



L'évaluation pour l'apprentissage à l'éducation supérieure : qu'en est-il?

Myriam Girouard-Gagné

Doctorante en sciences de l'éducation, Université de Montréal

Pour citer cet article :

Girouard-Gagné, M. (2021). L'évaluation pour l'apprentissage à l'éducation supérieure : qu'en est-il? *Initio*, 9 (1).

Résumé

L'évaluation aux études supérieures est, par tradition, axée sur la performance et les résultats. Or, la massification aux études supérieures amène une variété d'apprenant·es et de programmes professionnels qui remettent en question l'évaluation des connaissances par écrit telle qu'on la connaît. Une revue systématique de la littérature a été menée pour faire état de la recherche sur les pratiques vécues dans les universités à travers le monde. Il en ressort que bien que les pratiques d'évaluation pour l'apprentissage soient bénéfiques, les pratiques d'évaluation traditionnelles sont encore majoritairement utilisées et les enseignant·es universitaires ont une compréhension partielle de l'évaluation pour l'apprentissage. Des pistes de recherche sont ensuite proposées dans une perspective de développement professionnel des enseignant·es universitaires.

Mots-clés

Évaluation formative, évaluation pour l'apprentissage, éducation supérieure, enseignants universitaires, pratiques évaluatives

Summary

Assessment in universities has traditionally been performance and outcome oriented. However, the massification in higher education brings a variety of learners and professional programs, which question the written assessment of knowledge as known. A systematic review of the literature was conducted to report research on assessment practices experienced in universities around the world. It shows that while assessment for learning is beneficial, traditional assessment (assessment of learning) is still predominantly used and university teachers have a partial understanding of assessment for learning. Research perspectives are then proposed in relation with the professional development of university teachers.

Keywords

Formative evaluation, assessment for learning, higher education, university teachers, assessment practices



Introduction

Les universités dans le monde accueillent une clientèle d'étudiant·es de plus en plus nombreuse et diversifiée (Detroz *et al.*, 2020) comprenant des étudiant·es vivant avec un handicap ou des difficultés émergentes, ayant des repères culturels, une culture scolaire et des codes différents, ainsi que des étudiant·es à la fois travailleur·euses et parents. À Montréal particulièrement, ajoutons une forte augmentation des étudiant·es internationaux·ales depuis 2017, observée par le Bureau de coopération interuniversitaire. Plus récemment, un article de Fortier dans le journal *Le Devoir* (2018) relatait le recensement de plus de 14 652 étudiant·es ayant des besoins particuliers à l'université. Par ailleurs, les universités sont reconnues comme étant des institutions de savoirs et de formation de niveau supérieur. Les individus y sont formés pour innover et faire avancer le monde vers un avenir prometteur. Cette réalité vécue conséquemment à la massification aux études supérieures amène les professeur·es et les chargé·es de cours universitaires à repenser leurs pratiques évaluatives. En effet, il est nécessaire que le processus d'évaluation des apprentissages soit effectué pour soutenir les apprenant·es.

On reconnaît généralement deux rôles à l'évaluation : l'évaluation certificative (ou de l'apprentissage) pour rendre compte des acquis à la fin d'un parcours, et l'évaluation formative (ou pour l'apprentissage) pour soutenir les apprenant·es et réguler l'enseignement (Cizek *et al.*, 2019). La fonction d'évaluation en tant qu'apprentissage [*assessment as learning*] (Earl, 2012) est aussi présente dans la littérature. Néanmoins, nous adoptons une vision élargie de l'évaluation pour l'apprentissage, ou l'évaluation-soutien d'apprentissage [ÉsA] (Allal et Laveault, 2009), incorporant ces différentes fonctions dans une approche intégrée d'enseignement et d'évaluation dans laquelle les apprenant·es, les enseignant·es et les pairs-apprenant·es sont engagé·es. Cette approche s'intéresse tant à la façon d'apprendre qu'à l'état de l'apprentissage au moment actuel dans une visée de soutien à l'apprentissage. Elle s'effectue dans un environnement d'apprentissage où les apprenant·es sont actif·ives et indépendant·es, où ils et elles peuvent s'autoévaluer et évaluer leurs pairs par le truchement d'un processus de collecte de données et d'interprétation de celles-ci permettant de faire le point sur les acquis et d'établir les étapes à franchir pour atteindre les objectifs d'apprentissage (Sambell *et al.*, 2012; Klenowski, 2009; Swaffield, 2011; Allal et Laveault, 2009; Carless, 2015; Black *et al.*, 2004). Les pratiques principalement reconnues comme étant des façons de faire l'évaluation-soutien d'apprentissage sont : 1) la clarification des intentions et des critères de succès, 2) l'utilisation de la discussion en classe pour questionner et rendre explicite la compréhension des apprenant·es, 3) la rétroaction pour permettre aux apprenant·es de s'améliorer, 4) rendre les étudiant·es utiles pour les autres par la collaboration et l'évaluation par les pairs, et 5) rendre les étudiant·es responsables de leurs apprentissages par l'utilisation de l'autoévaluation (Black et Wiliam, 2009; Marshall et Drummond, 2006; Black *et al.*, 2004; Bove-Deegan, 2015; Wiliam, 2011; Wiliam et Thompson, 2017; Black et Wiliam, 2018). Dans une perspective où les étudiant·es universitaires possèdent une diversité de caractéristiques et que l'université vise la formation d'individus compétents et novateurs, l'ÉsA semble représenter une approche à privilégier dans l'enseignement universitaire pour faire progresser les étudiant·es et les amener à mobiliser des compétences complexes, tout en prenant en considération leurs caractéristiques diverses.

Toutefois, les façons de faire l'ÉsA sont nombreuses et variées. De plus, il est pertinent de se questionner sur la place qu'occupe ce type d'évaluation dans les institutions universitaires, car bien que la massification aux études supérieures soit un fait documenté, les changements de pratique dans les milieux prennent souvent plus de temps à être effectués. Ainsi, cet article présente une recension effectuée dans le but de répondre à la question suivante : « Comment l'ÉsA est-elle mise en œuvre dans l'enseignement universitaire? » Dans un premier temps, la démarche de recension empruntée sera exposée pour faire état, dans un deuxième temps, des apports et des limites des recherches recensées.

Revue systématique de la littérature sur l'évaluation-soutien d'apprentissage à l'université

La recension des écrits a pour principal objectif de faire état de la recherche sur un sujet donné (Rhoades, 2011) de manière « rigoureuse, transparente et reproductible » (Villeneuve-Lapointe et Moreau, 2019, p. 63). La démarche empruntée pour cette recension a été guidée par la revue systématique de la littérature en huit étapes, présentée par Rhoades (2011). Cette démarche a été privilégiée pour sa clarté dans l'explication des différentes étapes et pour sa complétude. Les étapes proposées par cette autrice sont les suivantes : 1) définir le sujet ou la question de recherche; 2) relever les informations importantes (critères d'inclusion et d'exclusion; mots-clés); 3) conduire la recherche de textes; 4) explorer les textes et exclure les recherches non pertinentes; 5) examiner les recherches intéressantes; 6) extraire les données et développer un organisateur graphique; 7) faire une synthèse; et 8) rapporter des conclusions et des recommandations.

Définir le sujet ou la question de recherche

L'ÉsA représente le sujet de la recension effectuée et, comme l'expose l'introduction, « Comment l'ÉsA est-elle mise en œuvre dans l'enseignement universitaire? », représente également la question qui a guidé la recension effectuée. Plus spécifiquement, cet article vise à répondre aux questions suivantes :

- Dans le cas où une évaluation pour apprendre est mise en œuvre, quels seraient ses effets sur les apprenant-es?
- Quelles sont les croyances et les pratiques évaluatives des enseignant-es universitaires?
- Dans quelle mesure les conceptions des enseignant-es universitaires en matière d'évaluation pour apprendre sont-elles alignées sur leurs pratiques évaluatives?

Relever les informations importantes (critères d'inclusion et d'exclusion; mots-clés)

Des mots-clés en anglais et en français ont été sélectionnés pour mener les recherches dans les différentes bases de données. Ceux-ci s'inscrivent dans deux grands thèmes, soit l'ÉsA et l'enseignement à l'université. Comme l'évaluation-soutien d'apprentissage est une terminologie qui n'est pas encore généralisée, l'évaluation pour l'apprentissage et l'évaluation formative ont aussi été retenues comme mots-clés, car elles constituent l'origine de l'ÉsA. Certain·es auteur·trices distinguent ces termes (Bowe-Deegan, 2015; Laveault et Allal, 2016; Sambell *et al.*, 2012; Swaffield, 2011), mais ils sont fréquemment utilisés comme synonymes (Cizek *et al.*, 2019; Black et Wiliam, 2009; Chao Chao, 2017; Cook, 2012; Harlen, 2005).

Tableau 1

Mots-clés utilisés

Anglais	Formative evaluation; assessment for learning; formative assessment; assessment as learning; higher education; undergraduate study; undergraduate students; universities
Français	Évaluation formative; évaluation pour l'apprentissage; éducation supérieure; enseignement supérieur; université

Outre les mots-clés, les recherches publiées de 2010 à 2020 ont été retenues afin d'obtenir un regard actuel sur la situation de l'évaluation pour l'apprentissage à l'université. De plus, seules des revues scientifiques révisées par les pairs ont été retenues ainsi que les recherches portant sur les premiers cycles universitaires, lors desquels de grands groupes sont plus fréquents (plus de 30 étudiant·es). Finalement, comme l'objectif de la recension était d'obtenir un portrait général de la mise en œuvre des pratiques d'évaluation pour l'apprentissage, seuls les textes empiriques (thèses et articles) ont été retenus.

Les critères d'exclusion touchent les différents types d'évaluation qui ne concernent pas nécessairement le soutien à l'apprentissage. Par exemple, l'évaluation ipsative et l'évaluation des cours par les étudiant·es n'ont pas été retenues. Il en va de même pour l'évaluation formative aux niveaux primaire, secondaire et collégial, l'évaluation en contexte de stage ainsi que l'évaluation aux cycles supérieurs (maîtrise et doctorat).

Conduire la recherche de textes

Explorer les textes et exclure les recherches non pertinentes

Le tableau 2 ci-dessous présente la démarche empruntée pour la recherche et la sélection des textes. Pour ce faire, les bases de données ERIC, PSYCInfo, ERUDIT, CAIRN et Dissertations and thesis ont été explorées. On constate d'ailleurs une certaine variabilité dans les termes utilisés d'une base de données à l'autre. Elle s'explique par la présence de thésaurus dans certaines bases de données, où les termes prévus sont regroupés sous des termes similaires, et par l'absence de certains termes dans les mots-clés recensés. Le nombre de textes qui résulte des recherches est inscrit, tout comme le nombre de textes retenus après un survol des résultats et l'exclusion de ceux qui ne correspondent pas aux critères.

Une relecture plus attentive des résumés retenus a été effectuée, et les textes ont été triés en fonction de leur pertinence perçue relativement à la mise en œuvre de l'évaluation pour l'apprentissage à l'université et des recherches quant aux pratiques et aux perceptions des enseignant·es universitaires. Or, plusieurs textes recensés (p. ex. Petit et Bélisle, 2018; Gikandi et Morrow, 2016; Han et Riazi, 2018) traitaient spécifiquement de certaines pratiques d'évaluation formative et non d'une approche intégrée d'évaluation pour l'apprentissage, ce qui restreignait leur apport quant à la compréhension générale de la mise en œuvre des pratiques d'ÉsA. Cet exercice a permis de sélectionner seize textes qui ont contribué à l'élaboration de la représentation graphique qui sera présentée.

Extraire les données et développer un organisateur graphique

Faire une synthèse

La lecture attentive des seize textes retenus a permis de brosser un portrait de l'état de la recherche empirique relativement à la mise en œuvre des pratiques d'ÉsA à l'enseignement universitaire. La section suivante permet d'illustrer la provenance de ces recherches, les méthodologies employées, les participant·es sollicité·es ainsi que les principaux constats qui se dégagent de ces recherches.

Analyse et discussion

La recension effectuée permet d'observer que, parmi les seize recherches retenues, sept ont été réalisées en Europe, trois en Océanie, quatre en Asie et deux en Amérique du Nord, dont seulement une au Canada. Quatre recherches ont été menées auprès des enseignant·es universitaires et d'étudiant·es, huit auprès d'enseignant·es seulement, trois auprès d'étudiant·es, et une auprès de coordonnateur·trices de cours et d'étudiant·es. Finalement, on remarque un certain équilibre entre les méthodologies employées, c'est-à-dire que quatre recherches ont utilisé une méthodologie quantitative, six ont préféré une méthodologie qualitative et cinq ont adopté une méthodologie mixte.

Par ailleurs, trois grands constats se dégagent des recherches recensées : 1) l'évaluation pour l'apprentissage semble avoir des effets positifs sur l'apprentissage, 2) les pratiques évaluatives employées dans les universités semblent demeurer majoritairement traditionnelles, et 3) la conceptualisation de l'évaluation pour l'apprentissage par le personnel enseignant universitaire ne semble pas cohérente avec les pratiques qu'il utilise. Ces trois constats, ainsi que les limites des recherches en question, seront approfondis dans la section suivante.

Tableau 2

Recherche dans les bases de données et première sélection des textes en fonction des critères d'inclusion et d'exclusion

Base de données	Stratégie de recherche	Nombre de résultats	Nombre de textes retenus
ERIC	Mots clés : MAINSUBJECT.EXACT("Formative Evaluation") AND (MAINSUBJECT.EXACT["Higher Education"] OR MAINSUBJECT.EXACT["Undergraduate Study"] MAINSUBJECT.EXACT["Undergraduate Students"] OR MAINSUBJECT.EXACT["Universities"]) AND (stype.exact("Scholarly Journals") AND pd(20100101-20201231)) AND ab("assessment for learning" OR "formative assessment" OR "formative evaluation" OR "assessment as learning")	75	37
PSYCInfo	Keywords: {higher education} AND Keywords: {Formative Assessment} OR {assessment for learning} OR {formative evaluation} OR {assessment as learning} AND Peer-Reviewed Journals only AND Year: 2010 To 2020	17	5
CAIRN	("évaluation pour l'apprentissage" OU "assessment for learning" OU "évaluation formative" OU "formative assessment") ET ("éducation supérieure" OU "enseignement supérieur") PARTOUT	102	10 (beaucoup de chapitres de livre, 1 article)
ÉRUDIT	(Title, abstract, keywords: "Évaluation pour l'apprentissage"OU"assessment for learning"OU"évaluation formative"OU"formative assessment") AND (Title, abstract, keywords: "higher education"OU"éducation supérieure"OU"enseignement supérieur") AND (Published between 2010 and 2020) AND (Collections : ['Érudit', 'UNB'])	82	1
Dissertations and Thesis	noft("Formative evaluation" OR "assessment for learning" OR "assessment as learning") AND noft("higher education" OR "universities") thèse de doctorat uniquement 2010 à 2020	38	5

Constat 1 : Les effets positifs de l'évaluation pour l'apprentissage (AfL)

Plusieurs auteur·trices ont analysé les effets de pratiques d'évaluation pour l'apprentissage sur les étudiant·es et les enseignant·es à l'enseignement supérieur (Ludvigsen *et al.*, 2015; McLean, 2018; Pastor, 2011; Romero-Martín *et al.*, 2014; Scalise *et al.*, 2018; Winstone et Millward, 2012).

En observant plus spécifiquement l'évaluation continue accompagnée de rétroactions, c'est-à-dire l'évaluation fréquente en cours d'apprentissage, qui permet ainsi de fournir de la rétroaction régulière, Romero-Martín et ses collaborateurs (2014) et Scalise et ses collaboratrices (2018) ont observé que cette pratique, lorsqu'elle était accompagnée d'une rétroaction, permettait aux étudiant·es d'obtenir de meilleurs résultats dans leurs cours.

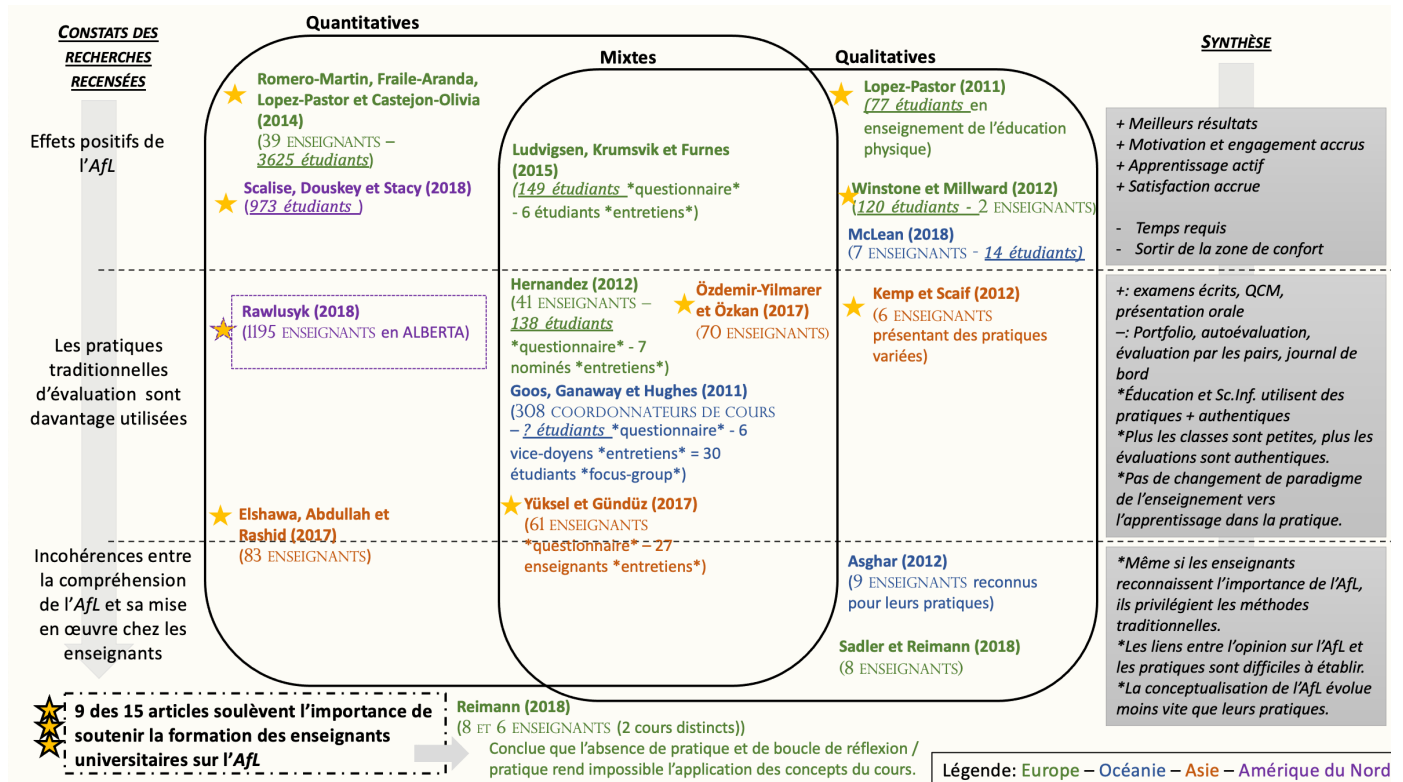


Figure 1. Synthèse graphique

Scalise et ses collaboratrices (2018) ont interrogé 973 étudiant·es d'un cours de chimie, répartis en trois groupes, soit : un groupe de comparaison suivant un cours dans lequel seuls des examens étaient utilisés comme instruments d'évaluation, accompagnés de devoirs hebdomadaires avec la rétroaction habituelle; un groupe conceptuel pour lequel des questions plus axées sur la pensée conceptuelle remplaçaient certaines des questions du groupe précédent; et finalement, un groupe différencié qui, en plus des devoirs hebdomadaires, recevait à quatre reprises par courriel une rétroaction personnalisée sur leurs réponses. Des ANOVA appliquées entre les trois groupes ont permis à ces autrices de constater que le groupe de comparaison réussissait toujours moins bien que les deux autres, et le groupe différencié a évolué plus rapidement dans la première partie du semestre, mais cette différence avec le groupe conceptuel s'est atténuée dans la deuxième moitié du semestre. L'interprétation des résultats de cette recherche doit par ailleurs être réalisée avec prudence. Peu d'indications sur la constance et la cohérence de la mise en œuvre des pratiques dans les différents groupes sont fournies. Le nombre d'enseignant·es participant·es dans ces trois groupes d'étudiant·es n'est pas mentionné, et la façon dont les rétroactions différenciées étaient fournies n'est pas explicite. De plus, à mi-chemin, les groupes conceptuels et différenciés ont été interchangés ce qui ne permet pas de vérifier l'effet des pratiques sur une session complète.

Quant au projet de recherche de Romero-Martin et ses collaborateurs (2014), 39 professeur·es universitaires de 15 universités espagnoles à la formation initiale des maîtres ont été interrogé·es à l'aide d'un questionnaire ainsi que 3625 étudiant·es qui étaient dans les cours de ces professeur·es. Ces auteur·trices ont constaté une différence significative entre les résultats des étudiant·es selon le type d'évaluation qu'ils subissaient. Les étudiant·es pouvaient choisir en début de semestre leur modalité d'évaluation, soit une évaluation continue, soit une évaluation basée seulement sur un examen final, soit une évaluation combinant les deux modalités (mixte). Les étudiant·es qui ont choisi l'évaluation continue ont obtenu des résultats supérieurs aux autres. Les étudiant·es qui ont opté pour une évaluation mixte ont obtenu des résultats justes derrière ceux de l'évaluation continue, alors que les étudiant·es qui ont été évalué·es par examens ont

abouti à des niveaux de performance plus bas. De plus, les résultats du questionnaire ont aussi permis à ces auteur·trices de constater que les étudiant·es ayant choisi une évaluation continue ou mixte considèrent avoir mieux appris et sont plus satisfaits des méthodes employées. Ce dernier élément corrobore les résultats de Lopez-Pastor (2011) et de Winstone et Millward (2012), qui sont abordés ci-dessous. Toutefois, il convient d'être prudent sur l'interprétation de tels résultats relativement à l'évaluation continue comme évaluation pour l'apprentissage. Des exemples concrets des instruments utilisés auraient permis de mieux évaluer la compréhension des auteur·trices, car ceux-ci peuvent représenter une variable non négligeable dans l'analyse des résultats des étudiant·es.

Lopez-Pastor (2011) a interrogé 77 étudiant·es participant à un cours d'un programme d'enseignement au primaire spécialisé en éducation physique. Il a utilisé, entre autres, le questionnaire d'évaluation du cours et des entretiens semi-dirigés, individuels et de groupe, ainsi que l'analyse des documents du cours. Il fait ressortir plusieurs avantages et désavantages de l'évaluation formative continue, dont l'engagement et la motivation des étudiant·es en plus d'un meilleur développement des compétences visées. Cependant, le fait que les étudiant·es ne soient pas familiarisés avec ces pratiques entraîne des difficultés organisationnelles et une charge de travail qui semble particulièrement lourde. Par ailleurs, cette recherche a permis aussi de mettre en évidence le fait que la mise en œuvre des pratiques d'évaluation continue permet aux enseignant·es de mieux connaître leur étudiant·e, ce qui améliore la communication et les relations entre l'enseignant·e et les étudiant·es. Il faut toutefois être attentif·ve au fait que cette recherche a été menée auprès d'étudiant·es dans le cadre d'un seul cours et que la possibilité de transférer les résultats dans des contextes différents est ainsi réduite.

Ces résultats sont appuyés par la recherche de Winstone et Millward (2012), qui ont interrogé 120 étudiant·es participant aux cours de deux enseignant·es. Ils ont utilisé deux questionnaires, un pour les enseignant·es et un pour les étudiant·es, qui comprenait chacun une question ouverte. L'analyse thématique réalisée à partir de ces réponses a permis aux autrices de constater que les étudiant·es estiment que l'évaluation formative sous forme de questionnaire et de discussion avec les pairs favorise leur développement personnel et leur compréhension, alors que l'aspect actif des activités d'apprentissage et d'évaluation favorise leur engagement et la rétention de la matière. Les enseignant·es participant·es, ont à leur tour soulevé que les étudiant·es semblaient plus motivé·es et engagé·es, qu'eux-mêmes et elles-mêmes étaient plus engagé·es, connaissaient mieux leurs étudiant·es et favorisaient leur développement personnel. Toutefois, les enseignant·es ont tout de même souligné que la mise en œuvre de telles pratiques requiert du temps et les sort de leur zone de confort. Par contre, tout comme la recherche de Lopez-Pastor (2011), les résultats découlent des réponses d'étudiant·es ayant participé aux cours de deux enseignant·es, ce qui réduit la décontextualisation des résultats.

Dans une autre perspective, McLean (2018) et Ludvigsen *et al.* (2015) mettent davantage de l'avant l'apport de la discussion entre les étudiant·es dans des modalités d'évaluation pour l'apprentissage afin de favoriser des apprentissages plus significatifs.

McLean (2018) a interrogé 7 enseignant·es expérimenté·es dans la mise en œuvre de pratiques de soutien à l'apprentissage et 14 de leurs étudiant·es en dernière année de formation. L'évaluation a été effectuée par des entretiens semi-dirigés et l'analyse des politiques et des documents curriculaires. Toutefois, le nombre total d'étudiant·es dans les cours de ces enseignant·es n'est pas mentionné dans l'article. L'analyse décrite comme intuitive et dans le « milieu des choses » a permis de faire ressortir 5 thèmes relativement à l'expérience de pratiques d'évaluation par les pairs et de rétroactions comme évaluations soutenant l'apprentissage, soit la présence de dialogues qui favorisent l'apprentissage, la plus grande confiance des enseignant·es, la meilleure motivation des enseignant·es, l'apprentissage favorisant un changement social et l'intégration de l'évaluation à l'apprentissage.

Ludvigsen et ses collaborateurs (2015) sont arrivés à des conclusions similaires au premier thème ressorti par McLean (2018) dans leur projet. Cette recherche mixte lors de laquelle ont été interrogés 149 étudiant·es par questionnaire et 6 étudiant·es par des entretiens semi-dirigés conclut que les étudiant·es apprécient les activités d'évaluation formative

en cours d'apprentissage, qui les aident à mieux se situer par rapport aux contenus du cours. Les entretiens ont corroboré ce constat en approfondissant l'importance de rendre les étudiant·es actif·ves durant leurs apprentissages grâce au temps de discussion avec leurs pairs, qui permet d'argumenter les points de vue et favorise une meilleure compréhension des contenus. Ce projet de recherche a été réalisé auprès d'étudiant·es dans un seul cours offert à un grand groupe. La pratique d'évaluation formative consistait à questionner les étudiant·es pendant le cours à l'aide d'un dispositif technologique pour leur permettre de se questionner sur les contenus abordés dans le cours, d'en discuter avec leurs collègues, puis de vérifier leur compréhension. Au même moment, l'enseignant·e pouvait sonder les étudiant·es pour vérifier leur compréhension en cours d'apprentissage et réajuster l'enseignement en cas de difficultés.

Ainsi, ce premier constat nous permet de répondre à notre première question spécifique de recherche, soit : dans l'éventualité où une évaluation pour apprendre est mise en œuvre, quels seraient ses effets sur l'apprentissage? Les recherches recensées permettent de constater une amélioration sur le plan de la motivation et de l'engagement des étudiant·es ainsi qu'une satisfaction accrue par rapport aux apprentissages réalisés. De plus, certaines recherches ont même pu démontrer une amélioration des résultats des étudiant·es ayant vécu des pratiques d'évaluation soutenant l'apprentissage. Plusieurs de ces recherches mentionnent des pratiques d'évaluation pour l'apprentissage utilisant la rétroaction (par l'enseignant·e ou par les pairs), faisant ainsi écho aux travaux de Hattie (2009) qui soulignent l'effet important de celle-ci sur l'apprentissage. Toutefois, ces différents projets de recherche ont permis de mettre en lumière que le temps requis pour mettre en œuvre des pratiques d'évaluation en soutien à l'apprentissage représente une préoccupation pour les enseignant·es et que ces pratiques sortent à la fois les enseignant·es et les étudiant·es de leur zone de confort. Par ailleurs, les recherches recensées ont étudié des pratiques d'évaluation pour l'apprentissage qui varient d'un contexte à l'autre, ce qui rend difficile la généralisation de ces résultats à l'ensemble des méthodes d'évaluation pour soutenir l'apprentissage. En ce sens, des recherches propres à certaines pratiques évaluatives permettraient certainement de nuancer les résultats présentés.

Constat 2 : La préséance des pratiques traditionnelles

Nonobstant la valeur pédagogique de l'évaluation en soutien à l'apprentissage démontrée précédemment, plusieurs auteur·trices confirment que les pratiques traditionnelles sont plus souvent utilisées (Elshawa et Abdullah, 2017; Goos *et al.*, 2011; Hernández, 2012; Kemp et Scaife, 2012; Rawlusk, 2018; Yilmazer et Özkan, 2017; Yüksel et Gündüz, 2017).

D'abord, la recherche d'Elshawa et ses collaboratrices (2017) réalisée en Malaisie auprès de 83 enseignant·es de 6 universités différentes, à l'aide d'un questionnaire à 67 questions, relève que les enseignant·es priorisent les formats « papier-crayon » pour les évaluations.

Cette observation est enrichie par le projet de recherche mixte de Yüksel et Gündüz (2017), qui ont interrogé 61 enseignant·es en Turquie par un questionnaire et qui ont approfondi les réponses en tenant 27 entretiens semi-dirigés parmi ceux et celles ayant répondu au questionnaire. Ces chercheur·euses sont arrivé·es à la conclusion que les questions à choix multiples, les questions à développement, et les questions vrai ou faux restent les méthodes les plus employées pour l'évaluation des cours théoriques. Pour apprécier les apprentissages réalisés dans les cours pratiques, les enseignant·es privilégient des tests d'habiletés et des examens de performance à l'utilisation de projets.

Quant à elles, Yilmazer et Özkan (2017) ont utilisé une méthodologie mixte et un questionnaire incluant des questions ouvertes pour interroger 70 enseignant·es d'anglais en Turquie afin de décrire les pratiques évaluatives d'enseignant·es universitaires en anglais. Elles ont pu constater par les analyses statistiques et thématiques menées que les portfolios, l'évaluation par les pairs, l'autoévaluation, et les journaux de bord sont peu utilisés comparativement aux examens écrits, aux dissertations et aux présentations orales. De plus, elles ont pu observer que les enseignant·es qui se considèrent comme efficaces relativement à leurs pratiques évaluatives justifient ce point principalement par leur capacité à donner des rétroactions aux étudiant·es. À l'inverse, ceux et celles qui ne se trouvent pas efficaces justifient leur réponse par le manque de formation.

Finalement, Rawlusk (2018) a interrogé 1 195 enseignant·es de 301 facultés réparties dans 12 institutions d'enseignement supérieur en Alberta, au Canada, pour examiner l'occurrence et la diversité des pratiques évaluatives dans l'éducation supérieure et leur relation avec l'apprentissage des étudiant·es. Les résultats ressortis des analyses quantitatives du questionnaire distribué ont permis de corroborer les résultats précédents relativement à la prépondérance de l'utilisation d'examens à réponses écrites, des choix multiples et des questionnaires pour apprécier les apprentissages des étudiant·es. Il en ressort aussi que plus de 50 % des répondant·es n'utilisent que rarement, voire jamais, l'autoévaluation ou l'évaluation par les pairs dans leurs cours. Quelques constats sont aussi mis de l'avant quant à l'utilisation d'évaluation authentique en sciences infirmières et en éducation en comparaison avec les facultés de sciences, de business, de kinésiologie et d'arts. De plus, il semblerait que les petits groupes d'étudiant·es favoriseraient l'utilisation de ces situations authentiques comme méthodes évaluatives.

Dans les quatre articles précédemment mentionnés ont été interrogés un grand nombre d'enseignant·es, ce qui permet de brosser un portrait général de la réalité des institutions universitaires. Toutefois, il s'agit de pratiques autorapportées par les enseignant·es. Or, il n'est pas possible de vérifier précisément la façon dont ces pratiques sont mises en œuvre et dans quelle mesure elles contribuent ou non à soutenir l'apprentissage des étudiant·es. Par ailleurs, la méthodologie mixte empruntée par deux de ces recherches a permis d'aller préciser certains éléments des questionnaires afin que les résultats se rapprochent de ce qui est réellement vécu.

En résumé, la présence des évaluations de type « papier-crayon », des examens écrits à choix multiples ou à réponse courte, des dissertations, et des présentations orales est appuyée par plusieurs recherches (Rawlusk, 2018; Yüksel et Gündüz, 2017; Elshawa *et al.*; Özdemir-Yilmarer et Özkan, 2017). De plus, certains confirment que les portfolios, les journaux de bord, les autoévaluations et les évaluations par les pairs sont encore très peu exploités (Özdemir-Yilmarer et Özkan, 2017; Rawlusk, 2018). Des nuances semblent toutefois se présenter entre les différentes disciplines étudiées. Il serait d'ailleurs pertinent de mener une analyse comparative des méthodes d'évaluation soutenant l'apprentissage déployées dans différentes facultés afin de déterminer si certaines disciplines sont davantage propices à leur mise en œuvre.

Dans une autre perspective, Goos et ses collaboratrices (2011) et Hernandez (2012) ont plutôt observé l'utilisation de la rétroaction en cours d'apprentissage et la possibilité des étudiant·es de se réajuster à la suite de la réception d'une rétroaction, élément indissociable des pratiques évaluatives en soutien à l'apprentissage.

D'abord, Goos et ses collaboratrices (2011), dans le but de décrire les perceptions des étudiant·es, des coordonnateur·trices de cours et des vice-doyen·nes des facultés relativement à l'évaluation, ont mis en commun l'évaluation des cours, les réponses à un sondage rempli par 308 coordonnateur·trices des cours, ainsi que les échanges lors de trois groupes de discussion avec dix étudiant·es de première année, et lors d'entretiens avec six vice-doyen·nes de la faculté responsable de l'enseignement et l'apprentissage. De l'ensemble de ces données, les coordonnateur·trices des cours et les vice-doyen·nes reconnaissent l'importance de donner de la rétroaction et sont confiant·es de son utilisation adéquate dans les cours, mais soulignent toutefois certaines difficultés à mettre en place des pratiques efficaces de rétroaction, entre autres à cause de la culture de l'école qui serait encore axée sur les résultats et la performance. En contrepartie, les étudiant·es estiment que les rétroactions en cours d'apprentissage sont peu fréquentes ou peu utiles, car elles ne préparaient pas à l'examen ou n'arrivaient pas assez rapidement dans le processus d'apprentissage.

Ayant un objectif de recherche similaire, soit celui d'examiner les perceptions des enseignant·es et des étudiant·es sur la façon dont les pratiques évaluatives supportent l'apprentissage des étudiant·es, Hernandez (2012) a sondé 138 étudiant·es et 41 enseignant·es au moyen d'un questionnaire, accompagné d'entretiens semi-dirigés auprès de 7 élu·es à la tête des départements. Elle constate que la grande majorité des étudiant·es accordent beaucoup d'importance aux notes de leurs travaux et que ces dernières favorisent la motivation et leur indique leur niveau de compréhension. Toutefois, environ 25 % des étudiant·es n'apprécient pas l'évaluation continue étant donnée la pression de performance

que cela impose. Aussi, environ 20 % des étudiant·es disent recevoir peu d'informations sur la façon de s'améliorer pour la suite du cours. Par contre, la satisfaction des étudiant·es relativement aux rétroactions reçues de l'enseignant·e sur les travaux réalisés est plutôt mitigée. Lorsque la satisfaction est élevée, les étudiant·es la justifient en mentionnant que les erreurs sont surlignées ou que des commentaires sont formulés, et des suggestions sont faites pour comprendre ce qui a été moins bien réussi. Les étudiant·es moins satisfait·es de la rétroaction mentionnent soit que les rétroactions ne sont pas suffisamment nombreuses, soit qu'aucun conseil ne leur est donné pour s'améliorer, soit qu'elle arrive trop tard. Néanmoins, seulement 16 % des étudiant·es prendraient des mesures pour améliorer leurs travaux à la suite des recommandations de leur enseignant·e. Toutefois, il est nécessaire d'être prudent·e dans l'interprétation de ces résultats étant donné qu'ils rapportent les perceptions des différent·es participant·es plutôt que les pratiques employées, tout comme la recherche précédente de Goos et ses collaboratrices (2011). Certaines précisions quant aux pratiques sur lesquelles se prononcent les étudiant·es et les coordonnateur·trices auraient permis aussi de mieux comprendre le contexte étudié.

Finalement, Kemp et Scaif (2012) ont effectué des entretiens semi-dirigés auprès de 6 enseignant·es universitaires en utilisant des stratégies pédagogiques variées et auprès de 6 étudiant·es (un de chaque enseignant·e). Ils ont pu observer qu'il était courant d'utiliser des évaluations sommatives à des fins de diagnostic et que les enseignant·es se servaient souvent d'indices informels en classe pour évaluer la progression des étudiant·es dans le cours. Toutefois, ils soulèvent que peu d'enseignant·es ajustent leurs interventions pédagogiques en fonction des données recueillies de manière formative sur l'apprentissage des étudiant·es. De leur côté, les étudiant·es apprécient les stratégies pédagogiques leur permettant d'être actif·ves dans leurs apprentissages et de situer leur appropriation des contenus, mais les stratégies restent limitées. Un seul des 6 étudiant·es affirme avoir obtenu de la rétroaction par ses pairs. Cette recherche réalisée auprès d'un petit nombre de participant·es permet de constater que les pratiques d'évaluation en soutien à l'apprentissage, même chez des enseignant·es reconnu·es pour varier leurs pratiques pédagogiques, ne sont pas encore exploitées à leur plein potentiel.

En somme, bien que les pratiques d'évaluation pour l'apprentissage soient reconnues comme étant bénéfiques pour les apprentissages, la motivation et l'engagement, les pratiques qui sont encore largement utilisées dans les institutions à travers le monde semblent ancrées dans un certain traditionalisme. Ainsi, en réponse à notre deuxième question spécifique de recherche (Quelles sont les croyances et les pratiques évaluatives des enseignant·es universitaires?), les recherches recensées soutiennent le fait que les pratiques évaluatives employées dans les institutions universitaires sont majoritairement traditionnelles. Toutefois, les enseignant·es semblent reconnaître l'importance de la rétroaction pour l'apprentissage alors que les étudiant·es auraient plutôt une perception mitigée de l'apport de celle-ci. Considérant que la mise en œuvre de pratiques évaluatives pour l'apprentissage, dont la rétroaction, représente une compétence complexe qui requiert de la formation, du temps et du travail collaboratif entre collègues afin d'être développée (Stiggins, 2005; Laveault et Allal, 2016; Harlen, 2005), il est possible de se demander dans quelles conditions sont réalisées les rétroactions dont les étudiant·es sont moins satisfait·es. En effet, la rétroaction requiert un savoir-faire particulier et une certaine expertise pour être efficace (Hattie et Temperley, 2007). Or, les recherches recensées présentent davantage un portrait descriptif des perceptions et des pratiques alors qu'il pourrait être pertinent de s'intéresser aux facteurs qui déterminent la mise en œuvre de telles pratiques et possiblement d'utiliser des méthodes de recherche participatives afin d'accompagner et de documenter le déploiement de pratiques nouvelles.

Constat 3 : Incohérences entre la compréhension et la mise en œuvre de l'évaluation pour l'apprentissage

Le dernier constat qui émane des recherches recensées est le fait que des incohérences sont présentes entre la compréhension des enseignant·es relativement à l'évaluation pour l'apprentissage et la mise en œuvre de celle-ci dans leur pratique (Asghar, 2012; Elshawa *et al.*, 2017; Sadler et Reimann, 2018; Yüksel *et al.*, 2017).

Elshawa et ses collaboratrices (2017), dont nous avons décrit brièvement la recherche plus tôt, constatent que les enseignant·es reconnaissent l'importance et la pertinence de l'évaluation en soutien à l'apprentissage. Le même constat est fait par Yüksel et Gündüz (2017), aussi décrit précédemment.

Le projet mené par Asghar (2012) auprès de 9 enseignant·es volontaires, déjà reconnu·es comme étant engagé·es dans le développement de leurs pratiques par des projets pédagogiques, a permis de comprendre que, bien que ces enseignant·es donnent de la valeur à l'évaluation formative, ils et elles expriment de la frustration quant à la taille des classes et du manque de temps. Les entretiens semi-dirigés réalisés ont aussi mis de l'avant que les enseignant·es en enseignement sont plus avancé·es dans leur compréhension de l'évaluation formative et de leurs pratiques, entre autres, car ils et elles veulent que leurs étudiant·es reproduisent ces stratégies dans leurs pratiques professionnelles. Ce dernier point confirme les conclusions de la recherche réalisée auprès de 1 195 enseignant·es par Rawlusk (2018), que nous avons décrite précédemment et qui veut que les enseignant·es en éducation et en sciences infirmières mettent davantage en œuvre des pratiques d'évaluation dans un contexte authentique.

Toutefois, la représentation de ce concept ou de ces pratiques reste difficile pour les enseignant·es et évolue plus lentement que leurs pratiques. En effet, la recherche de Sadler et Reimann (2018) porte un regard sur la compréhension des enseignant·es universitaires de l'évaluation et de l'évolution de leurs pratiques dans le temps. Pour ce faire, deux entretiens semi-dirigés ont eu lieu avec huit enseignant·es de deux universités et deux cartes conceptuelles de l'évaluation pour l'apprentissage par ces enseignant·es ont été réalisées, à deux moments différents, mais en laissant écouler une année entre les itérations des deux entretiens. Cette démarche, qui ne prenait pas en considération les activités de développement professionnel ou de formation continue entre les deux entretiens, a permis d'observer que les enseignant·es auraient plus de facilité à modifier certaines pratiques que leur conception de l'évaluation.

Ainsi, en réponse à notre dernière question spécifique de recherche (Dans quelle mesure les conceptions des enseignant·es universitaires en matière d'évaluation pour apprendre sont-elles alignées sur leurs pratiques évaluatives?), les recherches recensées permettent de comprendre que les enseignant·es reconnaissent l'importance de mettre en œuvre des pratiques évaluatives soutenant l'apprentissage. Toutefois, leur compréhension de ce que représentent ces pratiques ne semble pas exacte ou complète, ce qui n'est pas surprenant considérant que la définition du concept d'évaluation pour l'apprentissage ne fait pas consensus dans la littérature (Bowe-Deegan, 2015; Allal et Laveault, 2016; Carless, 2015; Sambell *et al.*, 2012; Taras, 2008). En ce sens, la recherche de Sadler et Reimann (2018) permet d'observer que les pratiques sont susceptibles d'évoluer plus rapidement que la compréhension grâce, entre autres, à l'influence de différents éléments contextuels forçant les enseignant·es à s'ajuster. Plusieurs auteur·trices soutiennent que les conceptions des enseignant·es influencent les pratiques mises en œuvre (Offerdhal et Tomanek, 2011; Demougeot-Lebel et Perret, 2011; Loiola et Tardif, 2001; Butler, 2005), alors que d'autres observent des incohérences entre les pratiques employées et les conceptions (Sadler et Reimann, 2008; Laksov *et al.*, 2008). Asghar (2012) soulève l'effet des éléments contextuels sur les pratiques utilisées, et ce, malgré les conceptions. Dans cette perspective, des analyses systémiques, inspirées du modèle écologique de Bronfenbrenner (1979), pourraient mettre en lumière les nombreuses variables impliquées dans le déploiement de pratiques d'évaluation pour l'apprentissage.

Conclusion

La revue systématique effectuée sur les pratiques évaluatives pour l'apprentissage a permis de constater que plusieurs recherches décrivent les pratiques employées par les enseignant·es universitaires et les perceptions des enseignant·es et des étudiant·es par rapport à ces pratiques. Ces dernières révèlent que l'évaluation pour l'apprentissage présente plusieurs avantages, mais que les pratiques traditionnelles restent largement utilisées. De plus, la compréhension des enseignant·es universitaires sur ces pratiques semble fragile et ne correspond pas aux pratiques qu'ils mettent en œuvre.

Toutefois, peu de recherches ont été effectuées au Canada, et la généralisation de résultats internationaux au contexte canadien, voire québécois, semble hasardeuse. De plus, bien que les méthodologies employées soient variées et touchent à la fois les étudiant·es et les enseignant·es, nous constatons qu'aucune n'a abordé le problème dans une perspective de recherche-développement ou dans une perspective collaborative en vue de promouvoir et de favoriser la mise en œuvre de telles pratiques dans le milieu universitaire. Pourtant, neuf des recherches recensées et décrites précédemment ouvrent la voie, dans leurs perspectives de recherche et dans leur conclusion, à la formation ou au soutien destinés aux enseignant·es universitaires afin de favoriser le développement de pratiques d'évaluation pour l'apprentissage (Romero-Martín *et al.*, 2014; Scalise *et al.*, 2018; Rawlusk, 2018; Elshawa *et al.*, 2017; Yüksel et Gündüz, 2017; Özdemir-Yilmazer et Özkan, 2017; Kemp et Scaif, 2012; Winstone et Millward, 2012; Lopez-Pastor, 2011).

Pour conclure, cet article a permis de constater certains éléments que la recherche a pu dégager sur les pratiques d'évaluation pour l'apprentissage dans le milieu universitaire à travers le monde. La présente recension démontre que les pratiques et les perceptions sur l'évaluation pour l'apprentissage ont beaucoup été étudiées. Toutefois, la façon dont on peut développer davantage de telles pratiques chez des enseignant·es a été peu explorée. Or, si les avantages de la mise en œuvre de ces pratiques sont reconnus, il serait intéressant de se pencher sur les dispositifs efficaces pour amorcer un changement de pratiques chez les enseignant·es qui préfèrent encore les méthodes d'évaluation traditionnelles.

Références

- Allal, L. et Laveault, D. (2009). Assessment for learning: évaluation-soutien d'apprentissage. *Measure et évaluation en éducation*, 32(2), 99-106. <https://www.erudit.org/en/journals/mee/1900-v1-n1-mee01394/1024956ar/abstract/>
- Asghar, M. (2012). The lived experience of formative assessment practice in a British university. *Journal of Further and Higher Education*, 36(2), 205-223. <https://doi.org/10.1080/0309877X.2011.606901>
- Black, P., Harrison, C., Lee, C., Marshall, B. et Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, 86(1), 8-21.
- Black, P. et Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education)*, 21(1), 5.
- Black, P. et Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, 25(6), 551-575.
- Bowe-Deegan, P. (2015). *Do assessment for learning practices act as a springboard or straitjacket for educationally worthwhile learning?: a case study in an irish higher education institute* [Thèse de doctorat, University of Sheffield, Royaume-Uni]. ProQuest. <https://search.proquest.com/docview/1810885055?accountid=12543>
- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Harvard University Press. <https://play.google.com/store/books/details?id=OCmbzWka6xUC>
- Butler, D. L. (2005). L'autorégulation de l'apprentissage et la collaboration dans le développement professionnel des enseignants. *Revue des sciences de l'Éducation*, 31(1), 55-78.
- Carless, D. (2015). Exploring learning-oriented assessment processes. *Higher Education*, 69(6), 963-976.
- Chao Chao, K.-W. (2017). *Les pratiques évaluatives visant l'efficacité de la rétroaction pour le développement de la production écrite des étudiants du français langue étrangère : le cas de l'Université de Costa Rica* [Thèse de doctorat, Université de Montréal].
- Cizek, G. J., Andrade, H. L. et Bennett, R. E. (2019). Formative assessment : History, definition, and progress. Dans H. L. Andrade, R. E. Bennett et G. J. Cizek (dir.), *Handbook of Formative Assessment in the Disciplines* (p. 3-19). Routledge.
- Cook, M. (2012). *An analysis of the effects upon itet students' learning and self-efficacy when adopting the principles and practices of assessment for learning (afl)* [Thèse de doctorat, Cardiff Metropolitan University]. ProQuest. <https://search.proquest.com/docview/1683358155?accountid=12543>

- Demougeot-Lebel, J. et Perret, C. (2012). Une formation pédagogique peut-elle modifier les conceptions de jeunes enseignants universitaires sur l'apprentissage et l'enseignement? *Revue des sciences de l'éducation*, 37(2), 327-354.
- Detroz, P., Tessaro, W. et Younès, N. (2020). Pour la relance d'une évaluation congruente à l'université. *Évaluer. Journal international de recherche en éducation et formation*, 1(Numéro Hors-Série), 111-119. <http://journal.admee.org/index.php/ejiref/article/view/224>
- Elshawa, N. R. M. et Abdullah, A. N. (2017). Malaysian instructors' assessment beliefs in tertiary ESL classrooms. *International Journal of Education and Literacy Studies*. <http://www.journals.aiac.org.au/index.php/IJELS/article/view/3403>
- Fortier, M. (2018, 7 mars). Le nombre d'étudiants souffrant de déficience bondit dans les établissements postsecondaires. *Le Devoir*. <https://www.ledevoir.com/societe/education/522005/enseignement-superieur-explosion-du-nombre-d-etudiants-en-difficulte>
- Gikandi, J. W. et Morrow, D. (2016). Designing and implementing peer formative feedback within online learning environments. *Technology, Pedagogy and Education*, 25(2), 153-170.
- 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*, 38(1), 95-107.
- Han, C. et Riazzi, M. (2018). The accuracy of student self-assessments of English Chinese bidirectional interpretation: a longitudinal quantitative study. *Assessment & Evaluation in Higher Education*, 43(3), 386-398.
- Harlen, W. (2005). Teachers' summative practices and assessment for learning-tensions and synergies. *The Curriculum Journal*, 16(2), 207-223.
- Hattie, J. (2009). *Visible learning: A synthesis of meta-analysis relating to achievement*. Routledge.
- Hattie, J. et Timperley, H. (2007). The power of feedback. *Review of educational research*, 77(1), 81-112.
- Hernández, R. (2012). Does continuous assessment in higher education support student learning? *Higher Education*, 64(4), 489-502.
- Kemp, S. et Scaife, J. (2012). Misunderstood and neglected? Diagnostic and formative assessment practices of lecturers. *Journal of Education for Teaching*, 38(2), 181-192.
- Klenowski, V. (2009). Assessment for learning revisited: An Asia-Pacific perspective. *Assessment in Education Principles Policy and Practice*, 16(3).
- Laveault, D. et Allal, L. (2016). *Assessment for Learning: Meeting the Challenge of Implementation*, Springer.
- Ludvigsen, K., Krumsvik, R. et Furnes, B. (2015). Creating formative feedback spaces in large lectures. *Computers & education*, 88, 48-63.
- Marshall, B. et Drummond, M. J. (2006). How teachers engage with Assessment for Learning: Lessons from the classroom. *Research Papers in Education*, 21(2), 133-149.
- McLean, H. (2018). This is the way to teach: Insights from academics and students about assessment that supports learning. *Assessment & Evaluation in Higher Education*, 43(8), 1228-1240.
- Laksov, K. B., Mann, S. et Dahlgren, L. O. (2008). Developing a community of practice around teaching: A case study. *Higher Education Research & Development*, 27(2), 121-132.
- Loiola, F. A. et Tardif, M. (2001). Formation pédagogique des professeurs d'université et conceptions de l'enseignement. *Revue des sciences de l'éducation*, 27(2), 305-326.
- Lopez-Pastor, V.-M. (2011). Best practices in academic assessment in higher education: A case in formative and shared assessment. *Journal of Technology and Science Education*, 1(2), 25-39.

- Offerdahl, E. G. et Tomanek, D. (2011). Changes in instructors' assessment thinking related to experimentation with new strategies. *Assessment & Evaluation in Higher Education*, 36(7), 781-795.
- Petit, M. et Bélisle, M. (2018). Démarche d'implantation du portfolio numérique en enseignement supérieur : le cas de deux programmes de formation à l'enseignement. *Revue internationale des technologies en pédagogie universitaire/International Journal of Technologies in Higher Education*, 15(3), 52-67.
- Rawlusk, P. E. (2018). Assessment in higher education and student learning. *Journal of Instructional Pedagogies*, 21. <https://eric.ed.gov/?id=EJ1194243>
- Reimann, N. (2018). Learning about assessment: The impact of two courses for higher education staff. *International Journal for Academic Development*, 23(2), 86-97.
- Rhoades, E. A. (2011). Literature Reviews. *The Volta Review*, 111(3), 353-368.
- Romero-Martín, R., Fraile-Aranda, A., López-Pastor, V.-M. et Castejón-Oliva, F.-J. (2014). The relationship between formative assessment systems, academic performance and teacher and student workloads in higher education. *Infancia y Aprendizaje: Journal for the Study of Education and Development*, 37(2), 310-341.
- Sadler, I. et Reimann, N. (2018). Variation in the development of teachers' understandings of assessment and their assessment practices in higher education. *Higher Education Research & Development*. https://www.tandfonline.com/doi/abs/10.1080/07294360.2017.1344199?casa_token=iyK5lzKuRuQAAAAA: XVz6X-9l6llacsUl5LM-OzRy9HxqRr0W3gNE5xkwzpOELeqVu9AzNX5VZ8 WMuuixfLQOz8st9u74w
- Sambell, K., McDowell, L. et Montgomery, C. (2012). *Assessment for Learning in Higher Education*. Routledge. https://play.google.com/store/books/details?id=7bieE6Uz_ssC
- Scalise, K., Douskey, M. et Stacy, A. (2018). Measuring learning gains and examining implications for student success in STEM. *Higher Education Pedagogies*, 3(1), 183-195.
- Stiggins, R. (2005). From formative assessment to assessment for learning: A path to success in standards-based schools. *Phi Delta Kappan*, 87(4), 324-328.
- Swaffield, S. (2011). Getting to the heart of authentic assessment for learning. *Assessment in Education: Principles, Policy & Practice*, 18(4), 433-449.
- Taras, M. (2008). Summative and formative assessment: Perceptions and realities. *Active Learning in Higher Education*, 9(2), 172-192.
- Villeneuve-Lapointe, M. et Moreau, A. C. (2019). Les méthodologies de recherche en orthographe : stratégies de rédaction d'une recension intégrative d'écrits scientifiques. *Revue canadienne des jeunes chercheuses et chercheurs en éducation, Hors Série*, 63-74.
- Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, 37(1), 3-14.
- Wiliam, D. et Thompson, M. (2017). Integrating assessment with learning: What will it take to make it work? Dans C. A. Dwyer (dir.), *The Future of Assessment* (p. 53-82). Routledge.
- Winstone, N. et Millward, L. (2012). Reframing perceptions of the lecture from challenges to opportunities: Embedding active learning and formative assessment into the teaching of large classes. *Psychology Teaching Review*, 18(2), 31-41. <https://eric.ed.gov/?id=EJ991406>
- Yilmazer, M. Ö. et Özkan, Y. (2017). Classroom assessment practices of English language instructors. *Journal of Language and Linguistic Studies*, 13(2), 324-345. <http://jlls.org/index.php/jlls/article/view/678>
- Yüksel, H. S. et Gündüz, N. (2017). Formative and summative assessment in higher education: opinions and practices of instructors. *European Journal of Education Studies*, 3(8).