement de la discipline ou profession voulue. Lorsqu'on harmonise cette évaluation avec les résultats, le degré auquel les tâches d'évaluation simulent les problèmes du « monde réel » et situations liées aux disciplines ou professions correspondantes permet de mesurer l'authenticité cette évaluation.

La taxonomie de la **structure des résultats d'apprentissage observés (SRAO) (niveau micro)** propose une approche parmi d'autres permettant d'exprimer la complexité de l'apprentissage désiré tandis que les étudiants progressent dans un programme (Biggs et Collis, 1982; Collis et Biggs, 1986). Cela procure une structure utile pour façonner des résultats d'apprentissage convenant à la quantité et la qualité d'apprentissage voulues à des stades particuliers d'un programme. On peut aussi s'en servir pour harmoniser les tâches d'évaluation aux résultats d'apprentissage, évaluer le degré de réussite des étudiants et faire cadrer les résultats d'apprentissage et les tâches d'évaluation avec les stratégies d'enseignement et d'apprentissage. Nous croyons que le cadre de la SRAO fournit une structure pratique et utile qui permet de guider la définition et l'évaluation des résultats d'apprentissage au niveau d'un programme, lesquelles peuvent aider les pédagogues et concepteurs de cours à élaborer intentionnellement des plans de programmes et d'évaluation faisant en sorte que les étudiants puissent décrocher un diplôme au niveau le plus élevé du cadre.

La SRAO représente l'apprentissage à cinq niveaux différents, qui vont du purement quantitatif (acquisition de nouvelles quantités d'information) au qualitatif (changement dans la compréhension et création de signification à partir de l'information) (figure 4). La préstructure est le stade qui précède l'apprentissage. Les stades, qui vont de l'unistructurel au relationnel, s'enchaînent à l'intérieur d'un cycle dans lequel la compréhension de l'étudiant s'accroît et s'approfondit. Les étudiants peuvent devoir passer plusieurs fois par un stade ou un autre du cycle d'apprentissage au fil de la transmission de nouvelles idées, mais, à terme, leur objectif est de quitter le cycle d'apprentissage en atteignant le stade « abstrait étendu ». Les étudiants peuvent en être à différents stades du cycle d'apprentissage selon la matière. Même lorsqu'un étudiant a atteint le degré abstrait étendu de compréhension d'une matière, il peut encore régresser si on lui transmet de nouvelles informations qui secouent ou ébranlent sa compréhension (Potter et Kustra, 2012). On axera plus adéquatement l'évaluation au niveau du programme sur le stade relationnel ou le stade abstrait étendu

Pour expliquer la SRAO, nous utilisons l'exemple de la rédaction d'essais, qui nous est familière à tous, et, dans certaines disciplines, ce type de devoir peut être adapté de manière authentique pour préparer les étudiants aux tâches qu'ils devront accomplir dans le cadre de leurs professions (voir Partie 2).



Évaluation authentique

Une évaluation fournissant la preuve directe d'une application efficace de l'apprentissage, souvent dans le contexte d'une discipline ou pratique professionnelle.

Elle comprend deux éléments clés :

- 1) l'évaluation du rendement : l'apprenant démontre sa capacité d'appliquer les compétences et connaissances requises;*
- 2) le contexte : les tâches et conditions de la discipline ou de la profession sont fidèlement représentées.*

Taxonomie de la SRAO de Bigg

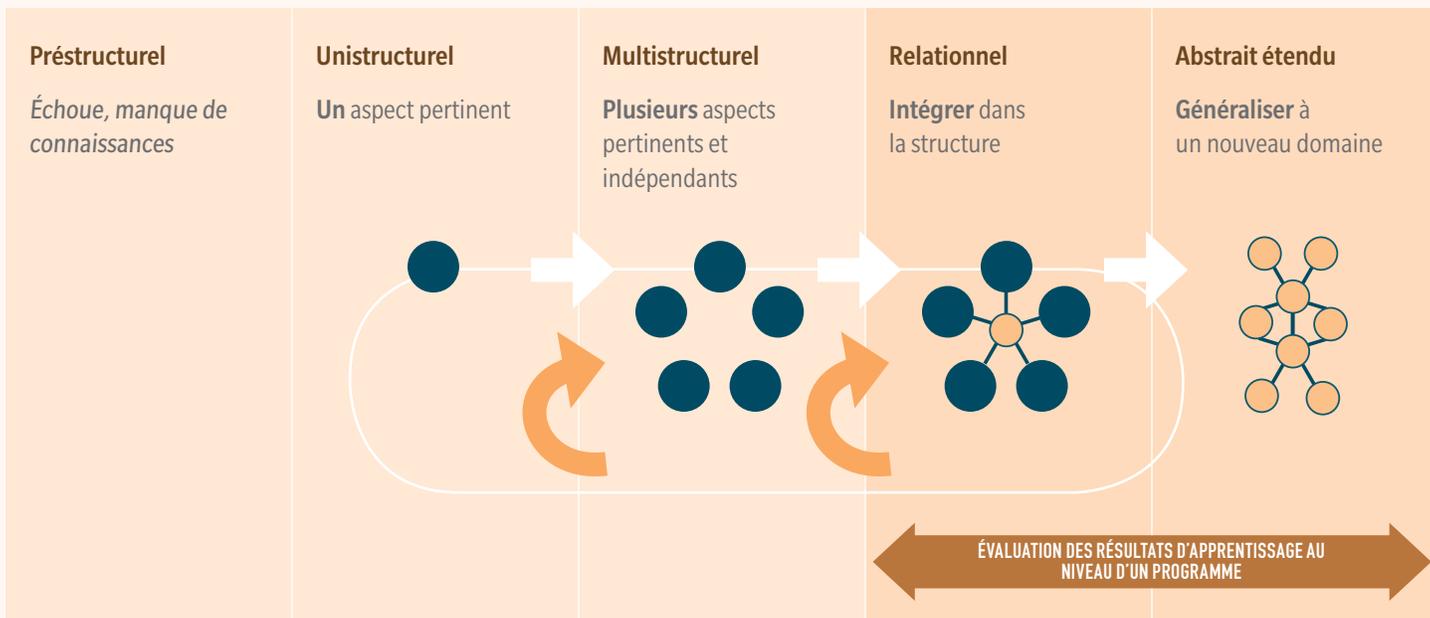


FIGURE 4

Préstructurel

Le stade préstructurel se produit avant le début du cycle d'apprentissage, lorsque l'étudiant ne comprend pas ce qu'on attend de lui (ou le pourquoi de ces attentes). Par exemple, les étudiants ignorent ce qu'est un essai. Ils sont susceptibles de confondre l'essai avec d'autres formes de rédaction, d'attribuer des caractéristiques non pertinentes ou triviales aux essais ou de ne considérer un essai que comme des mots qui ont été employés par ceux qui en font (« les essais sont ce que les professeurs nous disent d'écrire »).

Comme il s'agit du stade auquel l'étudiant est censé commencer un cours ou un programme, les seules tâches d'évaluation au stade préstructurel sont de nature diagnostique, ayant pour objet d'obtenir de l'information sur la nature et l'ampleur de la compréhension ou de l'incompréhension de l'étudiant au début d'un programme. L'évaluation réalisée à ce stade servirait de point de référence pour une comparaison avec l'évaluation des résultats d'apprentissage à la fin d'un programme.

Unistruclurel

Le stade unistruclurel est le premier niveau de ce que Biggs et Collis (1982) appellent le cycle d'apprentissage; il s'agit du stade auquel les étudiants commencent à apprendre. Au stade unistruclurel, les étudiants sont passés de la **non-compréhension** ou mauvaise compréhension des essais à la compréhension d'approches **très** simples. L'apprentissage, à ce stade, est de nature **quantitative** en ce que les étudiants ont ajouté un fait ou une idée à leur banque de connaissances. La compréhension – quelle qu'elle soit – que les étudiants auront des essais sera sans doute une connaissance réductrice. Ils peuvent être capables de faire des rapprochements d'ordre général simples et évidents (« rédiger sous

cette forme équivaut à faire un essai »), mais ils n'ont pas la capacité de créer une signification utile à partir de ces rapprochements.

Parmi les résultats d'apprentissage atteignables à ce niveau, pour employer les verbes familiers de la taxonomie de Bloom, il y a la détermination de ce que sont les essais, la répétition d'une explication de ce que sont les essais et la rédaction d'essais simples de types particuliers.

De nombreuses tâches d'évaluation (souvent inauthentique) standard peuvent servir à déterminer si l'étudiant a atteint les résultats d'apprentissage du niveau unistructurel, par exemple les questionnaires et examens à choix multiples, les questions où il faut compléter les réponses et des devoirs d'essai simples.

Multistructurel

Le stade multistructurel est le deuxième niveau du cycle d'apprentissage; il s'agit du stade auquel les étudiants ont appris de nombreux faits et idées au sujet des essais et/ou peuvent rédiger plusieurs formes d'essais. Toutefois, cette augmentation **quantitative** de l'apprentissage ne se traduit pas par une profondeur ou une complexité accrue dans la façon dont ils comprennent ces faits ou idées les uns par rapport aux autres. Les étudiants apprennent chaque fait ou idée isolément des autres faits ou idées, hormis les liens évidents et simples que l'on peut faire entre ces faits et idées. À ce stade, il y a peu ou pas d'intégration ou d'organisation du savoir.

Les résultats d'apprentissage atteignables à ce niveau comprennent la description des composantes de multiples formes d'essais, le regroupement de ses pensées en paragraphes pour créer un essai et l'exécution d'algorithmes ou de procédures simples.

Bien des tâches standard d'évaluation (souvent inauthentique cependant) peuvent servir à déterminer si l'étudiant a atteint les résultats d'apprentissage du stade multistructurel, comme des questionnaires et examens à choix multiples, les tests de complèment (espaces vides à remplir), les exercices d'application et les devoirs d'essai.

Relationnel

Au stade relationnel, l'étudiant passe d'une augmentation quantitative à une augmentation **qualitative** de l'apprentissage (approfondissement de la compréhension et création de sens plutôt que simple combinaison d'éléments à un ensemble sans rapprochements de ce qu'ils savent). Cet apprentissage dépend de la capacité de l'étudiant de généraliser et extrapoler à partir d'éléments particuliers d'information. À ce stade, les étudiants peuvent déduire des rapprochements (relations) et des distinctions de plus en plus complexes entre les faits, les idées, les théories, les actions et les buts. Ils peuvent forger leurs propres arguments au sujet de ce qu'ils apprennent et peuvent créer du sens et un but à partir de la matière enseignée, intégrer des parties pour créer un portrait d'ensemble et expliciter des rapprochements entre les faits et la théorie, l'action et le but. Au niveau relationnel, la capacité des étudiants d'appliquer de façon compétente ce qu'ils ont appris à des expériences ou problèmes familiers devient évidente. Face à des difficultés d'ordre contextuel, il se peut que des étudiants retournent au stade multistructurel ou unistructurel avant de revenir au stade relationnel. C'est la raison pour laquelle nous utilisons le terme « cycle d'apprentissage » – pour éviter de donner l'impression d'une progression strictement linéaire à travers les stades du SRAO.

Les résultats d'apprentissage atteignables à ce niveau comprennent l'analyse des raisons pour lesquelles les essais sont organisés de certaines façons et pour certains buts, l'établissement d'une relation entre la structure de l'essai et l'intention rhétorique et l'intégration de plusieurs types d'essais pour créer un document servant plusieurs buts. Évaluer si les étudiants ont atteint les résultats d'apprentissage du stade relationnel (ou du stade abstrait étendu) exige des méthodes plus complexes, ciblées et longues que celles utilisées pour l'apprentissage multistructurel, par exemple des projets réalisés pendant toute la durée du cours et des travaux de session à plusieurs étapes. La plupart des cours d'enseignement supérieur devraient commander à tout le moins l'obtention de résultats au stade relationnel; par conséquent, il devrait y avoir plus d'évaluations effectuées à ce stade et aux niveaux supérieurs.

Abstrait étendu

Au stade de l'abstrait étendu, l'augmentation **qualitative** de la compréhension de l'étudiant est difficile à prédire; les étudiants qui atteignent ce niveau sont sortis du cycle d'apprentissage envisagé dans cette taxonomie. Alors qu'au stade relationnel, les étudiants pouvaient théoriser et extrapoler à partir de la matière enseignée, au niveau abstrait étendu, ils théorisent et généralisent **au delà** de la matière du cours, en transférant et appliquant des idées à de nouvelles situations et expériences, en les combinant aux idées dans d'autres cours et à d'autres aspects de leur vie, en effectuant des tests et en émettant des hypothèses qui sortent du cadre d'un cours, et en élaborant des arguments et des théories de leur propre cru.

Naturellement, au chapitre de l'évaluation des résultats d'apprentissage au niveau d'un programme, on devrait s'attacher aux stades relationnel et abstrait étendu de la SRAO. Il ne serait pas approprié de planifier délibérément des tests pour l'évaluation d'un programme aux stades inférieurs, puisque les étudiants qui terminent nos programmes avec succès sont censés comprendre ce qu'ils ont appris et être capables de transférer cette connaissance à des situations nouvelles. Les tâches d'évaluation décrites dans la Partie 2, par conséquent, ne comprennent que les évaluations convenant aux stades relationnel et abstrait étendu de l'apprentissage.

AUTHENTICITÉ, VALIDITÉ ET FIABILITÉ

Dans l'harmonisation de l'évaluation aux résultats, la mesure dans laquelle les tâches d'évaluation simulent des problèmes et situations du monde réel associés à la discipline ou profession voulue représente leur degré d'authenticité. Il apparaît clairement que, intentionnellement ou non, les évaluations influent tant sur la façon dont les professeurs enseignent que sur la façon dont les étudiants apprennent (Boud et collègues, 2010). L'évaluation et l'apprentissage sont intimement liés. Les professeurs sont susceptibles d'« enseigner en fonction des examens », même si le contenu ou la forme de l'examen ne reflète pas ce qu'ils considèrent comme important (Fredericksen, 1984; Prodromou, 1995). En pareil cas, ce que les étudiants apprennent de l'examen dépend en grande partie de ce qu'ils perçoivent comme étant l'exigence de la tâche d'évaluation plutôt que de ce qui pourrait être le résultat d'apprentissage souhaité (Gibbs, 1992). Les principes et pratiques d'une évaluation efficace coïncident avec une évaluation axée sur les résultats d'apprentissage au niveau d'un programme (Barrie et al., 2012).

Ainsi qu'on les a définies plus haut, les tâches d'évaluation authentique demandent aux étudiants d'appliquer ce qu'ils ont acquis (compétences, connaissances, valeurs et attitudes) à des situations reproduisant le contexte de rendement de la discipline ou profession voulue. L'évaluation authentique comprend deux éléments clés : 1) l'évaluation du rendement : l'apprenant démontre sa capacité d'appliquer les compétences et connaissances requises; 2) le contexte : les tâches et conditions de la discipline ou profession sont représentées fidèlement. Collectivement, ces éléments requièrent des étudiants qu'ils utilisent les mêmes connaissances, compétences, jugement et attitudes que ceux qui sont escomptés des praticiens de ces disciplines ou professions. L'authenticité d'une tâche d'évaluation dépend de la mesure dans laquelle elle ressemble à la réalité de la pratique. Les instruments normalisés d'évaluation, comme les tests ou examens normalisés, sont souvent isolés des environnements pratiques, ce qui les rend en grande partie inefficaces pour évaluer la preuve d'apprentissage dans ces contextes (Barrie et al., 2012; Rhodes, 2012). Au lieu de cela, nous recommandons des méthodes d'évaluation authentique qui font partie intégrante du programme d'enseignement et sont propres à la discipline ou au contexte.

On mesure la **validité** des évaluations authentiques en déterminant si elles évaluent ce qu'elles sont censées évaluer, et donc si elles permettent d'atteindre les résultats d'apprentissage voulus. Les résultats d'apprentissage commandent des tâches d'évaluation qui tiennent adéquatement compte du degré d'habiletés et de compétences attendu des étudiants et qui exigent d'eux qu'ils démontrent les processus, approches et attitudes des praticiens du domaine. Les stratégies d'évaluation authentique qui tiennent compte non seulement des tâches mais aussi du contexte des situations réelles ont un haut niveau de validité notionnelle.

La **fiabilité** des évaluations se rapporte à leur cohérence et leur objectivité : deux étudiants ayant acquis de semblables connaissances et compétences obtiendraient-ils un résultat semblable à l'évaluation? On tente fréquemment d'établir la fiabilité des méthodes d'évaluation traditionnelles en soumettant les étudiants à une même évaluation, en même temps et dans un même endroit (souvent inauthentique). Du fait que ces méthodes ne tiennent pas compte de la plupart des variables qui influent sur ce que les étudiants apprennent, s'ils l'apprennent et la mesure dans laquelle ils l'apprennent, leur fiabilité peut nuire à la validité de la tâche d'évaluation.

En mettant l'accent sur les tâches d'évaluation authentique, il se peut qu'on accentue la tension entre fiabilité et validité : plus une évaluation reproduit des situations de la vie réelle, plus elle sera individuelle, située précisément, incertaine, imprévisible et variable. Comment, alors, peut-on concilier fiabilité et authenticité?

Une façon de le faire consiste à accentuer le degré de complexité des simulations (Petti, 2013) et des évaluations, par exemple en utilisant les examens cliniques objectifs structurés (ECOS), de manière que l'on puisse présenter aux étudiants des problèmes stylisés et réalistes d'une façon normalisée et contrôlée (Burn, Nestel et Gachoud, 2013). Une autre façon consiste à mettre l'accent sur la **fiabilité**, la **crédibilité** et l'**authenticité** de l'information que procurent les évaluations dans un contexte qualitatif. Cresswell (2009) a recensé huit procédures qualitatives d'établissement de la fiabilité et de la crédibilité que l'on peut utilement appliquer à l'évaluation. Un exemple consiste à utiliser concurremment diverses méthodes d'évaluation ainsi que l'évaluation par les pairs et l'autoévaluation pour obtenir une **triangulation** des données. En utilisant plusieurs perspectives et sources de données convergentes, on accroît la crédibilité des jugements. Les portfolios et les journaux de réflexion aux niveaux des programmes et des cours fournissent une information utile et détaillée (qu'appuie aussi le projet VALUE de l'AACU, 2009). De plus, comme ces types d'évaluation sont le produit d'un **engagement prolongé** de la part des étudiants, ils fournissent un meilleur contexte et ajoutent à la validité des tâches d'évaluation.

Des outils d'établissement de l'authenticité, de la fiabilité et de la crédibilité des évaluations peuvent aider à obtenir confirmation que les étudiants ont atteint les résultats d'apprentissage désirés et peuvent appliquer, de manière fiable, leurs connaissances et compétences dans le monde réel.



Validité

Mesure dans laquelle une tâche d'évaluation évalue ce qui est censé l'être, en tenant dûment compte du degré de connaissances et de compétences attendu d'un apprenant eu égard aux résultats d'apprentissage au niveau d'un programme.



Fiabilité

L'uniformité et l'objectivité des évaluations. Au niveau d'un programme, on pourrait l'atteindre au moyen d'évaluations multiples et/ou de tâches normalisées et contrôlées d'évaluation authentique.

PRINCIPES DE L'ÉVALUATION DES RÉSULTATS D'APPRENTISSAGE AU NIVEAU D'UN PROGRAMME²

En utilisant les notions et concepts susmentionnés – l'amélioration de la qualité, l'harmonisation constructive, l'évaluation authentique, la SRAO et les validité, fiabilité et crédibilité – et en faisant fond sur les travaux de Biggs et Tang (2011), Biggs et Collis (1982), Gulikers, Bastiaens et Kirschner (2004) et Boud et collègues (2010), nous proposons l'ensemble suivant de principes pour guider l'évaluation des résultats d'apprentissage au niveau d'un programme :

1. S'assurer de l'harmonisation constructive des programmes pour permettre l'utilisation d'évaluations arrimées aux cours.

Lorsque les cours d'un programme sont adéquatement harmonisés aux résultats d'apprentissage au niveau du programme, il devient possible de recourir à des évaluations arrimées aux cours pour évaluer l'apprentissage tout au long du cursus. On peut sélectionner des évaluations clés ancrées dans les cours qui démontrent de la meilleure façon les progrès d'apprentissage de l'étudiant par rapport aux résultats d'apprentissage voulus du programme. Un exemple d'une façon d'y parvenir au moyen d'un système de gestion de l'apprentissage (SGA) se trouve dans l'Étude de cas sur l'analyse de l'apprentissage à la section 1.3.

2. Concevoir l'évaluation des résultats d'apprentissage au niveau d'un programme en gardant à l'esprit le contenu et la complexité des résultats d'apprentissage.

La validité d'une tâche d'évaluation reposera sur la mesure dans laquelle cette évaluation évalue les implications de ce que les étudiants sont censés apprendre, si bien qu'il est important de porter une grande attention au contenu et à la complexité de l'apprentissage exprimés dans les résultats. Ainsi que l'a décrit la taxonomie de la SRAO, l'évaluation au niveau d'un programme devrait viser les résultats d'apprentissage aux niveaux relationnel et abstrait étendu.

3. Veiller à ce que l'évaluation des résultats d'apprentissage au niveau d'un programme soit authentique par rapport au programme d'études et/ou à la profession.

L'évaluation devrait requérir des étudiants qu'ils accomplissent des tâches ou créent des produits qui reflètent véritablement la pratique ou sont pertinents au programme d'études ou à la profession, et ce, dans un contexte qui simule la réalité de ce programme d'études ou de cette future profession (dans la mesure du possible).



Étude de cas :

Carnegie Mellon University
Évaluation pour améliorer les processus d'examen des programmes

Contexte : [Traduction] « En 2009, la Carnegie Mellon University [CMU] a démarré, à l'échelle de tout l'établissement, une initiative dans laquelle chaque programme menant à un diplôme devait systématiquement documenter ses résultats et processus d'évaluation. Cette initiative était directement liée au processus d'examen et de révision des cursus [...] appuyé centralement par l'Eberly Center [d'excellence de l'enseignement et d'innovation en éducation]. Cette documentation montre comment les programmes utilisent données et information pour améliorer continuellement leur qualité et leur efficacité. »³

Outils :⁴ Outre un graphique standard des résultats de programme⁵ la CMU publie une grille des résultats de maîtrise⁶ et une matrice des résultats pour cartographier les résultats depuis le niveau d'un cours jusqu'au niveau d'un programme. Toutefois, l'évaluation en vue de l'amélioration des processus à la CMU met l'accent sur les différences entre les disciplines et évite une approche « universelle »⁷.

[Traduction] « Nous appréhendons l'évaluation sous l'angle des données plutôt que des outils, notre choix des méthodes étant guidé par les questions que se pose le corps professoral, comme les suivantes :

- Qu'est-ce que ce processus m'apprendra au sujet des connaissances, compétences et croissance de mes étudiants?
- Qu'apprendrai-je au sujet des points forts et des points faibles de notre programme?
- Quelle information cela m'apportera-t-il sur la façon d'améliorer mon enseignement ou notre programme? »⁸

² Pour en savoir plus sur l'évaluation des résultats d'apprentissage au niveau d'un programme, consulter la page (en anglais seulement) <http://gotothexchange.ca/index.php/curriculum-overview/curriculum-models-and-design-principles>

³ <http://www.cmu.edu/teaching/assessment/assessprogram/index.html>

⁴ <http://www.cmu.edu/teaching/assessment/assessprogram/toolsfacilitateprocess.html>

⁵ <http://www.cmu.edu/teaching/assessment/assessprogram/components.html>

⁶ <http://www.cmu.edu/teaching/assessment/assessprogram/masterygrid.html>

⁷ <http://www.cmu.edu/teaching/assessment/history/ATFdocs/ATFpositionOnAssessment.pdf>

⁸ <http://www.cmu.edu/teaching/assessment/history/ATF.html>

1.3 CRÉATION D'UN PLAN POUR L'ÉVALUATION DES RÉSULTATS D'APPRENTISSAGE AU NIVEAU D'UN PROGRAMME

Dans cette section, nous formulons des recommandations sur la façon de créer un plan assurant une harmonisation constructive et permettant d'évaluer les résultats d'apprentissage au niveau d'un programme. Comme on l'a souligné dans les deux premières sections du manuel, l'évaluation des résultats d'apprentissage consiste à [traduction] « rendre nos attentes explicites et publiques, à fixer des critères appropriés et des normes élevées pour la qualité de l'apprentissage, à systématiquement colliger, analyser et interpréter les données probantes pour déterminer dans quelle mesure le rendement obtenu correspond à ces attentes et normes et à utiliser l'information qui en résulte pour documenter, expliquer et améliorer le rendement » (Angelo, 1995, p. 7).

Des plans d'évaluation au niveau d'un programme bien conçus doivent définir les résultats d'apprentissage au niveau du programme, les évaluations utilisées pour mesurer l'obtention de ces résultats ainsi qu'un plan pour utiliser cette information aux fins de l'amélioration du programme. Ce processus est le plus efficace lorsqu'il est entrepris par une équipe formée d'enseignants, de coordonnateurs de programme, d'étudiants et de pédagogues/concepteurs de cours. Les plans d'évaluation au niveau d'un programme visent à faire état des caractéristiques des programmes réussis et à relever les aspects à améliorer. De solides programmes, un leadership efficace et un bon soutien administratif sont indispensables au processus de planification et de mise en œuvre (voir la Partie 3).

Nous avons conceptualisé l'évaluation des résultats d'apprentissage au niveau d'un programme sous la forme d'un cycle à quatre stades :

Cycle à quatre stades pour l'évaluation des résultats d'apprentissage au niveau d'un programme

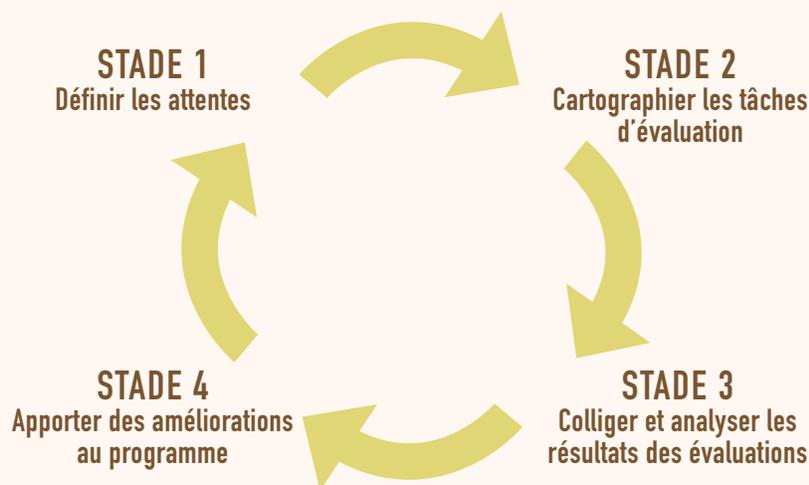


FIGURE 5

STADE 1 : DÉFINIR LES ATTENTES LIÉES AUX RÉSULTATS D'APPRENTISSAGE AU NIVEAU D'UN PROGRAMME

La première étape de l'évaluation d'un programme consiste à exprimer clairement ce que vous voulez que les étudiants apprennent et la façon dont vous vous attendez à ce qu'ils soient capables de démontrer cet apprentissage au moment où ils terminent le programme. Des attentes bien définies sont importantes en ce qu'elles fondent les futures décisions au sujet des types d'évaluation appropriés à retenir pour mesurer le degré d'obtention de ces résultats. En clarifiant vos résultats d'apprentissage au niveau du programme, vous serez aussi mieux à même de vous assurer que les expériences d'apprentissage et les évaluations sont centrées sur les compétences, idées, attitudes et valeurs les plus importantes de la discipline ou profession. Il ressort des travaux récemment publiés que nombre d'étudiants ne sont pas capables de définir ou d'exprimer les résultats d'apprentissage qu'ils ont atteints; plus de clarté pourrait aider les diplômés à mieux faire valoir leurs compétences et caractéristiques aux futurs employeurs (Barrie et al., 2009; Martini et Clare, 2014). Il est donc important de communiquer les résultats d'apprentissage aux étudiants tout au long de leur programme d'études. Il y a un certain nombre de facteurs qu'il vaut de prendre en considération au moment de définir des résultats d'apprentissage (figure 6), par exemple les connaissances préalables que les étudiants doivent posséder avant de commencer un programme, le cursus, les types d'enseignement et techniques d'apprentissage auxquels les étudiants seront exposés pendant le programme, les méthodes possibles d'évaluation et les attentes des organes de l'extérieur, parmi lesquels les organismes gouvernementaux et les associations professionnelles.

Il est souvent utile que tous les enseignants rattachés au programme contribuent à l'élaboration et à l'examen des résultats d'apprentissage au niveau du programme pour s'assurer de bien cerner les principaux résultats d'apprentissage et de comprendre en profondeur ces résultats et leur raison d'être. En outre, il est utile, à ce stade, de faire participer étudiants et employeurs, souvent dans le cadre de groupes de réflexion, d'entrevues ou de questionnaires. Une fois les résultats d'apprentissage voulus au niveau du programme clairement et adéquatement définis, l'étape suivante consiste à s'assurer qu'il y a harmonisation constructive aux niveaux des cours, du programme, de l'établissement et de la province.

Facteurs influant sur les résultats d'apprentissage au niveau d'un programme



FIGURE 6

STADE 2 : SCHÉMATISER (METTRE EN CORRESPONDANCE) LES TÂCHES D'ÉVALUATION TOUT AU LONG DU PROGRAMME

La schématisation (ou mise en correspondance) du cursus procure une stratégie efficace pour exprimer, harmoniser et intégrer les résultats d'apprentissage dans un enchaînement de cours ainsi que localiser précisément la façon dont les résultats d'apprentissage des étudiants sont obtenus dans un programme (Kopera-Frye, Mahaffy et Svare, 2008; Uchiyama et Radin, 2009). Il est important de déterminer comment et où se produit l'apprentissage dans le cursus et comment il est évalué. Les schémas de cursus catégorisent souvent l'apprentissage en fonction des cours et explicitent comment chaque cours contribue à l'obtention des résultats d'apprentissage au niveau d'un programme et, au surplus, comment ces résultats contribuent à la satisfaction des attentes de l'établissement ou de la province (figure 3). La schématisation est une méthode visuelle d'analyse des programmes que l'on peut effectuer au moyen de graphiques, de tableaux, de diagrammes ou d'autres processus permettant d'obtenir un portrait général et une analyse du programme dans son ensemble. La schématisation révèle des schèmes qui, autrement, seraient difficiles à déceler et rend transparentes les subtilités de l'harmonisation des évaluations et des résultats d'apprentissage au niveau du programme.

Ce processus peut être géré par une petite équipe qui collige les données initiales obtenues des enseignants et les incorpore à un schéma du cursus. L'équipe peut organiser un atelier ou une réunion pour passer en revue les données afin d'en vérifier l'exactitude et d'analyser le schéma et d'inviter les enseignants du programme, le coordonnateur du programme, le département responsable du programme ainsi qu'un concepteur de cours/pédagogue à agir comme facilitateurs. Plusieurs établissements, comme l'Université de Guelph⁹ et l'Université Ryerson¹⁰ (Mazurat et Schönwetter, 2009), ont conçu ou acheté un logiciel de facilitation de l'élaboration de schémas de cursus. Deux exemples de schémas simples de cursus sont présentés ci-dessous.

⁹ <http://www.uoguelph.ca/vpacademic/avpa/outcomes/curriculummap.php>

¹⁰ <http://www.ryerson.ca/lt/programs/curriculum/curriculumdevelopment>

Mise en correspondance des résultats d'apprentissage au niveau d'un programme avec l'enseignement, l'apprentissage et l'évaluation

Résultats d'apprentissage au niveau du programme	Attentes au niveau du diplôme	Activités d'enseignement et possibilités d'apprentissage	Évaluations et données probantes
D'ici la fin du programme, les étudiants qui obtiendront leur diplôme seront...	Avec quelle attente au niveau du diplôme ou quel résultat d'agrément chaque résultat d'apprentissage au niveau du programme s'harmonise-t-il?	Quelles activités précises d'enseignement et possibilités d'apprentissage aideront les étudiants à obtenir chaque résultat d'apprentissage au niveau du programme?	Pour chaque résultat d'apprentissage au niveau du programme, quelle information est précisément obtenue des étudiants comme preuve qu'ils peuvent atteindre ou ont atteint le résultat d'apprentissage avant l'obtention du diplôme?
1...			
2...			

EXEMPLE 1

Une fois les tâches d'évaluation mises en correspondance avec les résultats d'apprentissage au niveau du programme, il se peut que des lacunes dans l'évaluation de certains résultats d'apprentissage deviennent apparentes (Kopera-Frye et al., 2008; Uchiyama et Radin, 2009). Par exemple, une analyse du cursus australien a révélé que les résultats liés au développement éthique, à la compétence interculturelle et à la responsabilité sociale étaient rarement évalués (Barrie et al., 2012); cette tendance s'observerait probablement aussi au Canada. Alors que la triangulation et de multiples formes d'évaluation sont recommandées, il se peut que certains résultats soient exagérément évalués, ce qui alourdit le fardeau tant des étudiants que des enseignants. D'autres résultats peuvent être évalués à des niveaux de maîtrise inappropriés, par exemple aux niveaux inférieurs de la taxonomie de la SRAO (section 1.2).

La schématisation, ou mise en correspondance, fournit une excellente occasion de revoir les tâches d'évaluation incluses à tous les stades du programme et de s'assurer de leur harmonisation avec les résultats d'apprentissage voulus. Des tâches précises d'évaluation peuvent être mises au point pour mesurer des résultats au niveau des cours. Ces tâches d'évaluation seraient réglées au degré approprié de maîtrise, idéalement au niveau abstrait étendu de la taxonomie de la SRAO, où les étudiants sont capables de transférer des idées et de les appliquer à un contexte allant au delà du cours. En outre, des tâches d'évaluation devraient être ancrées aux cours et basées sur les principes d'évaluation authentique pertinents à la discipline (section 1.2). On peut examiner chaque cours individuellement pour s'assurer que les évaluations concordent et convergent vers les résultats d'apprentissage au niveau du programme qui sont définis. Cette approche à l'échelle d'un programme dans laquelle un cadre général guide l'évaluation individuelle des cours pour assurer la production systématique de données probantes au niveau du programme est fortement recommandée par les experts dans des projets tels que le projet VALUE du National Institute for Learning Outcome Assessment (NILOA) des États unis et le projet AAGLO (Assessing and Assuring Graduate Learning Outcomes) du conseil australien de l'apprentissage et de l'enseignement (Barrie et al., 2012). Dans la Partie 2 sont décrites des évaluations auxquelles on peut recourir pour divers types de résultats d'apprentissage au niveau d'un programme.

Schématisation des résultats d'apprentissage au niveau du programme selon les cours

Résultats d'apprentissage au niveau du programme	Cours de niveau inférieur			Cours de niveau supérieur			Résumé d'évaluation
	Cours 1	Cours 2	Cours 3	Cours 4	Cours 5	Cours 6	
Résultat 1	I		R	R		M E1	E1
Résultat 2		I		R E2	M E3		E2, E3

- I** NIVEAU DE L'INITIATION : le résultat est atteint au niveau de l'initiation, en présumant peu ou pas de connaissances préalables.
- R** RENFORCÉ : le résultat est renforcé, en présumant que l'étudiant avait déjà été initié à la matière dans un cours précédemment suivi.
- M** MAÎTRISÉ : le résultat est maîtrisé ou atteint; on présume habituellement que l'étudiant a été initié à la matière ou a renforcé ses connaissances dans des cours antérieurement suivis.
- E** ÉVALUÉS : moment – dans le déroulement du programme – où les résultats d'apprentissage au niveau du programme sont évalués.

EXEMPLE 2

STADE 3 : COLLIGER ET ANALYSER LES RÉSULTATS D'ÉVALUATION

Après avoir mis au point des tâches d'évaluation pour mesurer le degré d'atteinte des résultats d'apprentissage au niveau de l'ensemble du programme, vous devriez colliger ces données sur une base continue (par semestre et annuellement) et les évaluer aux fins de l'amélioration du programme. Plus précisément, vous pouvez utiliser l'information recueillie pendant votre évaluation des résultats d'apprentissage au niveau du programme pour :

- déterminer le degré auquel les étudiants atteignent les résultats d'apprentissage au niveau du programme qui ont été préalablement déterminés;
- prioriser des dimensions dans lesquelles une amélioration du programme est nécessaire;
- faire valoir la qualité du programme aux employeurs, donateurs et étudiants potentiels;
- documenter la preuve de la réussite des étudiants et de leur apprentissage aux fins d'agrément/attestation et de responsabilisation.

Il y a un certain nombre de facteurs à prendre en considération avant et pendant ce stade du processus, notamment le meilleur moment auquel colliger les données d'évaluation, la portée des tâches d'évaluation, les multiples façons de recueillir les données et la meilleure méthode d'en analyser les résultats. Ces considérations sont explorées plus avant ci-après.

Moment de l'évaluation

Idéalement, la collecte de l'information se produit au début, tout au long et à la fin de tout cours ou programme. Comme il vous faudra recueillir les données pendant des périodes chargées de la session, il est important de s'assurer de disposer d'un soutien suffisant. Les évaluations arrimées aux cours sont idéales pour évaluer l'apprentissage des étudiants au niveau d'un programme. Entre autres avantages, elles évitent de soumettre les étudiants à des tests pendant les périodes où ils sont très fatigués, elles incitent les étudiants à remplir les évaluations et elles donnent davantage aux enseignants l'assurance que l'évaluation sera authentique et pertinente (Ewell, 2009). Si le temps ou les ressources sont limités, vous pouvez vous concentrer sur les cours de fin de programme, puisque ce sont les cours les plus susceptibles de rendre compte du degré d'atteinte des résultats d'apprentissage au niveau d'un programme.

Si la collecte et l'analyse des données prennent plus d'un an, l'information sera moins utile pour peaufiner le programme avec souplesse. Là encore, la suffisance des ressources et du soutien du département et de l'établissement influent sur le moment d'exécution. Si vous ne disposez pas de ressources suffisantes, choisir de se concentrer sur un aspect particulier de l'évaluation (comme un résultat d'apprentissage en particulier ou un groupe d'intervenants en particulier) peut s'avérer utile.

Portée de l'évaluation

Utiliser diverses tâches d'évaluation aide à brosser un portrait général de la réussite des étudiants. La quantité de données recueillies tout comme leur type devrait être suffisant. On conseille aussi de conjuguer les méthodes directes et indirectes d'évaluation. Les méthodes directes permettent d'obtenir des données probantes sur ce que les étudiants apprennent, tandis que les méthodes indirectes (comme les enquêtes, sondages et groupes de réflexion) procurent de l'information à partir de laquelle des conclusions peuvent être tirées au sujet de l'apprentissage des étudiants. Il est important de recueillir les types de données qui sont le plus utiles/significatifs pour les valeurs de la discipline et qui sont le plus susceptibles d'amener professeurs et enseignants à parler de l'amélioration du programme.

Collecte des données d'évaluation

Il serait trop ambitieux de vouloir recueillir et analyser de l'information sur tous les résultats d'apprentissage au niveau d'un programme en même temps. L'établissement d'un plan réalisable, l'ancrage de l'évaluation au niveau du cours et l'utilisation d'évaluations qui représentent raisonnablement l'apprentissage des étudiants rendront le processus beaucoup plus efficace en ce qu'ils feront diminuer le temps consacré à la collecte et à l'analyse des données. Réfléchissez dès le début du processus à la façon dont les résultats seront communiqués et utilisés, car cela peut aider à déterminer quelles données sont les plus utiles à recueillir.

Nous vous recommandons de commencer à petite échelle et de vous concentrer sur quelques buts ou résultats d'apprentissage au niveau d'un programme importants chaque année aux fins d'examen et d'évaluation. Certains établissements recourent à des outils électroniques pour faciliter la collecte et l'analyse des données. Par exemple, certains systèmes de gestion de l'apprentissage comportent des éléments permettant aux évaluations d'être liées aux résultats d'apprentissage aux niveaux du cours et du programme. Les grilles ou notes qui sont entrées dans ce système peuvent être extraites et présentées sous la forme d'un rapport démontrant le niveau général d'atteinte des résultats d'apprentissage (p. ex., Kaupp, Frank et Watts, 2013). Les technologies utilisant des portfolios électroniques pour présenter les données recueillies au long de la durée du programme deviennent aussi de plus en plus efficaces pour appuyer ce processus (p. ex., Mentkowski, 2006).

D'autres établissements font appel aux services des bureaux d'analyse des recherches de l'établissement. Ces bureaux possèdent l'expertise voulue et ont accès à des systèmes de sondage en ligne et des outils analytiques de données. Divers types de données peuvent attirer l'attention sur les disparités entre les formes directe et indirecte de l'évaluation. Cela peut être particulièrement utile au moment de trianguler les résultats des évaluations, les perceptions que les enseignants ont de l'évaluation des résultats d'apprentissage et les déclarations que font les étudiants quant à leur atteinte des résultats d'apprentissage. On se limite lorsqu'on se fie uniquement à ce que l'étudiant autodéclare; il peut être



Moment de l'évaluation

1. Quel moment de l'année convient le mieux pour recueillir l'information?
2. De quel point dans le programme recueillerez-vous de l'information (P. ex., les cours de toutes les années, les cours de la fin du programme)?
3. Combien de temps allouerez-vous à la collecte de l'information?
4. Combien de temps allouerez-vous à l'analyse des données?



Portée de l'évaluation

1. À quelle diversité de tâches d'évaluation recourrez-vous?
2. Centrez-vous la portée sur un point de mire en particulier (p. ex., sur un résultat d'apprentissage particulier au niveau du programme)?
3. Disposerez-vous de ressources (personnes, temps, expertise)?

éclairant de demander aux étudiants de réfléchir à leurs expériences d'apprentissage. Les questions suivantes pourraient être posées aux étudiants :

- Dans quelle mesure ce cours/programme a-t-il contribué à l'atteinte des résultats d'apprentissage suivants?
- Quels résultats d'apprentissage vous ont pris le plus de temps à obtenir?
- Quels résultats d'apprentissage vous ont pris le moins de temps à obtenir? (adapté de Kenny et Desmarais, 2012)
- Quelles actions ou interventions des professeurs et du personnel ont fait le plus de différence dans votre apprentissage?
- Qu'est-ce qui rend une classe difficile à vos yeux? (adapté de Blaich et Wise, 2011)

Comme on l'a mentionné à la section 1.2, vous pourriez utiliser des évaluations de référence pour analyser les résultats d'apprentissage au niveau d'un programme au début et à la fin d'un programme afin d'en examiner les variations dans le temps. Depuis de nombreuses années, l'Alverno College effectue l'évaluation de résultats d'apprentissage au niveau d'un programme en utilisant des portfolios et d'autres évaluations ancrées pour déterminer si les étudiants ont atteint les résultats d'apprentissage (Mentkowski, 2006).

Analyser les résultats

À cette étape, le dialogue et la réflexion entre le personnel enseignant et les étudiants sont très utiles. Orienter les questions peut mener à une discussion fructueuse et d'utiles révélations.

Par exemple, s'il ressort de l'analyse que le degré de réussite des étudiants n'a pas répondu aux attentes pour une tâche ou une autre, vous devriez vous pencher sur les facteurs ayant pu donner lieu à cette constatation. En présentant quelques éléments essentiels tirés de l'analyse initiale des données aux groupes d'enseignants et de membres du personnel, cela peut mener à un dialogue plus nourri, à des cycles itératifs (en demandant plus de données) et à une discussion plus ciblée (Ewell, 2009).

Les récents développements en conception d'outils d'analyse de l'apprentissage ont produit de nouvelles méthodes que les établissements peuvent envisager d'utiliser pour lier l'évaluation du rendement des étudiants au niveau d'un cours à leur rendement au niveau d'un programme. Un exemple d'application d'un outil analytique de l'apprentissage est présenté dans l'étude de cas ci dessous.



Étude de cas : Analyse de l'apprentissage

Contexte : En 2013, l'Université de Guelph et Desire2Learn (D2L) ont commencé un projet pilote d'utilisation de l'outil analytique Insights de D2L pour harmoniser les résultats d'apprentissage au niveau d'un programme et au niveau d'un cours avec leur évaluation. Ce projet pilote est centré sur la mesure de l'atteinte des résultats d'apprentissage pour plusieurs cours de deux programmes – Génie et Arts et sciences – en intégrant un système de suivi des résultats d'apprentissage ainsi que l'évaluation de ces résultats directement dans le système de gestion de l'apprentissage de D2L. L'application d'outils d'analyse de l'apprentissage à l'évaluation des résultats étant un concept relativement nouveau, ce projet procurera à l'Université de Guelph un moyen exhaustif de suivre et évaluer les résultats d'apprentissage au niveau du cours et au niveau du programme.

À la deuxième année, les Universités McMaster, Wilfrid Laurier et de Waterloo ainsi que le Mohawk College ont été invités à se joindre au projet. Ces partenaires contribuent au processus d'intégration de l'outil d'analyse Insights à l'échelle de la province et d'adaptation de ses fonctionnalités aux divers contextes des disciplines et des établissements. Chaque établissement fait l'essai pilote de l'outil et l'utilise de diverses façons et selon les circonstances qui lui sont propres.

Un outil analytique tel qu'Insights comporte de vastes possibilités d'application aux fins d'évaluation et, à terme, procurera de la souplesse dans la façon dont les données sur le degré d'atteinte des résultats d'apprentissage sont recueillies en Ontario.



Collecte des données d'évaluation

1. Combien de types d'information et quels types de données d'évaluation examinerez-vous? Quel type d'information rendra le mieux compte de l'apprentissage des étudiants?
2. Y a-t-il une question ou un problème auquel vous voudriez apporter une réponse et qui pourrait être le point de mire de votre collecte de données?
3. Quels outils allez-vous utiliser pour recueillir les données?
4. Qui sera chargé de gérer la collecte des données? D'effectuer la collecte des données?
5. De quelle façon allez-vous stocker l'information?



Analyser les résultats

1. Les étudiants ont-ils atteint les résultats d'apprentissage au niveau du programme?
2. Quels schèmes ou tendances observez-vous?
3. Avez-vous eu des surprises?
4. Quels facteurs ont pu contribuer à vos constatations?
5. Qu'est-ce que cela signifie quant aux mesures à prendre?
6. Y a-t-il des changements que vous devriez apporter au processus de collecte et d'analyse des données pour les futurs cycles?

STADE 4 : APPORTER DES AMÉLIORATIONS AUX PROGRAMMES

L'utilisation de données d'évaluation pour améliorer les programmes contribue à l'instauration d'une culture d'établissement qui valorise l'amélioration de la qualité et mobilise en ce sens. Pourtant, la plupart des établissements ont de la difficulté à matérialiser l'information tirée des évaluations en plans d'amélioration et à mettre efficacement en œuvre ces plans. S'il apparaît que les établissements utilisent souvent les données d'évaluation au niveau d'un programme à des fins d'agrément et d'examen des programmes, seul un petit nombre se servent de cette information à des fins de planification stratégique, d'élaboration de politiques scolaires, de révision du cursus et d'amélioration de l'établissement (Kuh, Jankowski, Ikenberry et Kinzie, 2014). Pour que des changements soient opérés, il faut que le processus et les résultats de l'évaluation des résultats d'apprentissage soient partagés avec de multiples intervenants à chaque stade, ce qui façonnera une culture de campus qui mobilise pour une amélioration systémique et stratégique des programmes et qui valorise cette amélioration (Blaich et Wise, 2011).

Vous pouvez commencer par passer en revue les conclusions tirées au sujet de l'apprentissage des étudiants. Si des possibilités d'amélioration se présentent à la suite de préoccupations exprimées concernant le degré de réussite des étudiants dans un domaine, il demeure qu'il n'est pas toujours évident de prédire la façon dont certains changements sont susceptibles d'influer sur le rendement. Si le degré de réussite des étudiants est en deçà des attentes, vous devriez vous demander s'il existe des possibilités d'amélioration à l'intérieur d'un cours ou d'un programme ainsi qu'entre les cours et le programme. Lorsqu'on a cerné les dimensions qui, à l'intérieur du programme, commandent une amélioration, la prochaine étape consiste à poser des questions pouvant orienter la réflexion.

Il peut s'avérer utile de classer les suggestions de mesures à prendre en fonction des réponses données aux questions susmentionnées. Vous devriez aussi inclure dans ces discussions d'autres intervenants et parties concernées (p. ex., enseignants, étudiants, membres du personnel et employeurs). Il faudrait dresser un plan de suivi précisant les échéanciers et les personnes responsables. À ce stade, vous pouvez déterminer s'il y a lieu d'obtenir davantage d'informations avant d'amorcer la prise de mesures. Dans la plupart des cas, cela peut comprendre un examen de la littérature sur l'enseignement et l'apprentissage ou encore la réalisation d'un petit projet de recherche sur le résultat d'apprentissage au niveau d'un programme en question. Il pourrait être utile de procéder à une vérification (un audit) des données pour vérifier si des données d'évaluation que recueille l'établissement ne sont pas déjà utilisées pour évaluer le résultat d'apprentissage au niveau du programme. Nombre d'établissements ont de petits programmes de subventions ou des possibilités de financement stratégique qui pourraient leur procurer des ressources permettant de mobiliser des personnes sur un petit projet de recherche afin de mettre à l'essai et d'évaluer une amélioration aux programmes (Deepwell et Buckley, 2013).

Il est important de partager les résultats de l'évaluation des programmes étant donné qu'il faut souvent consacrer beaucoup de temps à l'amélioration des programmes et que la volonté de consacrer plus de temps à l'avenir peut dépendre de la perception que les cycles antérieurs ont eu un impact. L'information peut être communiquée de diverses façons, comme les suivantes : réunions du corps professoral ou de départements; séances/journées de réflexion; petites réunions ciblées; mises à jour envoyées par courriel; sites Web; rapports officiels; exposés de conférence; articles (Kuh et al., 2014). Tout comme plusieurs tâches d'évaluation fournissent une information plus exhaustive sur l'apprentissage des étudiants, plusieurs formes de communication de l'information élargissent la compréhension et conviennent à un plus large éventail d'auditoires. L'information peut être résumée sous forme de chiffres, de pourcentages ou de pointages, ou encore sous la forme de résumés qualitatifs. Les schèmes de rendement au fil du temps peuvent être comparés entre cohortes ou groupes, enseignants ou pairs. Cibler les questions qui influent sur la pratique des enseignants peut motiver le changement (Blaich et Wise, 2011). Là encore, la meilleure approche dépend des résultats d'apprentissage au niveau d'un programme, de la façon dont les résultats sont évalués, etc.; il est impératif d'avoir une discussion ouverte et créative. Il est utile de décider, pendant le processus de collecte de l'information, de la façon dont les résultats seront résumés et communiqués.

En dressant un plan détaillé d'évaluation, on s'assure plus facilement du caractère explicite des résultats d'apprentissage au niveau d'un programme pour les étudiants et les intervenants; les activités



Apporter des améliorations aux programmes

1. Quelles sont les constatations les plus importantes?
2. Quels aspects ou dimensions affichent les plus grands défis quant à l'apprentissage (et donc les meilleures occasions d'amélioration)?
3. Quelles mesures pourrait-on prendre pour relever chacun de ces défis?
4. Parmi ces mesures, lesquelles sont le plus susceptibles d'être efficaces?
5. Lesquelles sont possibles du point de vue du coût, du fonctionnement, etc.?
6. Quels changements est-il possible d'opérer immédiatement et quelles difficultés pourraient être résolues à l'avenir?
7. Qui serait capable de prendre ces mesures?
8. Comment mesureriez-vous l'efficacité des améliorations que vous aurez apportées (planification du prochain cycle d'évaluation)?



Communication transparente

1. Avec qui communiquerez-vous?
2. Quels seront les meilleurs moyens et moments de communiquer?
3. Quel mode de présentation des résultats permettra d'alimenter les conversations les plus susceptibles d'aboutir à une amélioration de l'apprentissage des étudiants dans le programme?
4. De quelle façon les modes de communication permettront-ils de mieux répondre aux questions touchant l'apprentissage des étudiants et leur degré d'atteinte des résultats d'apprentissage au niveau d'un programme?

d'enseignement et d'apprentissage s'harmonisent avec les résultats d'apprentissage; les tâches d'évaluation sont efficaces, coordonnées et structurées tout au long du programme; les étudiants atteignent les principaux résultats d'apprentissage; les investissements faits dans l'évaluation rapportent au chapitre de l'amélioration des programmes. Un exemple de cela est présenté dans l'étude de cas ci-dessous (Washington State University).



Étude de cas : Washington State University, Suivi des progrès dans l'évaluation aux fins des améliorations

Contexte : [Traduction] « Le sénat de la Washington State University [WSU] a instauré sa politique initiale d'évaluation des programmes en 1990. Dans la mise à jour de la politique, en 2009, il est précisé que le corps professoral associé à chaque programme de premier et de deuxième cycles menant à un diplôme doit élaborer un plan visant à évaluer [...] les étudiants sur le point de décrocher le diplôme [...]. Les départements doivent être capables de démontrer des améliorations dans le temps. »¹¹

Exemples d'évaluations ayant mené à des améliorations : L'Office of Assessment of Teaching and Learning (bureau d'évaluation de l'enseignement et de l'apprentissage de l'Université) fournit de nombreux exemples de la façon dont les programmes de la WSU utilisent des résultats d'évaluation.

- [Traduction] « L'École de sciences économiques a changé son programme d'enseignement afin d'y inclure une exigence de travaux de recherche [...] après que les évaluations ont révélé des points faibles chez les étudiants en outils économiques et quantitatifs appliqués. »¹²
- [Traduction] « Le Programme de développement humain a recueilli les notes que les superviseurs ont attribuées aux stagiaires afin de [...] renforcer les aspects où des lacunes de compétences sont relevées. Les changements apportés [au programme d'enseignement] ont entraîné une forte hausse des notes attribuées par les superviseurs. »¹³
- [Traduction] « L'École des sciences de l'alimentation s'est concentrée sur l'amélioration des compétences des étudiants en communication et en gestion du temps après avoir sondé d'anciens étudiants et des employeurs de l'industrie au sujet du rendement des diplômés de la WSU à leur emploi. »¹⁴

Suivi des progrès dans les évaluations aux fins d'amélioration : La WSU recourt aussi à l'autoévaluation par les facultés et les départements afin de suivre l'évolution des évaluations en vue d'améliorer les processus dans les programmes. Le tableau ci-dessous montre comment les établissements peuvent se servir d'une grille de cotation de la maturité des processus pour planifier le soutien à des programmes dans les évaluations aux fins d'amélioration.¹⁵ En réponse à ce rapport, l'organisme d'agrément a cité une « transformation notable de la culture d'évaluation »¹⁶ à la WSU.

Autoévaluation globale du système d'évaluation des programmes

Nombre de programmes	4 (7 %)	20 (37 %)	20 (37 %)	10 (19 %)
DEGRÉ DE MATURITÉ DU SYSTÈME D'ÉVALUATION ET DES PRATIQUES	AU COMMENCEMENT : Une itération de processus d'évaluation commencé; peut en être à l'étape de l'essai pilote; peut ne pas encore disposer de données, ou les données peuvent ne pas encore avoir été communiquées ou discutées.	EN DÉVELOPPEMENT : Rajuste activement les processus ou outils de base après une itération/essai pilote; un certain degré de communication et de discussion des données; système de participation en cours d'élaboration.	EN PEAUFINAGE : Données régulièrement communiquées et discutées au long de plus d'un cycle d'évaluation; résultats utilisés pour améliorer et valider l'apprentissage des étudiants; utilisation des résultats régulièrement documentée.	ÉTABLI : Plusieurs itérations de cycle d'évaluation; processus structurellement axé avec vaste participation; processus et outils établis, mais permettant aussi de répondre aux besoins changeants du programme; système cyclique et utilisé pour améliorer et valider l'apprentissage des étudiants.

¹¹ <http://facsen.wsu.edu/eppm/AssessmentStudentLearning.pdf>

¹² http://oai.wsu.edu/assessment_resources/assessment_highlights_long.html#assessment_data_for_change

¹³ http://oai.wsu.edu/assessment_resources/assessment_highlights_long.html#internship_and_field_experience

¹⁴ http://oai.wsu.edu/assessment_resources/assessment_highlights_long.html#professional_skills

¹⁵ <http://accreditation.wsu.edu/reports/WSU-2013-Year-Three-Self-Evaluation-Report-Final.pdf> p. 223

¹⁶ <http://president.wsu.edu/blog/perspectives/2013/10/17/accreditation-report-praises-assessment-progress-outlines-need-for-more/>

PARTIE 2

PRATIQUES D'ÉVALUATION

Résumé

Dans cette partie, nous présentons une variété de tâches d'évaluation que l'on peut sélectionner et organiser au sein d'un cadre afin d'évaluer l'apprentissage au niveau du programme. Elles sont catégorisées en fonction des types de résultats d'apprentissage qu'elles permettent le mieux d'évaluer, et l'on s'attend à ce que la plupart puissent être adaptées de manière à évaluer plus d'un résultat d'apprentissage. Chacune de ces tâches possibles a été choisie du fait qu'elle concorde avec les principes énoncés à la Partie 1. Bien que ces évaluations soient arrimées à des cours individuels, il est important de tenir compte de la façon dont on peut les organiser afin d'étager leur déroulement au long du programme. Les évaluations du niveau d'initiation qui sont placées tôt dans le programme peuvent être renforcées ultérieurement dans le programme. Dès lors que les cours sont harmonisés de façon constructive, on peut utiliser les données des évaluations ancrées dans les cours pour faire état des progrès et de l'atteinte des résultats d'apprentissage au niveau du programme. Dans la Partie 3, nous nous concentrons sur le développement de cultures d'établissement favorisant une évaluation authentique et harmonisée, ce qui accroît la probabilité tant de l'adhésion que de la viabilité.

Dans cette partie sont résumées les tâches d'évaluation afférentes à des ensembles précis de résultats d'apprentissage au niveau d'un programme susceptibles d'être authentiques dans plusieurs contextes de disciplines et de professions. Bien que catégorisée sous une forme de résultats d'apprentissage pour vous aider à trouver et sélectionner les tâches, chaque tâche peut être utilisée pour évaluer plusieurs types de résultats d'apprentissage. Bien des tâches peuvent être adaptées à l'évaluation individuelle, par paire ou par groupe et la plupart peuvent être exécutées dans un cours en personne, en ligne, bimodal ou hybride. Toutes ces tâches ont été sélectionnées parce qu'on peut s'en servir pour évaluer des résultats d'apprentissage au niveau d'un programme qui démontrent l'apprentissage aux stades relationnel ou abstrait étendu. Plus l'évaluation se produit tard dans un programme, plus l'étudiant est censé s'être rapproché du stade abstrait étendu.

Au moment de choisir les tâches d'évaluation, gardez à l'esprit les principes énoncés dans la Partie 1. Le besoin de **validité** restreint l'éventail des choix possibles – demandez-vous si cette tâche d'évaluation particulière permet d'évaluer ce résultat particulier. Par exemple, la validité dans l'évaluation de la capacité à critiquer des arguments exige que les étudiants critiquent des arguments. Dans une évaluation authentique, les étudiants sont censés critiquer des arguments d'une manière qui simule l'environnement de rendement voulu.

En bout de ligne, les jugements sur les mesures d'évaluation devraient être portés par les enseignants et appuyés mais non dictés par les pédagogues/concepteurs de cours, étudiants, administrateurs et représentants des professions visées ou d'organismes gouvernementaux. La raison en est que les évaluations ancrées dans un ensemble de cours et de programmes d'enseignement **harmonisés de façon constructive** se sont révélées comme le moyen le plus efficace et le plus pertinent d'évaluer des résultats au niveau d'un programme (Barrie et al., 2009; Barrie et al., 2012; Rhodes, 2012; Kuh et al., 2014). Sadler (2013) fait une mise en garde contre l'évaluation de compétences distinctes au moyen d'un modèle à grande échelle : [traduction] « Au lieu de cela, on devrait se concentrer sur la notion de compétence comme capacité à orchestrer les connaissances et les aptitudes de manière

indépendante, dans un éventail de contextes, sur demande et avec un degré élevé de maîtrise. Le point de mire complémentaire est sur la compétence, en ce que celle-ci est acquise et perfectionnée par les étudiants dans le cadre de leurs programmes scolaires normaux, ainsi que sur la façon dont cette compétence pourrait être enrichie et évaluée » (p. 26). Le choix des évaluations se produit dans le cadre d'un plan intentionnel d'évaluer les résultats d'apprentissage au niveau d'un programme en sorte que l'on puisse se servir d'un cadre stratégique pour examiner le programme dans son ensemble et générer de façon systématique des données probantes sur l'apprentissage (voir p. ex. la section 1.3 et Barrie et al., 2012).

Certaines évaluations se rattachent spécifiquement à des devoirs en particulier dans un cours, tandis que d'autres rendent compte de l'apprentissage pendant toute la durée des cours (c. à d., portfolios et projets de fin d'études). Garder les principes en tête facilite de beaucoup les décisions à prendre quant aux tâches d'évaluation appropriées. Divers critères de conception importent. Par exemple, on peut difficilement qualifier un outil d'évaluation de « portfolio »; il doit être conçu avec la complexité et le contenu appropriés. La tâche portfolio doit être bien conçue et clairement communiquée aux étudiants, et ce, dans un environnement qui favorise et soutient l'apprentissage. Ces mêmes considérations s'appliquent à l'évaluation et à la notation des travaux des étudiants; si les étudiants sont comparés à leurs pairs ou assujettis à des normes inconnues ou déraisonnables, les résultats de l'expérience d'évaluation ne seront pas fiables. Les étudiants doivent être soumis à des normes explicites et bien exprimées qui sont tirées de vos résultats d'apprentissage et leur sont clairement communiquées à l'avance. De plus en plus, les grilles de notation sont conçues de façon collaborative par des universités et collèges afin d'être utilisées par les établissements.¹⁷ De plus, il importerait d'adapter les évaluations en sorte de les rendre authentiques et pertinentes à la discipline une fois sélectionnées les tâches appropriées afférentes à un cadre d'évaluation.¹⁸

Les grandes catégories d'évaluations pour les types particuliers de résultats qui sont présentées dans les sections qui suivent s'inspirent abondamment des travaux de McMichael (2009), lesquels ont fait fond sur les travaux de Nightingale et al. (1996) et de Brown, Rust et Gibbs (1994).

¹⁷ <http://www.aacu.org/value/rubrics/>

¹⁸ http://www.itl.usyd.edu.au/projects/aaglo/pdf/AAGLO%20Summary%207%20Assessment%20tasks_Final.pdf

2.1 PENSÉE CRITIQUE, RÉOLUTION DE PROBLÈMES, JUGEMENT ET PERSPICACITÉ

« Pensée critique » est un terme général aux significations diverses dans le domaine des études supérieures, ces acceptions variant non seulement selon la discipline, mais aussi d'une sous-discipline à l'autre et parmi les personnes qui travaillent dans ces disciplines. C'est vrai de la discipline de la logique informelle (le domaine de la pensée critique) comme de tout autre système de pensée organisée. Pour être efficacement évalués, les résultats d'apprentissage doivent définir la pensée critique par rapport au contexte du programme particulier.

Dans la plupart des cas, la pensée critique comporte une ou plusieurs des activités cognitives suivantes : analyse, synthèse, évaluation, jugement, argumentation et critique. La pertinence de certaines de ces activités peut varier selon la discipline. Chaque activité peut être subdivisée en éléments qui la composent, lesquels peuvent être enseignés et évalués séparément. Le jugement et la perspicacité vont habituellement au delà de la pensée critique; le jugement exige l'expression claire des motifs de la préférence d'une solution à une autre, tandis que la perspicacité peut commander d'interpréter de l'information complexe et d'en tirer une signification imprévue mais défendable.

Les résultats relatifs à la résolution de problèmes sont propres aux processus particuliers de la résolution de problèmes, lesquels impliquent habituellement de relever, poser, définir, interpréter, analyser et résoudre des problèmes particuliers. Certaines disciplines incluent, dans la « résolution », les composantes de la planification et de l'élaboration de stratégies. À leur tour, ces composantes impliquent de dresser et réviser des plans, d'utiliser de l'information pour ébaucher des stratégies et mettre en œuvre des stratégies.

En plus d'être des outils utiles pour évaluer les aptitudes à la pensée critique et à la résolution de problèmes, les **études de cas** et les **problèmes ouverts** présentent la possibilité, pour les étudiants, de synthétiser et appliquer un large éventail de connaissances de contenu et de processus (Christensen et Hansen, 1987; Dunne et Brooks, 2004). Les cas sont souvent assignés à des groupes, bien qu'on pourrait s'en servir pour évaluer des étudiants individuellement à l'intérieur d'un groupe ou indépendamment de groupes. Les études de cas figurent probablement parmi les types d'évaluation les plus authentiques et efficaces dans nombre de disciplines et de professions; elles sont très pertinentes, motivationnelles et cognitivement exigeantes (quoique des études de cas mal conçues et administrées peuvent être superficielles et inauthentiques). Leur conception peut être simple ou complexe et elles sont faciles à utiliser simultanément comme stratégies d'apprentissage et tâches d'évaluation. Les stratégies d'apprentissage autonome (ou autodidacte), dans lesquelles les étudiants assument une plus grande responsabilité de l'apprentissage (comme dans l'apprentissage axé sur les problèmes¹⁹, par exemple), peuvent reposer sur des études de cas, mais elles peuvent aussi bien s'harmoniser avec des formes d'enseignement davantage dispensé par l'enseignant, comme le traditionnel cours magistral.



Apprentissage axé sur les problèmes

Un scénario authentique dans lequel le problème étudié suscite l'apprentissage de l'étudiant. Les étudiants définissent les questions, trouvent des ressources appropriées, analysent et synthétisent l'information pour mieux comprendre le problème, communiquent leurs constatations et autoévaluent leurs propres progrès.

Adaptation à l'évaluation de classes nombreuses :

<http://insight.glos.ac.uk/tli/resources/toolkit/resources/alcs/pages/pbltriplejump.aspx>

¹⁹ <http://cill.mcmaster.ca/resources/pbl.html>

Pour éprouver l'authenticité et l'harmonisation des études de cas, il est utile de solliciter la rétroaction de collègues des départements ou de praticiens de l'externe. Une fois les études de cas rédigées, les étudiants peuvent pratiquer la synthétisation et appliquer les connaissances pertinentes dans des micro-cas, en classe comme à l'extérieur. Le résultat de cet entraînement fournit une évaluation en soi

Les **questions d'essai modifiées** (QEM) sont une façon utile de combiner les points forts des études de cas et les essais dans un environnement d'examen en présentant aux étudiants une série structurée de questions d'essai (à développement) ouvertes auxquelles ils doivent répondre en rapport avec une étude de cas. Après que les étudiants ont répondu à une question au sujet du cas, on leur fournit davantage d'information et leur pose une deuxième question, etc. Idéalement, l'enchaînement devient de plus en plus complexe et exige au fur et à mesure une intégration, une analyse et une évaluation plus poussées et plus profondes. Les QEM peuvent aider certains étudiants à mesurer l'importance de suivre un enchaînement ou un processus pour évaluer l'information. Bien que les examens aient historiquement été utilisés de façons inauthentiques, si vous vous voyez forcé d'administrer des examens en raison d'un taux élevé d'inscriptions, les QEM peuvent fournir une approximation d'évaluation authentique de la pensée critique, de la compréhension écrite, de la résolution de problèmes, de la réflexion et de la souplesse. Du soin doit toutefois être appliqué pour s'assurer que les étudiants qui répondent mal aux premières questions ne soient pas indûment pénalisés pour cela ultérieurement dans la séquence, par exemple en leur permettant de revoir et réviser leurs réponses au fil des questions.

Les **ensembles de problèmes** – problèmes prédéterminés, souvent concis, à résoudre rapidement – peuvent être utilisés dans pratiquement n'importe quelle discipline ou profession et procurer un moyen utile, quoique habituellement inauthentique ou déconnecté, d'évaluer des aptitudes en résolution de problèmes comportant un éventail de contenus. Ils sont relativement faciles à concevoir et coter, peuvent s'appuyer sur l'utilisation de problèmes de plus en plus complexes qui s'échafaudent les uns sur les autres et sont assez rigoureux et fiables lorsqu'ils sont bien conçus.

Les **débats, simulations de séances de tribunal, arguments oraux** et autres expériences **simulées en résolution de problèmes** et d'argumentation peuvent aider les étudiants à transférer leur pensée critique dans des situations du monde réel. La structure artificielle de certains débats nuit à leur authenticité, si bien que vous devez prendre soin de ne pas les structurer exagérément. De la même façon, on peut évaluer l'aptitude à résoudre des problèmes en donnant aux étudiants un ensemble de principes et de l'information et en leur demandant de suivre un processus afin d'en arriver à une solution défendable, ou encore on peut rendre l'évaluation plus complexe en demandant aux étudiants de concevoir une stratégie en intégrant, évaluant et appliquant de multiples sources d'information pour mettre au point une stratégie réalisable. Les évaluations comportant la critique et l'amélioration d'un processus existant de résolution de problèmes permettent d'évaluer et de coter les compétences en résolution de problèmes et pourraient être facilement conçues en sorte de faire participer les étudiants à des travaux de recherche.

L'évaluation de la pensée critique est souvent conjuguée à celle des capacités rédactionnelles dans des devoirs de rédaction de textes argumentatifs, persuasifs ou évaluatifs, comme des essais, discours, poèmes, comptes rendus de lecture/critiques de livre, lettres à l'éditeur ou revues. Nous revenons plus en détail sur certaines de ces formes un peu plus loin. Pour évaluer la perspicacité et le jugement, des rubriques de conseils simulées axées sur des problèmes rencontrés par des personnes réelles ou fictives peuvent être utiles.

L'**analyse** – distinguer les parties d'un tout et distiller l'information de manière qu'elle soit clairement comprise – peut être évaluée en demandant aux étudiants de créer des diagrammes décrivant les principales étapes d'un processus ou de déconstruire un argument. Les évaluations dans lesquelles on demande aux étudiants de déconstruire les hypothèses posées dans des ensembles d'informations, d'analyser des arguments dans les affirmations et la logique qui les composent ou de problématiser un énoncé apparemment non problématique peuvent permettre aux étudiants de démontrer les aptitudes à la pensée critique qu'on juge essentielles à l'éducation de citoyens démocratiques.



Étude de cas

Une description d'une situation réaliste ou authentique. Elle peut être présentée sous diverses formes, comme un paragraphe descriptif, un article de journal ou un long cas descriptif.

Échantillon de processus :

http://ar.cetl.hku.hk/am_case_study.htm



Questions d'essai modifiées

Examen comportant une série structurée de questions d'essai ouvertes auxquelles l'étudiant doit répondre en rapport avec une étude de cas. Une fois que les étudiants ont répondu à une question au sujet d'un cas, on leur fournit davantage d'information.

Adaptation en ligne par l'University of Sydney (Australie) :

<http://sydney.edu.au/medicine/showcase/assessment/meq/index.php>



Débat

Une discussion structurée dans laquelle sont pris en compte divers points de vue et l'on en arrive à un jugement.

Variations de formes :

http://www.usma.edu/cfe/literature/vargo_12.pdf

La capacité d'évaluer et le jugement apparaissent aux stades relationnel et abstrait étendu de la SRAO et ne peuvent être évalués que dans ce contexte. Souvent, il faut prendre davantage de précautions et de temps pour créer des évaluations pour ces compétences, en particulier lorsqu'on commence à utiliser ce type d'évaluation pour la première fois. Les évaluations en deux temps dans lesquelles on demande aux étudiants de concevoir et défendre des critères pour juger de phénomènes avant de les appliquer comportent une évaluation poussée de la connaissance du contenu, de l'application des principes et de la logique précisée par le programme, ainsi que de multiples formes de pensée et de jugement. De la même façon, les débats, les argumentations orales et la rédaction de tâches qui demandent aux étudiants de juger de la valeur, de l'acceptabilité, du mérite ou de l'exactitude d'une idée ou d'un argument font intervenir des éléments subjectifs de la pensée critique qui poussent les étudiants à faire davantage que suivre les procédures du par cœur et de la répétition de ce qu'ils pensent qu'on veut entendre pour leur faire découvrir ce qu'implique réellement la résolution de problèmes particuliers.

Au bout du compte, lorsque les cours sont adéquatement harmonisés aux résultats d'apprentissage au niveau du programme, il devient possible d'utiliser les résultats des évaluations susmentionnées ancrées dans les cours pour informer la progression de l'apprentissage tout au long du programme.

2.2 TRAVAUX DE RECHERCHE ET D'ÉRUDITION

La « recherche », à l'instar de la « pensée critique », revêt bien des sens différents. Les résultats d'apprentissage axés sur la recherche peuvent être centrés sur des composantes du processus, comme trouver l'information et y accéder, gérer et organiser l'information telle qu'elle est recueillie, mesurer la crédibilité des sources, exécuter des techniques de recherche dans une discipline, interpréter de l'information, l'utiliser de façon appropriée, etc. Certaines disciplines ou professions sont plus à l'aise avec le terme « érudition ». Quel que soit le nom que nous lui donnons, nous reconnaissons les formes fondamentales des activités qui s'y rapportent.

Pour les programmes de premier cycle, les **comptes rendus de recherche** et les **bibliographies annotées** sont des choix populaires, notamment parce qu'en tant qu'expériences d'apprentissage, ils peuvent faire connaître aux étudiants une variété de recherches courantes dans la discipline, dans un laps de temps relativement court, et leur donnent la possibilité de partager ces connaissances entre eux. Cela fonctionne le mieux lorsque les étudiants doivent synthétiser les résultats de leur lecture plutôt que laisser les contributions individuelles déconnectées les unes des autres. L'efficacité des bibliographies annotées, qui sont habituellement plus courtes et plus faciles à produire et à évaluer que les comptes rendus de travaux de recherche, est au plus haut lorsque les étudiants précisent l'ordre dans lequel les autres devraient prendre connaissance des entrées et qu'ils joignent à cela une brève justification de ce jugement. Cela aide les étudiants à atteindre le stade relationnel de la SRAO, puisqu'une bonne explication doit comporter des liens entre les entrées.

De façon générale, et en particulier aux fins de l'évaluation de résultats d'apprentissage au niveau d'un programme, les **projets de recherche d'envergure** qui sont réalisés au long d'une ou de deux sessions sont la meilleure façon d'évaluer les aptitudes en recherche (en plus d'autres types de résultats). Selon la discipline et la profession, de tels projets peuvent exiger des étudiants qu'ils se servent de l'information pour échauffer et éprouver des hypothèses et des prédictions, qu'ils mesurent et comparent de multiples extraits et variables, qu'ils classifient des phénomènes en utilisant des schémas de principes ou une taxonomie, qu'ils mènent des expériences (virtuelles ou pratiques), qu'ils synthétisent de volumineux ouvrages savants et qu'ils accomplissent bien d'autres activités importantes pour le développement des chercheurs, des universitaires et des citoyens. La communication de résultats de recherches peut prendre telle ou telle forme, voire de multiples formes – rapports, thèses, mémoires/travaux de recherche, exposés/présentations, affiches, vidéos, chorégraphies, etc. Certains cours sont conçus en utilisant l'**apprentissage basé sur la recherche/l'enquête** et évalués en fonction d'un projet final (Healey, 2005). Pour maximiser le potentiel d'apprentissage des projets, on devrait les concevoir en sorte que les étudiants qui le réussissent doivent synthétiser, interpréter, évaluer et appliquer ce qu'ils ont appris dans



Bibliographies annotées

Une liste alphabétique de références pertinentes avec annotations.

Échantillon de processus :

<http://www.writing.utoronto.ca/advice/specific-types-of-writing/annotated-bibliography>

le cadre du programme et pas simplement de tel ou tel cours. Cette façon de faire, conjuguée au temps et aux efforts considérables que l'on doit déployer pour fournir une rétroaction régulière, est le garant, pratiquement, d'un apprentissage utile et d'une évaluation significative.

Si l'on donne le choix du sujet aux étudiants, ils sont plus susceptibles d'être motivés à mener le projet à bon terme. L'évaluation de projets de ce type comporte souvent une évaluation des résultats d'apprentissage liés aux habiletés de communication, aux aptitudes à gérer le temps et aux compétences en gestion de projet, de même qu'aux qualités interpersonnelles et aux compétences en résolution de problèmes dans le cas de projets de groupe.

Pour faciliter la réussite de l'exécution, nous recommandons que le projet soit soumis aux fins de rétroaction et de révision à plusieurs étapes; à tout le moins, les étudiants devront construire une rétroaction sur leurs premières propositions. L'échange d'ébauches pour rétroaction par les pairs peut procurer une rétroaction constructive supplémentaire aux étudiants et fournir en outre d'autres modèles de travail des pairs. Pour que l'évaluation par les pairs produise tous ses bénéfices, les enseignants doivent être clairs quant aux résultats d'apprentissage et fournir une formation sur la façon de donner une rétroaction constructive qui soit pertinente à la tâche d'évaluation particulière. Ces évaluations fondées sur les cours peuvent ensuite informer les résultats d'apprentissage au niveau du programme.



Apprentissage fondé sur une approche d'enquête

Une forme d'apprentissage autodirigé/ autonome dans laquelle les étudiants choisissent une question, l'affinent et apprennent par la recherche/l'enquête.

Processus d'enquête à McMaster :

http://fhs.mcmaster.ca/bhsc/inquiry_what.html

Échantillon de processus et impact :

<http://insight.glos.ac.uk/TLI/RESOURCES/TOOLKIT/Pages/default.aspx>

2.3 COMMUNICATION

La plupart des programmes escomptent l'atteinte de certains résultats liés à la communication. C'est la communication par écrit qui domine cette catégorie de résultats – et la rédaction peut prendre bien des formes. Dans certaines disciplines et professions, les textes argumentatifs et persuasifs sont la norme, tandis que dans d'autres, il faut des rapports de recherche, des notes de service professionnelles et des documents de stratégie, des résumés ou des plans de commercialisation. En fait, la communication écrite est habituellement le moyen par lequel on évalue indirectement la pensée critique.

Les **essais** demeurent la forme la plus courante d'évaluation de l'écrit. Mal conçus, les essais peuvent être une forme inauthentique d'évaluation. Cependant, il y a de nombreux types d'essais différents – certains textes sont critiques et argumentatifs, d'autres essais rendent compte de résultats de recherches, certains ressortissent à la réflexion et d'autres sont à caractère persuasif – que l'on peut utiliser de façon authentique pour évaluer l'apprentissage des étudiants. De nombreux enseignants ont fait l'expérience d'**essais visuels**, sous la forme de bandes dessinées, de films, de séquences photographiques ou d'essais multimédias que l'on peut téléverser dans YouTube ou dans un système de gestion de l'apprentissage. Ces autres formes d'essai peuvent sembler inorthodoxes à ceux qui ont l'habitude d'apparier les essais à des documents. Quoi qu'il en soit, ces formes de rechange cadrent avec la signification première d'un « essai », qui est une présentation soutenue du point de vue d'une personne. C'est à cause de ce sens historique persistant que, quel que soit le mode de composition et de communication, les bons essais comportent tous un **engagement profond et soutenu à l'égard d'idées**.

L'un des avantages de recourir aux essais est que la plupart des étudiants sont familiers avec au moins un type d'essai. Cela dit, ce degré de familiarité peut amener certains étudiants à faire fi de l'utilité des essais. Il se peut aussi que des étudiants aient mal appris à rédiger des essais (en utilisant par exemple la « forme à trois paragraphes ») et qu'ils aient de la difficulté à rajuster leurs choix de rédaction de façon stratégique, en particulier s'ils sont habitués à considérer les essais comme des recettes que l'on doit suivre scrupuleusement. Quoi qu'il en soit, il est facile pour les enseignants de structurer les devoirs d'essai en sorte que les étudiants progressent dans leur rédaction sur le plan développemental, pas à pas, en recevant de la rétroaction en cours de route. Lorsque le processus d'essais est bien structuré et qu'il permet l'autonomie, les étudiants apprécient l'occasion de se consacrer à un sujet d'intérêt personnel et d'exprimer leurs idées de façon cohérente.



Essais visuels

Présentation visuelle soutenue du point de vue d'une personne dans laquelle il y a un engagement soutenu à l'égard d'idées.

Comment composer un essai visuel :
<http://virginialynne.hubpages.com/hub/How-to-Write-a-Visual-Essay>

Les essais peuvent s'harmoniser à de multiples résultats d'apprentissage pour l'évaluation des compétences rédactionnelles, des arguments, des aptitudes à la pensée critique, de l'application des idées et des théories et de la démonstration de la capacité de recherche. Ils sont authentiques dans les disciplines qui préparent les étudiants à la vie universitaire dans le domaine des sciences humaines, du journalisme, des carrières littéraires et du discours public. Ils ne sont pas authentiques dans tous les contextes et peuvent nécessiter beaucoup de temps à rédiger et noter. La fiabilité entre évaluateurs et la rigueur exigent une certaine modération ainsi que l'élaboration collaborative d'une grille d'évaluation qu'utiliseront toutes les personnes prenant part à la notation.

Outre les essais, il y a de nombreuses autres formes de communication écrite, qui peuvent toutes être apprises et évaluées si elles conviennent à votre contexte : **rapports, nouvelles, poèmes, notes de service, propositions, mémoires et journaux en ligne**. Dès lors que la forme de rédaction est authentique pour la discipline et qu'elle permet d'évaluer de façon valide les résultats d'apprentissage, les choix sont illimités. Mentionnons aussi que nous ne considérons pas l'une ou l'autre de ces formes de communication écrite comme excluant nécessairement toutes les autres. Les rapports de recherche, par exemple, peuvent se conjuguer à des textes de réflexion et de persuasion – un rapport factuel sur ce qui a été constaté, par exemple, suivi d'une réflexion critique sur le processus de recherche, suivie de la défense d'un train de mesures que l'étudiant croit justifié de par les résultats de la recherche.

Nonobstant la dominance de l'écrit, de nombreux programmes escomptent aussi une communication visuelle, orale ou multimédia, laquelle peut également être évaluée de bien des façons. Bien que l'**évaluation orale** soit tombée en désuétude depuis quelques décennies, son utilisation en éducation est plus ancienne que toute forme de travail par écrit. Elle peut revêtir de nombreuses formes, mais elles ont toutes en commun le fait que les étudiants sont évalués selon leur capacité à répondre aux questions, à débattre ou à converser dans l'instant présent, à voix haute ou en réponse aux questions d'un enseignant ou d'autres étudiants. La plupart des formes d'évaluation orale s'effectuent aussi devant les pairs, et les qualités oratoires sont souvent évaluées par un comité ou un groupe plutôt que par un seul enseignant. On évalue souvent la qualité du discours autant que le contenu qui est communiqué.

Il est possible d'évaluer tout un éventail d'apprentissages oraux, en plus de fournir une occasion d'évaluer les facultés d'adaptation, la réflexion rapide et le calme sous la pression – importantes habiletés une fois hors du milieu universitaire, mais qui sont pourtant difficiles à évaluer par d'autres moyens. Les étudiants reçoivent une rétroaction immédiate de leur auditoire, de façon très semblable à ce qu'ils recevraient en pratiquant les arts de la scène, et cette rétroaction viscérale peut être suivie d'une rétroaction de votre part et des pairs qui n'a pas besoin d'être écrite, ce qui fait gagner du temps. La fiabilité, la rigueur et l'authenticité d'une évaluation orale peuvent être aussi élevées que celles de toute autre forme d'évaluation si l'on prend soin de communiquer les critères prédéfinis et de les utiliser. Le recours à un groupe ou un comité pour noter une évaluation orale peut aider à prévenir à la fois la perception et la réalité d'un jugement biaisé, pour autant que chaque membre suive une formation et fasse preuve de modération afin d'accroître la fiabilité entre évaluateurs.

L'avantage le plus évident, peut-être, est que l'évaluation orale fournit des occasions à ceux qui éprouvent de la difficulté en communication écrite, occasions dont ils ont besoin pour être évalués équitablement et s'assurer que ce soit leurs compréhensions réelles que l'on évalue plutôt que leur aptitude à rédiger. Cela dit, cette évaluation désavantage aussi les étudiants qui ont du mal à parler en public ou ont le trac. L'intimidation et l'anxiété sont des problèmes que certains étudiants ressentent plus que d'autres. Pour cette raison, nous recommandons de pratiquer de petites évaluations orales en classe (plus faciles à effectuer lors de cours ou d'ateliers basés sur la discussion) et, dans certains cas, de permettre aux étudiants de choisir l'évaluation orale parmi d'autres options.

Les exposés/présentations sont très populaires, en particulier dans les cours des années supérieures, où le gros de l'évaluation des résultats d'apprentissage au niveau d'un programme se produit. De nos jours, il s'agit habituellement de présentations multimédias comprenant une composante orale, visuelle, multimédia ou écrite, par exemple, dans des diaporamas et des vidéos. Les exposés ont le potentiel – habituellement non réalisé – d'être véritablement multimédia et donc d'offrir la possibilité d'évaluer

la capacité de l'étudiant à communiquer verbalement, par l'écrit et visuellement en même temps. Les présentations orales peuvent être enregistrées et comparées d'une année à l'autre d'un diplôme pour documenter et évaluer les changements observés au fil d'un programme. Par exemple, l'Alverno College a une longue histoire d'évaluation intentionnelle de l'acquisition de compétences orales au long du programme (Locker, 1981).

La création d'affiches et les **expositions d'affiches** peuvent être un moyen authentique d'évaluer la capacité des étudiants de communiquer des idées scolaires, en particulier dans les programmes qui visent surtout à préparer les étudiants à des carrières universitaires. Créer une affiche bien conçue pour une exposition exige des étudiants qu'ils résumant une tonne d'informations de façon succincte, qu'ils interprètent cette information et qu'ils la communiquent visuellement, puis répondent aux questions à son sujet tout en étant en concurrence avec les autres affiches pour le temps et l'attention du public. Par conséquent, la barre des normes à respecter et des critères d'évaluation à appliquer peut être très haute pour un devoir d'affiche, en particulier s'il est demandé aux étudiants de concevoir leurs affiches en étapes, avec rétroaction, même de la part des pairs, à chaque étape. Les présentations d'affiches dénotent les compétences requises dans bien des domaines des sciences, du génie, des sciences de la santé et des affaires. Avec suffisamment de possibilités d'obtenir de la rétroaction et de la révision, les étudiants peuvent atteindre des résultats d'apprentissage exigeants assujettis à des normes élevées. Les devoirs de synthèse et de représentation alternative d'une information (habituellement) écrite et numérique peuvent à eux seuls être exigeants.

Les devoirs d'affiches peuvent servir à évaluer des individus ou des groupes; ils se prêtent à de la rétroaction formative de la part des enseignants et des pairs, et l'on peut les noter avec fiabilité et rigueur en utilisant des critères relativement intuitifs. Les étudiants ayant du mal à interpréter et communiquer de l'information visuellement peuvent se sentir désavantagés, aussi devrait on tenir compte de considérations semblables à celles de l'évaluation orale, y compris une multiplicité d'évaluations. Il peut aussi s'avérer difficile de veiller à ce que les étudiants se concentrent sur ce qu'ils doivent apprendre et sur l'efficacité avec laquelle ils communiquent la matière enseignée plutôt que sur les aspects plus superficiels du style de présentation.

En déterminant une variété de devoirs arrimés à des cours existants qui permettent d'évaluer les habiletés de communication, on peut évaluer ou communiquer l'acquisition de ces compétences par les étudiants au niveau du programme. Cela est possible lorsque les cours et les devoirs qui s'y rattachent sont harmonisés de façon constructive avec les résultats d'apprentissage du programme. Les données tirées des devoirs liés aux cours peuvent être compilées aux fins de l'évaluation du programme.



Exposition d'affiches

Combine la représentation visuelle et la communication orale, souvent montrée à des pairs et des collectivités d'experts.

Un court guide et des exemples :

<http://www.ucd.ie/t4cms/UCDTLA0039.pdf>

2.4 CRÉATIVITÉ ET CONCEPTION

L'évaluation de la créativité est souvent associée aux arts, mais elle est pertinente pour la plupart – si ce n'est toutes – des disciplines et professions. Il va de soi que, dans les arts, de nombreux programmes sont principalement axés sur la conception, la création, la production et la représentation (exécution). La plupart des scientifiques penseraient d'abord à la créativité scientifique, qui alimente ce qu'ils font. Dans les programmes axés sur la recherche, les étudiants sont censés élaborer des interprétations neuves de l'information, avancer des arguments originaux et/ou créer des façons d'appréhender les problèmes. Lors de l'élaboration des tâches d'évaluation, de nombreuses caractéristiques peuvent être considérées comme des indications de créativité, par exemple la souplesse des interprétations et leur caractère inorthodoxe, l'établissement de liens obscurs mais défendables et le repérage de similarités et différences subtiles entre des phénomènes (Sedlacek, 2004).

La créativité est souvent plombée par des présomptions de mystère et de caractère ineffable qui portent à croire qu'il est difficile, voire impossible, d'effectuer une évaluation efficace (Sedlacek, 2004). Pourtant, il est commun de distinguer, par exemple, entre un travail qui dénote de la créativité et un travail qui se contente d'imiter (mimétisme). Reconnaître la dimension neuve de solutions à des problèmes exige une profonde connaissance du domaine et, du point de vue de l'enseignant, cela exige aussi de comprendre la perspective des étudiants pour être en mesure d'évaluer si le travail revêt un caractère neuf ou créatif compte tenu de ce que l'étudiant pourrait être censé connaître au sujet du problème et des autres tentatives de le résoudre. Un jugement humain réfléchi, perspicace est nécessaire.

La créativité en soi est rarement évaluée seule. Habituellement, on l'évalue comme une dimension d'une autre tâche, qu'il s'agisse d'un devoir de résolution de problèmes, de conception d'affiche, de représentation alternative de l'information (p. ex., un diagramme, une histoire numérique, une bande dessinée ou une vidéo) ou de conception de portfolio. Par exemple, les évaluations qui demandent aux étudiants de **créer des diagrammes de concepts ou d'événements**, de **concevoir des simulations** et de prendre part à des **jeux de rôles** peuvent être de puissants moyens d'évaluation de la créativité, ainsi que des aptitudes à la recherche, de la compréhension de la matière du cours et des habiletés de communication. Toutefois, cela n'est pas nécessaire si vous souhaitez seulement évaluer directement la créativité. Une façon de le faire consiste à demander aux étudiants de **créer un nouveau modèle visuel** pour expliquer une théorie ou un phénomène complexe. Cette sorte d'évaluation incite les étudiants à convertir l'information apprise (habituellement) par le texte et la voix en un médium complètement différent, ce qui permet aux étudiants de reconceptualiser et reformuler des idées et des concepts de façon créative.



Jeu de rôles

Les étudiants jouent un rôle afin d'être capables d'apprendre, souvent en simulant une situation authentique, que ce soit en ligne ou en personne.

EnROLE : Encourager les environnements d'apprentissage en ligne axés sur les rôles :

<http://www.uow.edu.au/cedir/enrole/>

Grille d'évaluation de la créativité de l'AACU : <http://www.aacu.org/value/rubrics/pdf/CreativeThinking.pdf>

2.5 AUTONOMIE ET COMPÉTENCE PROFESSIONNELLE

Pour que les programmes produisent des diplômés compétents prêts à pratiquer dans le « monde réel », deux sous-catégories de résultats devraient être prises en considération : l'application et l'autonomie (ou autodiscipline). L'éventail des résultats possibles d'application n'est circonscrit que par les limites de la profession ou discipline. Appliquer ces limites exige souvent de se livrer à certains types de calculs, d'utiliser des types d'équipement et de suivre certains protocoles et procédures qui sont souvent présentés sous la forme de codes de déontologie ou d'éthique professionnelle. Par exemple, les Sciences infirmières et de la santé utilisent l'examen clinique objectif structuré (ECOS) pour observer les étudiants agissant à titre professionnel (Pugh et Smee, 2013). En revanche, l'autonomie est centrée sur la démonstration d'habitudes que l'on considère comme indispensables à la réussite professionnelle, comme la gestion du temps, l'établissement de buts, le sens de l'organisation et la façon de se présenter.

La **rédaction réflexive (ou réfléchie)** – qu'il s'agisse de journaux d'apprentissage ou de travaux d'essai réfléchi – peut être centrée sur l'attention que les étudiants portent aux détails de leurs propres expériences dans la vie ou relatives au contenu du cours. D'aucuns considèrent cette forme de métacognition comme essentielle au développement de l'autonomie et, par extension, de la pensée critique. Dans un premier temps, les étudiants auront tendance à se concentrer sur ce qui apparaît immédiatement évident, mais, au fil du temps, les travaux de rédaction réflexive peuvent les aider à élargir et approfondir la portée de ce qu'ils remarquent, tandis qu'ils apprennent à déduire du sens de leurs expériences. La répétition de travaux de journal réfléchi peut aider à renforcer certains aspects de l'expérience des étudiants et leur procurer des possibilités jusqu'à ce que de la signification ait été établie.

Pour ces raisons, la rédaction réflexive présente la possibilité d'aider les étudiants à intégrer la matière du cours à leur propre pensée, à établir des liens entre les idées initialement perçues comme isolées et à progressivement voir le cours et la discipline comme étant pertinents à leur propre vie et au monde. Réaliser effectivement ce potentiel est, bien entendu, affaire de détails.

La rédaction réflexive (ou « rédaction réfléchie ») ne devrait pas être confondue avec la rédaction réflexive critique. En rédaction réflexive critique, les étudiants doivent aller au delà de l'établissement de liens et de la création de sens; ils doivent juger leurs propres conclusions et hypothèses en les soumettant à une analyse rigoureuse. Les étudiants dépassent ainsi le stade de la simple reconnaissance de schèmes pour acquérir des aptitudes à la pensée critique, du jugement et des habitudes d'humilité, d'honnêteté et d'intégrité, autant de qualités qui sont importantes dans nombre de disciplines.



Rédaction réflexive

La rédaction réflexive implique de réfléchir à sa propre expérience en établissant des liens et en créant du sens. Dans la rédaction réflexive critique, on dépasse ce stade pour en arriver à juger ses propres conclusions et hypothèses.

Exemples de rédaction réflexive et de rédaction d'un journal personnel de réflexion dans différentes disciplines :
<https://www.aub.edu.lb/ctl/activities/seminars/Documents/2012-13/Selfjournals.pdf>

L'un des avantages des tâches de rédaction réflexive et de rédaction réflexive critique est qu'elles sont faciles à créer et adapter. La plupart des étudiants n'ont besoin que d'une simple invite – souvent une question ou une demande qui peut être trouvée en ligne – pour se mettre en train. Les invites provenant d'une discipline peuvent souvent être adaptées à d'autres avec de légères modifications.

Pour les enseignants, l'un des dangers qu'il y a à évaluer le rendement des étudiants aux tâches de réflexion est de confondre l'évaluation de la rédaction réflexive avec l'évaluation de la pensée réflexive. À l'instar de toute tâche d'évaluation au moyen de laquelle on veut juger d'un état ou processus interne – comme la compréhension, la conscience, la valeur, l'attitude, la créativité et autres – il faut, dans cette évaluation, faire preuve de beaucoup de jugement. En pareils cas, nous tirons des conclusions au sujet de ce qui se produit à l'interne à partir de produits observables publiquement. Les étudiants qui ont du mal à s'exprimer de façon claire par écrit ne sont pas nécessairement moins capables de réfléchir de façon critique, si bien que les critères d'évaluation doivent être soigneusement définis. En apprenant aux étudiants à réfléchir et critiquer au moyen de divers exemples, de la discussion, d'une bonne rétroaction constructive et de possibilités de se réviser, et ce, dans un climat de respect et d'honnêteté, nombre de ceux qui rencontrent des difficultés au début peuvent s'améliorer substantiellement par la suite. Au surplus, ces mêmes outils procurent de multiples occasions d'évaluer la façon dont les étudiants pensent, ce qui peut vous conforter dans vos inférences.

Tout comme les travaux d'essai, les tâches de rédaction réflexive sont chronophages tant pour les étudiants que pour les enseignants, peuvent facilement être traitées superficiellement et sont authentiques dans un éventail étroit de contextes. À l'heure actuelle, elles sont probablement utilisées de façon inauthentique plus souvent qu'autrement, puisque dans bien des cas les liens à la discipline ou à des contextes futurs aux fins d'application ne sont pas posés clairement par les enseignants. Les devoirs que l'on donne à faire aux étudiants pourront alors leur sembler être des « acrobaties » à exécuter pour impressionner. Et la fiabilité entre évaluateurs ainsi que la rigueur de leur notation exigent une modération réfléchie et une élaboration collaborative des grilles de notation.

Dans les programmes professionnels, la norme par excellence de l'évaluation authentique de l'**application** au niveau d'un programme est le **stage**. À proprement parler, un stage représente toute évaluation qui exige des étudiants qu'ils démontrent leur compétence ou leur maîtrise dans un environnement simulé ou du « monde réel ». Les étudiants qui effectuent un stage sont habituellement observés et évalués par une combinaison d'enseignants, de coordonnateurs, de superviseurs et de pairs. Par définition, les stages sont plus susceptibles d'être authentiques que toute autre forme d'évaluation puisqu'on peut les utiliser pour évaluer presque tous les types de résultats d'apprentissage. Ils s'avèrent particulièrement utiles pour l'évaluation des résultats aux stades relationnel et abstrait étendu de la SRAO. Hormis leur évidente propension à l'authenticité, les stages se prêtent bien à une rétroaction immédiate à laquelle on peut donner suite au même moment, ils ont tendance à être faciles à noter, ils sont intrinsèquement formateurs, on peut les adapter à de multiples niveaux d'apprentissage et leur rigueur et leur fiabilité tendent à être élevées. Il n'en demeure pas moins que les stages exigent beaucoup de temps et de ressources, si bien qu'on les réserve souvent à l'élite (les étudiants qui ont tendance à bien performer dans les évaluations traditionnelles) dans de petites classes.

Des **microstages** et des **simulations**, qui sont centrés sur des tâches plus simples et de moindre envergure accomplies dans des approximations de conditions du monde réel, peuvent remplacer les stages/travaux pratiques dans les classes plus grandes ou servir de répétitions à des stages ultérieurs. Dans les grandes classes, les étudiants peuvent soumettre des enregistrements de leur travail pratique dans les simulations.

Les simulations et les microstages sont habituellement effectués dans un environnement plus contrôlé et sont à durée déterminée (par exemple au long de la durée d'un cours). Bien que moins authentiques que les stages complets, ils offrent quand même nombre des mêmes avantages pour l'apprentissage et l'évaluation. On peut créer une séquence ou un ensemble de simulations pour reproduire ce qui pourrait être exigé dans un stage complet. L'une des formes de microstages, qui est courante dans les études en sciences infirmières et que l'on désigne par examen clinique objectif structuré (ECOS), exige des étudiants qu'ils accomplissent une série de tâches sous la pression d'un environnement simulé.



Les stages

Les étudiants démontrent leur compétence dans un environnement simulé ou du monde réel, habituellement observés par d'autres.

Stages d'apprentissage en milieu de travail :

http://casn.berkeley.edu/resource_files/WBL_Definitions_Outcomes_Criteria_pg_120512_v2.pdf

Lorsque l'on conjugue l'évaluation par les pairs et/ou l'autoévaluation aux stages, il faut que les étudiants reçoivent une formation au préalable, à défaut de quoi leur rétroaction pourrait s'en trouver incohérente ou exagérée. Comme on peut s'y attendre, il se peut que la fiabilité d'une évaluation au moyen de stages soit proportionnelle au niveau de structure fourni – moins vous fournissez de structure, plus vous trouverez de disparités dans les résultats d'évaluation. Dans certains cas, ce peut être une situation anticipée et acceptable; dans d'autres, le but peut être d'en arriver à des résultats uniformes ou cohérents.

La création et la coordination de stages et de simulations peuvent nécessiter beaucoup de temps; les stages requièrent plus de coordination, mais il faut moins de temps pour les créer, tandis que les simulations exigent davantage de temps à la création, mais moins à la coordination.

Les **portfolios d'apprentissage**, que ce soit au niveau d'un cours ou au niveau d'un programme, continuent de croître en popularité. Ils requièrent des étudiants qu'ils combinent de multiples éléments de preuve pour démontrer l'atteinte de l'un ou de plusieurs des résultats d'apprentissage escomptés. Ces devoirs sont souvent accomplis au long d'une session ou d'une année scolaire, quoique aux fins de fidélité (la capacité d'assurer que l'on peut se fier à une note comme le reflet exact et authentique de ce qu'un étudiant a atteint), les éléments accomplis au début de la session devraient être révisés et affinés à la lumière de la rétroaction formative. Les portfolios comprennent souvent des essais faisant état de ce que les étudiants ont appris. Les explications convertissent chaque élément inclus en élément de preuve de la réussite, établissant explicitement le lien entre ce degré de réussite et l'atteinte des résultats d'apprentissage escomptés. Au niveau d'un programme, ces éléments devraient être glanés à partir de chacun des cours suivis par l'étudiant, et l'on devrait ensuite interconnecter ces éléments pour être à même de faire l'historique des progrès et de la réussite de l'étudiant.

Les portfolios peuvent être des classeurs ou filières de feuilles de papier, des médias électroniques ou encore une combinaison des deux. Les récents progrès technologiques permettent de créer des portfolios électroniques, qui sont essentiellement des sites Web représentant l'apprentissage de différentes manières à l'intention de différents auditoires et qui sont maintenant beaucoup plus accessibles à l'étudiant moyen qu'ils ne l'étaient il n'y a encore que quelques années. Le numérique permet aux étudiants de puiser à un bien plus grand bassin de sources ou d'artéfacts comme éléments probants de leur apprentissage que ne le leur permettaient les traditionnels systèmes axés sur le papier. Les éléments de preuve peuvent maintenir revêtir la forme de vidéos, d'enregistrements audio, d'œuvres d'art, de photographies, de programmes informatiques, de performances dans des simulations, de performances enregistrées dans des mondes virtuels, de quêtes sur le Web, de versions électroniques de devoirs écrits (y compris la rétroaction reçue), de blogues, d'histoires numériques et de bien d'autres éléments « non traditionnels » pour autant qu'ils soient pertinents et démontrent de façon authentique le degré d'atteinte des résultats d'apprentissage.

En plus de fournir la preuve de l'atteinte de résultats d'apprentissage, les portfolios peuvent efficacement permettre à l'étudiant de développer des habitudes de réflexion critique de l'esprit et la conscience de soi, en particulier s'ils sont introduits au début d'un programme puis régulièrement renforcés à chaque cours. L'accumulation de plusieurs éléments de preuve au fil du temps rend les portfolios plus susceptibles que toute autre forme d'évaluation de fournir un portrait complet et crédible du degré d'atteinte de résultats de l'étudiant. Pour cette même raison, les portfolios peuvent aisément être harmonisés avec les autres expériences et évaluations d'apprentissage, comme les journaux personnels. Bien que chronophages pour les étudiants comme pour les enseignants, les portfolios constituent l'un des plus solides moyens d'évaluer des résultats d'apprentissage au niveau d'un programme en ce qu'ils permettent la collecte et la communication de multiple éléments de preuve sur la période d'un cours et d'un programme. Des produits provenant de nombre des tâches d'évaluation mentionnées dans la présente section peuvent être inclus dans un portfolio, ce qui permet une triangulation des éléments probants, la démonstration d'un changement et une réflexion métacognitive sur les points forts et points faibles en lien avec les résultats d'apprentissage au niveau d'un programme.



Portfolio d'apprentissage

Une collecte systématique d'éléments de preuve de l'atteinte, par l'étudiant, de résultats d'apprentissage, avec une réflexion sur le travail.

Comparaison de portfolios au niveau d'un cours et au niveau d'un programme :

<http://manoa.hawaii.edu/assessment/howto/portfolios.htm>

Ressources du National Institute for Learning Outcome Assessment sur les portfolios :

<http://www.learningoutcomeassessment.org/publications.html>

PARTIE 3

DÉVELOPPER LA CAPACITÉ DE L'ÉTABLISSEMENT

Résumé

Dans cette partie, nous discutons de l'importance de faire en sorte que la culture d'établissement en vienne à valoriser l'évaluation à la fois pour satisfaire aux exigences externes liées à l'agrément et à la reddition de comptes et pour répondre aux exigences internes liées à l'amélioration de l'enseignement et de l'apprentissage. Parmi les éléments nécessaires que nous cernons, nous mentionnons un solide leadership, des ressources suffisantes et la participation du corps professoral. En outre, il faudrait que ce changement de culture se produise parallèlement à une utilisation généralisée de pratiques d'évaluation authentique arrimées aux programmes d'études et propres à des disciplines; et là aussi, la participation des professeurs est vitale. Finalement, nous comblons le fossé entre les évaluations formatives et sommatives en discutant du nouveau domaine de l'analyse de l'apprentissage et de la façon dont les données recueillies au moyen de mesures formatives peuvent servir à mener des évaluations sommatives. La participation des étudiants à tous ces processus est l'une des tendances futures qui visent à améliorer davantage encore la culture de l'établissement dans son approche en matière d'évaluation des résultats d'apprentissage au niveau d'un programme et d'amélioration continue.

3.1 MODIFIER LA CULTURE DE L'ÉTABLISSEMENT ET ACCROÎTRE LE RECOURS AUX PRATIQUES D'ÉVALUATION AUTHENTIQUE

Dans cette partie, nous aborderons les composantes de la culture d'établissement et de l'évaluation authentique. Nous présenterons aussi des exemples d'approches sommatives traditionnelles en évaluation des résultats d'apprentissage au niveau d'un programme, mais le point de mire de notre discussion résidera dans le domaine émergent des outils d'analyse de l'apprentissage et dans la façon dont les données tirées des approches formatives peuvent être agrégées afin de produire une information sommative aux niveaux de l'établissement et d'un programme.

S'engager dans le processus d'évaluer des résultats d'apprentissage au niveau d'un programme et utiliser cette information aux fins d'une amélioration continue est on ne peut plus difficile si l'on ne bénéficie pas de la participation du corps professoral et de l'adhésion des équipes de soutien et de direction de l'administration et de l'établissement. Cela requiert un processus systémique et inclusif de collecte, d'analyse et de partage de données et, il importe de le mentionner, cela se produit dans une culture d'établissement qui valorise l'évaluation et l'amélioration de la qualité. Une culture d'établissement qui considère l'évaluation comme un moyen d'améliorer l'enseignement et l'apprentissage incite les membres du corps professoral à naviguer aussi dans cette direction, ce qui les amène à élaborer et utiliser des instruments d'évaluation authentique et les résultats que ces instruments produisent, ce qui influe directement sur l'apprentissage des étudiants et améliore cet apprentissage (figure 7).

Si des pratiques d'évaluation au niveau du cours et propres aux disciplines sont indispensables pour améliorer l'enseignement et l'apprentissage, elles peuvent néanmoins ne pas produire les données nécessaires à la conduite d'analyses à grande échelle au niveau de l'établissement. Les évaluations au niveau d'un programme fournissent une information cruciale sur l'apprentissage des étudiants à l'issue d'un module de cours et, par là même, permettent d'évaluer l'efficacité du programme et de l'établissement dans son ensemble. Par conséquent, l'évaluation au niveau d'un programme doit se faire parallèlement à des approches formatives afin de répondre aux exigences tant externes qu'internes de l'établissement.

Facteurs influant sur la culture de l'établissement et sur l'évaluation des résultats d'apprentissage au niveau d'un programme

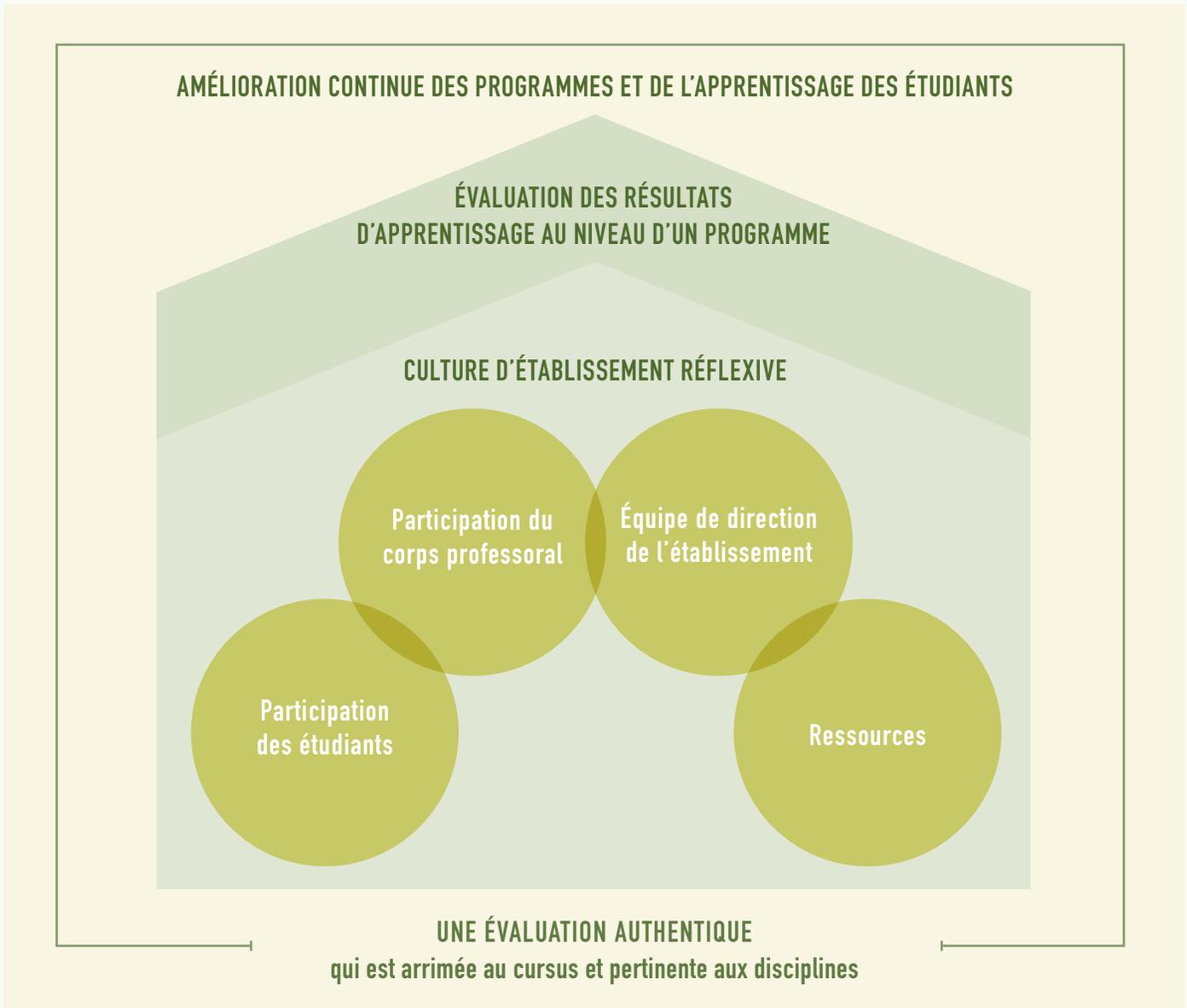


FIGURE 7

CULTURE RÉFLEXIVE DE L'ÉTABLISSEMENT

L'évaluation des résultats d'apprentissage au niveau d'un programme sera plus susceptible de se traduire par une amélioration des programmes, de l'enseignement et de l'apprentissage si elle est étayée par une culture d'établissement axée sur la réflexion. Les processus d'assurance de la qualité que sont les évaluations ne devraient pas simplement servir à répondre à des exigences ou demandes externes, mais aussi nourrir une réflexion sur les actuelles pratiques d'enseignement ainsi que la compréhension et l'amélioration de ces pratiques (Biggs, 2001). Pour édifier efficacement une culture d'établissement qui aborde l'évaluation comme un moyen d'améliorer l'enseignement et l'apprentissage, les résultats d'apprentissage devraient s'harmoniser de façon constructive avec les méthodes d'enseignement en classe, le cursus des programmes et les attributs ou caractéristiques des diplômés de l'établissement (sections 1.1 et 1.2), de même qu'avec les procédures et les politiques de l'établissement (Biggs, 2001).

Au Royaume-Uni (R. U.) et en Australie, les établissements ont officialisé ce traitement réflexif de l'évaluation en élaborant, au niveau de l'établissement, des cadres scolaires et/ou des plans stratégiques d'enseignement et d'apprentissage soutenus par des ressources et mandats externes et gouvernementaux. Au R. U., la Quality Assurance Agency (QAA) et la Higher Education Academy (HEA)²⁰ prônent des pratiques d'évaluation et des procédures d'assurance de la qualité novatrices qui permettent d'examiner comment l'évaluation a été désignée. L'Écosse veille à ce que l'assurance de la qualité soit intentionnellement liée à l'amélioration de la qualité en créant tous les trois ans un nouveau thème autour duquel d'abondantes ressources sont orientées vers la réalisation d'une priorité commune d'amélioration de l'enseignement supérieur.²¹

Baker, Jankowski, Provezis et Kinzie (2012) ont examiné les études de cas d'établissements des États-Unis qui font une utilisation efficace des données d'évaluation pour améliorer constamment leurs programmes et ils ont fait valoir que ces établissements :

1. recourent aux appels à la responsabilisation pour tirer le meilleur parti des efforts d'amélioration qu'ils déploient eux-mêmes à l'interne;
2. harmonisent les travaux d'évaluation aux structures organisationnelles;
3. centrent les efforts d'évaluation sur des questions ou problèmes particuliers afin de permettre un point de mire des données de façon gérable.

LEADERSHIP EN ÉVALUATION

Est-ce que l'évaluation permet d'améliorer l'enseignement et l'apprentissage, et de quelle façon si c'est le cas? La réponse à cette question dépend de la façon dont un établissement appréhende l'évaluation, en d'autres termes, de sa culture de l'évaluation. Si la culture de l'établissement prône et soutient l'évaluation dans le but d'améliorer l'enseignement et l'apprentissage, alors les pratiques d'évaluation se traduiront sans doute par une amélioration de l'enseignement et de l'apprentissage. De solides dirigeants et centres qui se passionnent pour l'enseignement et l'apprentissage sont deux éléments essentiels pour promouvoir une culture d'établissement qui valorise l'évaluation comme outil d'amélioration de l'enseignement et de l'apprentissage.

²⁰ <http://www.qaa.ac.uk/ABOUTUS/Pages/default.aspx>

²¹ <http://www.enhancementthemes.ac.uk/>



Leadership : centres pour l'enseignement et l'apprentissage

Ces centres pourraient offrir ce qui suit :

- faciliter le processus d'élaboration structurée et systématique d'un plan d'évaluation des résultats d'apprentissage au niveau d'un programme;
- développement/perfectionnement professionnel (évaluation des résultats d'apprentissage au niveau d'un programme sur l'arrimage d'une évaluation efficace au cursus);
- des ressources, comme de petites subventions, du temps du personnel, un accès à la littérature/aux travaux publiés et la connaissance de ceux-ci;
- du réseautage entre des personnes intéressées et possédant de l'expertise;
- le partage d'exemples de pratiques efficaces.

Centres d'enseignement et d'apprentissage

La plupart des établissements de l'Ontario ont des centres d'enseignement et d'apprentissage (ou de semblables services intégrés responsables de la qualité de l'enseignement). Au cours des 30 dernières années, ces centres ont évolué pour fournir tout un éventail de services, programmes, ressources et expertise aux professeurs, assistants d'enseignement de premier et deuxième cycles, aux chargés de cours et aux administrateurs scolaires (Grabove et al., 2012), dont beaucoup sont eux-mêmes devenus des chefs de file dans le domaine. Des membres du corps professoral qui ont déjà eu la motivation d'examiner des résultats d'apprentissage au niveau d'un programme peuvent se décourager s'ils ne disposent pas du soutien et des ressources nécessaires (Hersh et Keeling, 2013). Ces centres offrent un soutien en proposant des ateliers ponctuels d'initiation, des programmes enrichis d'une durée d'une semaine et d'une session, des consultations individuelles, du financement pour des subventions d'enseignement et d'apprentissage et l'organisation de conférences. Le personnel des centres fournit un leadership et facilite le processus d'élaboration des résultats d'apprentissage au niveau d'un cours et d'un programme et l'évaluation des résultats d'apprentissage, une représentation aux réunions de comités de l'établissement qui influent sur l'enseignement et l'apprentissage, ainsi que d'autres méthodes de mise en commun d'exemples de pratiques efficaces (Grabove et al., 2012).

Les centres pour l'enseignement et l'apprentissage peuvent fournir ressources et soutien conceptuel aux membres du corps professoral qui participent déjà au processus, tout en encourageant d'autres à se joindre à eux. Au moyen de ces efforts variés et systémiques, les centres peuvent contribuer à une culture d'établissement qui valorise la qualité de l'enseignement et l'évaluation authentique dans le but d'améliorer les programmes et l'apprentissage des étudiants. Nous faisons toutefois remarquer que le personnel des centres ne doit pas être considéré comme les seuls experts; nous recommandons plutôt la création d'un partenariat entre les centres et le personnel enseignant et les départements des universités en vue de développer une meilleure expertise et de faire participer les professeurs au processus d'évaluation, en soutenant un leadership réparti.



Étude de cas : Leadership et centres d'enseignement et d'apprentissage*

Un collège du New Hampshire a financé la création d'un programme systématique de perfectionnement professionnel à l'intention du personnel enseignant à temps partiel appelé l'Adjunct Teaching Forum. Cette tribune a incité les membres du corps professoral à prendre part à des conversations qui lient explicitement les évaluations à l'amélioration des pratiques d'enseignement et d'apprentissage. L'action du Forum se situait sur quatre fronts : développer des aptitudes cognitives de niveau plus élevé; faciliter l'apprentissage expérientiel et actif; repenser les cours; améliorer les stratégies d'évaluation.

Le collège a continué de tirer parti de ce travail en réalisant des recherches sur les stratégies d'évaluation et les pratiques d'enseignement. L'établissement a ensuite amorcé un processus global d'évaluation afin de suivre l'élaboration des programmes, de mesurer le degré de réussite des apprenants, d'améliorer l'enseignement et l'apprentissage et de rendre l'établissement efficace. Avec un financement renouvelé à l'échelle de l'établissement, le programme d'évaluation s'est poursuivi, et a fini par favoriser un leadership chez les professeurs, qui ont établi un processus d'évaluation arrimé aux cours et analysé les programmes d'évaluation en place.

De nombreux facteurs ont contribué à créer une culture d'évaluation à ce collège. Tout d'abord, la mise sur pied d'un comité chargé des évaluations a permis l'aménagement d'une tribune formelle dans laquelle les membres pouvaient examiner l'évaluation à l'établissement, en discuter et l'améliorer. En deuxième lieu, la création de postes de leadership destinés au corps professoral a accru l'engagement, le savoir et la participation du personnel enseignant. En dernier lieu, le rôle qu'on a pris soin de confier aux dirigeants de l'établissement pour qu'ils fassent partie de la discussion a aidé à l'intégration des dimensions d'évaluation aux examens budgétaires, aux évaluations des cours et à la création de programmes.

*Extrait de Zubrow (2012)

Dirigeants d'établissement officiels

Les dirigeants officiels d'établissement peuvent influencer directement sur le succès des efforts d'évaluation déployés dans un établissement (Shipman, Aloi et Jones, 2003). Pour promouvoir et favoriser une culture d'évaluation axée sur l'amélioration de l'apprentissage et de l'enseignement, il faut absolument pouvoir compter sur des dirigeants d'établissement efficaces. Il est important de faire observer que des dirigeants efficaces peuvent déclencher un changement de culture, mais que ce sont les dirigeants successifs partageant des vues semblables qui doivent maintenir ce changement. Des chercheurs ont cerné les qualités et attitudes de leadership qui sont indispensables à la mise en œuvre et au maintien d'une culture réussie d'évaluation (Shipman et al., 2003; Stayhorn, 2006; Baker et al., 2012). Les leaders efficaces :

- proposent une vision pour l'évaluation, en particulier au niveau des programmes et au niveau de l'établissement;
- participent aux processus d'évaluation des résultats d'apprentissage au niveau des programmes;
- encouragent des efforts collaboratifs d'évaluation entre le corps professoral, le personnel, les administrateurs, les étudiants et les employeurs;
- prônent des incitatifs que l'établissement peut offrir pour encourager et reconnaître la participation du corps professoral et du personnel aux efforts d'évaluation;
- opèrent des changements graduels et durables dans les pratiques et politiques d'évaluation;
- connaissent les forces et faiblesses des divers instruments d'évaluation et utilisent cette information pour fournir un cadre permettant d'améliorer l'enseignement et l'apprentissage;
- rendent disponibles des ressources (p. ex., pour le perfectionnement professionnel des membres du corps professoral et du personnel, pour l'évaluation au niveau des programmes et l'amélioration des programmes, ainsi que pour une infrastructure et des systèmes de TI qui permettent une analyse et une visualisation plus efficaces des données, etc.);
- offrent un espace pour la discussion et la collaboration et encouragent ces dernières;
- guident l'établissement pour que l'information soit organisée autour de questions d'intérêt particulières qui aideront à améliorer l'apprentissage des étudiants;
- prévoient du temps et des processus pour mener une réflexion et dégager du sens de l'information recueillie à partir des résultats d'apprentissage au niveau des programmes;
- encouragent la diffusion transparente de l'information, en mettant l'accent sur l'amélioration continue;
- utilisent l'information aux fins des priorités, de la planification stratégique et des décisions de l'établissement;
- célèbrent les réussites.

Les initiatives menées à l'échelle de l'établissement, en particulier au chapitre de la culture, des perceptions et des valeurs, requièrent la participation de représentants et de dirigeants de divers niveaux de l'établissement : membres du corps professoral, personnel administratif et étudiants possédant divers degrés d'expertise (Stayhorn, 2006). Cette approche multidimensionnelle de l'évaluation contribue à instituer une culture d'établissement globale et assure l'accès à d'importantes ressources (Baker et al., 2012).



Dirigeants d'établissement officiels

Un dirigeant efficace se livre à plusieurs activités essentielles :

- établir une vision;
- encourager la collaboration et l'engagement entre le corps professoral, le personnel, les administrateurs, les étudiants et les autres intervenants;
- prévoir une tribune encourageante et du temps pour la discussion et la réflexion;
- prévoir des ressources pour l'amélioration;
- encourager la diffusion;
- reconnaître le temps et les efforts déployés et célébrer les réussites.

RESSOURCES

Amener une culture d'établissement à valoriser l'évaluation exige habituellement une affectation de ressources, du temps, des locaux, du matériel, des politiques et du personnel, ainsi que l'expertise nécessaire pour faciliter le changement. Les études s'intéressant à ce que dépensent les établissements au chapitre de l'efficacité des pratiques d'enseignement font ressortir que ce n'est pas nécessairement le montant total des dépenses qui compte, mais plutôt la **façon** dont les fonds sont **alloués** (Wellman, 2010). Il vaut aussi de considérer que, si le financement de l'évaluation de résultats d'apprentissage au niveau d'un programme représente une dépense dont il faut rendre compte à court terme, cela constitue aussi un investissement qui offre des avantages à long terme de par l'amélioration du programme (Swing et Coogan, 2010). Comme l'ont fait valoir Swing et Coogan (2010), rien ne nuit plus au ratio coûts-avantages que la collecte de données qui ne sont jamais analysées, de sorte que l'on ne peut clore le chapitre de la mise en œuvre des améliorations (p. 12).

Pour utiliser vos ressources de façon optimale, il vous faut un but clairement énoncé, ainsi qu'un schéma et un plan pour l'analyse et l'utilisation des données avant d'entamer l'évaluation des résultats d'apprentissage au niveau d'un programme (section 1.3). Les ressources nouvelles ou réaffectées pourraient comprendre des congés pour activités professionnelles aux responsables de l'évaluation au niveau du programme, des consultations à l'externe, des logiciels pour recueillir, stocker et analyser les données ainsi que faciliter la représentation visuelle des données recueillies et des incitatifs à la participation pour les étudiants et le corps professoral (Cooper et Terrell, 2013). En outre, la façon sensée de veiller à ce que l'on dispose de suffisamment de ressources consiste à discuter de l'évaluation de résultats d'apprentissage au niveau d'un programme pendant les phases de planification, de budgétisation et d'examen des cursus (Shipman et al., 2003).

PARTICIPATION DU CORPS PROFESSORAL

Le soutien apporté par les centres pour l'enseignement et l'apprentissage, les chefs de campus et les ressources contribue à une culture qui valorise l'évaluation, et ces efforts doivent être complétés à leur tour par une adhésion et une participation du corps professoral (Hutchings, 2010). Une évaluation formelle peut devenir l'élan initial d'une évaluation des programmes; toutefois, il est indispensable de favoriser une culture qui valorise l'enseignement, l'apprentissage et l'évaluation pour améliorer l'apprentissage sur une base continue. Les universitaires ont constaté qu'une culture d'établissement qui valorise l'évaluation incite le corps professoral à prendre part à ce processus d'évaluation : plus simplement dit, les membres du corps professoral sont plus susceptibles de participer à une évaluation authentique si l'établissement prend aussi un engagement public et stratégique à l'égard de cette évaluation (Wang et Hurley, 2012). À la base, ce sont les membres du corps professoral qui fournissent l'expertise et le savoir propres aux disciplines qui sont nécessaires pour élaborer des instruments d'évaluation authentique, et ils sont indispensables à la mise en œuvre réussie de pratiques d'évaluation authentique (Baker et al., 2012; Van Dyke, 2013). À l'opposé, plusieurs facteurs peuvent dissuader les professeurs de prendre part à ces types d'initiatives (Hutchings, 2010), comme les suivants :

- l'évaluation est vue comme faisant partie de la culture de la direction de l'établissement;
- les membres du corps professoral ne sont pas formés à l'évaluation;
- souvent, l'enseignement, de façon générale, de même que le travail d'évaluation, est sous-évalué dans les processus d'embauche, d'avancement professionnel et d'obtention de la permanence;
- il n'y a pas de données probantes qui attestent indiscutablement que de participer à un processus d'évaluation des résultats d'apprentissage au niveau d'un programme se traduit par une amélioration de l'apprentissage chez les étudiants

Pour tenter de dissiper nombre de ces préoccupations, nous trouvons particulièrement efficace de considérer l'évaluation comme une activité universitaire qui est entreprise dans le but d'améliorer l'enseignement et l'apprentissage.



Ressources

Amener une culture d'établissement à valoriser l'évaluation exige habituellement une affectation de ressources. Cela comprend :

- *des centres pour l'enseignement et l'apprentissage;*
- *une formation professionnelle sur l'utilisation et l'évaluation des données;*
- *Integration of assessment information into revised curriculum, policies and practices;*
- *des incitatifs aux membres du corps professoral, comme du financement, des congés pour activités professionnelles;*
- *des logiciels pour l'analyse des données;*
- *du soutien au financement pour les projets et réseaux qui utiliseront l'information recueillie.*

L'évaluation en tant qu'activité universitaire

On peut accentuer la disposition des membres du corps professoral à prendre part à une évaluation des résultats d'apprentissage au niveau des programmes si on leur fait percevoir l'évaluation comme une activité intellectuelle ou savante relative à la mission professorale, une activité pouvant prendre [traduction] « la forme d'un bien communautaire dont on peut discuter, que l'on peut critiquer et échanger et dont on peut tirer parti » (Schulman, 1993). Schulman (1999) fait observer que quelque chose devient [traduction] « savant lorsque cette chose devient publique, un objet d'évaluation et d'examen critique par les membres de sa communauté, et lorsque les membres de cette communauté commencent à utiliser, mettre à profit et développer ces fruits de l'esprit et actes de création. » Si les professeurs perçoivent l'évaluation comme une entreprise intellectuelle, une dimension de leur rôle en tant que membres de l'établissement et un processus pouvant produire des données et un changement effectif, ils seront plus susceptibles de participer à ce processus et de le valoriser.

Évaluation dans le but d'améliorer l'enseignement et l'apprentissage

Les membres du corps professoral sont plus susceptibles de participer à des efforts d'évaluation des résultats d'apprentissage au niveau d'un programme s'il est clairement établi que les résultats peuvent influencer directement sur leurs propres pratiques d'enseignement, sur le programme et sur l'apprentissage des étudiants (Wang et Hurley, 2012). Des activités d'évaluation qui sont principalement entreprises à des fins de production de rapports ou d'agrément auront un effet dissuasif sur la volonté de participation des professeurs. Lorsque l'on inverse le point de mire de l'évaluation en le faisant passer d'une initiative venant d'en haut à une mesure à laquelle participent de multiples intervenants qui travaillent ensemble à l'amélioration de l'apprentissage des étudiants, on facilite une culture de l'inclusion, du leadership réparti et d'enseignement de qualité.

Faire participer le corps professoral : approches pratiques

Nous avons inclus trois approches pratiques que les établissements peuvent retenir pour aider le corps professoral à voir l'évaluation comme une activité savante qui peut améliorer l'enseignement et l'apprentissage.

1. Perfectionnement professionnel

Les centres pour l'enseignement et l'apprentissage peuvent offrir des ateliers et programmes sur l'évaluation authentique en expliquant en détail comment des mesures efficaces d'évaluation peuvent améliorer l'enseignement et l'apprentissage, comment concevoir et évaluer des résultats d'apprentissage et comment concevoir et mettre en œuvre des instruments d'évaluation, des grilles d'évaluation et des méthodes de collecte de données et d'analyse. Alors qu'une évaluation arrimée au département ou à la discipline peut être utile pour créer un contexte, les établissements du R. U. ont constaté qu'en créant des réseaux et des possibilités (Hutchings, 2010) de partager des idées avec le corps professoral venant d'autres disciplines et départements on pouvait fournir une tribune utile permettant d'apprendre les uns des autres et d'échanger des idées (Eales-Reynolds, communication personnelle, 2004). En tant que service central à l'échelle du campus, les centres pour l'enseignement et l'apprentissage peuvent organiser et soutenir ces initiatives.

Certains établissements peuvent aussi offrir des occasions de perfectionnement professionnel aux étudiants diplômés : de plus en plus, les départements offrent des programmes d'études avec options pour participer à des cours avec crédit ou des programmes traitant de sujets tels que l'enseignement et l'apprentissage, la conception de cursus/programmes d'études, l'évaluation et l'harmonisation (voir p. ex., le programme universitaire de certificat en enseignement de l'Université de Windsor, à l'adresse <http://www1.uwindsor.ca/ctl/utc>).



Étude de cas : *Virginia Polytechnic Institute et State University*
*Ateliers de perfectionnement du corps professoral**

Ray Van Dyke (2013) décrit le succès des ateliers graduels de perfectionnement du corps professoral qui ont été offerts à la Virginia Polytechnic Institute and State University. Plus particulièrement, il explique comment les ateliers d'initiation ont couvert des exigences externes de production de rapports tout en facilitant la satisfaction d'exigences internes d'amélioration. Il aborde en outre la façon dont, dans ces ateliers propres à une tâche, l'évaluation a été appréhendée comme une activité savante, en reconnaissant que les membres du corps professoral ont été plus susceptibles de prendre part à l'évaluation en la considérant à partir des perspectives qui leur sont propres et de leur plein gré.

* Extrait de Van Dyke (2013)

2. Reconnaissance de l'évaluation comme une forme de mission professorale : subventions et permanence

Si nous considérons l'évaluation comme une forme de mission professorale et tenons cela pour vrai de par nos structures de renouvellement et de récompenses au niveau de l'établissement, les membres du corps professoral seraient enclins à prendre part à des processus d'évaluation. Les établissements pourraient instaurer des politiques et procédures reconnaissant et récompensant les enseignants qui participent à une évaluation au niveau d'un programme. Ils peuvent aussi encourager la participation des professeurs en offrant du financement ou un soutien aux demandes de subventions à l'externe pour des projets examinant des méthodes efficaces d'évaluation des résultats d'apprentissage au niveau d'un programme, en affectant temps et ressources à la mise au point de techniques d'évaluation et en offrant aux membres du corps professoral des récompenses et des possibilités de publier leurs conclusions pour les mettre en commun. Ces changements démontrent l'engagement de l'établissement à l'égard de l'évaluation en vue d'améliorer l'enseignement et l'apprentissage ainsi que sa volonté d'instaurer et de maintenir une culture d'évaluation.



Étude de cas : *Instaurer un corps professoral « dirigeant »**

Un collège public multicampus du New Hampshire a mis sur pied un comité de professeurs « dirigeants » à qui l'on a confié la tâche de diriger les efforts d'évaluation avec le soutien d'un bureau central des évaluations. Les professeurs ont été sélectionnés dans différentes disciplines et différents lieux géographiques, et ont chacun reçu une formation sous différentes formes (en ligne et en classe). Ils étaient responsables de concevoir des instruments et des grilles d'évaluation dont se serviraient les équipes de professeurs chargées de noter les artefacts des étudiants, d'interpréter les données et de faciliter des discussions inter- et intrafacultaires sur les évaluations.

*Extrait de Zubrow (2012)

3. Leadership chez les membres du corps professoral

Les membres du corps professoral peuvent aussi assumer diverses fonctions de leadership dans le processus d'évaluation. Les chercheurs ont constaté qu'il était déterminant de retenir directement les services de professeurs dans les efforts d'évaluation, en particulier dans des postes de dirigeants, pour susciter la participation des autres membres du corps professoral (Zubrow, 2012).



Étude de cas : Valencia Community College

L'évaluation des résultats d'apprentissage comme compétence essentielle d'enseignement

Contexte : Chaque année, le Valencia Community College dispense à plus de 50 000 étudiants des programmes d'enseignement dans ses cinq campus de la région d'Orlando, en Floride. Sa feuille de route émaillée de succès en matière de résultats d'apprentissage, de degré de réussite des étudiants, d'équité et diversité, et d'effets sur le marché du travail lui a valu de remporter le prix Aspen d'excellence en enseignement collégial communautaire, décerné pour la première fois en 2011. [Traduction] « Les résultats au plan de la diplomation et de la main-d'œuvre sont clairs et particulièrement impressionnants lorsque l'on sait que la population étudiante du collège Valencia est variée, qu'une forte proportion d'élèves arrivent avec des lacunes à combler et qu'ils sont également nombreux à provenir de ménages à faible revenu [...] Des experts nationaux et visiteurs sur place ont attribué la note « excellent » aux processus d'évaluation et d'amélioration des résultats d'apprentissage au collège Valencia, et ils ont été tout particulièrement impressionnés par le programme de perfectionnement professionnel que le collège destine aux professeurs. »¹

Compétences essentielles pour les enseignants du Collège Valencia : L'un des aspects distinctifs de l'excellence de Valencia réside dans la définition des *compétences essentielles* à l'intention de son corps professoral et des autres enseignants. Tous les nouveaux membres du corps professoral doivent démontrer qu'ils possèdent ces compétences pendant le processus de transition qui passe d'un poste à durée limitée à la permanence. Un vaste programme² est proposé pour soutenir les professeurs dans leur perfectionnement et la démonstration des *compétences essentielles*. Ces compétences comprennent les méthodes d'enseignement et d'évaluation centrées sur l'apprentissage, la facilitation de l'atteinte de résultats à l'échelle de l'établissement et de l'acquisition de compétences liées au perfectionnement professionnel par les étudiants, ainsi qu'un engagement à réaliser des études savantes et professionnelles en enseignement et en apprentissage. Il y a notamment une compétence propre à la *pratique axée sur les résultats*, laquelle comprend la capacité d'« utiliser les données et éléments probants sur l'apprentissage des étudiants pour réexaminer et améliorer cours et programmes ».³ Cinq cours composés d'ateliers portent spécifiquement sur l'évaluation des programmes.⁴

Exemples de plans d'évaluation et d'amélioration des programmes : L'Institutional Assessment Office (bureau d'évaluation des établissements) du Collège Valencia possède une bibliothèque en ligne⁵ dans laquelle on peut trouver de nombreux exemples de la façon dont les programmes ont évalué des résultats d'apprentissage particuliers des étudiants et les améliorations qui en ont résulté. Les programmes d'enseignement donnés en exemple comprennent la danse, l'hygiène dentaire, les services d'urgence médicale, la production de films, l'infographie et les sciences infirmières. Le collège tient aussi de nombreux répertoires de grilles d'évaluation adaptées par les divers programmes,⁶ des instruments de collecte de données et des échantillons,⁷ et des ressources et documents de référence pour les résultats de programme des établissements, comme les communications interpersonnelles⁸ et la responsabilité éthique.⁹

¹ <http://www.aspeninstitute.org/sites/default/files/content/docs/pubs/AspenPrize021312.pdf> p. 7-8

² <http://valenciacollege.edu/faculty/development/>

³ <http://valenciacollege.edu/faculty/development/programs/tla/Candidate/documents/EssentialCompetenciesCurrentrevised9-11.pdf> p. 3

⁴ <http://valenciacollege.edu/academic-affairs/institutional-effectiveness-planning/institutional-assessment/saicc/FacultyWorkshops.cfm>

⁵ <http://valenciacollege.edu/academic-affairs/institutional-effectiveness-planning/institutional-assessment/LOA/PlanLibrary.cfm>

⁶ <http://valenciacollege.edu/academic-affairs/institutional-effectiveness-planning/institutional-assessment/LOA/RubricLibrary.cfm>

⁷ <http://valenciacollege.edu/academic-affairs/institutional-effectiveness-planning/institutional-assessment/LOA/DataLibrary.cfm>

⁸ <http://valenciacollege.edu/academic-affairs/institutional-effectiveness-planning/institutional-assessment/loa/ResourcesInterpersonalPLO.cfm>

⁹ <http://valenciacollege.edu/academic-affairs/institutional-effectiveness-planning/institutional-assessment/loa/ResourcesEthicalResponsibilityPLO.cfm>

3.2 ÉVALUATION AUTHENTIQUE

ÉVALUATION ARRIMÉE AU PROGRAMME D'ENSEIGNEMENT

Dans les premières initiatives d'évaluation de résultats d'apprentissage au niveau d'un programme, on a tenté de distinguer ce processus d'évaluation de l'habituelle évaluation des cours effectuée au moyen de tests normalisés. Les mesures d'évaluation normalisées sont souvent isolées à la fois des enseignants et du cursus et n'invitent pas à la participation du corps professoral ni ne facilitent une amélioration significative de l'enseignement et de l'apprentissage (Rhodes, 2012). Elles exigent un surcroît de travail de la part des étudiants et leur valeur (aux yeux des étudiants) est souvent difficile à justifier. Qui plus est, on commence à douter de plus en plus que l'évaluation générale normalisée puisse prendre précisément la mesure d'un apprentissage qui est contextualisé dans une discipline (Barrie et al., 2012).

Les méthodes d'évaluation qui sont arrimées aux programmes d'enseignement et pratiquées par le corps professoral harmonisent les résultats d'apprentissage au niveau d'un programme avec les résultats d'apprentissage déjà existants au niveau des cours et intègrent ces activités d'évaluation aux exigences de cours (Cummings, Maddux et Richmond, 2008). Ces approches arrimées au cursus présentent la valeur ajoutée d'une évaluation qui se produit à l'intérieur du contexte de la discipline. Ainsi, l'évaluation est réalisée en parallèle puis harmonisée avec les activités d'enseignement et d'apprentissage. Ces approches intégrées présentent les avantages suivants :

1. Participation directe du corps professoral

Les instruments d'évaluation ancrés au cursus placent l'évaluation et les améliorations qui en découlent dans l'enseignement et l'apprentissage directement dans les mains des enseignants (Garretson et Golson, 2005). Les professeurs tiennent un rôle central dans l'élaboration et la mise en œuvre de l'évaluation dans les cours. Par la suite, les données que permettent d'obtenir ces instruments d'évaluation sont adaptées aux enseignants et abordent un enseignement et un apprentissage propres aux enseignants et aux cours.

2. Des étudiants motivés

Les traditionnelles évaluations normalisées s'effectuent hors de l'environnement d'enseignement et d'apprentissage, si bien que non seulement elles sont isolées du cursus, mais en plus elles ne portent pas vraiment à conséquence pour les étudiants. Les instruments d'évaluation arrimés au cursus, en revanche, évaluent les artefacts que les étudiants produisent comme résultat direct de l'enseignement des professeurs en classe ou en ligne. Les étudiants sont donc plus motivés à bien faire dans ces tâches d'évaluation puisqu'elles sont déjà intégrées dans les activités de cours existantes (Cummings et al., 2008; Rhodes, 2012).

3. Une affectation de ressources minimisée

Le surcroît de temps que l'on requiert des professeurs pour administrer les instruments et recueillir les données est minimisé, vu que les activités d'évaluation sont incorporées aux exigences de cours existantes (Cummings et al., 2008). Les membres du corps professoral peuvent traiter les besoins en notation et en évaluation sans avoir à y consacrer énormément de temps et d'énergie, et ce, tout en pouvant procurer au département et à l'établissement de substantiels résultats d'évaluation détaillés (Hardt, 2010).

4. Implications pour l'enseignement et l'apprentissage

Une approche d'évaluation arrimée au cursus est menée par le corps professoral et explicitement lie aux programmes d'enseignement; ainsi, les résultats d'évaluation que l'on acquiert sont directement pertinents pour l'amélioration de l'enseignement et de l'apprentissage, et ils permettent de cerner clairement les besoins et lacunes des programmes d'enseignement et de les traiter (Cummings et al., 2008). De façon significative, l'évaluation de multiples artefacts produits par les étudiants au fil du temps fournit une mesure à valeur ajoutée qui permet de s'assurer de l'efficacité des actuelles pratiques d'enseignement et d'apprentissage (McCarthy, Niederjohn et Bosack, 2011). Qui plus est, l'évaluation des programmes au moyen d'instruments d'évaluation arrimés aux cours permet aux enseignants de suivre le cheminement d'apprentissage des étudiants (Garretson et Golson, 2005) et d'ajuster l'enseignement afin d'améliorer l'apprentissage.

5. Implications pour l'évaluation des programmes

En ce qui touche l'évaluation des programmes, les membres du corps professoral sont censés mener des activités d'évaluation parallèlement à leur charge de travail normale. Des contraintes de temps et les ressources limitées des départements limitent ainsi leur participation aux efforts d'évaluation (Hardt, 2010). Un processus d'évaluation arrimé aux programmes d'enseignement comportant des cours individuels peut procurer d'autres moyens de conduire une évaluation globale des programmes tout en y faisant participer les professeurs (Cummings et al., 2008). Les membres du corps professoral examinent souvent leurs cours isolément plutôt que dans le cadre d'une initiative plus large et sont moins réceptifs aux instruments d'évaluation des résultats d'apprentissage au niveau d'un programme. En amalgamant les instruments d'évaluation arrimés au cursus et aux cours, on devrait harmoniser les objectifs des cours aux résultats d'apprentissage au niveau des programmes et donc établir la façon dont chaque cours appuie l'atteinte de buts plus généraux au niveau des programmes. Non seulement cette approche encourage-t-elle la collaboration et la communication entre membres du corps professoral, mais aussi les résultats acquis au moyen des instruments d'évaluation arrimés aux cours deviennent-ils pertinents aux efforts d'évaluation des programmes.

6. Souplesse

Comme les efforts d'évaluation arrimés aux programmes d'enseignement sont guidés par les enseignants, cette approche procure de la souplesse quant au style pédagogique, au contenu des cours et au style d'évaluation, et elle peut être mise en œuvre dans diverses disciplines (Garretson et Golson, 2005).

L'efficacité des instruments d'évaluation arrimés au cursus pour améliorer l'enseignement et l'apprentissage tant au niveau des cours qu'au niveau des programmes est déterminée par les deux facteurs suivants :

Harmonisation adéquate des résultats d'apprentissage individuels au niveau des cours et des résultats d'apprentissage au niveau des programmes

Les résultats d'évaluation provenant d'instruments arrimés à des cours individuels comportent des implications directes pour l'évaluation des programmes lorsque les résultats d'apprentissage individuels au niveau d'un cours sont harmonisés à des résultats d'apprentissage plus généraux au niveau d'un programme. L'atteinte de résultats au niveau d'un cours peut ainsi correspondre à l'atteinte de résultats au niveau d'un programme. Les données obtenues des mesures d'évaluation formative peuvent ainsi être utilisées dans les efforts d'évaluation sommative; cette dimension est traitée plus en détail à la section 3.4.

Intégration appropriée des activités d'évaluation aux exigences de cours existantes

Les activités d'évaluation arrimées au cursus devraient être intégrées à des cadres de programmes d'enseignement déjà existants pour minimiser la charge de travail et l'investissement en temps des membres du corps professoral, et obtenir des résultats d'évaluation pertinents quant à l'amélioration de l'enseignement et de l'apprentissage.

Entre autres instruments d'évaluation arrimés à un cursus qui permettent d'évaluer des résultats d'apprentissage au niveau d'un programme, mentionnons les portfolios des étudiants, les projets collectifs, les simulations et les projets de fin d'études (décrits à la section 2), lesquels sont évalués au moyen de grilles.



Étude de cas : *Évaluation arrimée au cursus Portfolios**

Rhoda Cummings et al. (2008) décrivent l'incorporation d'une compilation de portfolios d'étudiants dans un cadre d'évaluation au niveau des programmes, à un établissement de Land Grant, dans l'ouest des États-Unis. Aux étudiants qui s'inscrivent à ce programme, il est demandé de tenir un portfolio d'artefacts qu'ils sélectionnent eux-mêmes et qui démontrent six domaines de compétences professionnelles avancées dans le champ d'études. Au moins deux membres du corps professoral évaluent les portfolios à mi-parcours et à la fin pour déterminer si les artefacts inclus sont représentatifs des résultats d'apprentissage escomptés. Les faiblesses dans les portfolios sont relevées et l'on y remédie en rencontrant les étudiants et en recommandant des changements à apporter ou en acceptant d'autres présentations d'artefacts.

Les auteurs attribuent la réussite de telles évaluations ancrées au cursus à une familiarisation croissante avec les pratiques exemplaires pour améliorer tant les résultats individuels de l'étudiant que l'efficacité des programmes. En outre, l'harmonisation des résultats d'apprentissage individuels au niveau des cours et des résultats d'apprentissage au niveau des programmes s'est traduite par une plus forte participation des professeurs et des étudiants et par de judicieuses modifications apportées aux programmes. L'intégration des évaluations aux cadres existants a permis de minimiser la sollicitation des ressources du corps professoral et de maintenir l'engagement des professeurs à l'égard des efforts d'évaluation.

*Extrait de Cummings et al. (2008)

ÉVALUATION PROPRE AUX DISCIPLINES

Les études ont montré que la connaissance du contenu est essentielle pour démontrer la possession des compétences en pensée critique et en résolution de problèmes (Banta et Pike, 2012). Les instruments d'évaluation authentique requièrent des étudiants qu'ils établissent des liens entre les aptitudes et compétences générales et les connaissances et contextes relatifs aux disciplines qu'ils ont acquis pendant leurs spécialisations et leurs cours reliés aux disciplines. Dans les bons instruments d'évaluation axés sur les disciplines, il est tenu compte de ces contextes. L'intégration d'une évaluation des programmes axée sur les disciplines aboutit à une compréhension meilleure et plus cohérente de la façon dont chaque cours est lié aux résultats d'apprentissage au niveau des programmes et, subséquemment, à l'harmonisation des techniques d'évaluation de cours aux résultats au niveau d'un programme. Les résultats au moyen de mesures d'évaluation propres aux disciplines sont sensiblement plus bénéfiques au corps professoral et peuvent donc plus facilement s'appliquer aux efforts déployés pour améliorer l'enseignement et l'apprentissage (Banta et Pike, 2012).



Étude de cas : le système PACT (*Performance Assessment for California Teachers*) et les *Embedded Signature Assessments* (ESA – évaluations signature intégrées)

Le Performance Assessment for California Teachers (PACT) désigne un ensemble de programmes de préparation des enseignants dispensé dans 30 universités, un programme de stages de district et un réseau d'écoles à charte. Ces établissements ont collectivement mis au point une évaluation du rendement des enseignants et, en Californie, il faut se soumettre à une évaluation du rendement de l'enseignement pour obtenir un titre préalable d'enseignement de plusieurs matières ou d'une matière.

Les évaluations signature intégrées (ESI) sont des tâches d'évaluation ancrées à un ou plusieurs cours. Les responsables de programmes individuels définissent des devoirs clés dans leur cursus et les développent en évaluations signature intégrées; par exemple, l'Université de Californie à San Diego a élaboré une ESI axée sur le contexte social de la classe. L'ESI a recueilli des données probantes dans le temps et dans plusieurs cours. D'autres établissements ont élaboré des ESI intégrées à un seul cours; entre autres exemples, mentionnons une étude communautaire, une observation de la gestion de la salle de classe, une étude de cas sur un enfant et une unité des programmes d'études.

3.3 TENDANCES FUTURES

ÉTABLIR UN PONT ENTRE L'ÉVALUATION SOMMATIVE ET L'ÉVALUATION FORMATIVE : OUTILS D'ANALYSE D'APPRENTISSAGE

Instituer une culture d'établissement qui soutient les résultats d'apprentissage au niveau d'un programme et leur utilisation aux fins d'amélioration continue consiste en partie à utiliser plus efficacement les liens existant entre les évaluations sommative et formative.

Traditionnellement, les évaluations sommatives sont perçues comme des tâches qui mesurent l'apprentissage de l'étudiant à l'issue d'une unité d'enseignement. Le plus souvent, elles sont administrées à la fin d'un cours ou d'un programme, et l'évaluation de l'apprentissage de l'étudiant repose habituellement sur une mesure de la performance de l'étudiant par rapport à des normes ou repères prédéterminés (Harlen et James, 1997). Toutefois, il importe de noter que, dans la pratique, les tâches sommatives et formatives forment une sorte de continuum et qu'une même tâche d'évaluation peut fournir de l'information tant pour les évaluations sommatives que pour les formatives (Harlen et James, 1997). Ainsi, il se pourrait que la distinction ne soit pas nécessairement entre une tâche sommative et une tâche formative mais plutôt entre une information sommative et une information formative : les tâches d'évaluation qui examinent *comment* l'apprentissage se produit fournissent une information formative, tandis que les tâches d'évaluation qui déterminent *si* un apprentissage s'est produit fournissent une information sommative. Au nombre des tâches courantes d'évaluation qui fournissent une information sommative, mentionnons par exemple les tests normalisés, les examens finaux, les projets finaux, les travaux de session et les portfolios cumulatifs des étudiants. Non seulement l'information sommative permet-elle aux établissements de situer les progrès de l'apprentissage des étudiants – ce qui est important pour les fins d'agrément et de production de rapports –, mais aussi est-elle nécessaire pour prendre des décisions de grande ampleur au sujet du contenu des programmes et améliorer globalement l'efficacité de l'établissement, des programmes et des départements (Benjamin et al., 2012).

Évaluation formative utilisée à des fins sommatives

Les approches formatives d'évaluation peuvent fournir une information sommative. Dans leur discussion de l'évaluation, Harlen et James (1997) définissent les conditions suivantes d'une utilisation fiable de l'information et des procédures d'évaluation formative pour conduire une évaluation sommative :

- a) [Traduction] « [L'information formative] est scrupuleusement examinée par rapport aux critères que les étudiants sont censés remplir à certains âges/stades. »
 - On devrait se servir de critères externes pour déterminer comment l'on peut utiliser de façon sommative des données recueillies à des fins formatives.
- b) [Traduction] « On applique les critères de façon holistique en portant des jugements sur ce qui convient le mieux. »
 - Les données formatives devraient être examinées et harmonisées avec des critères appliqués uniformément. Il faut également comprendre que les données formatives recueillies ne satisferont pas toutes aux critères susmentionnés; par conséquent, les données formatives ne peuvent pas toutes être utilisées à des fins sommatives.
- c) [Traduction] « Il y a des façons de s'assurer que les jugements passés par un enseignant soient comparables à ceux des autres enseignants. »
 - Une approche possible de normalisation des jugements entre les enseignants consiste à discuter collectivement des jugements qu'ils ont portés sur les ensembles de travaux des étudiants.

Ainsi qu'il ressort des conditions circonscrites par Harlen et James, l'utilisation sommative de données formatives exige une harmonisation adéquate. La simple collecte de données d'évaluation formative ne suffit pas à procéder à une évaluation sommative; une telle évaluation exige plutôt une sélection scrupuleuse, du jugement et la prise en compte de l'information formative. De la même façon, l'évaluation à grande échelle de résultats d'apprentissage au niveau d'un programme peut se faire par l'harmonisation des résultats d'apprentissage au niveau des cours à ceux au niveau des programmes.

Évaluation sommative utilisée à des fins formatives

À l'inverse, des approches sommatives d'évaluation peuvent aussi fournir une information formative et servir à des fins formatives. Par exemple, dans sa discussion de l'évaluation formative, Taras (2009) se penche sur l'utilisation formative de l'information sommative : les enseignants peuvent utiliser les résultats sommatifs obtenus à partir d'examens, de projets et de tests d'unité pour cerner les points faibles du rendement des étudiants et ainsi cibler leurs efforts subséquents d'enseignement.

Analyse de l'apprentissage et évaluation sommative

Dans une entrevue portant sur les actuelles innovations en évaluation, Ahmad a cité le domaine émergent de l'analyse de l'apprentissage comme présentant beaucoup de potentiel pour la conduite d'évaluations sommatives (communication personnelle, 29 novembre 2013). L'analyse de l'apprentissage consiste à [traduction] « mesurer, recueillir, analyser et communiquer des données sur les apprenants et leurs contextes afin de comprendre et d'optimiser l'apprentissage et les environnements dans lesquels il se produit » (Mattingly, Rice et Berge, 2012). Les données recueillies peuvent servir à prédire l'efficacité de cours et de programmes. Au moyen de la modélisation statistique et prédictive, l'analyse de l'apprentissage peut évaluer de grands ensembles de données non seulement pour communiquer des résultats sommatifs, mais aussi pour cerner des schèmes de données afin de formuler des recommandations sur l'amélioration de l'apprentissage. Les systèmes de gestion de l'apprentissage (SGA), les systèmes de gestion du contenu (SGC) et les systèmes de gestion de contenu d'apprentissage (SGCA) simplifient ce processus et le rendent plus uniforme.



Étude de cas : Analyse de l'apprentissage à l'Université de Wollongong (Australie)*

La Graduate School of Medicine (GSM – École supérieure de médecine) de la University of Wollongong se sert de l'analyse de l'apprentissage pour recueillir et analyser des données sur les stages cliniques des étudiants pendant leur formation à l'école de médecine. L'outil élaboré permet aux étudiants de consigner leurs expériences et a permis aux enseignants de mieux aider les étudiants à intégrer leurs expériences dans le programme d'études.

La GSM utilise un système de gestion de contenu d'apprentissage appelé Equella pour recueillir et stocker des données pertinentes sur l'enseignement et l'apprentissage, y compris des renseignements démographiques sur les patients, des études de cas liées au programme d'études et les emplacements des stages des étudiants. Il vaut de noter que les données ainsi acquises sont utilisées à des fins tant sommatives que formatives. Equella permet aux enseignants de suivre le degré d'implication des étudiants et l'assurance qu'ils déclarent eux mêmes pendant les stages, tandis que les étudiants cernent les auto performances qui leur manquent et les actes subséquents. À l'opposé, les administrateurs se servent des données sommatives pour assurer la qualité du programme d'enseignement.

*Extrait de Mattingly et al. (2012)

PARTICIPATION DES ÉTUDIANTS

Une autre tendance importante de l'amélioration de l'évaluation des résultats d'apprentissage au niveau d'un programme réside dans la mobilisation des étudiants dans ce processus (Barrie et al., 2009; Hutchings, 2010). Il ressort d'une revue des pratiques en Australie que l'enseignement des attributs souhaités chez les diplômés ne fonctionnera vraisemblablement pas, à moins que les étudiants ne perçoivent le processus comme quelque chose de valable. Barrie et al. (2009) suggère que la mobilisation des étudiants constitue l'un des huit éléments critiques qui influent sur les efforts déployés par un établissement pour renouveler le cursus et atteindre les résultats d'apprentissage au niveau de l'établissement ou les attributs souhaités chez les diplômés. En outre, Hutchings (2010) nous rappelle que si les enseignants peuvent ne pas être certains de vouloir prendre part à l'évaluation de résultats d'apprentissage au niveau d'un programme, ils se soucient quand même beaucoup de l'apprentissage des étudiants. Faire participer les étudiants pourrait en fait être le moyen le plus efficace de faire participer les enseignants. Nous avons constaté, dans les interactions de groupes de réflexion, que les explications que les étudiants fournissent des résultats d'évaluation peuvent être éclairantes. On observe aussi une tendance à l'intégration de la recherche et de l'enseignement dans les expériences de recherche des étudiants de premier cycle. Faire participer les étudiants à des projets évaluant l'apprentissage est une façon phénoménale d'intégrer l'apprentissage, l'enseignement et la recherche pour les étudiants, d'instaurer une culture d'établissement plus complète qui se soucie de l'apprentissage et procède à une amélioration continue.



Participation des étudiants

Les tendances futures mobiliseront les étudiants en leur donnant la possibilité d'être des collaborateurs et des cochercheurs :

- groupes de réflexion chargés de recommander des formes d'évaluation;
- groupes de réflexion chargés d'expliquer les résultats des évaluations;
- membres d'une équipe de recherche chargée d'examiner les résultats d'apprentissage au niveau d'un programme;
- membres d'une équipe de projet chargée d'appliquer de nouvelles idées;
- participation à des réseaux, des comités et des débats ouverts.

RÉSUMÉ

Dans ce manuel, nous avons discuté de la façon dont vous pouvez améliorer et enrichir votre programme en évaluant les résultats d'apprentissage au niveau des programmes. Dans la Partie 1, nous avons abordé l'authenticité, la validité et la fiabilité des pratiques et principes d'évaluation et avons décrit comment vous pouvez concevoir des évaluations en gardant ces éléments à l'esprit. En mettant à profit les concepts théoriques et cadres établis, nous avons décrit la façon dont l'on devrait harmoniser les résultats d'apprentissage au niveau des cours aux résultats d'apprentissage plus vastes au niveau des programmes. La Partie 1 s'est conclue sur un exposé de la façon dont vous pouvez créer un plan d'évaluation des résultats d'apprentissage au niveau d'un programme.

Dans la Partie 2, nous avons appliqué le cadre théorique décrit au début de ce manuel. Plus précisément, nous avons décrit une variété de tâches d'évaluation, lesquelles étaient catégorisées par résultat d'apprentissage possible. On sélectionnerait puis incorporerait ces tâches à un plan au niveau des programmes, en choisissant les tâches d'évaluation convenant le mieux aux résultats d'apprentissage d'un programme donné. Nous avons souligné l'importance de créer des tâches d'évaluation bien conçues et clairement communiquées aux étudiants, et ce, dans un environnement qui favorise et soutient l'apprentissage. Dans cette partie, nous avons décrit une gamme de tâches d'évaluation et les façons dont on peut les appliquer aux résultats d'apprentissage. En les rangeant dans de vastes catégories — pensée critique, résolution de problèmes, jugement

et perspicacité; travaux de recherche et des redditions; communication; créativité et conception; autonomie et compétence professionnelle —, nous avons décrit comment l'on peut harmoniser les diverses tâches d'évaluation aux objectifs d'apprentissage des cours et, à terme, comment on peut les utiliser pour mesurer le degré d'atteinte des objectifs d'apprentissage au niveau d'un programme.

Dans la dernière partie, nous avons discuté de l'importance d'instituer une culture d'établissement qui valorise l'évaluation à la fois pour satisfaire aux exigences externes liées à l'agrément et à la production de rapports et pour répondre aux exigences internes liées à l'amélioration de l'enseignement et de l'apprentissage. Plus précisément, nous avons cerné, comme composante nécessaire à ce changement de culture dans les établissements, un solide leadership, des ressources suffisantes et la participation des professeurs. Finalement, nous avons comblé le fossé séparant l'évaluation formative de l'évaluation sommative en discutant du domaine émergent de l'analyse de l'apprentissage et des façons dont les données recueillies au moyen de mesures formatives peuvent servir à mener une évaluation sommative des objectifs d'apprentissage d'un programme.

Nous espérons que ce manuel vous aidera à opérer ce changement de culture dans votre établissement, et ce, au moyen de l'élaboration d'une évaluation authentique des résultats d'apprentissage au niveau des programmes.

TRAVAUX CITÉS

Altbach, P. G., Reisberg, L. et Rumbley, L. E. (2009), *Trends in global higher education: Tracking an academic evolution*, Organisation des Nations Unies pour l'éducation, la science et la culture (UNESCO). En ligne : <http://www.uis.unesco.org/Library/Documents/trends-global-higher-education-2009-world-conference-en.pdf>

Angelo, T. (1999), Doing assessment as if learning matters most, *American Association for Higher Education*. En ligne : <http://www.aahea.org/aahea/articles/angelomay99.htm>

Angelo, T. A. et Cross, K. P. (1993), *Classroom assessment techniques: A handbook for college teachers*, deuxième édition, San Francisco, CA, Jossey-Bass.

Angelo, T. A. (1995), « Reassessing and defining assessment », *American Association for Higher Education Bulletin*, vol. 48(2), pp. 7-9.

Association of American Colleges and Universities [AACU] (2009), The VALUE project overview, *Peer Review: Emerging Trends and Key Debates in Undergraduate Education*, vol. 11(1), pp. 4-7.

Baker, G. R., Jankowski, N. A., Provezis, S. et Kinzie, J. (2012), *Using assessment results: Promising practices of institutions that do it well*, Urbana, IL., National Institute for Learning Outcomes Assessment.

Banta, T. W. et Pike, G. R. (2012), « The bottom line: Will faculty use assessment findings? » in C. Secolsky et D. B. Denison (éditeurs), *Handbook on measurement, assessment, and evaluation in higher education* (pp. 47-56), New York, Routledge.

Barrie, S., Hughes, C., Crisp, G. et Bennison, A. (2011), *AAGLO Summary 1: The ALTC AAGLO project and the international standards agenda*. En ligne : http://www.itl.usyd.edu.au/projects/aaglo/pdf/AAGLO_Summary%201_Final.pdf

Barrie, S., Hughes, C., Crisp, G. et Bennison, A. (2012), *Assessing and assuring Australian graduate learning outcomes: Principles and practices within and across disciplines*, rapport final. En ligne : http://www.itl.usyd.edu.au/projects/aaglo/pdf/SP10-1879_FINAL%20sydney%20barrie%20final%20report%20part%201.pdf

Barrie, S., Hughes, C. et Smith, C. (2009), *The national graduate attributes project: Integration and assessment of graduate attributes in curriculum*, Australian Learning and Teaching Council. En ligne : <http://www.itl.usyd.edu.au/projects/nationalgap/resources/gappdfs/national%20graduate%20attributes%20project%20final%20report%202009.pdf>

Benjamin, R., Miller, M. A., Rhodes, T. L., Banta, T. W., Pike, G. R. et Davies, G. (2012), « The seven red herrings about standardized assessments in higher education », publication hors série numéro 15 du NILOA, Urbana, IL., National Institute for Learning Outcomes Assessment. En ligne : <http://learningoutcomesassessment.org/documents/HerringPaperFINAL.pdf>

Biggs, J. B. et Collis, K. (1982), *Evaluating the quality of Learning: The SOLO Taxonomy*. New York, Academic Press.

Biggs, J. (2001), « The reflective institution: Assuring and enhancing the quality of teaching and learning », *Higher Education*, vol. 41(3), pp. 221-238.

Biggs, J. et Tang, C. (2007), *Teaching for quality learning at university*, troisième éd., Berkshire, R. U., McGraw Hill.

Biggs, J. et Tang C. (2011), *Teaching for quality learning at university*, Maidenhead, R. U., McGraw-Hill et Open University Press.

Blaich, C. et Wise, K. (2011), *From gathering to using assessment results: Lessons from the Wabash National Study*, Urbana, IL., National Institute for Learning Outcomes Assessment. En ligne : http://www.learningoutcomeassessment.org/documents/Wabash_001.pdf

- Boud, D. et Associates. (2010), *Assessment 2020: Seven propositions for assessment reform in higher education*, Sydney, Australian Learning and Teaching Council.
- Brown, S., Rust C. et Gibbs, G. (1994), *Strategies for diversifying assessment*, Oxford Centre for Staff Development, Oxford, R. U., Oxford Brookes University.
- Burn, C. L., Nestel, D. et Gachoud, D. (2013), « Program innovations abstract: Simulated patient as co-facilitators: Benefits and challenges of the interprofessional team OSCE simulation in healthcare », *Journal of the Society for Simulation in Healthcare*, vol. 8(6), pp. 455-456.
- Chalmers, D. (2007), *A review of Australian and international quality systems and indicators of learning and teaching*, Australia: Carrick Institute for Learning and Teaching in Higher Education. En ligne : <http://www.olt.gov.au/resource-review-indicators-teaching-learning-2007>
- Christensen, C. R. et Hansen, A. J. (1987), « Teaching with cases at the Harvard Business School », in C. R. Hansen et A. J. Hansen (éd.), *Teaching and the case method: Text, cases, and readings* (pp. 34-68), Boston, MA: Harvard Business School. En ligne : <http://www.olt.gov.au/resource-review-indicators-teaching-learning-2007>
- Collis, K. F. et Biggs, J. (1986), « Using the SOLO taxonomy », *Set: Research Information for Teachers*, vol. 2(4).
- Conseil des universités de l'Ontario (2013), *Ensuring the value of university degrees in Ontario*. En ligne : <http://www.cou.on.ca/publications/reports/pdfs/ensuring-the-value-of-university-degrees-in-ontari>
- Conseil ontarien des affaires collégiales (1990), *Vision 2000, quality and opportunity: The final report of Vision 2000, a review of the mandate of Ontario's Colleges*, Toronto, ministère de la Formation et des Collèges et Universités. En ligne : <http://www.thecouncil.on.ca/download/10938>
- Cooper, T. et Terrell, T. (2013). *What are institutions spending on assessment? Is it worth the cost?*, Urbana, IL., National Institute for Learning Outcomes Assessment. En ligne : <http://learningoutcomeassessment.org/documents/What%20are%20institutions%20spending%20on%20assessment%20Final.pdf>
- Cresswell, J. W. (2009), *Research design: Qualitative, quantitative, and mixed methods approaches*, troisième éd., Thousand Oaks, CA, Sage.
- Cummings, R., Maddux, C. et Richmond, A. (2008), « Curriculum-embedded performance assessment in higher education: maximum efficiency and minimum disruption », *Assessment & Evaluation in Higher Education*, vol. 33(6), pp. 599-605.
- Deepwell, F. et Buckley, C. (éd.) (2013), *SEDA small grants: Celebrating the scholarship of educational development*, SEDA Special 35.
- Dunne, D. et Brooks, K. (2004), *Teaching with cases*, Halifax: Society for Teaching and Learning in Higher Education.
- Ewell, P. T. (2009), *Assessment, accountability, and improvement: Revising the tension*. NLOA Occasional Paper 1. Urbana, IL: National Institute for Learning Outcomes Assessment. En ligne : http://www.learningoutcomeassessment.org/documents/PeterEwell_005.pdf
- Fredericksen, N. (1984), « The real test bias: Influences of testing on teaching and learning », *American Psychologist*, vol. 39, pp. 193-202.
- Garretson, H. et Golson, E. (2005), « Synopsis of the use of course-embedded assessment in a medium sized public university's general education program », *The Journal of General Education*, vol. 54(2), pp. 139-149.
- Gibbs, G. (1992), *Improving the quality of student learning*, Bristol, R.-U., Technical and Education Services, Ltd.
- Grabove, V., Kustra, E., Lopes, V., Potter, M. K., Wiggers, R. et Woodhouse, R. (2012), *Les centres d'enseignement et d'apprentissage : leur rôle en évolution dans les collèges et universités de l'Ontario*, Toronto, Conseil ontarien de la qualité de l'enseignement supérieur. En ligne : <http://www.heqco.ca/SiteCollectionDocuments/TL%20Centres%20ENG.pdf>
- Gulikers, J. T. M., Bastiaens, T. J. et Kirschner, P. A. (2004). « A five-dimensional framework for authentic assessment », *Educational Technology Research and Design*, vol. 53, pp. 67-87.
- Hardt, J. (2010), *Killing two birds with one stone: Using rubrics to make the grading and assessment of political science capstones easier*, Rochester, NY, Social Science Research Network. En ligne : <http://papers.ssrn.com/abstract=1555928>
- Harlen, W. et James, M. (1997), « Assessment and learning: Differences and relationships between formative and summative assessment », *Assessment in Education: Principles, Policy & Practice*, vol. 4(3), pp. 365-379.
- Harris, K. (2009), *International trends in establishing the standards of academic achievement in higher education: An independent report and analysis*, Australian Universities Quality Agency. En ligne : http://www.cshe.unimelb.edu.au/research/teaching/docs/Intl_trends_standards_Harris_Aug09.pdf
- Healey, M. (2005), « Linking research and teaching: Exploring disciplinary spaces and the role of inquiry-based learning », in R. Barnett (éd.), *Reshaping the university: New relationships between research, scholarship and teaching* (pp. 67-78). Maidenhead, R. U., McGraw Hill/Open University Press.
- Hersh, R. H. et Keeling, R. P. (2013), *Changing institutional culture to promote assessment of higher learning*, Urbana, IL., National Institute for Learning Outcomes Assessment. En ligne : <http://www.learningoutcomeassessment.org/documents/occasionalpaperseventeen.pdf>
- Hutchings, P. (2010), *Opening doors to faculty involvement in assessment*. Urbana, IL., National Institute for Learning Outcomes Assessment. En ligne : <http://www.learningoutcomeassessment.org/documents/PatHutchings.pdf>
- Kaupp, J., Frank, B. et Watts, C. (2013), *Evaluation of software tools supporting outcomes-based continuous program improvement processes*, présenté aux actes de la conférence de l'Association canadienne de l'éducation en génie en 2013 (ACEG13), document 57, pp. 1-8. En ligne : <http://egad.engineering.queensu.ca/wp-content/documents/Papers/software%20tools%20evaluation.pdf>
- Kenny, N. et Desmarais, S. (2012), *A guide to developing and assessing learning outcomes at the University of Guelph*, Office of the Associate Vice-President (Academic), University of Guelph. En ligne : http://www.uoguelph.ca/vpacademic/avpa/outcomes/KennyDesmarais_LearningOutcomesGuide_2012.pdf
- Kopera-Frye, K., Mahaffy, J. et Svare, G. M. (2008), « The map to curriculum alignment and improvement », *Collected Essays on Teaching and Learning*, vol. 1, pp. 8-14.
- Kuh, G. D., Jankowski, N., Ikenberry, S. O. et Kinzie, J. (2014), *Knowing what students know and can do: The current state of student learning outcomes assessment in U.S. colleges and universities*, Urbana, IL., National Institute for Learning Outcomes Assessment. En ligne : <http://www.learningoutcomeassessment.org/documents/2013%20Survey%20Report%20Final.pdf>
- Lennon, M. C. (2014), *Piloting the CLA in Ontario*. Toronto: Higher Education Quality Council of Ontario. En ligne : <http://www.heqco.ca/SiteCollectionDocuments/AHELO%20The%20Ontario%20Experience-ENG.pdf>

- Lennon, M. C. et Jonker, L. (2014), *AHELO : L'expérience de l'Ontario*, Toronto, Conseil ontarien de la qualité de l'enseignement supérieur. En ligne : <http://www.heqco.ca/SiteCollectionDocuments/AHELO%20The%20Ontario%20Experience-ENG.pdf>
- Loacker, G. S. (1981), *Alverno College's program in developing and assessing oral communication skills*, Paper presented at the annual meeting of the Speech Communications Association, Anaheim, CA.
- Maki, P. L. (2004), *Assessing for learning: Building a sustainable commitment across an institution*, Sterling, VA, Stylus Publishing.
- Maki, P. L. (2010), *Assessing for learning: Building a sustainable commitment across the institution*, deuxième éd., Sterling, VA, Stylus Publishing.
- Martini, T. S. et Clare, M. (2014), *Undergraduates' understanding of skill-based learning outcomes: Can e-portfolios help?* Toronto: Higher Education Quality Council of Ontario. En ligne : <http://www.heqco.ca/SiteCollectionDocuments/e-Portfolios%20ENG.pdf>
- Mattingly, K. D., Rice, M. C. et Berge, Z. L. (2012), « Learning analytics as a tool for closing the assessment loop in higher education », *Knowledge Management & E-Learning: An International Journal (KM&EL)*, vol. 4(3), pp. 236 247.
- Mazurat, R. et Schönwetter, D. (2009) « Electronic curriculum mapping: Supporting competency-based dental education », *Journal of the Canadian Dental Association*, vol. 74(10), pp. 886 889.
- McCarthy, M. A., Niederjohn, D. M. et Bosack, T. N. (2011), « Embedded assessment a measure of student learning and teaching effectiveness », *Teaching of Psychology*, vol. 38(2), pp. 78 82.
- McMichael, K. (2009), *Transforming the humanities classroom for the 21st century: Using course redesign to enhance student success*. Reno, NV, Bent Tree Press.
- Mentkowski, M. (2006), « Accessible and adaptable elements of Alverno student assessment-as-learning: Strategies and challenges for peer review », in C. Bryan et K. Clegg (éd.), *Innovative assessment in higher education* (pp. 48 63), London, R. U., Taylor and Francis.
- MTCU (2006a), *History of program standard*. En ligne : <http://www.tcu.gov.on.ca/pepg/audiences/colleges/progstan/history.html>
- MTCU (2006b), *What Does a Program Standard Contain?* En ligne : <http://www.tcu.gov.on.ca/pepg/audiences/colleges/progstan/contain.html>
- Nightingale, P., Wiata, I. T., Toohey, S., Ryan, G., Hughes, C. et Magin, D. (1996), *Assessing learning in universities professional development centre*, University of New South Wales, Australie.
- Ontario Council of Academic Vice-Presidents [OCAV] (2007), *Guidelines for university undergraduate degree level expectations*. En ligne : <http://www.cou.on.ca/publications/reports/pdfs/university-undergraduate-degree-level-expectations>
- Ontario Universities Council on Quality Assurance (2012), *Quality assurance framework*. En ligne : <http://oucqa.ca/resources-publications/quality-assurance-framework/>
- Petti, L. (2013), « Program innovations abstract: Development of an assessment rubric to measure clinical competencies in physician assistant education simulation in healthcare », *Journal of the Society for Simulation in Healthcare*, vol. 8(6), p. 481.
- Potter, M. K. et Kustra, E. (2012), « A primer on learning outcomes and SOLO taxonomy ». En ligne : <http://www1.uwindsor.ca/ctl/system/files/PRIMER-on-Learning-Outcomes.pdf>
- Prodromou, L. (1995), « The backwash effect: From testing to teaching ». *ELT Journal*, vol. 49, pp. 13 25.
- Pugh, D. et Smee, S. (2013), *Guidelines for the development of objective structured clinical examination (OSCE) cases*. Ottawa, Medical Council of Canada.
- Rhodes, T. L. (2012), « Show me the Learning: Value, accreditation, and the quality of the degree », *Planning for Higher Education*, vol. 40(3), pp. 36 42.
- Sadler, D. R. (2013), « Making competent judgements of competence », in S. Blömeke, O. Zlatkin-Troitchanskaia, C. Kuhn et J. Fege (éd.), *Modeling and measuring competencies in higher education* (pp. 13 27), Netherlands: Sense Publishers.
- Sedlacek, W. E. (2004), *Beyond the big test: Noncognitive assessment in higher education*. San Francisco, CA, Jossey-Bass.
- Service de l'assurance de la qualité des collèges de l'Ontario (2014), *Credentials validation*. En ligne : <http://www.ocqas.org>
- Shipman, D., Aloj, S. L. et Jones, E. A. (2003), Addressing key challenges of higher education assessment, *The Journal of General Education*, vol. 52(4), pp. 335 346.
- Shulman, L. S. (1993), « Teaching as community property », *Change*, (nov./déc.), pp. 6 7.
- Shulman, L. S. (1999), « Taking learning seriously », *Change*, (juillet/août), pp. 10 17.
- Strayhorn, T. L. (2006), *Frameworks for assessing learning and development outcomes*. Washington, DC, Council for the Advancement of Standards in Higher Education. Disponible (pour achat) auprès du Council for the Advancement of Standards in higher Education, à http://www.cas.edu/store_product.asp?prodid=33
- Swing, R. L. et Coogan, C. S. (2010), *Valuing assessment: Cost-benefit considerations*. NILOA Occasional Paper 5. Urbana, IL., National Institute for Learning Outcomes Assessment.
- Taras, M. (2009), « Summative assessment: The missing link for formative assessment », *Journal of Further and Higher Education*, vol. 33(1), pp. 57 69.
- Tremblay, K., Lalancette, D. et Roseveare, D. (2012), *Assessment of higher education learning outcomes AHELO feasibility study report, volume 1, Design and Implementation*, Paris, Organisation de coopération et de développement économiques. En ligne : <http://www.oecd.org/edu/skills-beyond-school/AHELOFSReportVolume1.pdf>
- Uchiyama, K. P. et Radin, J. L. (2009), « Curriculum mapping in higher education: A vehicle for collaboration », *Innovative Higher Education*, vol. 33, pp. 271 280.
- Van Dyke, R. (2013), « The scaling-up of assessment in one institution », *European Journal of Higher Education*, vol. 3(1), pp. 102 113.
- Wang, X. et Hurley, S. (2012), « Assessment as a scholarly activity?: Faculty perceptions of and willingness to engage in student learning assessment », *Journal of General Education*, vol. 61(1), pp. 1 15.
- Wellman, J. V. (2010, janvier), *Connecting the dots between learning and resources*. NILOA Occasional Paper 3. Urbana, IL., National Institute for Learning Outcomes Assessment.
- Wiggins, G. P. et McTighe, J. (2005), *Understanding by design*. Alexandria, VA, Association for Supervision and Curriculum Development.
- Zubrow, J. (2012), « Case study: Engaging adjunct faculty in program assessment », *Journal of Assessment and Institutional Effectiveness*, vol. 2(1), pp. 77 102.

AUTRES RÉFÉRENCES

ÉVALUATION

Bresciani, M. J., Gardner, M. M. et Hickmott, J. (2012), *Demonstrating student success: A practical guide to outcomes-based assessment of learning and development in student affairs*. Sterling, VA, Stylus Publishing.

Strayhorn, T. L., Creamer, D. G., Miller, T. K. et Arminio, J. L. (2006), *Frameworks for assessing learning and development outcomes*. Washington, DC, Council for the Advancement of Standards in Higher Education.

Gardner, W. S., Aloï, S. L. et Lusher, A. L. (2003), « A framework for assessing general education outcomes within the majors », *The Journal of General Education*, vol. 52(4), pp. 237-252.

Goos, M., Gannaway, D. et Hughes, C. (2011), « Assessment as an equity issue in higher education: Comparing the perceptions of first year students, course coordinators, and academic leaders », *The Australian Educational Researcher*, vol. 38(1), pp. 95-107.

McKittrick, S. A. et Barnes, S. M. (2012), « Assessment of critical thinking: An evolutionary approach », *Journal of Assessment and Institutional Effectiveness*, vol. 2(1), pp. 1-29.

THÉORIE ET PRATIQUES

Banta, T. W., Griffin, M., Flateby, T. L. et Kahn, S. (2009), *Three promising alternatives for assessing college students' knowledge and skills*. Urbana, IL., National Institute of Learning Outcomes Assessment.

Clouder, L., Broughan, C., Jewell, S. et Steventon, G. (2013), *Improving student engagement and development through assessment: Theory and practice in higher education*, New York, Routledge.

Meyers, N. M. (2009), « How to use (five) curriculum design principles to align authentic learning environments, assessment, students' approaches to thinking and learning outcomes », *Assessment & Evaluation in Higher Education*, vol. 34(5), pp. 565-577.

APPRENTISSAGE ET RÉSULTATS D'APPRENTISSAGE

Allais, S. (2012), « Claims vs. practicalities: Lessons about using learning outcomes », *Journal of Education and Work*, vol. 25(3), pp. 331-354.

Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J. et Wittrock, M. C. (2001), *A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of educational objectives*. Edited by L. W. Anderson et D. R. Krathwohl. New York, Longman.

Gilchrist, D. et Oakleaf, M. (2012), *An essential partner: The librarian's role in student learning assessment*. Urbana, IL., National Institute for Learning Outcomes Assessment.

Klein, S. P. (2005), « An approach to measuring cognitive outcomes across higher education institutions », *Research in Higher Education*, vol. 46(3), pp. 251-276.

Prineas, M. et Cini, M. (2011), *Assessing learning in online education: The role of technology in improving student outcomes*. Urbana, IL., National Institute for Learning Outcomes Assessment.

Sum, P. E. et Light, S. A. (2010), « Assessing student learning outcomes and documenting success through a capstone course », *PS: Political Science & Politics*, vol. 43(3), pp. 523-531.

CONTEXTE DE L'ÉTABLISSEMENT

Bers, T. H. (2008), « The role of institutional assessment in assessing student learning outcomes », *New Directions for Higher Education*, vol. 141, pp. 31-39.

Healey, M. (2010), « Reflections on engaging students in the process and product of strategy development for learning, teaching, and assessment: An institutional case study », *International Journal for Academic Development*, vol. 15(1), pp. 19-32.

Jankowski, N. et Makela, J. P. (2010), *Exploring the landscape: What institutional websites reveal about student learning outcomes assessment activities*. Urbana, IL., National Institute for Learning Outcomes Assessment.

Volkwein, J. F. (2011), *Gaining ground: The role of institutional research assessing student outcomes and demonstrating institutional effectiveness*. Urbana, IL., National Institute of Learning Outcomes Assessment.

1 rue Yonge, Bureau 2402
Toronto, Ontario M5E 1E5
Téléphone : 416-212-3893
Télécopie : 416-212-3899
Courriel : info@heqco.ca
Site Web : heqco.ca



Conseil ontarien
de la qualité de
l'enseignement supérieur