Évaluer les apprentissages avec les outils numériques

Pédagogies en pratique

Sous la direction d'Emmanuel Sylvestre et Jean-Michel Jullien

pédagogie universitaire

Quiz en ligne, examens à distance, e-portfolios, *feedback* automatisé : les outils numériques ont profondément renouvelé l'évaluation des apprentissages dans l'enseignement supérieur. Ces évolutions modifient les pratiques enseignantes et les expériences étudiantes. Elles soulèvent aussi des questions décisives d'équité, d'intégrité académique et d'usage des données.

Cet ouvrage analyse l'e-évaluation à partir de quatre dimensions complémentaires : pédagogique, docimologique, technologique et éthique. Il s'appuie sur des cadres théoriques récents et sur des études de cas menées dans des contextes universitaires variés. Les contributions interrogent la conception des dispositifs d'évaluation, l'exploitation des traces d'apprentissage, la qualité du *feedback* et l'engagement des étudiant·es. Elles examinent également les contraintes liées aux infrastructures, à l'accessibilité, à la sécurité des examens et à l'essor de l'intelligence artificielle.

Destiné aux enseignant·es du supérieur, aux ingénieur·es pédagogiques, aux responsables institutionnels et aux chercheur·ses en sciences de l'éducation, cet ouvrage propose des repères analytiques et des pistes concrètes pour repenser les pratiques évaluatives dans un environnement numérique en constante évolution.

Jean-Michel Jullien est responsable du service Educational Development Unit for Distance Learning (EDUDL+) d'UniDistance Suisse.

Emmanuel Sylvestre est professeur titulaire et directeur du Centre de soutien à l'enseignement à l'Université de Lausanne et rédacteur en chef de la revue *e-Jiref* de l'Association pour le développement des méthodologies d'évaluation en éducation (ADMEE-Europe).

ISBN 978-2-88915-753-2

épistémé

Évaluer les apprentissages avec les outils numériques

Pédagogies en pratique

« Pédagogies en pratique » est une série d'ouvrages consacrés à l'innovation pédagogique et à son impact concret sur l'enseignement et l'apprentissage. Chaque volume propose une sélection minutieuse de dispositifs pédagogiques éprouvés, provenant de divers contextes et institutions éducatives.

Rédigés par des praticien·nes passionné·es et engagé·es, ces ouvrages offrent un regard authentique sur les pratiques novatrices qui transforment l'enseignement supérieur. Nos auteur·rices, eux·elles-mêmes enseignant·es, partagent leurs expériences, leurs réussites et les défis auxquels ils et elles sont confronté·es, offrant ainsi des perspectives pratiques et inspirantes pour tous·tes celles et ceux qui cherchent à améliorer leur pratique pédagogique.

Guidé·es par une approche soutenue par la recherche, nos praticien·nes collaborent avec des équipes d'appui pédagogique pour s'assurer que chaque dispositif est ancré dans le Scholarship of Teaching and Learning (SoTL), garantissant ainsi sa pertinence et son efficacité.

Sous la direction d'Emmanuel Sylvestre (Université de Lausanne), cette série se veut un précieux outil pour tous·tes celles et ceux aspirant à créer des expériences d'apprentissage enrichissantes et significatives pour leurs étudiant·es.

Évaluer les apprentissages avec les outils numériques

Pédagogies en pratique

Sous la direction d'Emmanuel Sylvestre et Jean-Michel Jullien

L'édition de cet ouvrage a reçu le soutien de Swiss Universities.

Direction générale : Lucas Giossi
Directions éditoriale et commerciale : Sylvain Collette et May Yang
Direction de la communication : Manon Reber
Responsable de production : Christophe Borlat
Éditorial : Antoine Derouin, Alice Micheau-Thiébaud et Jean Rime
Graphisme : Kim Nanette
Comptabilité : Daniela Castan

Première édition, 2026
Épistémé, Lausanne
Épistémé est une maison d'édition de la fondation
des Presses polytechniques et universitaires romandes
ISBN 978-2-88915-753-2, version imprimée
ISBN 978-2-8323-2346-5, version ebook (pdf), doi.org/10.55430/8076EAONSJ

Imprimé en Tchéquie



Ce texte est sous licence Creative Commons : elle vous oblige,
si vous utilisez cet écrit, à en citer l'auteur, la source et l'éditeur original,
sans modifications du texte ou de l'extrait et sans utilisation commerciale.

Sommaire

Acronymes et sigles _____ 7

Introduction _____ 9
Emmanuel Sylvestre & Jean-Michel Jullien

Partie 1
Évaluer à l'échelle de l'enseignement
Innovations pédagogiques et dispositifs en contexte de classe _____ 19

1 Le e-portfolio comme outil pour l'évaluation des compétences: analyse d'usage dans une formation de bachelor universitaire de technologie _____ 21
Aurélie Dupré & Éric Giraudin

2 La vidéo comme outil d'évaluation de la production orale: cas d'un groupe d'étudiantes de première année en licence de français à l'ENSB _____ 45
Assia Belgheddouche

3 Pratiques évaluatives d'enseignants du supérieur _____ 67
Aurélie Dupré & Anne-Céline Grolleau

4 *Feedbacks* par intelligence artificielle en école d'ingénieurs: un premier regard _____ 89
Étienne Gibaud, François Stephan & Diana Griffoulières

Partie 2

Évaluer à l'échelle de l'institution

Infrastructures, stratégies et transitions numériques

5 Intégration des outils numériques dans l'évaluation des apprentissages
Étude de cas: Université Cadi Ayyad – Maroc 107
Hiba Asri & Abdelali Rochdi

6 Évolution des modalités et infrastructures des examens en ligne 137
Laurent Moccozet, Omar Benkacem, André Camacho, Allan Bowman & Patrick Roth

7 Évaluations en ligne post-Covid: quels accompagnements et quel alignement pédagogique? 159
Evi Belsack & Patricia Mercier

Conclusion
Évaluer à l'ère numérique, entre potentialités, risques et responsabilités 179
Bernadette Charlier & Jean-Luc Gilles

Bibliographie 195

Présentation des auteurs et autrices 207

Table des matières 213

Acronymes et sigles

AMC	*Auto Multiple Choice*
APC	approche par compétences
BUT	bachelor universitaire de technologie
BYOD	*Bring Your Own Device*
CAP	Centre d'appui pédagogique
CIP	Centre d'innovation pédagogique
CTP	conseiller technopédagogique
DEC	Digital Education Council
DEUST	diplôme d'études universitaires scientifiques et techniques
DU	diplôme universitaire
ERA	*Emergency Remote Assessment*
ERT	*Emergency Remote Teaching*
IA	intelligence artificielle
IAg	intelligence articielle générative
IP	Internet Protocol
LMS	*Learning Management System*
MCD	modèle conceptuel de la base de données
MOOC	*Massive Open Online Course* (MOOCs *Massive Open Online Courses*)
PEP	professeur d'enseignement primaire
PHP	code informatique
PO	production orale
PTO	pratiques et techniques de l'oral
QCM	questionnaire à choix multiple(s)
RAM	*Random Access-memory*
RGPD	règlement général sur la protection des données
SAE	situation d'apprentissage et d'évaluation
SEB	*Safe Exam Browser*
SoTL	*Scholarship of Teaching and Learning*
TARC	*Typology of Assessment Responses to COVID-19*
TIC	technologies de l'information et de la communication
TICE	technologies de l'information et de la communication pour l'enseignement

Emmanuel Sylvestre
& Jean-Michel Jullien

Introduction

L'intégration croissante des technologies numériques dans l'enseignement supérieur transforme les pratiques d'évaluation des apprentissages, mais elle offre également de nouvelles opportunités dans les pratiques d'enseignement et d'apprentissage (Endrizzi, 2012). Au cours de la dernière décennie, et tout particulièrement pendant la pandémie de Covid-19, les établissements ont massivement adopté des outils numériques pour assurer la continuité pédagogique, ce qui a accéléré la transition vers ce que l'on appelle désormais l'e-évaluation (Fawns, 2019 ; Lee *et al.*, 2022 ; Lafleur *et al.*, 2021), ou *e-assessment* pour reprendre la terminologie anglo-saxonne. L'e-évaluation consiste en « l'utilisation de plateformes et de technologies numériques pour évaluer les progrès des apprenants par le biais de nombreuses activités, notamment la préparation de questions, la réception des réponses, la notation et le retour d'informations » (traduit de Wahas et Syed, 2024, p. 14432). Les outils numériques offrent ainsi de nombreuses possibilités de variantes et de combinaisons de modalités (en présence, à distance, mixte), d'objets sur lesquels reposent les évaluations (productions, oraux, quiz), de temporalités (synchrone, asynchrone, en continu et/ou terminal), ou d'acteurs (autoévaluation, évaluation par les pairs, automatique) (Sylvestre *et al.*, 2020). En parallèle, ils posent des défis inédits en matière d'équité, de sécurité et d'éthique. Cet ouvrage propose d'observer différentes pratiques de l'e-évaluation dans des établissements d'enseignement

supérieur selon plusieurs dimensions complémentaires : pédagogique, docimologique, technologique et éthique, afin d'éclairer les enjeux actuels et d'envisager des perspectives.

Dimension pédagogique

Du point de vue pédagogique, les outils numériques enrichissent le répertoire des méthodes d'évaluation en enseignement supérieur (Lafleur *et al.*, 2021). Ils favorisent l'essor de l'évaluation formative continue, grâce à des quiz en ligne, des sondages instantanés et des activités interactives en salle de classe qui donnent un *feedback* immédiat aux étudiants, même en grands groupes (Lepage *et al.*, 2019). L'e-évaluation permet d'améliorer la rapidité des retours fournis, renforçant leur autorégulation et pouvant mener à de meilleurs résultats d'apprentissage (Hettiarachchi *et al.*, 2015). Par exemple, des analyses montrent que les évaluations en ligne contribuent à améliorer l'engagement et la motivation des étudiants (Lepage *et al.*, 2019), tout en soutenant potentiellement de plus hauts niveaux de réussite grâce à un suivi régulier (Guàrdia *et al.*, 2017 ; Holmes, 2014).

En outre, les outils numériques élargissent le champ des compétences pouvant être évaluées (Ridgway *et al.*, 2004 ; El Asame *et al.*, 2022). Ils autorisent des tâches plus authentiques et complexes, difficiles à organiser sur papier, comme des simulations virtuelles, des études de cas interactives ou des jeux sérieux (*serious games*) immersifs. Par exemple, l'utilisation de jeux numériques éducatifs en évaluation a montré des effets positifs sur la motivation, tout en renforçant la confiance dans le processus d'apprentissage (Arias-Calderón *et al.*, 2022).

L'essor de l'e-évaluation permet également d'élargir le spectre du *feedback* fourni à la population étudiante. En complément du *feedback* traditionnellement transmis par les enseignants, il est possible d'apporter du *feedback* par les pairs, du *feedback* automatique par l'outil numérique ou de l'obtenir par autoévaluation, comme l'ont souligné Heil et Ifenthaler (2023) dans leur revue de littérature. D'une part, certaines plateformes en ligne permettent d'intégrer l'évaluation par les pairs à grande échelle (notamment dans les MOOCs [*Massive Open Online Courses*]), stimulant ainsi les capacités de collaboration, de pensée critique vis-à-vis du travail d'autrui et de métacognition (Gamage *et al.*, 2021). D'autre part, l'e-évaluation peut faciliter l'autoévaluation et la réflexion des étudiantes et des étudiants sur leurs apprentissages :

par exemple, via des e-portfolios (Qvortrup et Keiding, 2014) ou des questionnaires d'autopositionnement (Seufert *et al.*, 2019). Chaque étudiant peut ainsi suivre l'évolution de ses compétences et identifier ses points forts/faibles de façon autonome. Il est à noter que ces différentes modalités de *feedback* ont encore plus d'impact si elles sont utilisées conjointement (Gilles et Charlier, 2020).

Sur le plan pédagogique, l'e-évaluation peut agir comme un levier d'innovation permettant de diversifier les modalités d'évaluation (quiz en ligne, devoirs vidéo, simulations, etc.), de favoriser l'implication active de la communauté étudiante et d'évaluer certaines compétences jusque-là peu faciles d'accès avec des méthodes d'évaluation traditionnelles sur papier. Néanmoins, il importe d'aligner étroitement ces nouveaux dispositifs avec les objectifs pédagogiques et le curriculum afin d'en garantir la cohérence et l'efficacité, et de ne pas réduire l'e-évaluation à un simple outil facilitant les aspects logistiques de l'organisation d'un examen (Detroz *et al.*, 2020).

Dimension docimologique

La docimologie, ou science de l'évaluation, bénéficie également largement de la transition numérique. Contrairement aux évaluations traditionnelles qui laissent peu de traces, les évaluations numériques génèrent une quantité considérable de données sur les interactions, les réponses, les progrès et les parcours des étudiantes et des étudiants (Barana *et al.*, 2019). Chaque clic, chaque tentative de quiz, chaque soumission en ligne constitue une trace numérique potentiellement exploitable. Cette masse de données offre des opportunités sans précédent pour analyser et améliorer la qualité des évaluations elles-mêmes, ainsi que pour suivre finement le progrès de la communauté étudiante dans le temps (Ifenthaler et Greiff, 2021). On parle alors de *learning analytics*, ou analyse des traces d'apprentissage, appliquée à l'évaluation. Concrètement, les données issues des e-évaluations permettent d'identifier rapidement la difficulté et la pertinence des questions (en calculant, par exemple, des taux de réussite, des indices de discrimination, etc.), d'évaluer la fiabilité des épreuves (la cohérence des résultats, la détection d'items problématiques) et d'adapter en temps réel le parcours pédagogique proposé à chaque étudiant. Ainsi, dans un cours donné, l'analyse des résultats à un quiz en ligne peut révéler de mauvaises interprétations ou des idées fausses répandues et inciter

l'enseignant à y revenir en classe pour réguler les apprentissages. À l'échelle institutionnelle, l'agrégation des données d'évaluation peut servir à comparer des cohortes, à repérer des tendances longitudinales et à alimenter l'amélioration des programmes dans le cadre de démarches qualité. Chaque étudiant peut recevoir des tableaux de bord détaillés de sa progression, lui permettant d'observer l'acquisition de ses connaissances et compétences par rapport aux attentes de l'enseignant (Alam *et al.*, 2023).

Un domaine émergent, l'*assessment analytics* (Sahin et Ifenthaler, 2024), prône justement l'exploitation systématique de ces données d'évaluation pour fermer la boucle du *feedback* et enrichir les pratiques évaluatives. En mobilisant à la fois les données formatives (issues par exemple d'exercices d'entraînement) et sommatives (notes finales, examens) dans une même analyse, on peut obtenir une vision plus holistique du parcours d'apprentissage et prendre des décisions éclairées pour ajuster l'enseignement en continu. Ce potentiel est encore sous-exploité (Heil et Ifenthaler, 2023 ; Sahin et Ifenthaler, 2024) : trop souvent, l'évaluation est traitée comme une finalité du cursus plutôt que comme une source de données rétroactive au service de l'apprentissage. Un défi docimologique important consiste donc à intégrer pleinement ces analyses de traces dans la régulation des apprentissages, sans alourdir la charge pour les enseignants. Enfin, notons que l'essor de l'intelligence artificielle (IA) dans le domaine de l'évaluation, avec la possibilité de fournir des corrections automatisées de devoirs ou de proposer des *chatbots* tuteurs, soulève l'espoir d'une personnalisation plus accrue en apportant du *feedback* individualisé (Escotet, 2024 ; El Kamoun *et al.*, 2025).

Dimension technologique

Sur le plan technologique, l'évaluation numérique requiert une infrastructure fiable et accessible, ce qui représente un enjeu de taille pour les établissements et les usagers. Devant la diversité d'outils soutenant l'e-évaluation présents dans l'écosystème éducatif (Choi-Lundberg *et al.*, 2023), il est essentiel pour les universités de disposer de réseaux robustes, de serveurs capables de supporter des pics de connexions (par exemple, lors d'examens synchrones à grande échelle), et d'un service de support technique réactif. Or, le recours rapide et massif aux examens en ligne durant la pandémie a révélé des lacunes dans l'infrastructure

numérique de nombreuses institutions, qui ont dû passer en quelques semaines d'une adoption partielle du numérique à une évaluation totalement en ligne (Mate et Weidenhofer, 2021), et cela a pu occasionner un stress important tant dans la population enseignante qu'étudiante (Nolla *et al.*, 2022 ; Barras, 2020). En parallèle, des inégalités d'équipement et d'accès au sein de la communauté étudiante ont été aussi des facteurs de stress tels que l'accès à un ordinateur performant, une webcam de bonne qualité ou un lieu calme avec une connexion internet haut débit. L'enquête conduite en 2020 (Lee *et al.*, 2022) pendant le confinement imposé par le Covid-19 a montré que, pour plus de 70 % des étudiants interrogés, les examens en ligne étaient beaucoup plus sujets à des problèmes que les examens traditionnels. Plus de la moitié de la population étudiante a identifié les pannes techniques et la difficulté d'accès aux plateformes comme le principal obstacle rencontré pendant ses évaluations en ligne. Ces inégalités posent une question d'équité en creusant les écarts de performance indépendamment des compétences académiques réelles : certaines personnes perdent du temps ou du texte lors de déconnexions imprévues, par exemple. Les recherches dans le domaine rapportent que l'évaluation en ligne peut exacerber les désavantages socio-économiques ou géographiques en raison des différences d'accès aux technologies et de maîtrise de celles-ci (Mate et Weidenhofer, 2021). Assurer une égalité d'accès aux évaluations numériques est donc un impératif technique et social qui peut passer par le prêt d'ordinateurs, l'aménagement de salles informatiques en libre-service, ou l'adoption de l'approche BYOD (*Bring Your Own Device*) encadrée par l'institution pour que chacun puisse composer dans des conditions équitables.

La sécurité des examens en ligne constitue l'autre grand volet technologique. Il s'agit d'une part, de sécuriser les contenus (prévenir les fuites de sujets, la copie non autorisée des questionnaires, etc.) et d'autre part, d'authentifier l'identité de la personne étudiante ainsi que l'exactitude de sa composition. De nombreux systèmes d'e-évaluation intègrent aujourd'hui des modules de surveillance (*proctoring*) proposant un enregistrement vidéo pendant l'examen, le verrouillage du navigateur, la détection automatique de comportements suspects via l'intelligence artificielle (regards détournés, présence d'une tierce personne…), etc. Différents fournisseurs technologiques offrent des solutions visant à réduire les comportements frauduleux, mais leur efficacité autant que leur acceptabilité font débat comme le souligne

Dawson (2024). En effet, la technologie ne peut pas à elle seule garantir l'intégrité et la conception même des épreuves joue un rôle clé pour décourager la tricherie (par exemple, en posant des questions différentes et aléatoires, en limitant le temps de réponse pour empêcher les recherches externes, ou en évaluant des productions plus personnelles et réflexives). Par ailleurs, les dispositifs de surveillance en ligne soulèvent de nouveaux défis techniques et organisationnels (faux positifs de détection, besoin d'une connexion stable pour le flux vidéo, protection des données enregistrées) qui requièrent une amélioration continue. Les établissements doivent ainsi trouver un équilibre entre des mesures de sécurité renforcées et le respect de l'expérience enseignante et étudiante. Enfin, notons que l'accessibilité pour les personnes étudiantes en situation de handicap doit rester une priorité dans le déploiement de ces technologies. La conception des évaluations numériques ne doit pas exclure certaines personnes et des principes universels d'accessibilité doivent être déployés tels que la compatibilité avec les lecteurs d'écran, les sous-titres, les interfaces navigables au clavier, etc. Le volet technologique de l'évaluation numérique oblige donc à prendre en compte simultanément des considérations d'infrastructure (disponibilité et fiabilité des outils), d'accessibilité (prise en main facile et inclusive pour tous les publics) et de sécurité (intégrité des examens), sans quoi les plus-values pédagogiques de l'e-évaluation peuvent être compromises par des obstacles purement techniques.

Dimension éthique

La généralisation des outils numériques d'évaluation s'accompagne de profondes interrogations éthiques. La question de l'équité et de la justice sociale est centrale, car, si les évaluations en ligne ne sont pas déployées de façon à compenser les inégalités préexistantes, elles peuvent au contraire les accentuer (Mate et Weidenhofer, 2021). Comme nous l'avons vu dans la section précédente, les écarts dans l'accès aux technologies ou dans les conditions de passation défavorisent certaines personnes. Un dispositif d'évaluation numérique plus inclusif doit veiller à prendre en compte l'ensemble des profils des apprenants comme les personnes en situation de handicap, celles expatriées dans des fuseaux horaires différents, etc., faute de quoi apparaissent de nouvelles formes d'exclusion. Les institutions accordent donc une attention croissante aux principes d'évaluation inclusive en contexte

numérique, par exemple en offrant des alternatives équivalentes en cas de besoins spécifiques (Noguera *et al.*, 2018), mais aussi en formant la population étudiante à l'usage des outils en amont pour réduire la fracture numérique. Paradoxalement, certaines recherches notent que les « personnes natives du numérique » ne sont pas forcément mieux préparées face à des outils d'évaluation inhabituels (Roy *et al.*, 2018 ; Lemieux, 2021), et qu'elles peuvent paniquer lorsque la technologie ne fonctionne pas parfaitement. Assurer l'équité, c'est donc accompagner l'ensemble de la population étudiante, quel que soit son bagage technologique, afin qu'elle aborde l'évaluation dans des conditions optimales.

L'intégrité académique demeure également un enjeu majeur à l'ère du numérique. La facilité d'accès à Internet et aux ressources en ligne pendant une évaluation à distance augmente la tentation et les opportunités de tricher (Hébert *et al.*, 2024) comme une consultation non autorisée de documents, la communication entre pairs via la messagerie, la sollicitation de tiers pour réaliser l'épreuve à sa place, etc. Plus récemment encore, l'émergence de l'intelligence artificielle générative (IAg) soulève de nouvelles préoccupations quant à l'authenticité des productions rendues par les étudiants (Many *et al.*, 2024 ; Audran, 2024). Par exemple, une personne étudiante demandant à un *chatbot* de rédiger une dissertation ou de résoudre un problème de programmation, qui brouille ainsi la frontière de la paternité du travail. D'un côté, des solutions techniques de plus en plus sophistiquées apparaissent pour vérifier l'identité et l'originalité (suivi des adresses IP [Internet Protocol], analyse des plagiats en ligne, outils de détection d'IA dans les textes), de l'autre, les moyens de contournement évoluent tout aussi vite. Pour tenter de réguler ces pratiques non souhaitées, certaines institutions ont fait le choix de mettre le focus sur la sensibilisation de la communauté étudiante en proposant, par exemple, un code d'honneur ou une charte éthique (Many *et al.*, 2024), ou en demandant au corps enseignant de concevoir des activités d'évaluation moins vulnérables à la tricherie comme des devoirs personnalisés, des questions faisant appel à l'expérience personnelle ou à la réflexion critique plutôt qu'à la récitation brute de connaissances facilement trouvables (Lafleur *et al.*, 2021 ; Béchard, 2023) et d'installer un climat de confiance qui valorise l'apprentissage plutôt que la seule performance chiffrée (Béchard, 2023).

Enfin, la protection des données et de la vie privée constitue une dimension éthique incontournable. L'évaluation numérique entraîne la production d'une multitude de données personnelles sur la population

étudiante (identifiants, vidéos de surveillance, enregistrements audio, localisation IP, copies d'examen, résultats, etc.). Si le potentiel d'utilisation de ces traces produites lors des e-évaluations offre la possibilité aux établissements d'améliorer la qualité des évaluations et à la population étudiante de progresser plus efficacement, il est nécessaire de s'assurer des bases légales et réglementaires applicables (RGPD [règlement général sur la protection des données] en Europe par exemple, règlement des études ou des examens dans les établissements) : ce qui est possible dans un établissement ne l'est pas forcément dans un autre qui ne dispose pas du même cadre légal. Sur le plan éthique, ces précautions légales s'accompagnent d'une information claire transmise à la population étudiante et enseignante sur le traitement des données et le respect des droits individuels.

Que propose cet ouvrage ?

Comme nous venons de le voir, l'e-évaluation dans l'enseignement supérieur recèle un formidable potentiel d'impulsion pédagogique, tout en soulevant des enjeux docimologiques, technologiques et éthiques qu'il est crucial d'interroger. Actuellement, les institutions d'enseignement supérieur et la population enseignante naviguent entre enthousiasme et prudence. Les retours d'expérience montrent que les évaluations numériques bien orchestrées améliorent les apprentissages en offrant plus de *feedback*, de flexibilité et en mobilisant davantage la population étudiante dans son parcours, mais les difficultés techniques et les questions d'intégrité ont parfois entaché la confiance de la communauté dans ces dispositifs. Les enjeux actuels consistent à fiabiliser les outils et les pratiques (formation de la population enseignante et étudiante, support technique, cadre institutionnel clair) afin de pérenniser les gains observés tout en réduisant les écueils. Il s'agit notamment de développer une culture partagée de l'évaluation numérique dans laquelle l'ensemble de la communauté est conscient des pratiques possibles et des limites de ces nouveaux modes d'évaluation.

Dans le cadre de cet ouvrage, notre volonté est de présenter une diversité de dispositifs mis en place au niveau institutionnel ou au niveau de l'enseignement individuel afin d'illustrer des pratiques actuelles. Ces dispositifs, ancrés dans des contextes variés, permettent d'apprécier tant la diversité des finalités poursuivies, les modalités mobilisées, que les conditions de mise en œuvre pratique. Ils offrent

un aperçu concret des réponses apportées par les actrices et acteurs de terrain aux défis soulevés par l'e-évaluation et visent à nourrir une réflexion collective sur son développement futur.

Cet ouvrage est structuré en deux parties : la première présente quatre dispositifs mis en place au niveau d'un enseignement ou d'une formation, et la seconde, trois dispositifs mis en place au niveau institutionnel.

Dans le chapitre 1, Aurélie Dupré et Éric Giraudin analysent l'usage du e-portfolio pour évaluer les compétences dans un bachelor universitaire de technologie (BUT) en France. Leur étude de cas interroge l'appropriation de l'outil e-portfolio par la population étudiante, la construction des preuves de compétence, et le processus de démarche réflexive associé. En croisant autoévaluations et évaluations enseignantes, ce chapitre révèle les atouts et les limites de ce dispositif, et invite à réfléchir à ses usages pour favoriser l'apprentissage et l'autopositionnement.

Dans le chapitre 2, Assia Belgheddouche explore l'usage de la vidéo asynchrone pour évaluer la production orale en première année d'étude de français à l'École normale supérieure de Bouzaréah (ENSB) en Algérie. L'étude analyse l'impact de ce mode d'évaluation sur la performance, l'anxiété et l'autonomie de la population étudiante. Ce chapitre met en évidence les apports et les contraintes de la vidéo comme outil d'évaluation en langues.

Le chapitre 3, rédigé par Aurélie Dupré et Anne-Céline Grolleau, présente une analyse des pratiques d'e-évaluation auprès d'une population enseignante engagée dans un diplôme universitaire (DU) en pédagogie universitaire en France. À partir de portfolios réflexifs, elles mettent en évidence une évolution vers des évaluations centrées sur l'apprentissage, principalement formatives ou diagnostiques.

Dans le chapitre 4, Étienne Gibaud et ses collègues analysent l'usage de ChatGPT pour fournir un *feedback* personnalisé à des étudiantes et étudiants en école d'ingénieur en France. En comparant les effets de *feedbacks* donnés par un enseignant et par une intelligence artificielle, ils révèlent le potentiel formateur de l'IA générative, tout en concluant que l'impact du *feedback* humain reste plus marqué. Ce chapitre ouvre des pistes de réflexion sur l'intégration raisonnée de l'IA dans les dispositifs d'évaluation formative.

Le chapitre 5 présente une étude de cas conduite à l'Université Cadi Ayyad au Maroc par Hiba Asri et Abdelali Rochdi. Les auteurs comparent deux modalités d'e-évaluation, l'une entièrement menée dans

Moodle et la seconde, numérique et analogique, avec *Auto Multiple Choice* (AMC) pour faire face à la massification des effectifs. L'approche adoptée propose d'automatiser les évaluations, d'améliorer la transparence et de réduire la charge sur la population enseignante, tout en garantissant des pratiques équitables et fiables dans un contexte d'évaluation à grande échelle.

Le chapitre 6 retrace, sous la plume de Laurent Moccozet et ses collègues, la transformation des pratiques évaluatives à l'Université de Genève en Suisse, avant, pendant et après la pandémie de Covid-19. Le chapitre interroge la transition forcée vers les examens en ligne et les évolutions post-crise, révélant comment l'infrastructure et les usages se sont adaptés, entre héritage du numérique imposé et innovations durables.

Dans le septième et dernier chapitre, Evi Belsack et Patricia Mercier interrogent la persistance des examens en ligne à l'Université libre de Bruxelles après la pandémie de Covid-19. À partir d'une analyse quantitative, elles examinent l'usage de Moodle, le rôle du Centre d'appui pédagogique, et les tensions entre autonomie enseignante et alignement pédagogique. Ce chapitre invite à repenser l'accompagnement nécessaire dans un contexte d'évaluation numérique pérenne.

L'ouvrage se conclut par une analyse transversale proposée par Bernadette Charlier et Jean-Luc Gilles, qui revisitent les quatre dimensions clés : pédagogique, docimologique, technologique et éthique, à la lumière des chapitres présentés. Leur contribution met en perspective la diversité des pratiques observées et interroge les conditions d'une évaluation numérique pertinente et soutenable dans un contexte en constante mutation, marqué par les transformations post-Covid et l'émergence rapide de l'intelligence artificielle dans l'enseignement supérieur.

Partie 1

Évaluer à l'échelle de l'enseignement

Innovations pédagogiques et dispositifs en contexte de classe

Aurélie Dupré & Éric Giraudin

1 Le e-portfolio comme outil pour l'évaluation des compétences : analyse d'usage dans une formation de bachelor universitaire de technologie

Résumé

Ce chapitre s'intéresse à l'usage du e-portfolio comme outil d'évaluation des compétences. Les auteurs proposent d'analyser un cas d'usage en première année de bachelor universitaire de technologie. L'étude s'intéresse à la manière dont les apprenants se sont saisis de la démarche et de l'outil pour rendre compte d'un niveau de développement d'une des compétences visées par la formation. Il met en perspective les productions des étudiants et les évaluations portées d'une part par les apprenants eux-mêmes dans une démarche d'autopositionnement et d'autre part par les enseignants évaluateurs. Les questions posées sont alors les suivantes : comment les étudiants ont-ils construit la preuve de leur compétence à travers cet e-portfolio ? En quoi l'outil a-t-il pu soutenir ou, à l'inverse, nuire à leur démarche réflexive ? L'analyse de cet usage du e-portfolio ouvre des perspectives très prometteuses qui devraient stimuler les réflexions au sein des communautés universitaires.

Introduction

L'approche par compétences (APC) est une manière d'organiser les programmes de formation largement promue par les discours politiques

et institutionnels au sein de l'enseignement supérieur (Kaddouri et Loiola, 2022). Les communautés universitaires (enseignants du supérieur, conseillers et ingénieurs pédagogiques, responsables de formations, formateurs, tuteurs, chercheurs, décideurs politiques, étudiants, etc.), cherchant à se saisir de la démarche, s'engagent progressivement dans ce mouvement de transformation pédagogique en fonction de leurs objectifs propres. Si les manières de mettre en œuvre l'APC sont multiples, l'évaluation des compétences apparaît comme un enjeu essentiel de cette démarche. Le e-portfolio est un outil numérique visant à soutenir les pratiques d'évaluation des compétences. Nous proposons, dans cette contribution, d'analyser un cas d'usage de cet outil pour l'évaluation des compétences des étudiants d'une formation de bachelor universitaire de technologie (BUT).

Le contexte et la problématique

Le programme Avenir(s) – France 2030, piloté par l'Université Savoie Mont-Blanc et développé avec ses partenaires (Université de Lille, Aix-Marseille Université, Université de Bretagne occidentale, Fondation partenariale Polytech, Association nationale ESUP-Portail), vise à soutenir le déploiement de l'APC dans l'enseignement supérieur. Il s'organise autour de trois grandes actions : (1) la professionnalisation des enseignants du supérieur sur l'APC et l'accompagnement d'équipes pédagogiques (relevant de différents contextes du point de vue du champ disciplinaire, de la taille de cohorte…) dans la mise en œuvre de la démarche ; (2) l'outillage à modélisation des parcours de formation en APC au sein des systèmes d'information universitaires ; (3) le développement d'un outil e-portfolio pour l'évaluation des compétences. L'approche se veut collaborative, les outils pédagogiques et numériques développés sont coconstruits avec la communauté universitaire. Le développement d'un outil e-portfolio est un attendu majeur du programme. Pour ce faire, il a été décidé de partir d'une version existante de l'outil e-portfolio prototype Karuta (KAPC+), actuellement non satisfaisante au regard des enjeux d'une APC et d'une utilisation à grande échelle. Les fonctionnalités du nouvel outil sont définies à partir des retours d'expériences des équipes pédagogiques sur l'usage de la solution KAPC+, de séances de codesign organisées avec différents acteurs en lien avec la communauté universitaire (enseignants, étudiants, personnels non enseignants d'établissements, entreprises et collectivités) lors

de journées événementielles et à partir d'ateliers de travail de l'équipe Avenir(s) avec un prestataire spécialisé dans la traduction des expressions de besoins en spécifications fonctionnelles, puis techniques.

À ce stade, nous constatons, au sein des formations suivies dans ce programme, différents niveaux d'engagement dans la démarche. Certaines sont engagées dans un travail de fond visant à repenser leurs pratiques pédagogiques dans le but de placer l'apprenant dans une démarche d'acquisition de compétences. D'autres sont soucieuses de répondre à la demande en affichant un référentiel de compétences, mais elles tendent à vouloir conserver leurs pratiques actuelles. D'autres encore ont bien engagé un travail autour de l'élaboration d'un référentiel de compétences, mais elles semblent plus démunies pour poursuivre le travail vers l'entraînement et l'évaluation des compétences au sein de leur parcours de formation (questionnement sur la méthode, manque de moyens, tensions dans l'équipe, contraintes externes, etc.). Certaines, enfin, avancent en perdant le sens pédagogique de la démarche et se retrouvent à faire des choix problématiques au regard de leurs propres objectifs. À l'échelle du programme, nous constatons que les concrétisations effectives dans les pratiques pédagogiques et dans les méthodes d'évaluation sont encore timides. Peu d'utilisateurs, de fait, recourent au e-portfolio pour le moment. Toutefois, certaines initiatives semblent très prometteuses, notamment en raison de l'évolution qu'elles apportent dans l'expérience d'apprentissage vécue par les apprenants. Ainsi, l'expression des compétences peut favoriser la perception de la valeur d'une formation pour les futurs diplômés, les pratiques pédagogiques actives et collaboratives sous-jacentes à l'entraînement des compétences tendent à stimuler l'engagement des apprenants. À travers le portfolio, ces derniers analysent leurs démarches, explicitent leurs manières de faire, se positionnent sur le développement de leurs compétences et développent des capacités réflexives qui nourrissent leur parcours d'apprentissage.

Nous nous intéressons, dans cet article, à l'une d'elles et nous proposons d'analyser l'usage d'un e-portfolio pour l'évaluation des compétences visées par une formation de première année de bachelor universitaire de technologie. Notre étude se penche sur la manière dont les apprenants se sont saisis de la démarche et de l'outil pour rendre compte d'un niveau de développement d'une des compétences visées par la formation. Nos questions seront les suivantes : comment les étudiants ont-ils construit la preuve de leur compétence à travers

cet outil e-portfolio ? En quoi l'outil a-t-il pu soutenir ou, à l'inverse, nuire à leur démarche réflexive ?

Éléments de cadrage théorique

Nous nous attachons, dans un premier temps, à expliciter la notion de compétence et l'approche pédagogique qui lui est associée. Nous nous arrêtons ensuite sur l'évaluation des compétences et explicitons la démarche portfolio.

L'approche par compétences

Le terme de compétence est largement utilisé dans le langage courant et est, de fait, devenu polysémique dans le vocabulaire commun. Pour Coulet (2011, p. 17), c'est « une organisation dynamique de l'activité, mobilisée et régulée par un sujet pour faire face à une tâche donnée, dans une situation déterminée ». Roegiers (2000, p. 66) évoque, quant à lui, « la possibilité, pour un individu, de mobiliser un ensemble intégré de ressources en vue de résoudre une situation-problème qui appartient à une famille de situations ». Pour Le Boterf (1994), la compétence est de l'ordre du savoir-mobiliser pour agir en situation. Jacques Tardif (2006, p. 22), quant à lui, définit la compétence comme « un savoir-agir complexe reposant sur la mobilisation efficace de ressources internes et externes ancrées dans des familles de situations ». Cette dernière définition, largement relayée dans les contextes d'enseignement supérieur, met l'accent sur la notion de savoir-agir complexe qui n'est pas réductible à la capacité à effectuer une tâche, mais qui demande d'être en mesure de donner une réponse appropriée à un problème jamais rencontré (Barras, 2023). Dans une perspective de formation, il s'agit alors de permettre à l'apprenant d'agir dans des environnements complexes adossés à des situations authentiques (Tardif, 2006). Cela suppose de dépasser une approche traditionnelle centrée sur l'acquisition de connaissances dispensées au sein de champs disciplinaires pour concevoir des dispositifs dans lesquels l'apprenant sera en action et mettra en œuvre les compétences visées. Les connaissances, toujours essentielles dans le programme de formation, viennent alors contribuer au développement des capacités d'agir des apprenants aux côtés d'autres ressources complémentaires (approches méthodologiques, témoignages de professionnels, applications pratiques, etc.).

Les travaux du LabSET[1] (Poumay *et al.*, 2017 ; Poumay et Georges, 2022) proposent un cadre méthodologique, documenté par des publications scientifiques, pour soutenir dans cette démarche le travail des équipes pédagogiques. Il part de l'identification des compétences visées par la formation et de la constitution d'un référentiel en plusieurs étapes : la définition des critères d'exigence attendus pour chaque compétence, l'identification de familles de situations professionnelles, la formalisation de niveaux de développement de chaque compétence par paliers et des apprentissages essentiels associés. Il s'attache ensuite à considérer les pratiques pédagogiques pour organiser des activités d'apprentissage, dites « situations d'apprentissage et d'évaluation » (SAE, comme des séminaires, des études de cas, des projets et mini-projets, des travaux pratiques, des simulations, des *serious games*, etc.), problématisées de manière à soutenir le développement d'une ou plusieurs compétences visées par le programme de formation. Si d'autres démarches sont possibles, celle-ci est reprise par un certain nombre d'équipes pédagogiques dans l'enseignement supérieur et a été structurante pour la création des BUT.

L'évaluation des compétences et la démarche e-portfolio

Pour apprécier le développement des compétences, Wiggins (1990, 1991, cité par Barras, 2023) pose les principes d'une évaluation authentique fondée sur un examen où l'apprenant est mis en situation. Il ne s'agit alors plus seulement de mesurer l'acquisition ou la compréhension de contenus disciplinaires, mais plutôt d'apprécier la manière dont l'apprenant va les mobiliser pour agir dans différents contextes, différentes situations (Tardif, 2017). Cette capacité d'agir est appréciée au regard de performances et de réalisations d'apprentissage, mais aussi à travers la réflexivité que l'apprenant développe sur ses manières de faire : « Une personne ne pourra vraiment être reconnue comme compétente que si elle est capable [...] de comprendre pourquoi et comment elle s'y prend pour agir » (Le Boterf, 2002, p. 21). Pour l'étudiant, il s'agit alors d'être capable d'identifier comment les apports des contenus de

[1] Le LabSET-ULiège est un centre de recherche spécialisé en pédagogie de l'enseignement supérieur. Ses domaines de recherche ont trait aux réformes de programmes dans l'enseignement supérieur, à l'apprentissage et à l'évaluation de performances complexes (compétences) et au conseil pédagogique.

formation guident les manières d'agir et de s'inscrire dans une trajectoire de développement des capacités d'agir (Wiggins, 1990, 1991, cité par Barras, 2023). Apprentissage et évaluation sont enchâssés, par l'association de temps de mises en situation et d'évaluation formatives et certificatives (autoévaluation, évaluation par les pairs, *feedbacks* individuel et collectif, grilles critériées, etc.). Les apprenants ont alors la possibilité d'agir en situation proche du réel, d'apprendre de leurs erreurs, d'estimer leur progression par rapport aux attendus, de se réguler et de se projeter afin de s'adapter à de futures situations (Scallon, 2015). À travers la démarche réflexive de l'apprenant, l'enseignant apprécie le développement de ses compétences.

Poumay et Georges (2022) proposent d'évaluer le développement d'une compétence au travers de quatre dimensions : (1) les productions, c'est-à-dire la mise en valeur des résultats obtenus par l'apprenant et donc des performances accomplies au cours de son parcours de formation ou de ses expériences extracurriculaires ; (2) la justification des manières d'agir, notamment au regard des savoirs théoriques et pratiques qu'il a mobilisés ; (3) la régulation, c'est-à-dire sa capacité à analyser, de manière critique, les actions effectuées, à proposer des pistes d'amélioration ; par exemple, (4) la transférabilité et donc sa capacité, enfin, à se projeter dans d'autres situations, d'autres contextes et à adapter ses manières de faire pour y répondre adéquatement. La démarche portfolio est présentée comme une démarche permettant de structurer et d'accompagner la réflexivité de l'apprenant. Le e-portfolio rassemble une collection de traces sélectionnées et commentées par l'étudiant pour illustrer les compétences qu'il a développées. Il peut être utilisé à différentes fins : portfolio d'apprentissage, de réussite, de développement professionnel, de présentation (MESRI, 2013), c'est un outil pour appuyer l'évaluation des compétences (Poumay, 2017 ; Romainville et Fischer, 2020).

Présentation du dispositif

Nous nous intéressons ici à la mise en œuvre d'une démarche e-portfolio en première année de BUT Métiers du multimédia et de l'Internet (MMI) pour l'évaluation de la compétence présentée dans le tableau 1.1.

TABLEAU 1.1 Extrait du référentiel de compétences.

Compétence	Développer pour le Web et les médias critiques
Critères d'exigence (ou composantes essentielles)	• en se conformant aux standards du Web et aux normes d'accessibilité ; • en s'appuyant sur des concepts théoriques issus de l'informatique et des sciences de l'information ; • en produisant du code fonctionnel, sobre et réutilisable.
Niveau attendu	Développer un site web simple et le mettre en ligne
Tâche confiée aux étudiants (SAE 2.03 – MMI)	Développer un site web simple et le mettre en ligne. Le travail est réalisé en petits groupes, les étudiants disposent d'un outil numérique de gestion des tâches

Les étudiants utilisent l'outil e-portfolio dont l'interface est présentée ci-après, figure 1.1. Ils déposent des traces dans la zone 5 « Mes traces significatives » et rédigent un écrit réflexif dans la zone 4 « Analyse et réflexivité sur vos actions » pour justifier de l'atteinte du niveau de compétence attendu. Ils doivent également se positionner sur les apprentissages critiques associés au niveau de la compétence visée dans la zone 1 « Compétences et niveau associé » et sur une grille critériée fondée sur les composantes essentielles de la compétence. Une soutenance orale est organisée, les enseignants ont accès aux e-portfolios des étudiants en amont de la présentation orale. Sur la base de ces éléments produits par l'étudiant et de ses positionnements, et en s'appuyant sur la grille critériée dans la zone 7 « Évaluation de la SAE » qui a été partagée en amont avec les étudiants, les enseignants évaluent le niveau de la compétence visée.

La méthodologie

L'étude porte sur 54 productions d'étudiants évaluées systématiquement par deux enseignants (que nous appellerons E1 et E2) dans un enseignement type SAE de première année du BUT MMI visant la compétence « Développer un site web simple et le mettre en ligne » explicitée dans le tableau 1.1. Le corpus est constitué des extraits de e-portfolios qui permettent d'accéder aux preuves déposées par les étudiants pour justifier du développement de leur compétence (interface de la figure 1.1) ; à leur autopositionnement ; aux évaluations portées par les enseignants. Nous préciserons, dans les paragraphes qui suivent, les variables retenues pour leur analyse. L'ensemble des données recueillies sont anonymisées.

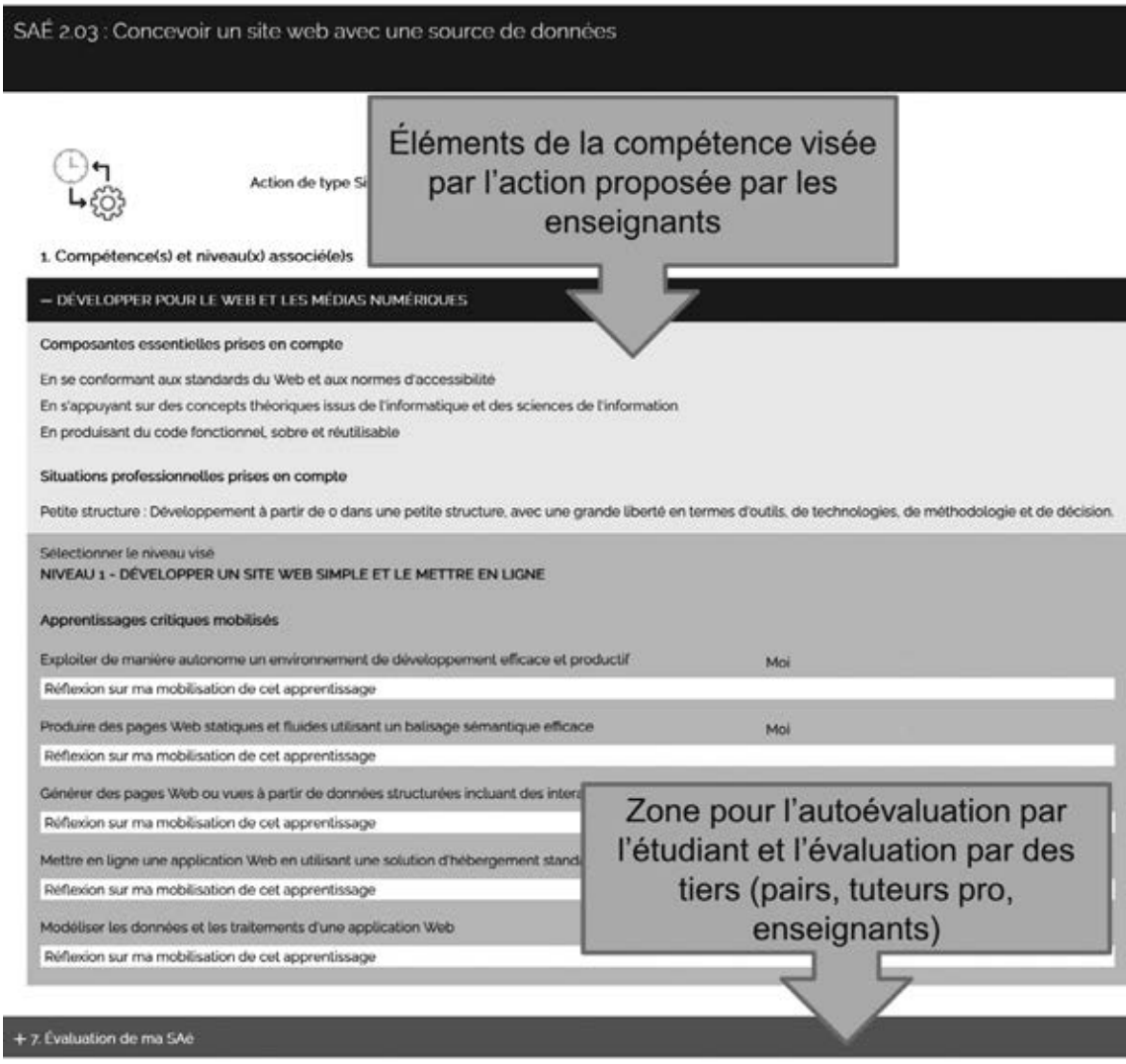
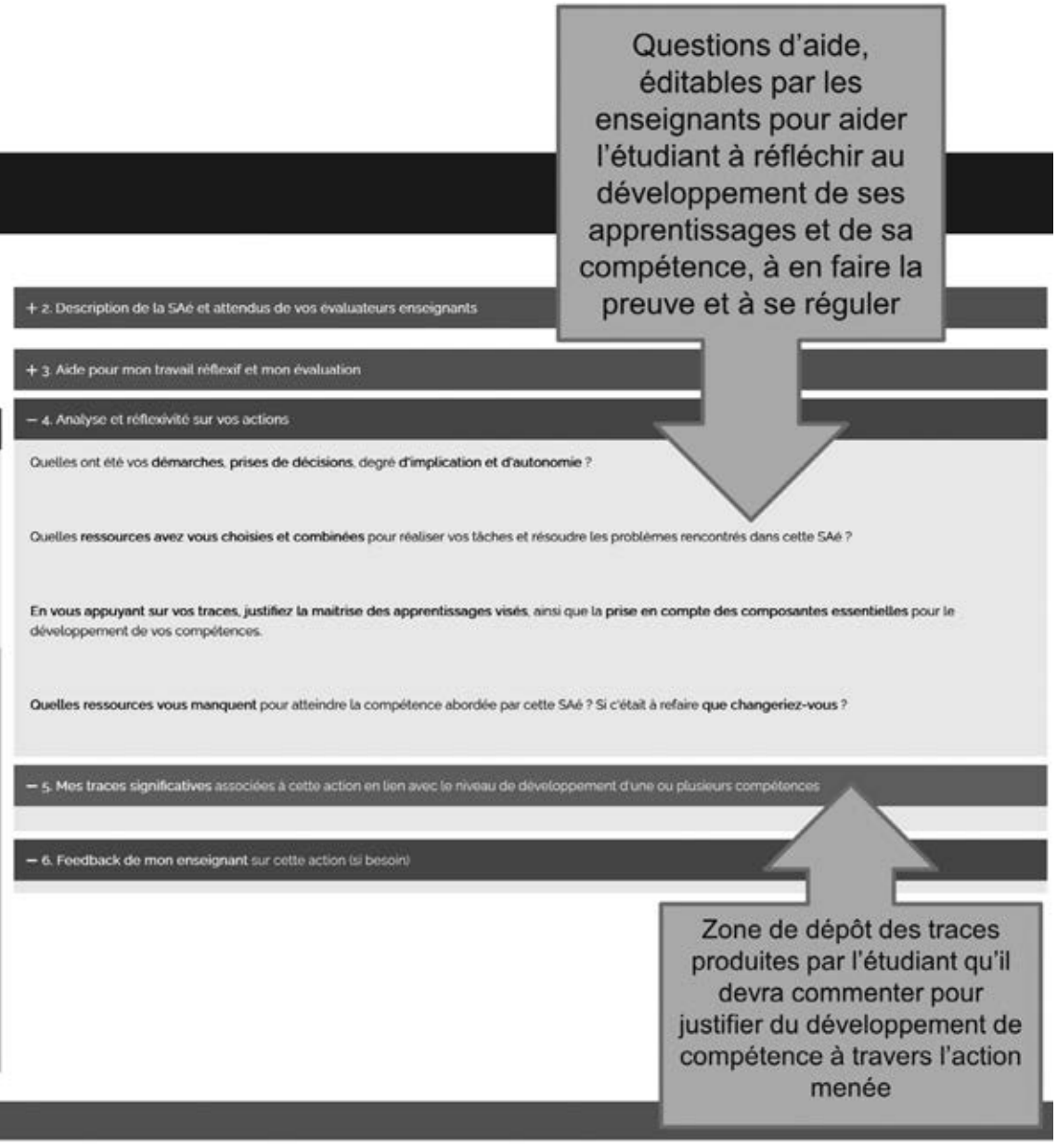


FIGURE 1.1 Présentation de l'interface de dépôt de preuve (Karuta KAPC+).


Questions d'aide, éditables par les enseignants pour aider l'étudiant à réfléchir au développement de ses apprentissages et de sa compétence, à en faire la preuve et à se réguler
+ 2. Description de la SAé et attendus de vos évaluateurs enseignants
+ 3. Aide pour mon travail réflexif et mon évaluation
– 4. Analyse et réflexivité sur vos actions
Quelles ont été vos démarches, prises de décisions, degré d'implication et d'autonomie ?
Quelles ressources avez vous choisies et combinées pour réaliser vos tâches et résoudre les problèmes rencontrés dans cette SAé ?
En vous appuyant sur vos traces, justifiez la maitrise des apprentissages visés, ainsi que la prise en compte des composantes essentielles pour le développement de vos compétences.
Quelles ressources vous manquent pour atteindre la compétence abordée par cette SAé ? Si c'était à refaire que changeriez-vous ?
– 5. Mes traces significatives associées à cette action en lien avec le niveau de développement d'une ou plusieurs compétences
– 6. Feedback de mon enseignant sur cette action (si besoin)
Zone de dépôt des traces produites par l'étudiant qu'il devra commenter pour justifier du développement de compétence à travers l'action menée

Variables d'analyse des productions étudiantes

L'analyse des productions étudiantes s'est attachée à apprécier la qualité de la preuve de compétence faite par les étudiants dans ces e-portfolios au regard de trois dimensions : (1) la mise en valeur des résultats ; (2) la qualité de la justification des manières d'agir ; (3) l'analyse critique des actions effectuées ou régulation. Les variables retenues pour cette analyse sont précisées dans le tableau 1.2.

TABLEAU 1.2 Variables retenues pour l'analyse des productions étudiantes.

	Variables étudiées	Modalités de la variable
Production	Présentation des résultats produits	Non ; oui, mais résultats peu convaincants ; oui et résultats convaincants
	Présentation de contre-performances, situations insatisfaisantes ou échecs	Oui/Non
Justification des manières d'agir	Justification apportée	Pas de justification ; justification narrative ou descriptive ; justification explicative sans lien avec la compétence ; justification explicative en lien avec la compétence
	Usage du « Je »	Oui/Non
	Type de ressources mobilisées pour la justification	Pas de ressource mobilisée ; mobilisation des ressources fournies dans la SAE ; mobilisation de ressources apportées par la SAE et appui sur d'autres enseignements de la formation ; mobilisation de ressources externes
	Manière de mobiliser les ressources dans l'écriture	Pas de ressource mobilisée ; citation dans le texte ; citation dans le texte avec explicitation de son apport, mais sans lien avec la compétence (« pour ») ; citation dans le texte pour justifier l'action réalisée (« parce que »)
	Pertinence de la justification apportée	Pas convaincante ; convaincante pour la maîtrise des apprentissages visés ; convaincante vis-à-vis des composantes essentielles identifiées pour la compétence
Régulation	Niveau d'analyse critique portée par l'étudiant	Non ; oui, mais décontextualisée de la compétence visée ; oui, sur la démarche procédurale suivie ; aboutie en lien avec la compétence
	Preuve convaincante du niveau de compétence dans cette situation	Pas du tout (––) ; pas vraiment (-) ; oui (+) ; oui de façon très convaincante (++)

Variables d'analyse des autopositionnements des étudiants

L'analyse s'est focalisée ici sur la zone pour l'autoévaluation des étudiants. Les étudiants devaient se positionner sur une grille d'évaluation à quatre niveaux (voir tableau 1.3). Ce positionnement, bien que demandé, n'était pas obligatoire. Les variables considérées sont précisées dans le tableau 1.4.

TABLEAU 1.3 Grille critériée d'appui au positionnement étudiant.

Non validé (0 à 9)	Validé (10 à 12)	Validé avec brio (13 à 16)	Félicitations du jury (17 à 20)
• Mauvaise explicitation • Ne répond pas correctement à la problématique visée • Ne prend pas en compte toutes les composantes essentielles associées • Certaines traces sont contestables : validité, authenticité… • Les traces n'abordent pas tous les attendus visés (compétences, apprentissages, ressources…)	• Répond bien à la problématique visée par la SAE • Démonstration explicite • Traces qui attestent bien les niveaux de compétences, apprentissages et ressources incontournables visés • Prise en compte de toutes les composantes essentielles associées	• Présente des composantes bien intégrées • Explicitation très claire de la démarche empruntée • Choix pertinents et combinaisons efficaces des ressources • Fait preuve d'auto-régulation	• Met bien en relation toutes les composantes essentielles entre elles (combinaison, compromis, dynamique) • Évalue bien son action et propose des pistes d'amélioration et d'autorégulation réalistes et adaptées

TABLEAU 1.4 Variables retenues pour l'analyse des autopositionnements des étudiants.

Variables étudiées	Modalités
Autopositionnement de l'étudiant	Oui/Non
Note portée par l'étudiant sur son travail	
Type de commentaire apporté par l'étudiant	Pas de commentaire ; commentaire sur l'action réalisée ; commentaire sur les apprentissages réalisés ; commentaire sur les difficultés éprouvées ; autre

Variables d'analyse des évaluations portées par les enseignants

Notre analyse s'est enfin portée sur les évaluations des enseignants. Chaque enseignant devait indiquer sa note et pouvait ajouter un commentaire dans une zone de texte dédiée et partagée entre eux. Ils décidaient ensemble de la note finale (voir tableau 1.5).

TABLEAU 1.5 Variables retenues pour l'analyse des évaluations portées par les enseignants (E1 et E2).

Variables étudiées	Modalités
Note attribuée par E1	*Note chiffrée*
Commentaire de E1	Pas de commentaire ; commentaire centré sur les résultats ; commentaire centré sur les démarches ; commentaire centré sur les apprentissages ; commentaire centré sur la démarche réflexive ; commentaire centré sur l'acquisition de la compétence
Note attribuée par E2	*Note chiffrée*
Commentaire de E2	Pas de commentaire ; commentaire centré sur les résultats ; commentaire centré sur les démarches ; commentaire centré sur les apprentissages ; commentaire centré sur la démarche réflexive ; commentaire centré sur l'acquisition de la compétence
Note finale	*Note chiffrée*

Présentation des résultats

Nous présentons, dans les paragraphes qui suivent, les résultats des analyses en les illustrant, quand cela est possible, de verbatim extraits des corpus. Nous nous focaliserons d'abord sur les productions des étudiants pour démontrer la compétence visée. Nous présenterons ensuite les résultats des analyses des démarches évaluatives des étudiants et des enseignants. Nous mettrons enfin en perspective ces différents résultats pour discuter de la pertinence de la preuve apportée par le e-portfolio.

Les productions des étudiants dans le e-portfolio

Il s'agit ici d'expliciter la manière dont les étudiants ont construit la preuve du développement de la compétence visée en reprenant les trois dimensions étudiées : les résultats présentés (tant du point de vue des performances que des contre-performances), les justifications apportées sur les manières d'agir et les postures de régulation.

Mise en valeur des résultats

Tous les étudiants ont déposé des traces dans leur e-portfolio (cinq en moyenne). Ces traces sont diversifiées : modèle conceptuel de la base de données (MCD) ; maquette de pages du site web réalisé ; peuplement de la base de données ; captures d'écran d'une page produit ou autre ; code informatique (PHP) d'un élément du site ; lien vers le site web en ligne (URL). 65 % des étudiants font référence à leurs traces dans l'écrit réflexif. Toutefois, dans la majeure partie des cas, ils se contentent de les citer en référence dans l'écrit, ce n'est que plus rarement qu'ils les commentent de manière à justifier leur démarche (voir le tableau 1.6) :

TABLEAU 1.6 Exemples de justification de démarches apportées dans le portfolio.

Référence étudiant [Etu] (et type de trace)	Justification apportée en appui aux traces
Etu 26 (trace MCD)	« [...] Pour ce qui est de la génération de pages web ou de vues à partir de données structurées incluant des interactions simples, j'ai commencé par modéliser les données et les traitements à l'aide de diagrammes, tels que le modèle conceptuel de la base de données (MCD) créé avec Workbench. J'ai pris des décisions sur la structure de ma base de données, en identifiant les tables, les clés primaires et les clés étrangères nécessaires pour représenter les données de manière cohérente. [...] »
Etu 34 (Trace fournie Code page produit SQL [*Structured Language Query*] + Tableau)	« [...] Une autre préoccupation importante pour moi était de rendre mon code compréhensible et facile à maintenir, notamment pour X, au cas où elle devrait le retravailler ultérieurement. J'ai donc accordé une attention particulière à l'organisation et à la structure de mon code. J'ai utilisé des noms de classes significatifs, des commentaires clairs et des espaces entre les sections logiques. [...] »

Cinq étudiants osent évoquer des contre-performances et reviennent sur des erreurs, voire des échecs pour justifier du développement de leur compétence. Les travaux de Masson (2020) soulignent que le retour réflexif sur les erreurs commises est un véritable levier pour l'apprentissage et le développement de compétences. L'analyse que ces étudiants de première année en font est, à ce stade, souvent encore légère pour appuyer la preuve du développement de leur compétence. Prenons l'exemple de cet étudiant qui écrit : « L'aspect *back* du site est un domaine que je ne maîtrise pas encore, mais que je renforce grâce à mes collègues qui m'expliquent lorsque je leur demande un approfondissement du sujet. » Le commentaire ne permet pas de saisir si l'étudiant a dépassé cette lacune, mais il traduit déjà une prise de conscience particulièrement intéressante dans le processus d'apprentissage. Les enseignants, dans leurs annotations, identifient ces démarches comme des marques d'honnêteté : « reconnaît ses faiblesses en PHP », « fait preuve d'honnêteté, a essayé de faire des choses, les traces et le résultat le montrent, cela reste faible, on sent que le dev n'est pas son dada, mais il a eu une tentative d'y arriver ». Ils ne les associent toutefois pas, ici, au processus d'apprentissage. L'évocation d'échecs ou de résultats insatisfaisants dans le processus de développement des compétences reste une tendance marginale. On voit ici, à travers la démarche portfolio, qu'un changement de rapport vis-à-vis de l'erreur est amorcé, mais qu'il nécessite encore beaucoup d'accompagnement pour qu'il puisse faire ses preuves au sein du e-portfolio.

Justification des démarches

En APC, l'évaluation de la compétence est par principe individuelle. Dans notre étude, près de 85 % des étudiants utilisent le « je », parfois complémentairement au « nous », pour expliciter leurs manières d'agir. Cela peut être interprété comme le signe d'une certaine appropriation de la démarche réflexive. L'espace dédié à la réflexivité était structuré autour de plusieurs questions, l'une d'elles invitait explicitement les étudiants à justifier leurs démarches : « Quels ont été vos démarches, vos prises de décisions, votre degré d'implication et d'autonomie ? ». Il apparaît, dans l'analyse des réponses, que les étudiants s'attachent plus à justifier leur organisation dans le travail de groupe que leurs démarches vis-à-vis de la compétence. Les exemples suivants en témoignent : « En autonomie, nous devons nous séparer les tâches

à plusieurs tout en ayant tous une partie chacune pour chaque étape, dont par exemple le *backend.*»; « Nous nous sommes réparti les tâches au fur et à mesure de la SAE. On avait une idée de qui ferait quoi au début du projet, mais on a dû changer le planning à certains moments, vu l'avancée de certaines personnes du groupe ou des difficultés d'autres.» La perception de la compétence visée est alors totalement évacuée. Le tableau 1.7 suivant catégorise le niveau de justification apporté par les étudiants observés.

TABLEAU 1.7 Analyse des justifications apportées par les étudiants quant au développement de leur compétence.

Niveaux de justification	Exemples de verbatim associés	Présence dans les écrits
Pas de justification	« Je pense avoir pris beaucoup de décisions lors de cette SAE 2.03. J'ai essayé de faire face à toutes mes difficultés même si cela était parfois compliqué. »	8 (15 %)
Justification descriptive	« [...] nous avons réalisé une analyse des besoins du projet en identifiant les fonctionnalités clés, [...] J'ai choisi de créer cinq tables fonctionnant uniquement avec des clés primaires et étrangères. »	33 (63 %)
Justification explicative	« [...] *pour peupler la base de données avec les produits*, j'ai d'abord dû identifier les informations nécessaires pour chaque produit, telles que le nom, la description, les caractéristiques et le prix. Ensuite, j'ai collecté ces données et les ai organisées de manière cohérente dans un fichier .csv (produit.csv) *afin de pouvoir l'importer dans la base de données*. »	6 (11 %)
Justification explicative en lien avec la compétence	« Une autre préoccupation importante pour moi était de rendre mon code compréhensible et facile à maintenir, notamment pour Élodie, au cas où elle devrait le retravailler ultérieurement. *J'ai donc accordé une attention particulière à l'organisation et à la structure de mon code*. J'ai utilisé des noms de classes significatifs, des commentaires clairs et des espaces entre les sections logiques. »	5 (9 %)

Les justifications apportées par les étudiants restent très descriptives. Ces derniers décrivent ce qu'ils ont fait, souvent de manière chronologique. On a parfois l'impression qu'ils cherchent à prouver la réalité de leur travail dans la production de groupe. Ce n'est que plus rarement qu'ils explicitent pourquoi ils ont agi comme cela, pourquoi ils ont fait ce choix, etc. Très peu d'étudiants font le lien avec les

composantes essentielles identifiées pour la compétence. C'est peut-être ici une limite ergonomique de l'outil qui met plus en évidence les apprentissages essentiels que les composantes. Cette caractéristique sera à interroger dans nos futurs développements.

Régulation des actions

Il s'agit ici de considérer la posture critique des étudiants sur leurs manières d'agir, c'est-à-dire leur capacité à identifier des pistes d'amélioration dans leurs démarches. Le tableau 1.8 suivant présente une analyse des niveaux de régulation observés chez ces étudiants.

TABLEAU 1.8 Analyse des niveaux de régulation des étudiants quant à leurs démarches.

Niveaux de régulation	Exemples de verbatim associés	Présence dans les écrits
Pas de régulation		7 (13 %)
Régulation décontextualisée	« Si la SAE était à refaire, je m'assurerais d'avoir bien compris toute la théorie derrière avant le démarrage, et je m'organiserais de manière à en faire un petit peu tous les jours au lieu de gros blocs de travail de trois heures. »	13 (25 %)
Régulation sur la démarche	« Je suis conscient que la modélisation des données et des traitements est une compétence clé pour le développement d'une application web fonctionnelle et sécurisée. Je suis déterminé à acquérir une compréhension plus approfondie de cette compétence et à m'assurer que mes projets répondront aux attentes des clients. [...] En prenant en compte cette expérience, si j'avais l'opportunité de refaire une tâche similaire, j'accorderais une attention particulière à une compréhension précise de la demande du client. J'investirais davantage de temps dans l'analyse approfondie de ses attentes, en posant les bonnes questions et en établissant une communication claire pour éviter toute confusion. »	25 (48 %)
Régulation aboutie avec pistes d'amélioration en lien avec la compétence	« Si c'était à refaire, je respecterais donc davantage les normes d'accessibilité. Je coderais plus de JavaScript et de CSS pour que mon site soit plus agréable visuellement et à l'utilisation. J'organiserais mon code pour pouvoir le comprendre plus facilement. Et je prendrais le temps de l'améliorer pour qu'il soit le plus court et compréhensible possible. »	7 (13 %)

Les résultats sont ici encourageants, les étudiants ont développé de réelles postures critiques sur leurs manières de faire même si, pour l'instant, leur analyse est encore souvent décorrélée de la compétence visée et de ses attendus.

La démarche e-portfolio a apporté des changements intéressants dans les postures d'apprentissage des étudiants. Au-delà de l'affichage d'un résultat ou d'une performance, tous font la démarche de chercher à justifier leurs résultats, de décrire leurs manières de faire, de réguler leurs actions. Même si ces postures sont encore fragiles, certaines sont particulièrement intéressantes. La question de l'accompagnement des étudiants à la démarche portfolio semble ici bien centrale.

L'évaluation du niveau de compétence visé

Nous présenterons dans un premier temps les démarches d'autopositionnement des étudiants et l'évaluation portée par les deux évaluateurs, puis nous discuterons des écarts entre ces positionnements.

L'autopositionnement des étudiants

L'interface e-portfolio incitait les étudiants à se positionner vis-à-vis du niveau de développement de la compétence. Comme précisé plus haut, une échelle à quatre niveaux était proposée : non validé, validé, validé avec brio, félicitations du jury. Cette grille est formatée dans l'outil, il n'est pas possible d'en modifier les niveaux. Si les descripteurs de performance sont quant à eux personnalisables, les enseignants ont fait ici le choix de garder ceux proposés de manière générique. Environ deux tiers des étudiants se sont positionnés (34 étudiants sur 52) et un peu plus des deux tiers d'entre eux ont commenté leur positionnement. Dans 57 % des cas, le niveau estimé par les étudiants est en adéquation avec le niveau d'évaluation porté par les enseignants. Dans près de 30 % des cas, l'étudiant a surestimé son niveau. Quatre étudiants (11 %) se sont sous-estimés. La démarche d'autopositionnement des étudiants semble donc ici relativement sincère.

Il nous est toutefois difficile de déterminer comment les enseignants se sont saisis de l'autopositionnement de l'étudiant. Ils n'ont pas fait de commentaire à destination de l'étudiant, même lorsque celui-ci a expliqué son positionnement. De rares fois, ils ont échangé entre eux en privé, sur un chat proposé par l'interface numérique, sur le positionnement réalisé par l'étudiant.

L'évaluation portée par les enseignants

Le niveau de compétence de l'étudiant (noté « Etu » dans le tableau 1.9) était évalué par deux enseignants (noté « Ens » dans le tableau 1.9). Un espace dédié permettait aux enseignants d'indiquer la note qu'ils attribuaient et de la commenter. Leurs évaluations sont en général cohérentes (moyenne +/– à 13 et médiane également à 13). Pour cinq étudiants, il y a toutefois une divergence de plus de deux points d'écart, illustrée dans le tableau 1.9 suivant.

TABLEAU 1.9 Verbatim associés aux écarts de jugement des deux évaluateurs.

Référence étudiant	Justification apportée en appui aux traces
Etu 03	Ens 01 (14/20) souligne une analyse réflexive très correcte alors que Ens 02 (11/20) juge que le résultat est faible.
Etu 25	Ens 01 (18/20) juge comme exceptionnel le travail de l'étudiant (résultats, démarche et analyse réflexive). Ens 02, quant à lui, trouve le travail très positif (15/20), mais souligne un manque dans les résultats.
Etu 27	Ens 01 (6/20) et Ens 02 (10/20) sont d'accord sur la faiblesse des résultats, mais Ens 02 souligne que l'étudiant « a essayé de faire quelque chose » et semble le prendre en compte.
Etu 49	Ens 01 (16/20) pense que l'étudiant a « été moteur et qu'il a couvert ses camarades ». Pour Ens 02 (12/20), l'analyse est complète et le résultat fonctionnel, mais « les traces ne montrent pas grand-chose » et un « code PHP qui ne sert à rien ».
Etu 33	Ens 01 (16/20) souligne le bon travail, les apprentissages en programmation, les bonnes démarches et l'analyse réflexive, mais Ens 02 (13/20) s'attarde sur des manquements techniques concernant les résultats.

On sent, à travers ces commentaires, que l'attention des évaluateurs ne se porte pas toujours sur les mêmes indicateurs. Notre analyse a cherché à objectiver ce sentiment : le tableau 1.10 catégorise les commentaires adossés aux notes attribuées par les enseignants sur l'ensemble du groupe et en présente les pourcentages.

TABLEAU 1.10 Type de commentaires portés par les enseignants et pourcentages relevés.

Type de commentaire enseignant	Présence chez Ens 01	Présence chez Ens 02
Centré sur les résultats	90 %	92 %
Centré sur la démarche	38 %	23 %
Centré sur les apprentissages	13 %	6 %
Centré sur la démarche réflexive	52 %	6 %
Centré sur la compétence	0 %	2 %

On constate que les enseignants restent très focalisés sur les résultats. Leur attention se porte, dans une moindre mesure, sur les démarches des étudiants à travers l'analyse des justifications fournies sur leurs manières de faire. Quelques commentaires soulignent des apprentissages réalisés, ou non. Il n'est presque jamais question de la compétence visée ni des composantes essentielles identifiées. L'enseignant Ens 01 porte une attention particulière à la posture réflexive adoptée par les étudiants comme le montrent les commentaires suivants :

> Les traces ne sont pas explicitées dans l'analyse réflexive.
>
> [...] explique le fonctionnement du site au lieu de faire l'analyse, manque analyse critique, tous sur PHP. Analyse réflexive très acceptable, mais MCD trop complexe et souvent inadapté.
>
> L'analyse réflexive est aboutie, la participation au groupe est très correcte, mais les aptitudes à la programmation et à la conceptualisation restent trop fragiles par rapport aux objectifs.
>
> Analyse personnelle plutôt que sur la démarche suivie, pas de pistes d'amélioration pour le *backend*.
>
> Analyse réflexive : substantielle et de bonne qualité, honnête et utilisation à bon escient de ChatGPT et Copilot (pour passer du langage naturel au code), de l'IDE Visual Studio Code. Il manque néanmoins une analyse critique sur la démarche suivie et les pistes d'amélioration.
>
> Attention également à l'analyse réflexive qui est correcte, mais qui ne semble pas justifier tous les choix faits pour le projet.
>
> Analyse réflexive détaillée (liste des ressources externes), honnête, avec pistes d'amélioration.

> Analyse réflexive : explique qu'ils se passent les tâches, reconnaît ses faiblesses en PHP.

Il semble donc plus engagé dans une démarche d'évaluation des compétences que son collègue. Cette évaluation via le e-portfolio reste encore exploratoire et on voit ici que l'évolution des postures des enseignants vis-à-vis de l'évaluation est quelque chose qui se construit progressivement. Il est en effet souvent difficile pour les enseignants de se détacher de la qualité de la production et de l'atteinte du résultat attendu qui gardent souvent une incidence majeure dans l'appréciation portée sur les compétences de l'apprenant. Des dispositifs de formation et d'accompagnement à l'entrée dans ce processus existent et sont souvent portés par les services d'appui à la pédagogie au sein des établissements (on remarquera, par exemple, les diplômes universitaires [DU] de pédagogie universitaire).

La preuve du développement de la compétence

Cette dernière partie tend à croiser (1) l'analyse portée sur les productions des étudiants dans le e-portfolio, objectivée avec les trois indicateurs détaillés précédemment ; (2) l'évaluation portée par les enseignants ; et (3) celle portée par les étudiants. En d'autres termes, il s'agit de voir en quoi les notes attribuées correspondent à la perception donnée par le travail d'analyse systématique des portfolios des étudiants.

Il apparaît que, dans 64 % des cas, l'analyse de la preuve de la compétence à travers le e-portfolio est en adéquation avec l'appréciation de l'enseignant. Cette preuve de la compétence perceptible dans les e-portfolios correspond alors à la manière dont l'enseignant a évalué le travail. Comme cela peut être le cas dans d'autres types d'évaluation, il y a toutefois des écarts qui peuvent être importants. L'analyse du portfolio de Etu 40, par exemple, fait apparaître des niveaux de justification et de régulation peu avancés. La preuve de la compétence semble donc très incertaine. Pourtant, les enseignants attribuent la note de 18/20 et indiquent :

> [...] il s'agit d'un travail exceptionnel pour des étudiants de première année. [...] Analyse réflective, traces, MCD et gestion de projet tout à fait à la hauteur des attentes. Soutenance innovante et agréable consistant en une démonstration théâtralisée durant laquelle les étudiants

> jubilent, suivie d'un tour de piste technique durant lequel X reprend tout son sérieux et a réponse à tout.

Le commentaire laisse penser qu'il peut s'agir de la qualité de la prestation orale de l'étudiant, ou encore de celle de la production du travail de groupe. Invités à comparer, pour chaque étudiant, leurs appréciations avant la soutenance (donc uniquement sur la base du e-portfolio) et après la soutenance, les enseignants déclarent que dans 32 % des cas la soutenance a eu un impact sur l'appréciation qu'ils avaient eue à la lecture du portfolio. On sent ici l'importance, dans une démarche d'évaluation des compétences, de dépasser l'évaluation de performances isolées en situation pour donner à l'étudiant des occasions de s'exprimer sur son parcours d'apprentissage, sa progression, ses difficultés, ses perspectives, etc. Cela peut changer la perception qu'a un enseignant de la solidité des compétences de l'apprenant. Ces considérations marquent l'existence de différents niveaux de réflexivité (justification d'un résultat, preuve d'un niveau de compétence…) : il importera de réfléchir à la manière dont l'outil peut les matérialiser.

Mis en perspective avec les autopositionnements, les résultats font apparaître une légère tendance des étudiants à se surestimer par rapport à la qualité de la preuve apportée par leur e-portfolio (pour près d'un quart d'entre eux). C'est le cas de Etu 09 qui se positionne entre 17 et 20 avec le commentaire suivant : « Travail demandé respecté et rendu à temps. Code probablement améliorable, ce qui peut me faire perdre un peu de points tout de même. »

Toutefois, l'analyse de la preuve apportée dans le e-portfolio n'est pas si convaincante. Le propos est très descriptif et complètement dépersonnalisé :

> Nous avons décidé de tous toucher à un peu de tout, ainsi, nous avons chacun progressé dans plusieurs domaines. Après avoir réparti nos tâches, nous avons travaillé de notre côté en nous entraidant, puis nous avons regroupé notre travail et avons recommencé le même schéma. Nous avons fixé des *deadlines* et, si le temps en cours manquait, je continuais de mon côté. Ainsi, nous avons tous avancé correctement et nous avons produit un site fonctionnel.

On n'observe pas non plus de posture de régulation, c'est-à-dire de posture critique de l'apprenant sur ses démarches, ses manières de faire.

On voit ici, à travers ces différentes considérations, que l'évaluation reste un processus relativement subjectif (De Vecchi, 2021). On peut se demander si les écarts entre les étudiants et les enseignants ne dénoteraient pas une incompréhension ou un manque d'explications. On peut aussi s'interroger sur l'apport de *feedbacks* formatifs, complémentairement aux évaluations certificatives, pour aider l'étudiant à mesurer le chemin parcouru et celui restant à parcourir pour atteindre le niveau d'exigence attendu. L'usage d'une grille d'évaluation personnalisée aurait peut-être permis de réduire ces écarts.

Conclusion et conseils de mise en œuvre pratique

Dans ce cas d'usage, des étudiants de première année de cycle universitaire sont entrés dans une démarche d'évaluation de leur compétence en utilisant un e-portfolio. Le changement de posture des apprenants dans cette pratique d'évaluation est manifeste. On note ici qu'ils ne se contentent pas de déposer le projet fini, mais qu'ils s'engagent dans la démarche de « faire la preuve de leur compétence ». Pour cela, ils réalisent une introspection de leur participation dans le groupe de travail, de leurs démarches, de leurs difficultés. Même si la preuve apportée n'est pas toujours convaincante, le changement semble bien amorcé et on observe des exemples remarquables de démarches réflexives, d'analyse de contre-performance ou encore d'adoption de posture critique. Ces premiers résultats, pour un public de première année, sont encourageants quant à l'intérêt d'une démarche portfolio à ce niveau d'études. Les étudiants ont toutefois encore du mal à articuler leur posture réflexive avec la compétence qu'ils veulent démontrer, tant dans leur mode de justification des manières d'agir que dans les stratégies de régulation. Si cela est peut-être lié à des biais induits par l'ergonomie et les fonctionnalités de l'outil, un travail d'accompagnement méthodologique des étudiants semble nécessaire et complémentaire à la mise en place d'une démarche e-portfolio, notamment en premier cycle.

Du côté des enseignants, les constats vont dans le même sens. Un changement de posture vers l'évaluation des compétences semble bien amorcé, mais les évaluateurs restent encore très focalisés sur le résultat en occultant parfois presque complètement l'évaluation de la compétence visée au regard des critères définis. Là aussi, on peut pointer un besoin de formation pour accompagner ces changements de posture.

Ce cas d'usage illustre, de façon concrète et pragmatique, une utilisation du e-portfolio pour l'évaluation des compétences. Il permet à la fois d'en identifier de premiers atouts, mais aussi de souligner certains écueils dans la mise en œuvre de la démarche. On remarquera notamment, sans être exhaustifs :

- l'importance de communiquer, de manière lisible et compréhensible pour les étudiants, tant les compétences visées que les critères d'évaluation à prendre en compte, les contextes d'application attendus et les apprentissages associés. L'ergonomie d'un outil e-portfolio aura ici une incidence forte sur la mise en valeur de ces éléments et de leurs liens, et donc sur la compréhension des attendus par les étudiants ;
- l'apport des *feedbacks* formatifs (par l'enseignant, entre pairs ou encore par un professionnel externe à l'établissement), complémentairement aux évaluations certificatives, pour soutenir le positionnement de l'étudiant dans le développement de ses compétences et situer sa progression. Sur ce point, l'outil numérique peut faciliter la gestion et le suivi des interactions enseignants – étudiants ;
- la difficulté à accompagner la démarche réflexive des étudiants sans pour autant la standardiser et, de fait, la brider. L'interface e-portfolio permettait ici à l'enseignant de personnaliser des espaces de réponse à des questions explicitement posées. On peut s'interroger sur les effets des formulations utilisées sur les justifications apportées par les étudiants ;
- l'intégration, dans un souci de considération des inégalités, d'une réflexion sur le rapport à l'écrit et, dans le cadre d'une mise en œuvre appuyée sur un e-portfolio, d'une réflexion sur l'accès aux outils numériques par les étudiants ainsi que sur les questions d'accessibilité ;
- la nécessité d'accompagner les changements de postures des étudiants. Poumay et Georges (2022) proposent de réaliser des « séminaires de traces » accompagnant la construction de leur posture réflexive et intégrés dans le temps de formation ;
- la nécessité d'accompagner, complémentairement, les changements de postures attendus des enseignants. L'offre des services d'appui à la pédagogie est ici un levier intéressant. La création de communautés d'échanges de pratiques, tant sur le plan pédagogique que technopédagogique, en est une autre.

Cette étude ouvre, dans ce sens, des perspectives prometteuses quant à la pertinence d'une démarche e-portfolio pour l'évaluation des compétences, et ce dès le premier cycle universitaire, tout en dégageant des éléments de réflexion intéressants pour l'accompagnement des étudiants et des enseignants. L'outillage de la démarche apparaît comme un élément structurant du dispositif. Il s'agira, dans le cadre du programme Avenir(s)-volet ESR, de tirer parti de cette étude et de la confronter à d'autres usages analysés dans les terrains suivis par le consortium pour capitaliser les apports et les limites ergonomiques et fonctionnelles observés dans Karuta KAPC+ dans le développement d'un nouvel outil e-portfolio dans une version industrialisée et à très grande échelle dans l'enseignement supérieur : CoFolio.

Assia Belgheddouche

2 La vidéo comme outil d'évaluation de la production orale: cas d'un groupe d'étudiantes de première année en licence de français à l'ENSB

Résumé

Dans ce texte, nous présentons les résultats d'une expérience d'évaluation asynchrone par vidéo réalisée avec des étudiantes en langue française au département de français de l'ENSB (École normale supérieure de Bouzaréah, Alger). Nous avons tenté d'expliquer l'impact de l'utilisation de la vidéo comme outil dans une évaluation asynchrone de la production orale sur les performances de ces étudiantes. Nous nous sommes interrogée sur la fiabilité de ce type d'évaluation, sur le degré d'adhésion des étudiantes et sur la plus-value qu'apporte la vidéo dans cette évaluation. Les résultats de notre recherche mettent en relief l'efficacité de ce type d'évaluation dans les indicateurs liés à la prosodie et à la diction et soulignent ses limites pour les indicateurs liés à la compétence morphosyntaxique, lexicale et discursive. Nous avons également relevé son intérêt manifeste dans la diminution de l'anxiété d'évaluation et dans le développement de l'autonomie des apprenantes.

Introduction

De nombreux chercheurs (Garcia-Debanc et Delcambre, 2001; Laparra, 2008; Nonnon, 2016) ont souligné la difficulté d'enseigner et d'évaluer l'oral, mais les pistes pour son enseignement et son évaluation

demeurent insuffisantes. Le caractère éphémère de l'oral, sa quasi-absence des classes et les doutes qui ont plané sur la légitimité même de son enseignement l'ont mis à la marge des travaux de recherche, qui, même s'ils connaissent un bond sur le plan qualitatif et quantitatif ces dernières années, restent en deçà de ceux consacrés à l'écrit. Les nouvelles technologies commencent, néanmoins, à atténuer ce constat, d'une part en redorant le blason des discours oraux qui prennent plus de place dans la communication numérique, et d'autre part en offrant de nouveaux outils pour travailler cet oral insaisissable en classe.

L'apport de ces technologies s'est fait sentir durant le confinement imposé par la pandémie de Covid-19. Tous les établissements d'enseignement supérieur ont dû changer leurs pratiques et l'enseignement en ligne a été l'alternative qui a sauvé les étudiants d'une année blanche et démocratisé de nouvelles méthodes d'enseignement et d'évaluation qui n'avaient alors qu'un timide succès auprès des enseignants universitaires. L'évaluation en ligne – plus que l'enseignement – a été difficile à réaliser en raison de sa nouveauté certes, mais surtout à cause de tout ce qu'elle soulève comme questionnements sur sa légitimité, sa fiabilité et ses enjeux éthiques (Lafleur *et al.*, 2021). La difficulté peut se décupler quand l'objet évalué apporte son lot de défis, comme c'est le cas de l'oral.

C'est pourquoi nous tenterons, à travers ce travail, de montrer la plus-value de l'introduction des nouvelles technologies dans une situation d'évaluation de la production orale. Nous explorerons plus particulièrement l'impact d'une évaluation asynchrone à travers la vidéo sur les performances orales des étudiants. Pour arriver à nos fins, nous articulerons notre texte autour de quatre axes principaux : le contexte de notre recherche et notre problématique, notre cadre théorique et méthodologique, et enfin les résultats et leur discussion.

Le contexte et la problématique

Nous avons réalisé notre recherche au département de français de l'École normale supérieure de Bouzaréah (ENSB)[2] avec des étudiantes[3]

[2] L'École normale supérieure de Bouzaréah-Alger forme des enseignants des différents niveaux (primaire, collège et secondaire) et dans plusieurs disciplines des sciences humaines et sociales : langues (arabe, français, anglais, tamazight), histoire/géographie, philosophie.

[3] Notre groupe est constitué d'étudiants de sexe féminin, le corps enseignant au primaire s'étant considérablement féminisé (82 % de femmes en 2022 selon l'Office national des statistiques [https://www.ons.dz/spip.php?article2990, consulté le 14.01.2026]) et les étudiants de sexe masculin se font rares dans nos classes.

de première année du profil PEP (professeur d'enseignement primaire). Précisons qu'avant d'intégrer cette école, ces étudiantes ont suivi leur scolarité en langue arabe tout en étudiant le français comme langue étrangère[4], mais durant leur cursus à l'ENSB, tous les enseignements se font en langue française, qui devient de ce fait objet et outil d'apprentissage.

Après trois ans de formation à l'ENSB, elles vont enseigner cette langue à des élèves du primaire et elles doivent rejoindre leurs postes en ayant une bonne maîtrise du français. La formation initiale dans cette école vise donc à renforcer les compétences linguistiques de ces futurs enseignants de français, et à leur offrir les fondamentaux théoriques et pratiques nécessaires à leur réussite dans leur carrière d'enseignants. Elle cible le développement de leurs compétences dans la discipline à enseigner ainsi que leurs compétences didactiques et pédagogiques. Dans le département de français, les deux premières années du cursus sont essentiellement consacrées au renforcement des compétences communicatives et linguistiques des étudiants. En effet, les modules au programme développent avant tout leurs compétences en production/compréhension de l'oral et de l'écrit ainsi qu'en pratique systématique de la langue. La dernière année de leur cursus est principalement réservée à des modules à caractère professionnalisant.

Nous nous intéresserons à l'évaluation de la production orale (PO) dans le module « Pratiques et techniques de l'oral » (PTO). Notre choix s'est porté sur cette compétence en raison des difficultés rencontrées par les étudiants durant cette épreuve particulièrement anxiogène pour eux. Cette dernière peut, en effet, être un lieu de manifestation de l'anxiété langagière et de l'anxiété de performance puisque l'apprenant est confronté au stress lié à l'évaluation et à celui lié à la prise de parole. Nous convoquons ici les travaux de Horwitz *et al.* (1986), Young (1990) et Yale-Soulière *et al.* (2023) qui soulignent le caractère anxiogène de ces situations. Nous nous appuyons également sur notre expérience de plus de quinze ans dans l'enseignement du module de l'oral. Nous avons constaté que nos étudiants avaient du mal à prendre la parole en classe, et cette difficulté se décuplait durant les épreuves d'évaluation de la production orale. L'anxiété se traduit par des manifestations « physiologiques intenses avant et pendant la situation d'évaluation »

4 En Algérie, les élèves font leur scolarité en arabe et ils étudient le français comme langue étrangère de la troisième année primaire à la troisième année secondaire, c'est-à-dire durant dix ans de leur scolarité.

(Yale-Soulière *et al.*, 2023, p. 139) qui aboutissent à des performances insuffisantes, à des blocages dans l'expression, voire à un abandon de la tâche à réaliser.

Nous avons par conséquent testé avec nos étudiants une évaluation de PO en asynchrone à travers la vidéo dans le but d'explorer les apports de cette modalité d'évaluation et les défis qu'elle représente pour l'enseignant et pour l'étudiant. En tant qu'enseignante, nous avons proposé cette modalité d'évaluation que nous pensons moins anxiogène pour offrir des conditions optimales de passation à nos apprenantes. En tant que chercheuse, et dans une perspective de généralisation ou de régulation de la pratique expérimentée, nous souhaitons explorer l'intérêt et les limites de cette pratique évaluative. L'objectif étant de proposer *in fine* aux enseignants et aux étudiants des alternatives à l'évaluation en présentiel et d'intégrer l'évaluation à distance de l'oral comme forme complémentaire aux évaluations existantes.

Pour réaliser notre recherche, nous avons axé notre réflexion autour d'une question centrale : quel impact peut avoir une évaluation en asynchrone à travers la vidéo sur les performances orales de nos étudiants ? Cette interrogation nous a menée à nous poser les questions suivantes :

- est-ce que l'utilisation de la vidéo dans une évaluation asynchrone apporte une plus-value pour ces apprenantes ? Si oui, laquelle ?
- l'utilisation de la vidéo permet-elle aux étudiants de réaliser une autoévaluation plus ou moins fiable de leur performance orale ?
- quel impact a la présence de l'enseignant à distance sur les performances orales des apprenantes ?

Nous pensons que la vidéo est un outil qui pourrait renforcer le potentiel réflexif d'une autoévaluation et qu'elle pourrait permettre aux apprenantes d'avoir une perception plus ou moins objective de leurs erreurs, ce qui favoriserait l'autorégulation et le développement de leur autonomie. Nous pensons également que l'examen à distance en asynchrone peut réduire l'impact de la présence de l'enseignant, généralement anxiogène dans les examens en présentiel. Mais avant d'aborder le volet méthodologique qui nous permettra de vérifier ces hypothèses, nous allons, dans les lignes qui suivent, essayer de dessiner un cadre théorique pour notre recherche afin d'en faciliter la compréhension.

Le cadre théorique

L'enseignement et l'évaluation de l'oral

Les différentes méthodologies et approches d'enseignement des langues, y compris celles qui prônent une revalorisation de la place de l'oral (SGAV[5], communicatives ou actionnelles), ne sont pas parvenues à hisser l'enseignement de l'oral au même niveau que celui de l'écrit, qui, fort d'une longue tradition dans les classes de langue, occupe toujours une place centrale dans l'enseignement. La question de l'enseignabilité de l'oral s'est posée assez tardivement en didactique des langues, ce qui a donné une longueur d'avance à l'écrit. L'oral demeure en effet un « objet de recherche émergeant » (Dias-Chiaruttini et Dolz, 2023) qui connaît des « mises en veilleuse » dues au fait que ce ne soit pas « une pratique facile à mettre en œuvre dans les conditions ordinaires de l'enseignement » (Nonnon, 2011, 2016). Cette difficulté en a fait « le maillon faible de l'enseignement des langues » (Lafontaine et Messier, 2009).

Le décalage entre l'enseignement des deux compétences s'est répercuté sur l'évaluation qui, dans le cas des apprenantes, sujets de notre recherche, est exclusivement faite à l'écrit. En Algérie, les élèves des différents cycles scolaires ne sont évalués que sur leurs compétences en production et compréhension de l'écrit durant toute leur scolarité prébaccalauréat, soit durant les dix ans d'étude du FLE (français langue étrangère). Tous les examens officiels ou tous les examens qui conditionnent leur passage d'un niveau à un autre prennent en charge l'évaluation de l'écrit alors que l'oral est évalué uniquement à travers les *feedbacks* de l'enseignant dans les séances d'apprentissage, c'est-à-dire sans réel enjeu pour les apprenantes. L'évaluation de l'oral en tant qu'objet à part entière ne se fait qu'à l'université quand il est enseigné en tant que module dans les formations de langue. Force est de constater que les modalités d'évaluation dépendent largement des choix d'enseignement, ce qui nous amène à cette relation enseignement-évaluation.

Il est désormais connu que, de la question de l'enseignabilité de l'oral, se sont dégagées deux options : un enseignement intégré (Nonnon, 1999 ; Maurer, 2001 ; Halté et Rispail, 2005) et un enseignement autonome (Dolz et Schneuwly, 1998). L'établissement dans lequel nous avons

5 Méthodologie structuro-globale audiovisuelle.

réalisé notre recherche se positionne dans la deuxième optique puisque l'oral est considéré en tant qu'« objet d'enseignement spécifique [et] en tant qu'élément d'une discipline spécialisée dans le développement des capacités langagières » (Erard et Schneuwly, 2005). L'oral faisant partie des compétences à développer pour renforcer les compétences communicatives des étudiants de l'ENSB, un module lui est consacré durant les deux premières années du cursus universitaire. Il s'agit dans notre cas d'un objet d'enseignement autonome qui appelle une organisation des contenus d'enseignement en genres plus ou moins standardisés. Contrairement à l'évaluation de l'oral intégré qui ne peut être séparée de la discipline à laquelle il est lié, l'enseignement autonome appelle indéniablement une évaluation systématique et ciblée de l'oral pour vérifier l'atteinte des objectifs d'enseignement/apprentissage.

Que son enseignement soit intégré ou autonome, l'évaluation de l'oral n'est pas simple à mettre en place à cause des caractéristiques de cet objet qui n'est pas clairement défini. Les chercheurs (Garcia-Debanc, 1999 ; Garcia-Debanc et Delcambre, 2001 ; Dumais, 2010) qui s'y sont intéressés dessinent les contours d'un objet insaisissable, complexe, particulièrement difficile à observer, à analyser, à enseigner et donc à évaluer. Selon eux, plusieurs éléments peuvent s'ériger en obstacles devant l'évaluation de l'oral : son aspect éphémère, le flou qui entoure les indicateurs de maîtrise de l'oral, le caractère chronophage de l'activité d'évaluation elle-même, mais aussi les aspects psychologiques dus à la prise de parole devant les autres. Parmi les facteurs psychologiques les plus cités dans la littérature, il y a l'anxiété d'évaluation.

L'anxiété d'évaluation

L'oral, et plus particulièrement la production orale, « implique l'ensemble de la personne » (Garcia-Debanc, 1999 ; Garcia-Debanc et Delcambre, 2001) et met l'apprenant dans une situation où il doit se dévoiler et s'exposer au jugement de l'autre. Cette situation peut générer du stress, notamment s'il s'exprime dans une langue étrangère pas encore maîtrisée, comme c'est le cas de notre public pour qui la langue française est outil et objet d'apprentissage. L'évaluation des performances orales de ces étudiants primo-arrivants à l'université, en phase d'apprentissage de leur métier d'étudiant (Coulon, 1997), et non initiés aux épreuves orales, est une situation génératrice d'anxiété.

Wine (1980) avance que « l'évaluation est, parmi les éléments de l'environnement scolaire, celui qui est le plus susceptible de créer de l'anxiété chez les élèves » (cité par Viau, 1995, p. 382-383). Certains chercheurs parlent d'une anxiété de performance ou anxiété d'évaluation qu'ils définissent comme une « tendance à considérer les évaluations, telles que des examens, comme une menace » (Beidel, 1988 ; Mashayekh et Hashemi, 2011 ; Spielberger, 1972, cités par Yale-Soulière *et al.*, 2023, p. 138). C'est une anxiété d'état, donc non chronique, qui surgit dans des situations où un individu réalise une performance devant des personnes qui peuvent l'évaluer ou le juger.

Les évaluations notées et comptabilisées dans le passage aux niveaux supérieurs du cursus constituent un enjeu de taille pour tout apprenant et le simple fait de savoir qu'une activité est notée peut engendrer de l'anxiété chez lui (Prokofieva *et al.*, 2017) ; mais d'autres facteurs peuvent la provoquer comme « le fait de savoir que le temps est limité » (*ibid.*). Dans le cas des épreuves orales, à l'anxiété générée par l'activité d'évaluation s'ajoute celle engendrée par la prise de parole devant un évaluateur. Ainsi, « la pression perçue par des élèves venant des enseignants constitue une des causes les plus importantes du stress scolaire par rapport à d'autres types de pression venant des parents ou des pairs » (*ibid.*). Cette pression est d'autant plus importante lorsque la performance évaluée est réalisée sous les yeux de cet évaluateur qui observe et juge l'apprenant au moment même de la production.

Toutes ces difficultés liées à l'évaluation de l'oral sont désormais étudiées à l'aune d'un développement technologique sans précédent qui peut contribuer à fournir des solutions viables pour les différents acteurs de la situation d'évaluation (les évaluateurs et les évalués) : c'est ce que nous proposons d'explorer dans ce qui suit.

L'évaluation de l'oral à l'ère du numérique

Aujourd'hui, certaines des propriétés de l'oral, qui constituaient un obstacle à son enseignement et à son évaluation, peuvent être neutralisées grâce aux nouvelles technologies qui sont plus accessibles et plus présentes dans l'entourage des apprenantes et des enseignants. Il est en effet possible d'évaluer plus facilement l'écoute compte tenu de la pléthore de documents authentiques que l'on peut se procurer et que l'on peut didactiser grâce à des outils gratuits et pratiques (GarageBand, Audacity, CapCut, etc.). Il est par ailleurs plus aisé de garder une trace

de l'oral vu que les enregistrements ne sont plus « techniquement exigeants » comme le soutenaient Garcia-Debanc et Delcambre (2001, p. 7), ce qui simplifie le travail en production et en compréhension.

Pour les apprenantes, par exemple, l'enregistrement vidéo « permet l'observation et l'écoute asynchrone de ses propres performances orales et, par conséquent, une analyse plus fine de l'expression orale : ses aspects verbaux, non verbaux, paraverbaux, ainsi que leur rapport particulier à la communication » (Lavoie et Bouchard, 2017). Le même intérêt peut être cité du côté de l'enseignant qui dispose, grâce à la vidéo, d'un moyen efficace pour observer, analyser et évaluer les performances de ses apprenantes. Il peut également mesurer leur évolution en comparant leur performance à des moments différents de l'apprentissage.

Les nouvelles technologies interviennent également pour réduire l'anxiété de performance en diminuant le poids du regard de l'évaluateur durant les épreuves orales. À ce niveau, ces outils (la vidéo et les plateformes, par exemple), qui permettent aux apprenantes de réaliser leur évaluation en ligne, sont à même d'atténuer l'impact de la présence de l'enseignant en instaurant une certaine distance entre lui et l'apprenant. Cette distance, qui appelle une forme de présence particulière de l'enseignant, peut être salvatrice dans les situations d'évaluation de l'oral. Il convient néanmoins de préciser qu'une évaluation à distance ne signifie pas l'absence de l'enseignant, car, quand l'apprenant enregistre une activité, il est tout à fait conscient qu'elle sera visionnée et évaluée par l'enseignant.

Blandin (2004) cite quatre grandes catégories de présence : « le mode présentiel », la « présence virtuelle », la présence « à distance synchrone » et la présence « à distance asynchrone » (p. 365). Cette dernière pourrait être l'une des solutions à proposer pour réduire l'anxiété d'évaluation chez les étudiants qui passent des examens oraux, puisque toutes les autres formes de présence s'apparentent à un face-à-face, bien que ce soit réalisé par écrans interposés : le même poids psychologique peut survenir puisque le regard de l'enseignant évaluateur est omniprésent durant la réalisation de la performance. C'est le fait que la communication soit différée qui pourrait contribuer à soulager l'apprenant du regard de l'enseignant évaluateur ou surveillant.

Nous sommes par ailleurs consciente que l'évaluation de l'oral à travers la vidéo pose certainement problème techniquement en raison de « la variabilité des formats des fichiers sonores, de leur poids, des contraintes techniques liées à leur acheminement, de la nécessité

d'une parfaite qualité sonore, de la limitation drastique de la durée des fichiers afin qu'ils restent "maniables", des limites au stockage » (Abou Haidar, 2021), mais il nous semble que ces problèmes sont surmontables, car les apprenantes dont il est question dans nos universités ont une bonne maîtrise des outils numériques et elles sont à même de s'adapter assez rapidement aux contraintes techniques. De plus, chaque jour, nous voyons l'émergence de solutions techniques performantes (nouvelles fonctionnalités dans les plateformes, nouveaux logiciels, plus d'espace de stockage, etc.) qui rendent insignifiants certains obstacles qui étaient considérés comme insurmontables.

Ce qui pourrait constituer un noyau de résistance devant l'adoption des évaluations en asynchrone serait davantage axé sur la fiabilité et la légitimité de cette pratique pour les enseignants et son utilisation à bon escient par les étudiants. Nous ne pouvons, par exemple, écarter d'un revers de la main le problème éthique que soulève l'évaluation asynchrone, mais ce n'est pas une problématique récente puisque, depuis que l'évaluation existe, il a toujours été question de réaliser la performance à évaluer sous les yeux d'un enseignant surveillant qui veille à ce que l'apprenant accomplisse son travail seul dans une durée limitée. À moins de se situer d'entrée de jeu dans une optique d'évaluation formative sans notation, il serait naïf de croire que les apprenantes peuvent effectuer les activités qui conditionnent leur réussite seules et sans recours à des aides extérieures, d'où la nécessité d'utiliser ces outils avec beaucoup de précautions.

Le dispositif et le scénario pédagogique

Notre intervention a été menée auprès d'un groupe d'étudiantes inscrites en première année de langue française à l'ENSB. L'enseignement dans cette école est organisé sous forme de modules annuels dont l'évaluation est programmée en fin de chaque semestre. Pour le module de PTO (« Pratiques et techniques de l'oral »), l'évaluation se fait en deux temps : un temps réservé à la compréhension de l'oral et un autre à la production orale, c'est-à-dire que les apprenantes passent deux épreuves en production orale (PO) et deux épreuves en compréhension de l'oral (CO) pour avoir la moyenne annuelle du module.

L'évaluation de la CO se fait en classe à travers des activités qui mobilisent les stratégies d'écoute avec comme supports des documents authentiques généralement en provenance d'Internet. Le deuxième

moment d'évaluation est celui dédié à la PO et il consiste en une passation individuelle en présentiel avec l'enseignant.

Comme cette épreuve de PO est généralement mal vécue par ces étudiantes peu habituées à prendre la parole dans une situation d'évaluation, nous leur avons proposé de réaliser une évaluation en asynchrone dont les résultats ont été comptabilisés avec la note du deuxième semestre. Cette activité était facultative, mais une majorité d'étudiantes (30 sur 35) ont fait le choix d'y participer pour améliorer leurs résultats dans le module. Pour les cinq étudiantes qui n'ont pas souhaité faire l'évaluation à distance, et pour ne pas les pénaliser, nous n'avons pris en compte que les notes des épreuves faites en présentiel dans le calcul de leur moyenne.

L'activité de production orale

Nous avons demandé aux étudiantes d'enregistrer une vidéo dans laquelle elles devaient raconter un conte en respectant les différentes étapes du récit vues en classe durant le deuxième semestre (en compréhension, plusieurs activités de compréhension sur des contes écoutés et visionnés ; et en production, des activités de production spontanée/préparée de contes). Nous avons précisé dans la consigne qu'elles devaient :

- raconter un conte avec les différentes phases du récit (situation initiale/déroulement des événements/situation finale) dans une vidéo de trois à cinq minutes ;
- respecter les temps de la narration ;
- raconter et non pas lire ou réciter le conte ;
- respecter le délai d'envoi de la vidéo.

Notre choix pour le conte comme genre oral est dicté par sa présence dans le programme de la première année et par le fait que ce soit un genre universel dont la structure discursive est plus ou moins maîtrisée par les étudiants puisqu'elle est étudiée sous sa forme écrite durant leur scolarité, en arabe et en français. De plus, c'est un genre qui peut laisser une certaine liberté d'expression aux étudiants et qui ne les pénalise pas en cas de manque d'informations sur des thèmes peu maîtrisés.

La méthodologie

Pour mesurer l'impact de l'utilisation de la vidéo dans l'évaluation sur les performances orales de nos apprenantes, nous leur avons fait

passer un questionnaire et une grille d'autoévaluation qu'elles ont renseignés après avoir posté l'enregistrement sur Classroom. Au début de notre recherche, nous avions prévu un questionnaire avec des items qui visaient l'évaluation de la production par les étudiantes, mais nous nous sommes rendu compte que cela ne suffisait pas pour donner une appréciation précise de leur travail. Les items du questionnaire qui orientaient le regard des apprenantes vers leur performance étaient noyés dans les autres items et restaient trop généraux pour répondre à nos objectifs de recherche. C'est pourquoi nous leur avons présenté une grille d'autoévaluation qui permettait non seulement de mieux cibler les points à évaluer dans la PO, mais aussi de comparer les résultats de l'autoévaluation avec ceux obtenus par notre propre évaluation.

À travers le questionnaire, nous aspirons à découvrir comment les étudiantes s'y sont prises pour réaliser l'activité (choix du conte, préparation de la PO, conditions matérielles de l'enregistrement, nombre de tentatives avant la version définitive, le contexte[6]). Nous voulons également savoir si nos témoins ont eu des moments d'analyse et de réflexion sur leur production et si ces moments ont donné lieu à des modifications dans leur production. Trois items ont été consacrés à leur appréciation sur leur production puisque nous leur avions demandé de donner leur avis sur leur performance, de citer des exemples de leurs problèmes à différents niveaux (prosodie, diction, syntaxe, lexique), de s'attribuer une note et de la justifier. Enfin, nous avons interrogé les opinions des témoins sur cette modalité d'évaluation comparée à l'évaluation en présentiel pour mesurer leur degré d'adhésion à cette nouvelle pratique et leur avis sur une possibilité de prévoir d'autres examens selon cette modalité.

Comme précisé plus haut, la grille d'autoévaluation vient compléter les items consacrés à l'appréciation de la performance des apprenantes. Notre grille est inspirée de celle de Dumais *et al.* (2013) et elle a été adaptée au genre du conte. Elle comporte quatre axes qui permettent d'évaluer la production sur le plan de la diction (articulation/portée de la voix/prononciation), de la prosodie (accentuation/rythme/intonation), de la langue (morphosyntaxe/lexique) et du discours (organisation/pertinence). Ainsi, les étudiantes devaient s'attribuer une note pour chaque indicateur entre 1 et 5, comme suit : 1 point pour insuffisant ;

6 Pour avoir ces informations, nous avons proposé aux étudiantes des questions à choix multiples, tout en leur laissant un espace pour saisir d'autres réponses qui représenteraient des actions que nous n'aurions pas prévues.

2 points pour moyen ; 3 points pour assez bien ; 4 points pour bien ; 5 points pour très bien. En ce qui concerne les indicateurs en lien avec le discours, nous avons préféré opter pour une échelle dichotomique qui permet aux apprenantes d'attester la présence ou l'absence des différentes phases.

Sur les 30 étudiantes qui ont participé à l'activité de production orale, la totalité a répondu au questionnaire, mais seulement 14 étudiantes[7] ont complété la grille d'autoévaluation. Nous allons dans ce qui suit présenter les résultats obtenus grâce aux 30 questionnaires renseignés et essayer de faire ressortir les modalités de réalisation de l'évaluation en asynchrone, les appréciations des apprenantes sur leur prestation et leur avis sur cette modalité d'évaluation. Les réponses des étudiantes au questionnaire, la comparaison des résultats obtenus dans les deux évaluations (la nôtre et celle des 14 étudiantes) ainsi qu'une analyse des vidéos nous permettront de vérifier à la fois l'impact de la vidéo et du travail en différé sur la performance des apprenantes et sur la pertinence de leur évaluation.

Les résultats et leur analyse

Ce que le questionnaire a révélé

Nous avons analysé les 30[8] questionnaires renseignés en les regroupant selon les principaux axes des items, à savoir : les modalités de réalisation de l'activité, les retours des apprenantes sur leur performance, les actions de révision et les modifications, les opinions des témoins sur l'évaluation en ligne en asynchrone.

Les modalités de réalisation de l'activité de production orale

En ce qui concerne les modalités de réalisation de la production orale, nous avons recensé plusieurs modes opératoires pour préparer la production et pour réaliser l'enregistrement. Sur les 30 étudiantes interrogées, 13 déclarent avoir choisi un conte qui leur a été raconté par un membre de leur famille (grand-mère/mère/tante/grand-père), une

7 Les étudiantes se sont plus ou moins désengagées dès qu'elles ont terminé l'année universitaire : 16 étudiantes qui ont participé à l'évaluation n'ont pas réalisé l'activité d'autoévaluation.

8 Nous avons jugé intéressant d'exploiter les réponses de toutes les étudiantes, y compris celles qui n'ont pas participé à l'activité d'autoévaluation.

partie non négligeable (10 étudiantes) a cherché un conte sur Internet et les autres étudiantes n'en ont pas précisé la provenance.

Une grande majorité (86,7 %) dit s'être raconté l'histoire avant de l'enregistrer, un peu moins de la moitié (43,3 %) déclare avoir raconté l'histoire à une autre personne, 43,3 % ont écrit le conte sur une feuille et enfin 10 % ont lu silencieusement le conte plusieurs fois. Quant à l'enregistrement, il est généralement réalisé par l'étudiante elle-même (73,3 %) sans aucune présence à proximité (56,7 %). Une partie des étudiantes (26,7 %) a montré l'enregistrement à une tierce personne.

Les appréciations des apprenantes sur la qualité de leur production

Les appréciations globales faites par les étudiantes sont équitablement réparties entre appréciations positives et négatives. Nous avons obtenu cependant des retours négatifs plus précis que les retours positifs qui restent d'ordre général, comme le montrent les exemples suivants :

> Je ne suis pas satisfaite de mon oral, car je bégaie trop et je travaille sur ça.
>
> Je n'étais pas vraiment satisfaite de ma prononciation.
>
> Ma langue française ne me satisfaisait pas. J'ai encore besoin de plusieurs façons de faire « sonner » ma langue.
>
> J'ai aimé, mais j'espère que le professeur a aimé.
>
> J'ai aimé.
>
> C'est pas mal.
>
> Je pense que c'est bon et acceptable.

Mais quand il leur est demandé de s'attribuer une note sur 10, les étudiantes s'octroient presque toutes (29 étudiantes) une note supérieure à la moyenne, c'est-à-dire entre 5 et 10.

En ce qui concerne leur capacité à identifier leurs lacunes, la majorité des étudiantes arrivent plus ou moins à citer deux exemples de leurs problèmes de prononciation, de syntaxe, de vocabulaire et certaines (entre 7 et 10 pour chaque type de difficultés) disent ne pas en avoir rencontré. Il est néanmoins possible de noter quelques erreurs de catégorisation des fautes commises.

Les actions de révision

Tous nos témoins ont déclaré avoir fait l'enregistrement plus d'une fois (entre deux et 25 fois), mais la moyenne est de trois à cinq enregistrements. Une majorité (76,7 %) dit avoir modifié les enregistrements et les principales modifications citées par ces étudiantes concernent le son, l'image, la prononciation. Une seule étudiante a évoqué un changement de la structure de son récit après l'avoir visionné, c'est-à-dire qu'il n'y a, selon elle, pratiquement pas de modifications au niveau de la morphosyntaxe, du lexique ou au niveau discursif.

Les avis des étudiantes sur cette modalité de passation de l'examen

Cette modalité d'évaluation semble avoir obtenu l'adhésion de la majorité de nos témoins (80 %) qui seraient d'accord pour que l'évaluation à distance soit adoptée pour ce module. Et les arguments les plus évoqués par ces étudiantes sont, dans l'ordre, l'absence ou la diminution du stress, le temps de préparation de l'activité plus important que celui dont elles bénéficient durant l'examen en présentiel, et la possibilité de refaire l'activité plusieurs fois avant de poster le produit final.

Les 20 % qui s'opposent à l'idée de passer l'examen de PO à distance en asynchrone avancent les arguments suivants : la possibilité de tricher à distance, l'efficacité d'un examen en face-à-face pour vaincre sa timidité et s'habituer à parler, les problèmes techniques (enregistrement et envoi du travail). Notons que les étudiantes n'ont pas soulevé de difficultés particulières pour l'enregistrement de la vidéo, mais plusieurs nous ont contactée pour des difficultés à poster leur vidéo sur Classroom (problèmes liés au type de fichier et au volume de la vidéo). Cela a pu être réglé grâce à un travail collaboratif avec leurs camarades.

Ce que les évaluations ont révélé

Les scores globaux obtenus dans l'autoévaluation des productions orales corroborent les réponses des étudiantes qui se sont toutes attribué des notes supérieures à la moyenne dans les items du questionnaire. Les résultats de la grille d'autoévaluation confirment leur appréciation générale sur leur production et se rapprochent des résultats obtenus par notre évaluation, comme c'est indiqué dans la figure 2.1[9].

[9] 70 points est la note maximale qu'il est possible d'attribuer à l'étudiant : elle correspond à la somme des points attribués aux 14 indicateurs (5 points pour chaque indicateur).

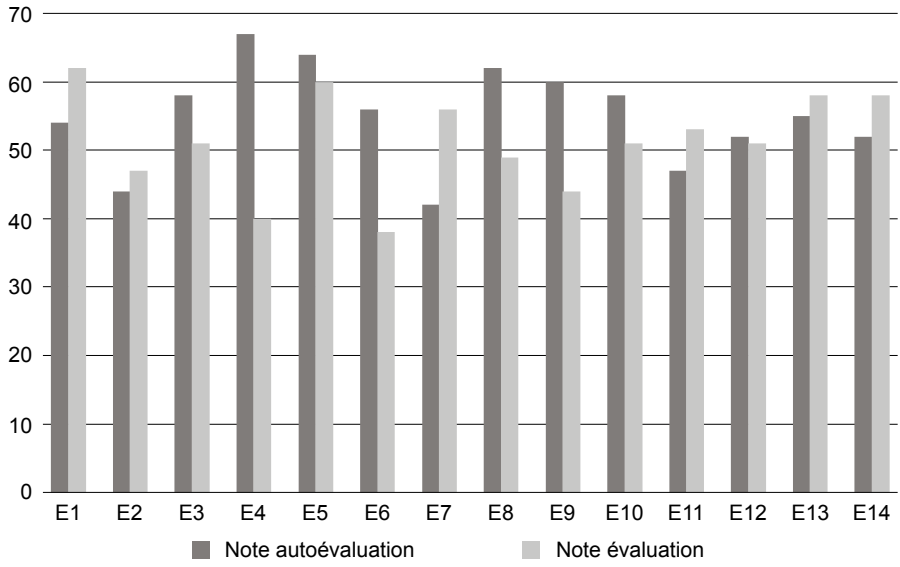


FIGURE 2.1 Score global de l'autoévaluation et de l'évaluation des productions orales.

Nous remarquons qu'il n'y a qu'une légère différence entre les résultats de notre évaluation et celle des étudiantes[10]. L'écart demeure très peu prononcé dans le score global, car les résultats des évaluations sur certains critères peuvent en compenser d'autres. À ce stade, nous pouvons voir quelques différences dans le sens de la surestimation et dans celui de la sous-estimation. Les écarts nous ont semblé plus palpables quand nous avons abordé les résultats par critères : la diction, la prosodie, la langue (morphosyntaxe/lexique) et le discours.

Commençons par ce dernier point qui semble faire l'unanimité chez les étudiantes qui estiment toutes avoir réussi à présenter un récit bien structuré, cohérent et contenant les différentes étapes d'un récit. Ce résultat est conforme à notre évaluation puisque nous avons constaté l'excellence du travail présenté au niveau discursif.

Les résultats obtenus pour les indicateurs liés à la diction et à la prosodie[11] sont représentés dans la figure 2.2 et ils mettent en relief des différences entre l'évaluation des étudiantes et la nôtre :

[10] « E1 » désigne l'étudiante 1.

[11] 30 points est la note maximale dédiée à la diction et 16 points est la note maximale dédiée à la prosodie.

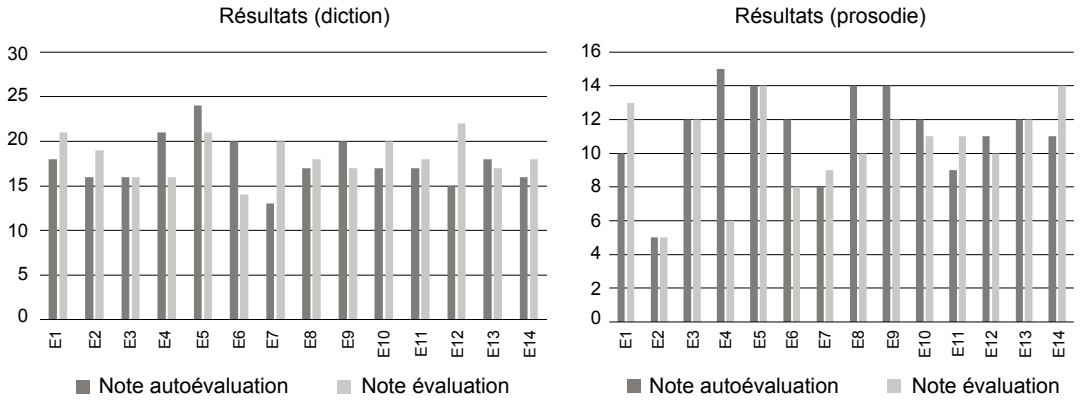


FIGURE 2.2 Les résultats des évaluations en prosodie et en diction.

En ce qui concerne la diction, il est possible de distinguer une surestimation de la part de quatre de nos témoins, une sous-estimation chez six étudiantes et une appréciation plus ou moins similaire à la nôtre pour quatre d'entre elles. Sur les faits prosodiques, nous constatons également un décalage entre les deux évaluations puisque nos apprenantes ont tendance à surestimer (cinq étudiantes) ou à sous-estimer (quatre étudiantes) leurs compétences pour ce critère. Ce qui signifie que l'écart se creuse entre l'autoévaluation et notre évaluation des productions à ce niveau.

Les résultats qui concernent les critères liés à la langue (morphosyntaxe et lexique), présentés dans la figure 2.3, montrent des décalages dans les appréciations, mais cette fois-ci avec une légère supériorité des surestimations pour sept étudiantes, contre trois qui sous-estiment leur performance et quatre qui présentent des notes identiques aux nôtres, ou qui s'en rapprochent.

Les étudiantes s'attribuent de bonnes notes qui sont en partie conformes à la prestation réalisée, mais celles qui s'octroient des notes supérieures à leur performance n'arrivent pas à détecter certaines erreurs dans la construction de leurs phrases. Ces erreurs ont été décelées chez les étudiantes qui n'ont pas mémorisé le conte et qui ont tenté de le raconter avec leurs propres mots, mais aussi, à une moindre échelle, chez les étudiantes qui ont mémorisé le conte et qui, en perdant le fil de la narration, font des fautes de langue qui ne sont pas corrigées malgré les nombreux enregistrements réalisés. Celles qui se

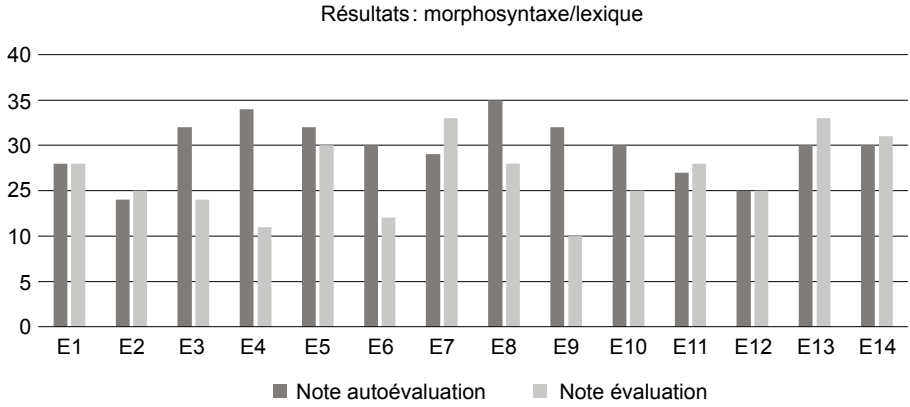


FIGURE 2.3 Résultats des évaluations en langue.

sous-estiment dans ce critère semblent s'interdire de s'octroyer une bonne note même dans les contes mémorisés, très corrects sur le plan morphosyntaxique et lexical.

L'évaluation des travaux des étudiantes a mis en relief certaines difficultés à noter leurs performances, car, même si le score global montre des résultats proches dans les deux évaluations, les écarts se creusent considérablement quand nous analysons les résultats par critères. Nous sommes souvent confrontée à des sous-estimations ou à des surestimations des performances dans l'autoévaluation.

Ce que les vidéos ont révélé

Le visionnement des vidéos a fait ressortir un certain nombre d'éléments qui ne sont pas pris en compte par la grille d'évaluation ou le questionnaire et qui pourront nous servir à compléter nos résultats. Nous avons en premier lieu constaté une réelle implication des étudiantes qui ont respecté la consigne puisque toutes ont présenté des contes en essayant de jouer le jeu de la narration et certaines ont même ajouté des effets supplémentaires à leur vidéo (par exemple, une musique de fond qui accompagne la narration). La majorité d'entre elles ont fait de réels efforts sur l'aspect non verbal dans leur présentation, mais nous n'avons malheureusement pas pu analyser ce côté de la communication, car ce n'est pas l'objectif de la présente recherche même si la vidéo peut avoir un réel intérêt dans ce cas. Il faudra cependant préciser que les contes étaient essentiellement racontés de mémoire : il

est très facile de voir que la narration s'appuie sur une mémorisation des récits avant de les raconter (contes parfaitement structurés, langue correcte, fluidité). Il n'y a que trois étudiantes qui ont présenté un travail personnel non mémorisé.

Ce que nous avons également constaté, mais sans avoir un autre point de comparaison que notre expérience d'évaluation du premier semestre et les activités faites en classe sur le même genre oral, c'est une prononciation plus soignée et un débit mieux maîtrisé. Les enregistrements étaient tous d'une durée supérieure à trois minutes (de 3 minutes à 5 min 48 s) : un temps que les étudiantes ne parvenaient jamais à atteindre durant leur production en présentiel, que ce soit durant l'examen ou les séances d'apprentissage.

Le visionnement des vidéos a enfin mis en relief la présence d'énoncés qui interpellent directement l'enseignant (« Bonjour, madame, j'espère que vous allez bien »/« merci » – « merci d'avoir écouté »/« j'espère que vous avez aimé le conte », etc.), ce qui dénote un sentiment de présence du destinataire qui est l'enseignant évaluateur, malgré la distance et le caractère asynchrone de l'activité.

Discussion des résultats

Le questionnaire, le visionnement des vidéos et l'évaluation ont permis de faire ressortir un premier constat : à savoir, le respect de la consigne qui a donné lieu à des récits généralement corrects sur le plan morphosyntaxique et lexical et sur le plan de la diction et de la prosodie. La réalisation de l'évaluation à travers un enregistrement vidéo a permis d'obtenir des productions complètes de toutes les étudiantes qui ont participé à l'activité alors qu'en présentiel, une grande majorité d'entre elles ne terminent pas leurs présentations ou font des productions très courtes et peu développées même si elles disposent d'un temps de préparation.

Les items du questionnaire qui interrogent les modalités de réalisation de la vidéo dévoilent une réalisation « en solo » de l'enregistrement, ce qui pourrait nous renvoyer vers la gêne qu'éprouvent les étudiantes à présenter leur production orale devant d'autres personnes, mais aussi vers ce souci d'avoir une meilleure qualité de son : élément qui a été souvent cité dans les modifications opérées sur les enregistrements. La préparation de la PO se fonde sur une narration faite pour soi ou pour un autre, sur une lecture ou une écriture du conte avant de le raconter, ce qui prouve l'omniprésence d'un entraînement fait en amont de

l'enregistrement. Ce sont autant d'activités réalisées en autonomie, qui contribuent au développement d'une compétence orale. Les étudiantes ont géré leurs façons de préparer l'activité et de l'enregistrer ; elles ont décidé de refaire ou non les enregistrements, effectué des choix sur les éléments à modifier ou à maintenir, et pris le parti de partager ou non leurs productions avec un autre destinataire. Ce dernier point a attiré notre attention sur le fait qu'une partie des étudiantes n'a pas été gênée de s'essayer à l'exercice de la narration devant une personne qui n'a pas le pouvoir de les noter alors qu'elles évoquaient presque toutes l'anxiété vécue lors d'un examen passé devant l'enseignant.

La vidéo nous a permis de détecter un travail de mémorisation qui n'a pas été cité par les étudiantes, peut-être de peur d'être sanctionnées parce qu'elles ne présentaient pas une production personnelle. Cette mémorisation a empêché l'évaluation des productions sur le plan discursif, morphosyntaxique et lexical. Cela nous amène à la nature même du conte qui autorise cette oralisation par mémorisation, mais qui présente la limite de ne travailler qu'une compétence d'oralisation de textes écrits comme le précise Laparra (2008, p. 124) dans le passage suivant : « Lire une histoire doit être différencié de raconter une histoire. Le "contage" des histoires en réception ou en production est trop souvent un entre-deux instable entre l'oralisation d'un texte et la reformulation de ce même texte. »

Nos résultats montrent les limites d'une évaluation asynchrone sur les indicateurs liés à la morphosyntaxe et au lexique puisque les apprenantes ne sont pas évaluées sur la production d'un discours oral, mais sur son oralisation. Il n'est en effet possible d'évaluer les compétences morphosyntaxiques et lexicales des apprenantes que pour les étudiantes qui ont fait l'effort de le raconter avec leurs propres mots. Un travail est peut-être à faire pour sensibiliser les apprenantes à la nécessité de présenter des productions personnelles et, dans ce cas, il serait plus raisonnable de miser sur une évaluation formative où l'enjeu de la notation ne viendrait pas compromettre l'investissement des apprenantes dans une vraie production orale.

Nos hypothèses n'ont pas été toutes confirmées. Certes, et sans surprise, la vidéo permet de fixer la production orale et un/des retour(s) des étudiantes sur leur production, mais ces retours ne sont pas toujours pertinents vu les positionnements inadéquats notés chez nos étudiantes, qui surestimaient ou sous-estimaient leur performance. Les écarts entre notre évaluation et celle des étudiantes peuvent révéler une

certaine difficulté à se positionner par rapport à leurs performances, mais ces écarts peuvent également être dus aux biais de l'autoévaluation socialisée soulignés dans les travaux de Saussez et Allal (2007). Selon ces auteurs, « il y aurait un pattern de stratégies servant d'analyseur d'une composante des représentations de l'autoévaluation associée à la sauvegarde de l'image positive du soi professionnel » (*ibid.*, p. 116). Ces stratégies permettraient aux apprenantes de préserver leur image de soi, mais aussi leur image devant l'évaluateur qui décidera de leur réussite. Ainsi, nos étudiantes peuvent avoir proposé une appréciation positive ou négative de leur prestation en fonction de leurs perceptions de l'usage qu'il sera fait de leur évaluation et en fonction du rôle qu'elles attribuent à cette évaluation dans leur réussite.

Cela dit, l'existence de la production à portée de main a du moins eu le mérite de proposer à l'apprenante une possibilité de regard sur son action, car le fait de réaliser l'enregistrement à plusieurs reprises montre qu'il y a un moment de réflexion sur une pratique langagière et, quand elle donne lieu à une modification, aussi minime soit-elle, elle témoigne d'une prise de conscience qui est l'une des composantes de la métacognition (Lafortune et Saint-Pierre, 1998). Nous pouvons de ce fait avancer que la vidéo n'est pas un simple support : il s'agit d'un outil qui contribue à la réalisation de la production en autorisant et en encourageant la régulation. Cet outil prend place dans le processus de production puisqu'il permet de concrétiser une de ses étapes les plus importantes, à savoir la révision et le contrôle.

Ce qu'il y a lieu de souligner en effet sur cette modalité d'évaluation, c'est la préparation qu'elle nécessite et l'entraînement qu'elle déclenche chez les apprenantes qui réalisent plusieurs fois leur activité, ce qui constitue un bon exercice, notamment pour l'amélioration de la diction et de la prosodie. Cette préparation n'a cependant pas autant d'intérêt pour les compétences morphosyntaxiques et lexicales puisque l'apprenante mémorise le discours et le récite au lieu de le produire dans la majorité des cas.

Nous pensons que l'évaluation, quel que soit son type et quelle que soit sa modalité de passation, a, certes à des degrés différents, une valeur formatrice pour les apprenantes. Cette valeur est néanmoins plus importante en situation d'autoévaluation à distance comme les apprenantes peuvent voir et revoir leur prestation, refaire l'activité et réguler leur action si besoin, ce qui laisse supposer l'existence d'une réflexion sur le processus de production orale et sur son produit.

L'avantage qui nous a semblé le plus prégnant dans cette expérience est la diminution significative de l'anxiété d'évaluation, car les productions longues et assez développées ne montrent pas ou peu de signes d'anxiété, et nous pensons que la présence de l'enseignant qui n'intervient qu'en amont, c'est-à-dire après la réalisation de l'activité, est la raison principale de cette différence. Notons aussi que le temps de préparation, fort apprécié par les étudiantes – qui avaient soulevé cela comme l'un des points forts de l'évaluation en ligne –, a permis de réaliser l'activité avec plus de sérénité et sans anxiété.

Nous pensons que cette expérience a été plus ou moins concluante pour ces étudiantes à plusieurs niveaux. Selon leurs témoignages, elle leur a permis un travail en autonomie, une réflexion sur l'action, une régulation de l'action et un travail moins stressant. Elle a toutefois donné l'occasion de souligner certaines limites qui touchent non seulement à l'évaluation des critères liés à la langue et au discours, mais aussi aux enjeux éthiques soulevés par une évaluation asynchrone. À ce stade de notre recherche, nous n'avons pas les réponses à ces problématiques, mais nous présenterons quelques conseils de mise en œuvre qui pourraient atténuer ces contraintes.

Conseils de mise en œuvre pratique

Ce que nous pouvons donner comme conseil de mise en œuvre pratique, ce serait de compléter le dispositif par une coévaluation ou une comparaison des résultats de l'évaluation de l'enseignant et celle des étudiants suivie ou accompagnée d'une négociation. La vidéo peut avoir plus d'intérêt s'il y a explicitation et négociation après complétion de la grille ou s'il y a une possibilité de coévaluation (évaluation par les pairs).

En ce qui concerne les difficultés techniques rencontrées durant l'activité, nous proposons de revoir les programmes du module d'informatique, qui, pour notre public, est essentiellement centré sur les outils de traitement de texte, et de formuler le besoin de les former à l'utilisation des nouvelles technologies dans leur apprentissage (compression et conversion de fichiers, montage, utilisation des plateformes d'apprentissage, etc.).

Nous pensons que la vidéo peut avoir plus d'impact dans une évaluation asynchrone si cette dernière est exclusivement formative, sans enjeux qui pousseraient les apprenants à travailler l'oralisation plus que la production orale. Il serait envisageable de penser la distance comme

un choix pédagogique volontaire et non pas comme une contrainte avec laquelle nous devons composer. Il s'agira de présenter une évaluation à distance pour les éléments qui peuvent pousser les apprenants à s'améliorer (un entraînement intensif sur la diction, la prosodie et le non-verbal) et de penser à miser sur une évaluation formative pour les éléments difficiles à contrôler à distance (l'aspect morphosyntaxique, lexical et discursif).

Proposer un terrain neutre sans danger pour les apprenants peut également réduire les biais de l'autoévaluation socialisée et donner lieu à des appréciations plus « sincères » sur leurs performances, et ainsi provoquer des régulations plus précises. Ce terrain rassurant peut en outre inciter les apprenants à soumettre des travaux personnels sans crainte de conséquences sur leur réussite.

Aurélie Dupré
& Anne-Céline Grolleau

3 Pratiques évaluatives d'enseignants du supérieur

Résumé

Ce chapitre analyse les pratiques d'évaluation avec le numérique d'enseignants du supérieur inscrits dans le diplôme universitaire (DU) en pédagogie universitaire de l'Université de Lille en 2022-2023. À partir des portfolios réflexifs de 13 enseignants sur leur développement professionnel pédagogique, il explore leurs représentations du numérique pour l'évaluation et les dispositifs qu'ils mettent en œuvre. L'analyse du corpus montre une évolution vers des pratiques centrées sur l'apprentissage, avec un usage du numérique principalement pour des évaluations formatives ou diagnostiques. Des outils comme Moodle facilitent la gestion logistique et le suivi des étudiants, tandis que l'e-portfolio émerge comme instrument prometteur pour la validation des compétences. Toutefois, l'analyse fait ressortir un certain mimétisme entre les pratiques de formation vécues par les enseignants et celles qu'ils mettent en œuvre avec leurs étudiants.

Introduction

L'évaluation des apprentissages est une composante essentielle des pratiques pédagogiques dans l'enseignement supérieur. Avec l'intégration croissante des outils numériques, en particulier de l'intelligence

artificielle, cette dimension clé de l'enseignement est fortement requestionnée, offrant de nouvelles opportunités pour mieux accompagner les apprentissages des étudiants, mais soulevant également des défis importants. Entre transparence, personnalisation et gain d'efficience, le numérique ouvre un éventail de possibilités pour adapter les évaluations aux besoins diversifiés des apprenants, tout en posant des questions sur la charge de travail des enseignants ou encore l'équité entre les étudiants.

Ce chapitre s'intéresse à la manière dont les enseignants du supérieur conçoivent la place du numérique dans leurs pratiques d'évaluation. Soucieuses de mettre en avant des méthodes pédagogiques émergentes en la matière, notre intérêt s'est porté sur un panel d'enseignants engagés dans un diplôme universitaire (DU) en pédagogie universitaire à l'Université de Lille. Notre étude, fondée sur une méthode mixte combinant analyse textuelle et analyse des dispositifs décrits, s'appuie sur un examen de leurs 13 portfolios réflexifs. Ces portfolios permettent d'explorer les représentations des enseignants sur l'évaluation et le numérique, tout en identifiant les pratiques qu'ils développent ou adaptent pour évaluer les apprentissages et les compétences de leurs étudiants.

Contexte et problématique

Soucieuse d'engager ses enseignants dans un développement professionnel sur le long terme et d'en concrétiser les apports dans leurs pratiques effectives, l'Université de Lille a ouvert, en 2020, un DU, à destination d'enseignants déjà en poste dans l'établissement et souhaitant s'engager dans une formation diplômante en pédagogie universitaire. L'ouverture de ce diplôme s'inscrit dans la stratégie politique de l'établissement qui s'investit, d'une part, dans la formation et l'accompagnement des enseignants dans leurs pratiques d'enseignement et, d'autre part, dans la reconnaissance et la valorisation de l'engagement pédagogique des enseignants-chercheurs. La formation, qui dure un an et demi à raison d'un regroupement d'une journée par mois, est organisée autour de quatre compétences : (1) concevoir des enseignements dans une approche par compétences ; (2) encadrer des activités d'apprentissage ; (3) évaluer les apprentissages et les compétences des étudiants ; (4) collaborer en équipe pédagogique. Fondées sur le principe d'homomorphisme, repris de Loisy (2012, 2017) par Lison et Paquelin (2019, p. 68), selon qui « [...] l'expérience vivante d'une

situation formative qui présente un homomorphisme avec la situation à mettre en œuvre dans sa pratique, doublée d'une démarche réflexive, contribue au développement de nouvelles pratiques », les pratiques pédagogiques de ce DU ont été pensées de manière à former les enseignants en leur faisant vivre une formation organisée en approche par compétences (APC). Ainsi, chaque compétence est travaillée à travers des apports théoriques et pratiques (ateliers de formation, interventions de spécialistes en pédagogie universitaire, ouvrages et articles scientifiques, retours d'expérience d'enseignants, etc.) et des situations d'apprentissage et d'évaluation (Georges et Poumay, 2020) permettant le développement des compétences visées. Au cours de la formation, les enseignants réalisent un projet pédagogique qui consiste à faire évoluer, de manière significative, un de leurs enseignements. Ce projet est accompagné par un conseiller pédagogique référent qui suit l'enseignant au cours des trois semestres. Le premier semestre de la formation est ainsi consacré à la conception du scénario ; l'enseignement est réalisé et animé au deuxième semestre et, au troisième semestre du DU, l'enseignant évalue s'il a atteint ses objectifs pédagogiques. L'évaluation finale du diplôme se fait à travers un portfolio dans lequel les enseignants ont à démontrer le développement de chacune des compétences du référentiel en explicitant leurs actions et en s'appuyant sur les traces concrètes de leur mise en œuvre.

Souhaitant analyser leur utilisation du numérique dans leurs dispositifs d'évaluation, nous nous sommes posé les questions suivantes : quelles sont les représentations des enseignants sur les usages du numérique pour l'évaluation des apprentissages ? Quelle place est donnée au numérique dans les pratiques évaluatives qu'ils conçoivent et mettent en œuvre dans le cadre de leurs cours ? Quels outils numériques sont effectivement intégrés aux pratiques d'évaluation et avec quelles intentions ? Les enseignants ont-ils conscience des apports des outils numériques pour leurs pratiques d'évaluation ?

Notre expérience dans le conseil et la formation pédagogique nous permet de dégager quelques hypothèses. On peut en effet penser que ces enseignants, fortement engagés dans la pédagogie, ne serait-ce que par leur inscription à ce DU, seront sensibles à la mise en place de pratiques pédagogiques innovantes intégrant les outils numériques et qu'ils en auront des usages riches et diversifiés, notamment pour l'évaluation des apprentissages de leurs étudiants. Il pourrait s'agir, par exemple, de scénarios de classe inversée appuyés par des évaluations

formatives automatisées (quiz, sondages, etc.) ou encore d'usages de technologies pour transmettre des *feedbacks* aux étudiants. Notre expérience nous amène toutefois à penser que les outils numériques seront moins mobilisés par les enseignants pour les évaluations certificatives que formatives en raison de l'augmentation ressentie des risques de tricherie, du manque d'assurance des enseignants, etc.

Éléments de cadrage théorique

Le développement professionnel en pédagogie des enseignants du supérieur

Les enseignants du supérieur sont, pour la plupart, des spécialistes dans un domaine disciplinaire particulier. Ils ont dû faire état de leurs compétences en matière d'activités de recherche, mais ils ne sont que très rarement préparés à enseigner (Lanarès et Poteaux, 2013). Sans formation pédagogique initiale, l'enseignant-chercheur s'adapte plus ou moins seul à sa mission d'enseignement (Altet, 2013). Depuis quelques années, une attention particulière est portée à leur développement professionnel en pédagogie (Rege-Colet et Berthiaume, 2015). Ainsi, des structures d'appui à la pédagogie sont apparues au sein des établissements afin de soutenir la montée en compétence des enseignants dans leurs pratiques pédagogiques et de renforcer la qualité des apprentissages des étudiants. Ces structures proposent des offres diversifiées : ateliers de formation, échanges de pratiques, entretiens collectifs ou individualisés, soutien aux évaluations des enseignements par les étudiants, accompagnement à la démarche de *Scholarship of Teaching and Learning* (SoTL)[12] (Rege Colet *et al.*, 2011), etc. En France, un décret de 2018 introduit la mise en place d'un dispositif de formation obligatoire à la pédagogie pour les maîtres de conférences nouvellement recrutés.

Ce mouvement s'accompagne d'une réflexion sur les conditions qui permettent aux enseignants d'acquérir des connaissances et des compétences en matière d'enseignement, de partager celles-ci avec leurs collègues et d'être reconnus et valorisés pour leur travail (Frenay *et al.*, 2010). Si un référentiel de compétences existe bien pour le métier d'enseignant-chercheur (MESRI, 2019), il n'existe pas de cadre de référence

[12] « Processus de développement et de valorisation de l'expertise des enseignant·e·s de l'enseignement supérieur en matière d'enseignement et à propos de l'apprentissage des étudiant·e·s » (Biémar *et al.*, 2015, p. 1).

communément accepté ayant défini les concepts spécifiques au développement pédagogique dans l'enseignement supérieur (Frenay *et al.*, 2010). Frenay *et al.* (*ibid.*) soulignent que l'accent est mis sur les aspects liés à la conception d'un enseignement : sélection de contenus, définition et formulation d'objectifs d'apprentissage, choix de méthodes d'enseignement et développement d'outils d'évaluation (Biggs, 2003). À cela s'ajoutent des apports technopédagogiques liés aux usages d'outils numériques pour l'enseignement.

Comme le soulignent Demougeot-Lebel et Perret (2011), les études sur l'impact des actions de soutien au développement professionnel en pédagogie des enseignants du supérieur sont relativement rares. Les travaux effectués dans ce sens semblent s'accorder à montrer une évolution des conceptions des enseignants formés. Leur vision serait moins focalisée sur la transmission de contenus d'apprentissage et plus orientée vers les apprentissages étudiants (Kugel, 1993 ; Stes et Van Petegem, 2011 ; Bailly *et al.*, 2015 ; Demougeot-Lebel, 2017). Au-delà d'une sensibilisation à de nouvelles pratiques et d'un retour sur la satisfaction des enseignants ayant participé, il est difficile de mesurer les effets des actions de formation sur l'évolution des pratiques effectives des enseignants du supérieur. Du point de vue des usages du numérique pour l'enseignement, il est encore difficile aussi de savoir dans quelle mesure et surtout avec quelles intentions ils sont intégrés dans les pratiques pédagogiques des enseignants.

Représentations et numérique

Migne (1994, p. 23) définit la représentation comme « un modèle personnel d'organisation des connaissances par rapport à un problème particulier ». Or, Lac *et al.* rappellent en 2010 que, pour Moscovici (1961) :

> [l]'individu n'est ni isolé, ni fondu dans un grand tout sociétal. Il navigue dans des états intermédiaires que sont les groupes « réels ». C'est à partir de, et dans, ces groupes que vont se construire (dans un rapport de communication et d'interactions langagières en particulier) des connaissances permettant à l'individu de comprendre et de se mouvoir dans le monde qui l'entoure. Ces formes de connaissances particulières, car naïves et collectivement produites, concerneront des objets spécifiques dont une des particularités est leur polysémie, et ayant valeur d'enjeu pour le groupe. (Lac *et al.*, 2010)

Ainsi, Richardot (2022, p. 379-380) en vient à écrire que les représentations sociales « peuvent se définir comme l'ensemble des connaissances et des opinions *produites* et *partagées* par les individus d'un même groupe à l'égard d'un objet social donné » et a précisé que « opinions, croyances et informations » sont « les éléments qui la composent ».

Marty et Thomas Vasquez (2021, p. 2) soulignent que les injonctions à l'innovation pédagogique avec « les outils numériques » constituent « un puissant générateur de représentations ». Leur étude des représentations des usagers d'un dispositif pédagogique universitaire fait ressortir des représentations, technophiles et technophobes, optimistes et « technosceptiques », qui évoluent au fur et à mesure de la propre évolution du dispositif. Elles concluent en signalant

> la nécessité d'appréhender les usages et les représentations dans une dynamique interdépendante, dans laquelle les représentations peuvent faire évoluer les usages, au même titre que les usages peuvent faire évoluer les représentations. Notre contribution se rapproche en cela d'une perspective à coloration constructiviste, dans laquelle l'Usage et la Représentation exercent, l'un sur l'autre, une dépendance et une influence mutuelles et réciproques. (*Ibid.*, p. 15)

Méthodologie

Notre travail s'est attaché à étudier les portfolios réalisés par les enseignants pour rendre compte de leurs parcours de développement professionnel dans le DU. La partie relative à la démonstration de la compétence 3, « Évaluer les apprentissages et les compétences des étudiants », a été spécifiquement considérée afin de voir la place accordée aux outils numériques dans la manière dont les enseignants évoquent leurs pratiques d'évaluation. Notre analyse s'est déroulée en deux temps. Nous avons d'abord reconstruit, à travers ces écrits, les représentations des enseignants sur les usages du numérique. Nous nous sommes ensuite attachées à étudier la place octroyée au numérique dans les dispositifs d'évaluation qu'ils ont décrits dans leurs portfolios.

Analyse des représentations des enseignants

Pour analyser les représentations des enseignants sur le numérique, en particulier sur son utilisation dans le cadre d'une évaluation des apprentissages, nous avons procédé à une analyse textuelle des 13 portfolios.

Nous avons commencé par mener une analyse lexicométrique manuelle en recherchant la présence des mots « évaluation » et « numérique » dans les textes. Cette première recherche nous a permis d'identifier d'autres mots-clés et de circonscrire les acceptions pertinentes. Ainsi, la liste définitive est composée de :

- évalu*[13] : évaluation des apprentissages et des compétences uniquement ;
- contrôl*, modalités ;
- numérique, digital, hybrid*, présent*, présenc*, distanc*, en ligne ;
- Wooclap, Kahoot ! ;
- Moodle, plateforme ;
- test, questionnaire, quiz* ;
- devoir, atelier (au sens d'activités Moodle).

Finalement, les termes « modalités », « digital » et « questionnaire » ont été retirés, car non représentatifs. Le premier était forcément présent du fait de son utilisation dans une composante essentielle de la compétence 3. « Digital » n'apparaissait qu'une seule fois dans l'expression *digital natives*. Enfin, « questionnaire » était toujours associé à l'évaluation des enseignements par les étudiants. Afin d'assurer une analyse homogène entre les portfolios, à chaque nouveau terme retenu, nous avons repris tous les textes pour réaliser une nouvelle recherche sur celui-ci. Cette étape a également permis d'identifier les passages permettant de reconstruire des représentations sur l'évaluation, le numérique et le numérique pour l'évaluation.

Dans un second temps, une analyse textuelle automatisée a été menée à l'aide de l'application en ligne Hyperbase. Après intégration du corpus dans l'application, l'analyse lexicométrique a permis de faire ressortir 11 nouveaux mots-clés. Une analyse des cooccurrences a également été réalisée.

Analyse des dispositifs d'évaluation

Notre analyse s'est ensuite portée sur les dispositifs pédagogiques décrits par les enseignants dans leurs portfolios. Ceux-ci sont de deux types. Il peut s'agir : (1) de pratiques effectives non observées directement, mais

[13] L'astérisque remplace les différentes formes possibles et permet donc d'élargir la recherche et l'analyse. En utilisant « évalu* », l'outil recherche « évalue », « évaluation », « évaluer », etc.

mises en œuvre dans le cadre du projet du DU et accompagnées par un conseiller pédagogique ; ou (2) de pratiques déclarées rapportées par les enseignants pour étayer la justification du développement de leur compétence. Trente-huit pratiques d'évaluation ont ainsi été reconstruites du corpus constitué. Une analyse factorielle multivariée à visée de classification a été réalisée avec le logiciel Sphinx IQ3 au regard des variables présentées dans le tableau 3.1.

TABLEAU 3.1 Variables retenues pour l'analyse des dispositifs d'évaluation.

Variables étudiées	Exemples de modalités
Domaine d'enseignement	ST, SHS, ALL, Santé, STAPS, SJPS, SEST, IUT[14]
Modalités d'évaluation	Diagnostique, formative, sommative, certificative
Objectifs d'apprentissage visés	Rétention, Compréhension, Réflexion, Perception, Reproduction, Perfectionnement, Réception, Valorisation, Adoption, Développement d'une compétence
Type d'évaluation	Quiz, devoir sur table, projet, oral, journal de bord, portfolio, simulation…
Fonction de l'évaluation	Entraîner les étudiants, vérifier l'acquisition de connaissance ou de savoir-faire, mettre les étudiants en situations authentiques, etc.
Temporalité de l'évaluation	En classe, hors classe
Usage du numérique	Oui, non, lequel ?
Points d'attention associés par l'enseignant	Barème de notation, validité de l'évaluation, *feedback*, formulation des questions, subjectivité de la notation…

Résultats et analyse

Représentations des enseignants sur les usages du numérique pour l'évaluation

La première phase de l'analyse textuelle des portfolios nous a donné l'occasion d'identifier le nombre de portfolios contenant des passages permettant de reconstruire les représentations de leurs auteurs (figure 3.1).

[14] ST : sciences et techniques ; SHS : sciences humaines et sociales ; ALL : arts, lettres et langues ; Santé, STAPS : sciences et techniques des activités physiques et sportives ; SJPS : sciences juridiques, politiques et sociales ; SEST : sciences économiques, sociales et des territoires ; IUT : institut universitaire de technologie.

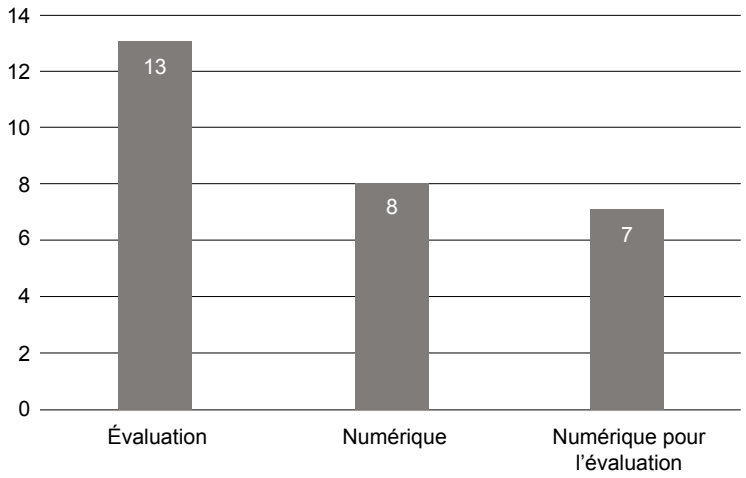


FIGURE 3.1 Nombre de portfolios intégrant des passages sur l'évaluation, le numérique et le numérique pour l'évaluation permettant de reconstruire des représentations.

Nous avons également comptabilisé le nombre de portfolios dans lesquels apparaît chaque mot-clé (figure 3.2).

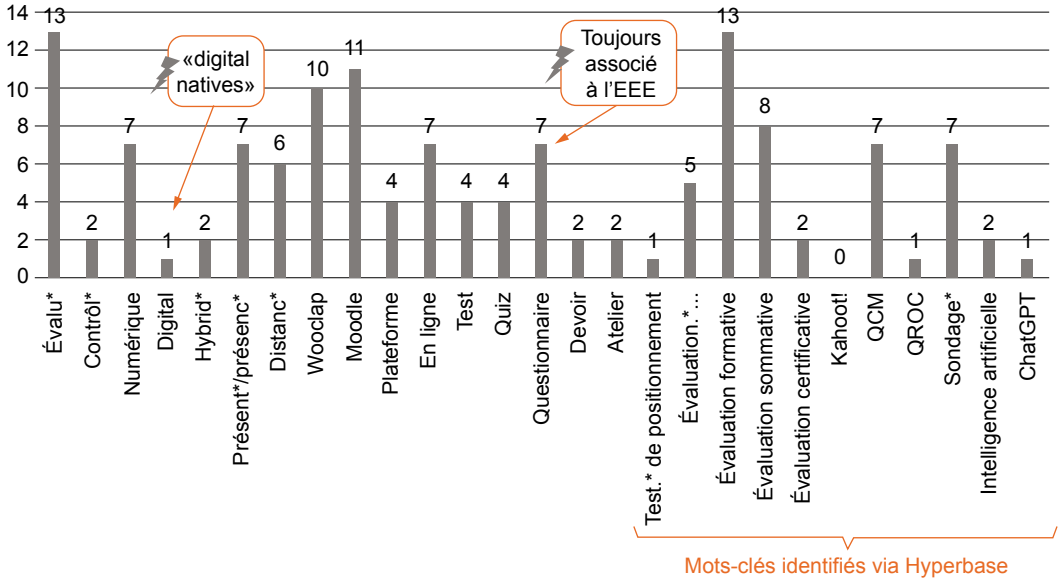


FIGURE 3.2 Nombre de portfolios dans lesquels apparaît chaque mot-clé.

Il n'est pas étonnant que le lemme « évalu.* » soit présent dans tous les portfolios étant donné la troisième compétence visée par le DU : « Évaluer les apprentissages et les compétences des étudiants ». Les auteurs ont donc eu à utiliser ce mot, en dehors de la simple copie du titre de la compétence. Nous pouvons également observer que, bien que « Moodle » apparaisse dans 11 portfolios sur 13, la moitié seulement utilise le terme « numérique ».

L'analyse des fréquences de cooccurrences fait ressortir une absence globale de lien entre l'évaluation et le numérique (figures 3.3 et 3.4).



FIGURE 3.3 Cooccurrences les plus fréquentes avec le lemme « évalu.* », avec Hyperbase.

Il est à noter que, dans le cas des cooccurrences avec « numérique.* », l'analyse est faussée. En effet, le projet d'un des participants impliquait des calculs numériques et c'est cette activité qui ressort en premier dans les cooccurrences avec « numérique.* », ainsi que l'analyse de résultats numériques (figure 3.4).

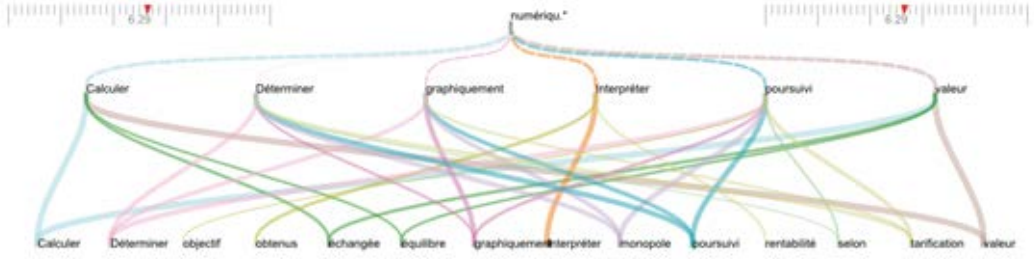


FIGURE 3.4 Cooccurrences les plus fréquentes avec le lemme « numérique.* », avec Hyperbase.

Ces différents résultats ainsi que la lecture des passages identifiés comme relevant de l'un des trois volets de notre étude nous ont permis à leur tour d'identifier des représentations partagées entre les auteurs des portfolios du corpus.

Représentations sur l'évaluation

Les enseignants indiquent que leur posture (Lameul, 2016) d'évaluateur a évolué avec leur participation au DU. Ainsi, l'un d'entre eux explique :

> J'ai évolué au cours du DU sur ma conception de l'évaluation et les différentes manières de la mettre en œuvre. J'ai pris conscience :
> - de l'importance de l'alignement pédagogique : veiller à ce que les questions posées aux examens aient bel et bien fait l'objet d'activités d'apprentissage adaptées et approfondies (je me rends compte que ce n'est pas toujours le cas pour les autres cours dans lesquels j'interviens) ;
> - que je mettais déjà en œuvre des évaluations formatives et que celles-ci sont bénéfiques pour soutenir l'apprentissage ;
> - qu'une note n'est pas indispensable ;
> - qu'engager les étudiants dans l'évaluation s'avère être une bonne option à approfondir. (Enseignant-chercheur [EC]-2)

Ils font également ressortir qu'ils adaptent leur évaluation pour rester alignés avec les profils des étudiants, les objectifs et les évolutions sociétales :

> J'essaie toujours de mettre en œuvre un bon alignement pédagogique (même si rien n'est jamais parfait) afin de m'assurer que les étudiants et étudiantes ont toujours les clés pour réussir et qu'ils comprennent les objectifs d'apprentissage et comment je les évalue. Cette question d'alignement pédagogique n'est pas une chose à laquelle je pensais auparavant, car pendant mon parcours en tant qu'étudiante, cela m'arrivait de suivre des cours qui n'avaient pas d'alignement pédagogique et je n'avais pas vraiment interrogé cela. Désormais, je comprends l'importance de cela, mais aussi l'importance d'être clair avec les étudiants et étudiantes et de les informer sur les objectifs d'apprentissage de leur cours et pourquoi cela va leur être utile. J'ai également appris beaucoup de choses sur les sortes d'évaluations que l'on peut mettre en place telles que les évaluations plutôt réflexives. (EC-9)

Comme l'indique l'EC-1, on observe à la lecture des textes une évolution de « l'évaluation sanction » vers une évaluation formative, voire une évaluation de la progression des étudiants : « Aujourd'hui, "évaluer les étudiants" ne rime plus pour moi avec "sanctionner", mais avec "aider les étudiants dans leurs apprentissages". » L'EC-3 précise :

> L'évaluation d'un projet professionnel ne vient pas à proprement parler sanctionner l'acquisition de connaissances, mais vise plutôt une démarche… celle de la capacité de l'étudiant à réfléchir à son projet, à s'évaluer et à mettre en place des pistes de progression.
>
> L'idée est d'aider les étudiants à comprendre que des compétences sont déjà acquises ou en cours d'acquisition […]. (EC-3)

Nous avons également relevé que les enseignants considèrent qu'une évaluation formative régulière favorise un apprentissage en profondeur et aide au positionnement des étudiants. D'où, selon eux, l'importance des *feedbacks* :

> Avec l'expérience, je me sens de plus en plus à l'aise dans ce rôle d'évaluatrice. Initialement, j'ai trouvé cette posture très difficile à prendre et je trouvais particulièrement désagréable de leur communiquer les points négatifs. Après beaucoup d'évaluations passées, j'envisage les choses différemment. Tout en restant empathique et constructive, j'essaie de leur donner des pistes d'amélioration si nécessaire. (EC-5)

Ils voient aussi le besoin de diversifier les méthodes pour évaluer les compétences des étudiants, tout en envisageant les modalités d'évaluation elles-mêmes comme outils pour développer d'autres compétences :

> Évaluation par les pairs : elle a l'intérêt d'être formative avec un retour des autres étudiants. Elle permet également aux étudiants de travailler leur communication. […]
>
> Nous l'avions mis[e] en place pour encourager la réflexion critique des étudiants nécessaire au bon déroulement de leur projet. (EC-4)

Enfin, les enseignants parlent de la nécessité de former les étudiants aux instruments et méthodes d'évaluation et de jongler avec les contraintes organisationnelles.

Représentations sur le numérique

Le numérique est vu par les enseignants auteurs des portfolios sous une dimension ludique : « J'avais déjà l'habitude d'utiliser les Wooclaps [...], alors je savais que les étudiants et étudiantes aiment bien et cet aspect compétitif d'avoir le plus de points a bien marché avec ce public. [...] » (EC-9) ; mais également comme une aide à l'apprentissage pour favoriser la motivation et la mémorisation à long terme. Ils envisagent également le numérique comme un outil leur permettant de rendre leurs étudiants plus actifs et de développer leur autonomie. Ils indiquent cependant être vigilants afin de ne pas surestimer cette « autonomie et l'engagement des étudiants dans leurs études en me projetant dans la manière dont j'aurais vécu ce cours en tant qu'étudiante » (EC-6). La question des disparités d'accès des étudiants au numérique et leurs impacts sur l'apprentissage ne ressortent pas dans notre étude.

Pour certains, c'est surtout en lien avec des questions logistiques qu'ils voient la valeur ajoutée du numérique. La plateforme Moodle est vue comme un espace de partage et de centralisation des informations, facilitant notamment la collaboration entre enseignants. Le numérique est utilisé pour hybrider les enseignements afin de dégager du temps pour mieux préparer à l'évaluation en présentiel, par exemple, comme pour assurer un accompagnement individualisé dans une grande cohorte.

Représentations sur l'usage du numérique pour l'évaluation

Concernant l'utilisation du numérique pour l'évaluation des apprentissages, les enseignants ne l'envisagent que pour des évaluations diagnostiques et formatives, principalement sous forme de tests via Wooclap ou Moodle, permettant l'autoévaluation en autonomie. Ils y voient l'avantage de pouvoir fournir une remédiation immédiate, qu'elle soit automatisée ou donnée par l'enseignant, ainsi que la possibilité d'adapter leur enseignement en cours de session. Enfin, ils considèrent que le numérique permet d'alléger la charge organisationnelle de l'évaluation avec un effectif important :

> Pour réduire le travail de correction, l'interrogation en séance prendra la forme d'un test hors ligne Moodle. Cette correction automatique permettrait d'obtenir un tableau avec les prérequis maîtrisés et non

> maîtrisés par chaque étudiant, ce qui est plus utile pour moi comme pour les étudiants qu'une note globale. (EC-6)

L'influence du numérique sur la validité et la fiabilité des évaluations n'est pas abordée par les auteurs.

Cette analyse des représentations montre que ces enseignants envisagent la mise en place de pratiques pédagogiques innovantes intégrant les outils numériques. Dans une logique évaluative, il apparaît que l'usage du numérique est plutôt associé à des pratiques formatives que certificatives, par crainte de comportements contraires à l'intégrité académique. Nous allons voir maintenant comment celles-ci se concrétisent dans les pratiques mises en œuvre par les enseignants.

Usages des outils numériques dans les pratiques d'évaluation des enseignants

L'analyse des pratiques évaluatives décrites dans le corpus permet de dégager différents types d'usages du numérique qui sont développés ici.

Des outils numériques pour la mise en place de QCM lors d'évaluations formatives ou diagnostiques

L'usage des outils numériques pour les évaluations de type questions à choix multiples (QCM) est largement répandu dans les pratiques de ces enseignants. Cela passe soit directement par la plateforme Moodle de l'établissement soit par des plateformes dédiées. Ces pratiques sont menées dans une démarche formative pour entraîner les étudiants et soutenir leur engagement dans leurs apprentissages ou dans une démarche diagnostique afin de faire le point sur des acquis d'apprentissage, notamment dans les enseignements de langue. Cette modalité n'est jamais utilisée dans un but certificatif dans les pratiques rapportées ici. L'évaluation, autocorrective, est menée soit directement en cours pour renforcer l'interactivité de l'enseignement, soit proposée en soutien du travail hors classe des étudiants qui les réalisent de façon autonome.

Dans leurs portfolios, les enseignants soulignent différents enjeux dans la mise en place de ces pratiques d'évaluation. Par exemple, la difficulté à rédiger les questions. Une enseignante, qui s'essaie nouvellement à cette pratique, explique sa démarche avec prudence : « Ayant

peu d'expérience dans l'élaboration de tests QCM, je me suis appuyée et inspirée de questions tirées d'un manuel dédié à la biologie végétale de niveau licence. Je souhaitais couvrir l'ensemble des notions vues en TD en évitant les ambiguïtés d'interprétation ou de syntaxe. » Les apports de la formation sur le principe d'alignement constructif (Biggs, 2003) amènent les enseignants à s'interroger sur la validité de ces évaluations au regard des objectifs pédagogiques identifiés pour leurs enseignements. Ainsi, un enseignant explique avoir pris conscience de l'écart entre ce qu'il demandait aux étudiants dans les évaluations formatives (vérification des connaissances) et ce qu'il attendait lors des évaluations sommatives (résolution de problèmes complexes). Il pointe alors le fait que ces évaluations formatives peuvent tromper le sentiment de compétence des étudiants et nuire à leur réussite si elles ne sont pas alignées avec les évaluations certificatives. Le barème à adopter pose également question aux enseignants. L'un d'entre eux explique avoir finalement décidé de supprimer « des points de pénalités liés aux réponses aléatoires de l'étudiant et cela, afin d'optimiser sa réflexion [...] le but de ces QCMs ne consistait pas à tester la "compétence assurance" de l'étudiant ». Utilisés dans une visée diagnostique, c'est alors les remédiations proposées à l'issue des tests et les pratiques pédagogiques adaptatives à mettre en place qui interrogent les enseignants.

Si l'usage des outils numériques a manifestement permis de développer le recours à des évaluations type QCM à visées formatives ou diagnostiques, l'analyse réflexive des enseignants de ce DU montre une prise de conscience des enjeux associés à ces pratiques d'évaluation. Alors qu'elles pouvaient paraître relativement intuitives, la justification de leur intégration dans le scénario pédagogique, notamment au regard de la cohérence avec les objectifs pédagogiques et des modalités d'évaluation certificatives, fait émerger des questionnements particulièrement intéressants.

Des pratiques d'évaluation certificatives sans usage du numérique

Un certain nombre de pratiques d'évaluation sont évoquées sans lien avec un usage des outils numériques. C'est le cas notamment pour les pratiques certificatives ayant pour objectif d'évaluer le niveau de connaissance et de compréhension des étudiants sur des points de cours spécifiques (niveaux de rétention, compréhension, réflexion dans

la taxonomie de Daele et Berthiaume [2009]). Il s'agit, par exemple, d'interrogations écrites, de devoirs surveillés de milieu ou de fin de semestre, etc. Les évaluations visant la maîtrise d'un geste entrent également dans ce cadre (niveaux de perception, reproduction ou perfectionnement dans la taxonomie citée précédemment). On trouve par exemple la réalisation d'une dissection, des coupes d'organes, l'observation microscopique, la reconnaissance de tissus, etc. Ces évaluations se passent « sur table » et sous surveillance, l'usage des outils numériques n'est pas attendu, voire interdit par souci de préserver la validité et la fiabilité de l'évaluation.

Pour les enseignants du DU, ces pratiques évaluatives sont souvent complémentaires à d'autres ; seuls deux enseignants utilisent cette pratique comme modalité d'évaluation principale de leur enseignement. Dans les justifications apportées dans les portfolios, les enseignants soulignent la nécessité de maintenir ce type de pratiques d'évaluation dans certains cas, pour certains apprentissages spécifiques. L'écriture réflexive amène les enseignants à pointer deux limites : les étudiants progressent peu d'un examen à l'autre et ils ont tendance à ne pas revenir sur leurs erreurs. Pour y faire face, un enseignant explique avoir mis en place une démarche d'autopositionnement avec ses étudiants. Avant de leur rendre leur copie, il leur distribue le corrigé et leur demande de s'évaluer. Selon lui, cette pratique leur permet de mieux s'approprier les critères d'évaluation et d'identifier activement ce qu'ils ont traité correctement ainsi que les erreurs qu'ils ont commises (erreurs de calculs, d'inattention, de raisonnement, absence d'une unité, d'une justification, etc.). Les retours des étudiants, recueillis dans le cadre d'un *focus group* animé par un conseiller pédagogique, vont dans ce sens : « J'ai beaucoup aimé les exercices d'autocorrection. Avant, j'avais l'habitude d'avoir la correction juste après, donc je faisais toujours les mêmes fautes, car ça ne me conduisait pas à réfléchir. » ; « Ça nous force à chercher nos erreurs, à les comprendre. Quand on regarde juste la correction, finalement on ne comprend pas vraiment. » ; « Ça me pousse à avoir un regard autre sur ma copie. »

L'usage du numérique comme outil de dépôt de productions support à l'évaluation

La plateforme Moodle est largement utilisée comme support au dépôt de productions réalisées par les étudiants dans le cadre d'activités

pédagogiques faites en classe ou hors classe (10 enseignants sur 13 l'utilisent à cette fin) : la résolution d'un *serious game* sur la gestion d'une entreprise en marketing, la création d'un poster de communication scientifique en sciences et technologies, la rédaction d'un sketch humoristique en anglais sur la culture et la civilisation britannique, le développement d'un algorithme de *machine learning* en informatique, le scénario d'animation d'une séance d'activité physique et sportive, un retour de stage en médecine, etc. Les dispositifs pédagogiques associés s'appuient sur des mises en situation authentiques et les enseignants, dans leurs justifications, renvoient parfois explicitement à la notion de « situation d'apprentissage et d'évaluation » (SAE) (Poumay *et al.*, 2017 ; Georges et Poumay, 2020). Ce résultat est toutefois bien à contextualiser dans le cadre de notre étude, car c'est un attendu de la formation : les enseignants doivent inscrire leur projet pédagogique au regard de la notion de SAE.

Les enseignants, dans leurs écrits réflexifs, expriment leurs prises de conscience dans leur tentative de mise en œuvre d'une situation d'apprentissage et d'évaluation, que nous reprenons ici :

- la difficulté du choix de la situation « authentique », pour qu'elle ne favorise pas plus un étudiant qu'un autre, qu'elle soit reproductible, qu'elle ne soit pas trop difficile, etc. ;
- la gestion des biais liés à la subjectivité de l'évaluation : on note la volonté, pour la plupart des enseignants, d'utiliser une grille critériée pour apprécier la performance des étudiants ;
- la nécessité de communiquer clairement les exigences aux étudiants en amont de l'activité ;
- l'ambiguïté du sens de la note et notamment les difficultés à pondérer des critères d'évaluation pour aboutir à une note chiffrée dans le cadre de tels dispositifs ;
- l'augmentation de la charge de travail liée à la rétroaction et aux *feedbacks*. La nécessité d'un travail collaboratif entre enseignants dans l'évaluation est pointée à plusieurs reprises ;
- la capacité à apprécier une compétence individuelle dans un travail de groupe.

Les enseignants soulignent également le changement de posture induit par l'animation de ces situations d'enseignement et d'apprentissage. Une enseignante explique : « Mon activité d'enseignante a donc pris différentes formes. Je les ai aidés à décrire et expliquer leurs

choix et leurs actions, puis aidés à apporter des solutions qui étaient finalement les contenus de mon cours. » Invités à expliciter les objectifs associés à chacun de leurs dispositifs, les enseignants pointent, derrière cette pratique, un objectif double : évaluer des compétences et vérifier l'acquisition de connaissances. Il est intéressant de constater que l'appréciation du développement des compétences et la vérification de la maîtrise de contenus essentiels ne sont pas dissociées, mais considérées comme entrechâssées dans les justifications apportées par les enseignants. Certains l'explicitent d'ailleurs comme une prise de conscience au cours de leur parcours de développement professionnel dans ce diplôme universitaire. Si le numérique est bien utilisé à des fins de dépôt de documents, il apparaît ici comme une interface nécessaire à la transmission de productions issues de situations d'apprentissages authentiques à des fins d'évaluation. Ces situations appellent, presque systématiquement, l'usage d'outils numériques (*serious game* en ligne, recherche documentaire, rédaction de livrables…).

Le e-portfolio, un usage émergent du numérique dans une visée certificative

Trois enseignants mettent en place, avec leurs étudiants, un journal de bord pour soutenir une démarche réflexive sur le développement de leurs compétences. Cette pratique est transposée de leur expérience dans le DU où ils ont, eux-mêmes, eu à réaliser un journal de bord. Chaque enseignant s'est toutefois approprié la démarche différemment, selon son propre contexte, selon ses propres objectifs. Pour l'un,

> [c]'est un livre de bord professionnel physique sous la forme d'un portevues contenant les différents documents travaillés (en lien avec les objectifs d'apprentissage du semestre). La construction du livre de bord permet de matérialiser une réflexion professionnelle, en cohérence avec ses propres finalités et son environnement professionnel [économie].

Pour l'autre, c'est un support d'analyse de l'expérience de stage : « l'étudiant rappelle d'abord ses objectifs de stage puis présente ses deux traces d'apprentissage. Le tuteur l'amène à prendre conscience du raisonnement observé, et à avoir une réflexion sur son action » (médecine). Un troisième explique :

> Je les invite à prendre un cahier de brouillon pour noter, pour mémoire, les actions entreprises, les réponses apportées individuellement et en groupe, les échecs, les réussites, les questionnements, les schémas, les tableaux, les renvois vers les documents concernés. Le cahier de brouillon est utile, car il permet de rayer les pistes qui se sont avérées de mauvaises pistes, de recommencer, de rajouter… Je suis convaincue que le cahier de brouillon permet d'accompagner le travail en cours d'élaboration de façon collective, personnelle et il permet la prise de recul. (BUT)

Deux enseignants introduisent l'e-portfolio (version Karuta – KAPC+) pour la certification du développement des compétences de leurs étudiants. L'un, en sciences et techniques des activités physiques et sportives, l'utilise pour évaluer la compétence « Participer à la gestion administrative et financière d'une structure d'accueil ». Il s'agit de dépasser l'analyse d'une situation et de présenter, au cours de son alternance, ce qui justifie d'avoir acquis les compétences du diplôme de DEUST (diplôme d'études universitaires scientifiques et techniques). Des preuves et des justifications sont attendues de la part des étudiants, elles sont déposées dans l'interface numérique. Une soutenance orale est organisée pour défendre la compétence visée devant le tuteur académique, le responsable de formation, le maître d'apprentissage, un autre apprenti et son équipe pédagogique. Les membres du jury ont accès à ce qui a été déposé par l'étudiant dans son e-portfolio. L'apprenti doit, quant à lui, répondre à la question : « En quoi votre année d'alternance sera-t-elle bénéfique pour un recrutement dans la poursuite de vos études ou dans le domaine professionnel ? » et doit s'appuyer sur son e-portfolio. L'autre l'a intégré dans le cadre d'un bachelor universitaire de technologie (BUT). Les compétences présentées dans le portfolio sont évaluées directement dans l'interface numérique sur la base d'une grille d'appréciation établie d'après des indicateurs de performance. Les étudiants doivent autoévaluer, également dans l'outil, leurs compétences au moyen de cette grille en amont de l'évaluation par l'enseignant. Cela leur permet de prendre conscience des attentes et d'autoréguler l'alimentation de leur portfolio. Grâce à l'outil numérique, l'enseignant dispose d'une interface lui permettant de suivre les dépôts des étudiants, d'y apporter des *feedbacks* et d'accéder à l'autopositionnement de l'apprenant.

Cette pratique reste émergente, y compris parmi les enseignants engagés dans le DU. Les prises de conscience évoquées dans les portfolios relativement à cette pratique font apparaître trois enjeux :

- l'accompagnement nécessaire des étudiants dans la démarche réflexive (souligné notamment par les enseignants de BUT 1). Analyser ses pratiques, justifier du développement de ses compétences, porter un regard critique sur ses choix sont des postures nouvelles pour les étudiants ;
- un changement du statut de l'erreur. Un enseignant explique que, dans cette démarche, « l'erreur devient constructive ». Elle devient un levier de progression et d'apprentissage ;
- la difficulté, dans un contexte universitaire, à dépasser la performance dans une situation donnée pour évaluer la compétence au regard de différentes situations. Un enseignant s'interroge : « l'évaluation des compétences, difficile… En effet, les compétences privilégient l'action et un apprenant compétent est opérationnel dans des situations complexes diversifiées […] Je n'ai la possibilité de mettre les étudiants en action que dans une seule SAE, qui leur propose une semi-autonomie. Est-ce suffisant ? »

L'apport d'une interface numérique venant outiller la démarche portfolio est manifeste à plusieurs niveaux. D'abord, elle permet à l'étudiant de conserver dans un même espace des traces de ses expériences et de ses productions dans une logique de valorisation sur le long terme (portfolio de présentation) ou de certification afin de collecter et de sélectionner les traces les plus pertinentes pour justifier d'un niveau attendu. Ces traces peuvent alors être de différentes formes : documents texte, fichiers multimédias, lien vers des sites web… Elle facilite également les interactions asynchrones entre l'étudiant et son ou ses évaluateurs. Par exemple, pour solliciter des *feedbacks* sur ses productions. Elle permet encore à l'étudiant de suivre et de matérialiser sa progression dans le développement de ses compétences en mettant en perspective des logiques d'autopositionnement et d'évaluation formatives et certificatives tout au long d'un parcours de formation.

Conclusion et conseils de mise en œuvre pratique

L'analyse des pratiques et des représentations des enseignants inscrits dans le DU de pédagogie universitaire met en évidence une transformation progressive de leur posture évaluative. Sous l'effet de leur développement professionnel, ces enseignants évoluent vers une approche plus formative et centrée sur l'apprentissage des étudiants. Cette

évolution s'accompagne d'une prise de conscience accrue des enjeux liés à l'alignement constructif, à la diversification des types d'évaluation et à l'intégration réfléchie des technologies dans les dispositifs.

Les outils numériques, bien qu'encore principalement utilisés pour des évaluations formatives ou diagnostiques, apparaissent comme des leviers efficaces pour dynamiser les apprentissages, faciliter le suivi des étudiants et réduire certaines contraintes organisationnelles. Cependant, leur utilisation reste limitée dans les pratiques certificatives, en raison de freins tels que le manque de formation, la complexité technique et la crainte de dénaturer le processus d'évaluation. Parmi les innovations observées, le recours au portfolio numérique apparaît aux enseignants comme un instrument prometteur pour évaluer les compétences de manière authentique et réflexive.

Notre étude s'appuyait sur l'analyse de 13 portfolios pour identifier les représentations des enseignants auteurs et leurs pratiques du numérique pour l'évaluation. Des perspectives de poursuite ressortent :

- sur le plan pédagogique : identifier les défis concrets rencontrés par les enseignants dans l'intégration des outils numériques (résistance au changement, formation insuffisante, surcharge de travail) pour proposer des solutions concrètes ; analyser les contributions pédagogiques des outils numériques, comme Moodle et Wooclap ;
- sur le plan docimologique : explorer de façon plus approfondie les effets des outils numériques sur le développement de compétences transversales, comme la réflexion critique et la collaboration ; mener une analyse critique des limites des outils numériques pour évaluer des compétences complexes ; étudier dans quelle mesure le numérique pourrait dénaturer le processus d'évaluation ;
- sur le plan éthique : interroger l'équité entre étudiants dans le cadre d'une évaluation via des outils numériques.

Un point mis en lumière par cette étude est le mimétisme observé entre les expériences formatives vécues par les enseignants dans le cadre du DU et les pratiques qu'ils mettent en œuvre dans leurs enseignements. Les activités expérimentées, comme les évaluations formatives avec Wooclap ou le recours au portfolio, sont souvent transposées dans leurs propres cours, témoignant de l'influence directe du DU sur leur développement professionnel pédagogique. Cette dynamique montre l'importance de proposer aux enseignants des formations ancrées sur le principe d'homomorphisme, favorisant une réflexion critique sur

leurs pratiques. Ainsi, le mimétisme devient un levier pour essaimer des pratiques pédagogiques innovantes. Il serait ainsi intéressant d'intégrer plus explicitement dans le DU une réflexion sur l'utilisation des outils numériques dans l'enseignement-apprentissage, et plus particulièrement dans l'évaluation, afin de contribuer de cette façon au développement des compétences des enseignants sur ce sujet et d'observer l'évolution de leurs pratiques.

Étienne Gibaud, François Stephan
& Diana Griffoulières

4 *Feedbacks* par intelligence artificielle en école d'ingénieurs: un premier regard

Résumé

L'influence du *feedback* sur la réussite des étudiants a été largement démontrée à travers de nombreuses études empiriques. Ces travaux soulignent le potentiel du *feedback* pour favoriser les processus d'enseignement et d'apprentissage à tous les niveaux éducatifs. Cependant, pour que le *feedback* soit efficace, il doit répondre à de nombreuses exigences, ce qui peut rendre difficile sa mise en pratique régulière. L'émergence de l'intelligence artificielle générative (IAg) connaît actuellement un développement sans précédent dans le domaine de l'éducation. Parmi les principales potentialités offertes par les IA génératives pour favoriser l'apprentissage des étudiants en enseignement supérieur, les études soulignent sa capacité à fournir un *feedback* personnalisé. Cette étude vise à analyser l'impact du *feedback* sur le travail de groupes d'étudiants dans deux classes d'une école d'ingénieurs, l'une recevant un *feedback* de l'enseignant et l'autre de l'intelligence artificielle ChatGPT. Elle évalue la qualité des travaux produits et les perceptions des étudiants sur le *feedback* reçu. Les résultats montrent que le *feedback* fourni par l'enseignant a eu un impact plus significatif sur l'amélioration des travaux. Cependant, le *feedback* fourni par les IA génératives a également montré un effet positif sur les réalisations des étudiants, soulignant le potentiel de l'intégration des outils

d'intelligence artificielle dans l'évaluation formative. Les résultats de cette étude offrent des pistes de réflexion et des considérations importantes pour la mise en pratique d'un processus de *feedback* fourni par les IA génératives.

Introduction

Le *feedback* joue un rôle crucial dans l'évaluation formative. Cependant, du fait du nombre important d'étudiants et des contraintes de temps auxquelles sont soumis les enseignants, la mise en place d'une approche personnalisée est difficile. Les nouvelles technologies peuvent contribuer au développement de modèles d'évaluation flexibles et adaptables, facilitant ainsi la mise en œuvre d'évaluations automatisées (Lafleur *et al.*, 2021).

Depuis 2022, l'émergence des intelligences artificielles (IA) génératives (IAg), notamment sous la forme d'agents conversationnels, a considérablement influencé les pratiques éducatives dans l'enseignement supérieur (Kamalov *et al.*, 2023). Cela nous a conduits à étudier d'une part, les usages de ces technologies par les étudiants et, d'autre part, les effets de ces usages sur les apprentissages des étudiants.

L'EPF École d'ingénieurs forme des ingénieurs généralistes. Tous les étudiants de troisième année participant au dispositif présenté ont suivi un cours de mathématiques intitulé « Méthodes numériques », où les aspects théoriques et pratiques de résolution numérique de problèmes scientifiques et d'ingénierie sont abordés. Dans le cadre de cet enseignement, nous nous intéressons à l'analyse et à la comparaison des effets du *feedback* donné soit par l'enseignant soit par l'intelligence artificielle générative ChatGPT sur un compte rendu de travail. Nous étudions la qualité des travaux produits par les étudiants, leurs perceptions de l'utilité de ces outils ainsi que la qualité du *feedback* reçu.

Le contexte et la problématique

Cette étude fait partie d'une initiative visant à mieux accompagner les enseignants dans leurs pratiques de *feedback*, afin de les rendre plus efficaces et durables dans le temps. En effet, la Cellule d'innovation pédagogique numérique (IPN) de l'EPF, au cours de l'année 2021-2023, a créé un groupe de travail autour du *feedback* avec cet objectif. L'enquête sur l'évaluation des pratiques de *feedback* au sein de l'EPF (Cellule

IPN, 2022) auprès des étudiants et des enseignants a permis d'obtenir un état des lieux de ses pratiques. Parmi les constats relevés, il a été observé que la mise en œuvre systématique du *feedback* individuel durant le processus d'apprentissage des étudiants est très difficile pour les enseignants. La raison évoquée principalement est le temps d'implication que cela nécessite pour fournir des informations personnalisées, constructives et opportunes. L'analyse des résultats a également indiqué que les étudiants ont besoin de recevoir un *feedback* en cours d'apprentissage pour pouvoir s'améliorer.

L'utilisation des IA génératives par les étudiants de l'EPF est importante. L'analyse des résultats d'une enquête sur les usages des IA génératives au sein de l'EPF (Cellule IPN, 2024) a révélé qu'environ 70 % des étudiants les utilisent régulièrement et que 75 % d'entre eux affirment avoir une bonne compréhension de leur fonctionnement et des prompts nécessaires à leur utilisation. L'enquête mondiale sur l'utilisation des IA génératives par les étudiants d'éducation supérieure, réalisée par le Digital Education Council (2024), a relevé que 86 % des étudiants utilisent l'IA dans le cadre de leurs études. Selon cette enquête, ChatGPT est l'outil d'IA le plus utilisé, avec 66 %. En raison de la croissance des usages des outils des IA et des résultats de nombreuses recherches qui indiquent les opportunités offertes par ces nouvelles technologies pour faciliter les pratiques éducatives, ChatGPT 4.0 a été choisi pour générer le *feedback* aux étudiants.

Cette étude a été réalisée dans le cadre du module « Méthodes numériques » destiné aux deux classes d'environ 30 étudiants de troisième année de formation du cycle ingénieur sur le campus de Montpellier, qui a pour objectif de fournir aux étudiants une compréhension des techniques et des outils numériques utilisés pour résoudre des problèmes mathématiques complexes.

Dans ce contexte, comment le *feedback* fourni par l'enseignant et celui fourni par ChatGPT influencent-ils, dans le cadre d'une évaluation formative, la qualité des travaux des étudiants et leurs perceptions sur le *feedback* reçu ?

Le cadre théorique

Le *feedback* est un élément clé dans la réussite des étudiants (Hattie et Timperley, 2007). Des recherches récentes indiquent que le *feedback*, qui a longtemps été considéré comme un processus unidirectionnel,

centré sur l'élaboration et la transmission des informations, devrait être considéré comme un processus dialogique (Guasch *et al.*, 2013 ; Yang et Carless, 2012 ; Espasa *et al.*, 2018) entre l'enseignant et les étudiants ou entre les étudiants eux-mêmes. Ce processus prend en compte le rôle actif des étudiants dans la construction de leurs connaissances, en s'appuyant sur les informations reçues du *feedback* (Winstone *et al.*, 2017). À travers les interactions liées au *feedback*, les étudiants sont guidés et encouragés à réfléchir, à raisonner et à partager leurs interprétations (Nicol, 2010). Cela leur permet d'utiliser de manière constructive les informations et les échanges générés lors du *feedback* (Yang et Carless, 2012).

Les travaux de recherche actuels sur le *feedback* se concentrent davantage sur la manière dont ce *feedback* est donné et comment ses effets peuvent être influencés par les différents agents qui le fournissent, ainsi que sur la manière dont les étudiants utilisent les informations du *feedback* pour leur propre bénéfice (Guasch et Espasa, 2015 ; Carless et Boud, 2018). Dans cette perspective, Guasch et Espasa (2015) ont présenté le cycle de *feedback* dialogique en trois étapes essentielles : (1) donner ou recevoir du *feedback* ; (2) comprendre les informations contenues dans le *feedback* ; et (3) appliquer les informations pour améliorer sa performance.

Pour mettre en place un *feedback* durable, il est nécessaire de faciliter et d'impulser des processus de *feedback* continu et autonome. Cela permettra aux étudiants de développer eux-mêmes des capacités de suivi et d'évaluation de leur propre apprentissage (Carless, 2013). En ce sens, Winstone (2022) affirme que l'étudiant est l'agent principal du processus de *feedback* et que la source de ce *feedback* peut provenir de plusieurs agents : de l'enseignant, d'un camarade de classe ou de l'étudiant lui-même.

Les plateformes numériques ont révolutionné de nombreux aspects des pratiques d'évaluation (Astuti et Retnawati, 2022). Ainsi, elles permettent aux enseignants de suivre et d'évaluer plus aisément les progrès des étudiants, de repérer leurs erreurs et de les corriger rapidement. L'émergence de l'intelligence artificielle générative, plus concrètement de ChatGPT en 2022, un agent conversationnel qui permet de générer du contenu et des réponses en simulant une conversation humaine, à travers des interactions écrites et plus récemment, vocales, influence considérablement le système éducatif et met en évidence la nécessité de repenser nos pratiques éducatives pour les rendre plus efficaces et

plus significatives. Une des raisons pour lesquelles ces technologies montrent un potentiel dans le domaine de l'éducation réside dans le fait que ces outils ont été entraînés sur une grande base de données pédagogiques telles que des articles scientifiques, des livres, des exercices, des plans de cours, etc. (Wang *et al.*, 2024).

L'évolution rapide des outils fondés sur les IA a suscité la réalisation de plusieurs études qui reconnaissent le potentiel de ces technologies pour améliorer l'efficacité des enseignements, notamment pour la génération de tests et de *feedbacks* personnalisés et automatisés en temps réel (Wang *et al.*, 2024 ; Qadir, 2023). Selon Mollick et Mollick (2023), les outils d'IA générative peuvent être utilisés pour générer des commentaires de *feedback*. De plus, ils sont capables de fournir des réponses rapides et personnalisées sur le travail des étudiants. Cao et Zhong (2023) ont examiné les effets du *feedback* fourni par ChatGPT sur les performances des étudiants dans la résolution des problèmes de programmation et la rédaction argumentative. Les résultats de cette étude soulignent le potentiel de ChatGPT pour améliorer les apprentissages des étudiants. ChatGPT pourrait alléger la charge de travail des enseignants, et leur permettre d'offrir un accompagnement psychologique et émotionnel aux étudiants (Roose, 2023 cité dans Wang *et al.*, 2024 ; Guo *et al.*, 2023). L'expérience menée par Sung *et al.* (2023) indique que ChatGPT-3 pourrait être utilisé de manière régulière pour rédiger des commentaires de *feedback* encourageants et en temps réel aux étudiants en difficulté. Cependant, les auteurs mentionnent que ces commentaires peuvent donner parfois des retours inappropriés. D'autres études montrent que les *feedbacks* fournis par des IA sont souvent génériques et parfois incorrects comparés à ceux des enseignants, plus précis et adaptés aux besoins des étudiants (Wang *et al.*, 2024). Selon Bosc-Miné (2014), l'automatisation du *feedback* pour des travaux complexes est plus difficile, car ils nécessitent un *feedback* qualitatif, individualisé et adapté aux prérequis de l'activité.

Malgré l'essor de l'intelligence artificielle en tant qu'outil d'évaluation formative, en particulier pour le *feedback*, ces pratiques ne sont souvent pas généralisables (Cao et Zhong, 2023). Il est nécessaire de mener davantage d'études sur l'utilisation des IA génératives au service des *feedbacks* pour mieux comprendre leurs potentiels et leurs performances.

Le dispositif et le scénario pédagogique

Le dispositif déployé dans cette étude implique deux classes constituées d'étudiants en troisième année de la formation généraliste de l'EPF École d'ingénieurs, suivant un cours de mathématiques intitulé « Méthodes numériques ». Ce cours aborde les aspects théoriques et pratiques de la résolution numérique de problèmes scientifiques et d'ingénierie. Chaque classe était composée de 32 étudiants, répartis en 11 groupes de travail. Ces groupes, généralement composés de trois étudiants, se sont formés volontairement sans aucun critère prédéfini.

La première classe correspond à la classe témoin (groupes E, pour « Enseignant ») et reçoit, comme d'ordinaire, un *feedback* général de la part de l'enseignant adressé à la classe entière. La deuxième classe (groupes IA, pour « Intelligence artificielle ») correspond à la classe qui a utilisé ChatGPT pour solliciter des *feedbacks* sur ses travaux (voir tableau 4.1).

Tous les groupes des deux classes ont réalisé la même tâche qui consistait en la production d'un compte rendu dit de « TP » (travaux pratiques) en trois fichiers distincts à rendre sur la plateforme Moodle : (1) modélisation et mise en forme du problème, impliquant un travail mathématique précis ; (2) résolution par programmation sur Matlab® ; et (3) analyse et interprétation des résultats. Une semaine après la soumission, les groupes ont reçu un *feedback* ou sollicité un *feedback* auprès de ChatGPT en fonction de leur classe et ont élaboré une seconde version améliorée de leur rendu. Cette activité de *feedback* a duré une heure trente, incluant la rédaction de cette version améliorée.

TABLEAU 4.1 Tâches réalisées par chaque classe et types de *feedbacks* fournis.

Classe	Tâche 1	Tâche 2 (*feedback*)	Tâche 3
Groupes E	Compte rendu v1	*Feedback* global de l'enseignant	Compte rendu v2
Groupes IA	Compte rendu v1	*Feedback* de ChatGPT	Compte rendu v2

Afin de profiter des modèles les mieux entraînés et des réponses les plus qualitatives, un espace dédié sur la plateforme ChatGPT version 4.0 a été proposé à chaque groupe de la classe IA pour la sollicitation de *feedbacks* (Tâche 2). Une présentation générale succincte de l'utilisation de ChatGPT a été fournie par l'enseignant, sans donner d'instructions particulières sur la nature des questions et prompts posés à l'IA. Il est important de préciser que les étudiants n'ont pas été formés à l'utilisation de ChatGPT dans le cadre institutionnel et qu'aucun exemple de prompts ne leur a été fourni. La rédaction des prompts était donc totalement libre et reposait exclusivement sur l'expérience personnelle des étudiants.

La plateforme Moodle a servi à centraliser l'ensemble des consignes et des instructions de l'activité. De plus, les espaces de remise de devoir ont été configurés pour permettre aux groupes d'étudiants, de deux classes différentes, de déposer leurs travaux durant les phases de « prétest » et de « post-test ».

La méthodologie

Afin de mesurer la qualité des travaux réalisés par les groupes d'étudiants de deux classes, une matrice d'évaluation qualitative a été élaborée. La perception des étudiants a été analysée à l'aide d'un questionnaire.

La collecte de données s'est déroulée en trois phases distinctes (voir figure 4.1). La phase 1 correspond au « prétest », durant lequel les groupes des deux classes ont réalisé le même travail de groupe. Au cours de cette phase, l'enseignant a effectué une évaluation des comptes rendus (version 1), mais ces résultats n'ont pas été communiqués aux étudiants. La phase 2 concerne l'intervention où les groupes IA et les groupes témoins ont reçu un *feedback* en fonction de leur classe. La phase 3 correspond au « post-test » durant lequel les groupes des deux classes ont soumis un compte rendu amélioré (version 2) à la suite du *feedback*. À la fin de cette phase, un questionnaire a été distribué à chaque étudiant de chaque groupe de travail, afin de mesurer sa perception de l'utilité du *feedback* reçu. Les deux itérations des comptes rendus ont fait l'objet d'une analyse comparative par l'enseignant.

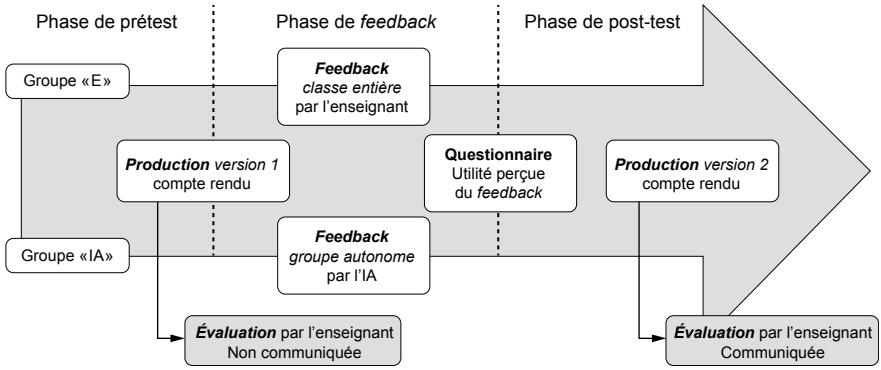


FIGURE 4.1 Phases de la collecte de données.

Matrice d'évaluation

L'évaluation des comptes rendus pour les étapes de prétest et de post-test repose sur une matrice d'évaluation conçue par l'enseignant lui-même. Elle revient sur les trois parties du compte rendu et les savoirs et savoir-faire que l'enseignant souhaite voir maîtrisés par les étudiants (tableau 4.2). La matrice d'évaluation est additionnée de commentaires qui permettent à l'enseignant de bien saisir l'évolution de la qualité des comptes rendus entre le prétest et le post-test.

TABLEAU 4.2 Matrice d'évaluation du compte rendu.

1. Modélisation et mise en forme du problème

	À revoir	À retravailler	Acquis
Discrétisation du problème		✓	
Explication de la démarche			✓

2. Résolution par programmation sur Matlab

	À revoir	À retravailler	Acquis
Prise en compte des conditions initiales et des conditions limites du problème	✓		
Implémentation de la méthode		✓	

3. Analyse et interprétation des résultats			
	À revoir	À retravailler	Acquis
Mise en évidence et analyse de l'instabilité numérique	✓		
Influence des paramètres numériques			✓
Analyse des résultats au regard du problème physique étudié		✓	

Le questionnaire destiné à mesurer la perception des étudiants sur le *feedback* reçu, que ce soit de la part de l'enseignant ou de ChatGPT, se compose de 11 questions fermées utilisant une échelle de Likert (de « Tout à fait d'accord » à « Pas du tout d'accord »), ainsi que de quatre questions ouvertes. Ces dernières ont pour objectif de capturer en détail les aspects du *feedback* qui ont permis d'améliorer le compte rendu dans chaque partie du travail. Tous ces éléments contribuent à un corpus de résultats dont l'analyse est présentée ensuite.

Les résultats et leur analyse

Les résultats obtenus à travers la matrice d'évaluation sont compilés pour établir l'évolution des deux groupes sur les différents aspects du compte rendu. Les résultats du tableau 4.3 montrent les moyennes générales des deux groupes. Les groupes IA (*feedback* des IA) sont passés d'une moyenne de 8,82 au prétest à 10,09 au post-test (+10 %), tandis que les groupes E (*feedback* global de l'enseignant) ont augmenté de 8,18 à 11,09, soit une amélioration plus importante de 20,71 %.

TABLEAU 4.3 Comparaison des moyennes générales.

Parties du compte rendu	Moyenne groupes IA	Moyenne groupes E
Prétest (noté sur 14)	8,82	8,18
Post-test (noté sur 14)	10,09	11,09
Moyenne des différences (%)	+10 %	+20,71 %

Les résultats du tableau 4.4 indiquent l'évolution des performances dans les différentes parties du compte rendu. Les groupes E montrent des améliorations significatives dans les trois parties du travail, notamment dans la partie « Modélisation et mise en forme » (+ 50 %) et « Interprétation et analyse » (+ 45 %) en comparaison avec les groupes IA qui ont progressé avec un pourcentage plus bas dans la partie « Interprétation et analyse » (+ 30 %) et « Modélisation et mise en forme » (+ 18 %).

TABLEAU 4.4 Comparaison des moyennes des évolutions pour chaque partie du compte rendu.

Parties du compte rendu	Moyenne groupes IA	Moyenne groupes E
Modélisation et mise en forme (%)	+ 18 %	+ 50 %
Programmation Matlab (%)	0 %	+ 27 %
Interprétation et analyse (%)	+ 30 %	+ 45 %

La figure 4.2 présente la distribution des amplitudes de l'amélioration entre le prétest et le post-test pour les deux catégories de groupes d'étudiants.

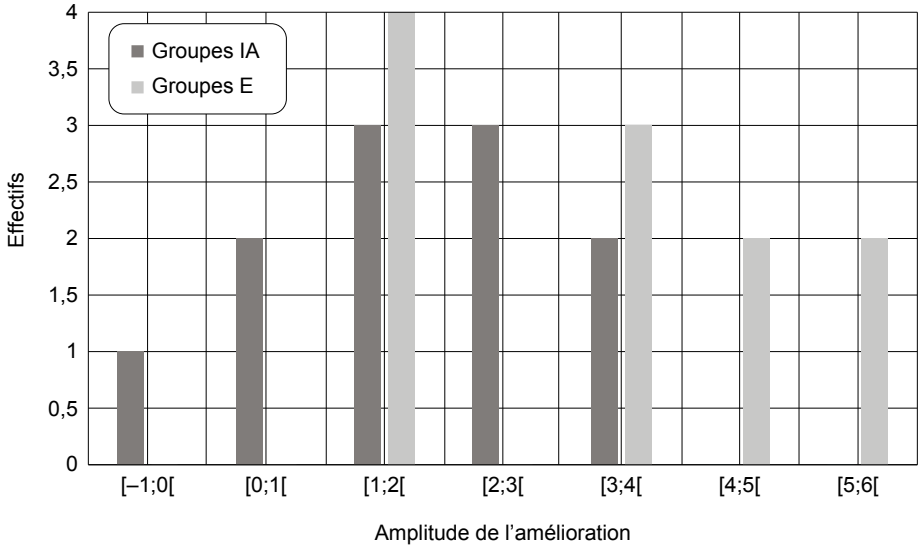


FIGURE 4.2 Distribution des amplitudes d'amélioration de notes de compte rendu.

Pour les deux classes, on calcule la taille d'effet à travers le d de Cohen. On obtient respectivement d_{IA} = 0,45 et d_{E} = 1,65. Cette amplitude montre que l'amélioration entre le prétest et le post-test a été plus significative dans les groupes E qui ont reçu le *feedback* de l'enseignant.

Les résultats du questionnaire montrent des différences dans la perception des étudiants concernant l'efficacité du *feedback* reçu. En général, les groupes E ont des moyennes plus élevées pour la majorité des questions posées par rapport aux groupes IA. Cela indique que le *feedback* fourni par l'enseignant a été perçu comme plus utile pour leur apprentissage (voir figure 4.3).

Les questions Q1 et Q2 portant sur les perspectives d'amélioration du travail préparatoire, notamment en ce qui concerne la discrétisation du problème et la démarche de résolution numérique, ont reçu des réponses plus favorables de la part des groupes E, qui ont bénéficié du *feedback* de l'enseignant. L'écart de réponses entre les deux groupes est d'environ un point.

Également, les réponses aux questions Q3 et Q4, visant à déterminer si le *feedback* a favorisé la réflexion pour élaborer la partie « Programmation sous Matlab » du compte rendu, indiquent que les groupes E ont une perception plus positive du *feedback* fourni par l'enseignant.

Concernant les questions Q5, Q6 et Q7, destinées à évaluer la perception du *feedback* sur la troisième partie du compte rendu, « L'interprétation et l'analyse des résultats », les réponses des groupes E montrent que le *feedback* de l'enseignant leur a permis de mieux réfléchir sur les différentes tâches demandées, par rapport aux *feedbacks* des IA génératives.

Les questions Q8, Q9, Q10 et Q11 ont été destinées à mesurer la perception générale du *feedback* reçu sur la qualité du travail rendu, les axes d'amélioration et la compréhension du sujet après le *feedback*. On observe un écart plus important sur la question 10, concernant le sentiment de confiance des étudiants au regard de la compréhension du sujet, en faveur des étudiants des groupes E.

En ce qui concerne les questions ouvertes, notamment la question Q14 : « Qu'avez-vous appris sur votre propre travail grâce au *feedback* donné par l'enseignant ou les IA génératives ? », les retours soulignent globalement que le *feedback* donné par l'enseignant a permis aux groupes d'étudiants de mieux comprendre les consignes de l'activité et les attendus, et de mieux structurer les trois parties du compte rendu, ainsi que d'améliorer la clarté des explications, notamment sur les démarches, les justifications et les analyses.

Concernant le *feedback* fourni par ChatGPT, les étudiants ont indiqué que ChatGPT les a aidés à mieux rédiger le travail, notamment en améliorant la rédaction des explications et en proposant des corrections pour des erreurs mineures. Certains étudiants ont signalé également avoir appris des concepts qu'ils n'avaient pas abordés en classe, comme des éléments liés à la syntaxe et au codage. Cependant, plusieurs réponses ont signalé que ChatGPT a parfois prodigué des conseils hors sujet dans le cas de consignes complexes, et que les analyses effectuées par ChatGPT ont manqué parfois de précision et de détails, notamment dans l'explication des démarches ou des résultats.

Finalement, les groupes IA ont remarqué que les fichiers complexes comme les PDF ne sont pas entièrement analysés, et les réponses sont plus efficaces avec des prompts simples et concis.

Il est à noter que les étudiants des groupes IA ont répondu plus favorablement à une seule question (Q9). Cette question concerne la perception de l'utilité générale du *feedback* de l'IA en matière de perspectives d'amélioration. Les questions portant sur les perspectives d'amélioration, cette fois contextualisées à la tâche à réaliser (Q1 et Q2), reçoivent des réponses moins favorables des groupes IA. On observe également un large écart à la question sur le sentiment de confiance des étudiants au regard de la compréhension du sujet, en faveur des étudiants des groupes E.

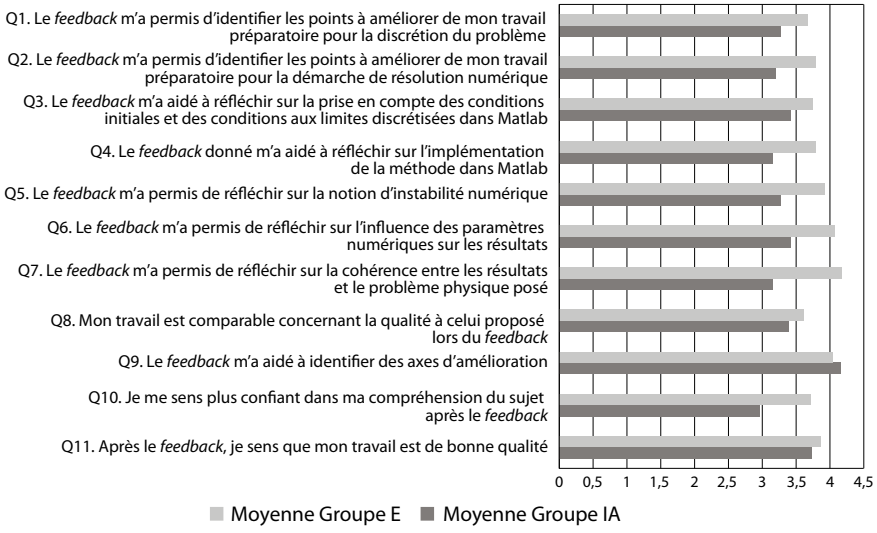


FIGURE 4.3 Comparaison des moyennes des réponses par groupe.

Bien que les interactions entre les étudiants et ChatGPT dans les conversations n'aient pas pu être étudiées en détail, nous avons pu remarquer qu'une majorité des étudiants pose des demandes très simples du type : « fais le TP », « fais-moi un retour » ou « qu'est-ce que je peux améliorer ? », accompagnées de l'ensemble des fichiers. Ces demandes ont généralement tendance à provoquer des réponses vagues et superficielles de l'IA sur les travaux proposés. Une proportion marginale d'étudiants se sont appuyés sur la matrice d'évaluation fournie pour alimenter et structurer les *feedbacks* de l'IA, et dans ce cas, les *feedbacks* générés s'avèrent plus précis et plus pertinents.

Discussion et conclusions

L'objectif de cette étude est de comparer les effets du *feedback* donné soit par l'enseignant, soit par l'intelligence artificielle générative ChatGPT sur l'apprentissage, et de mesurer la perception des étudiants de ce dispositif.

En matière d'apprentissage, on peut conclure de l'analyse des résultats que le *feedback* de l'enseignant a eu un effet plus important sur l'amélioration du travail des groupes E, par rapport à celui fourni par ChatGPT aux groupes IA. Cependant, l'utilisation de ChatGPT a permis également aux groupes d'étudiants d'évoluer positivement de la version 1 du compte rendu à la version 2. Cela signifie que le *feedback* fourni par les IA génératives, dans certains cas, pourrait être utilisé comme stratégie formative et favoriser l'apprentissage des étudiants. De plus, ce *feedback* pourrait être plus immédiat et individualisé que celui donné par l'enseignant (Wang *et al.*, 2024).

En matière de perception, les étudiants qui ont reçu le *feedback* de l'enseignant ont jugé celle-ci plus favorablement que les étudiants qui ont reçu le *feedback* de ChatGPT. Cette différence montre d'une part que le *feedback* donné par l'enseignant dans certaines situations d'apprentissage est plus apprécié, car il est plus adapté aux besoins spécifiques des étudiants en comparaison avec celui fourni par ChatGPT. Cela pourrait indiquer également que les étudiants n'ont pas encore pleinement adopté ces technologies fondées sur l'intelligence artificielle, en raison d'un manque de confiance lié principalement à la fiabilité des données utilisées pour traiter et analyser les requêtes, ainsi qu'aux préoccupations concernant la confidentialité et la sécurité des données (Polyportis et Pahos, 2024). À ce sujet, Chang et Hu (2023) ont ajouté qu'en plus des préoccupations des étudiants concernant la

précision des réponses fournies par ChatGPT, des questions éthiques sont également soulevées. En outre, ils ont indiqué que ces perceptions influencent considérablement les résultats d'apprentissage.

Les premiers éléments d'analyse des prompts utilisés par les étudiants peuvent faire penser que leur utilisation principale de ChatGPT est orientée sur la réalisation des tâches, la correction technique des éléments apportés et la production. Les formes d'utilisation ne semblent pas participer, à l'échelle des classes, à un processus plus général de réflexion métacognitive et à l'élaboration d'une pensée critique sur leurs travaux. C'est une des raisons qui justifierait un accompagnement des étudiants sur l'usage de l'IA dans ce cas, avec la mise à disposition de prompts adaptés et testés en amont par l'enseignant.

Pour ces raisons, il est nécessaire de mener plus d'études empiriques pour confirmer et compléter les résultats de cette expérience dans le cadre du *feedback* formatif fourni par les intelligences artificielles génératives.

Cette étude a donné des éclairages sur les effets du *feedback* des IA génératives, notamment de ChatGPT pour l'apprentissage dans le cadre de l'évaluation formative. Cependant, il serait pertinent d'analyser plus précisément les interactions des étudiants avec les agents conversationnels comme ChatGPT et d'évaluer comment, et dans quelle mesure, ils utilisent les *feedbacks* fournis par l'IA pour améliorer leurs travaux. Il est également important d'élargir la recherche afin de valider et de généraliser ces résultats dans divers contextes éducatifs.

Conseils de mise en œuvre pratique

Lors de cette étude, la formation des étudiants à l'usage efficace de l'IA n'a pas été interrogée. Même si les étudiants utilisent les IA génératives dans le cadre de leurs études, le contexte ici était celui d'un enseignement scientifique de haut niveau, pour lequel les instructions données à l'IA exercent une influence sur les réponses fournies.

Afin de créer une expérience enrichissante qui aide les étudiants à améliorer leurs connaissances et compétences grâce aux *feedbacks* fournis par les IA génératives, nous recommandons de définir clairement les attentes vis-à-vis des étudiants concernant leur interaction avec ce type d'outils et le *feedback* qu'ils sont capables de fournir pendant l'activité souhaitée. Cela signifie que l'enseignant doit évaluer les possibilités et les limites de ces outils avant de les utiliser en classe, afin de comprendre

le type de réponses que les IA génératives peuvent engendrer lors d'une activité précise, ainsi que les aspects qui doivent être pris en compte, car ces réponses peuvent être parfois imprécises. L'enseignant doit donc être vigilant afin de guider efficacement les étudiants.

Il semble également important d'expliquer comment les utiliser, en fournissant des exemples concrets de prompts, pour obtenir des retours plus précis sur leurs travaux, car les caractéristiques des *feedbacks* peuvent être très variées en fonction des demandes et des prompts utilisés. Par ailleurs, il pourrait être intéressant que l'enseignant configure et entraîne un modèle d'agent conversationnel spécifiquement dans le but de générer des *feedbacks*, de façon générale ou même pour cette discipline et ces travaux en particulier. Cela permettrait d'espérer une plus grande régularité de la qualité du *feedback* fourni aux étudiants au regard de la tâche à réaliser.

Concevoir une activité où les étudiants ont la possibilité de refaire leur travail en au moins deux versions remplirait les conditions du *feedback* dialogique avec les IA génératives. Il est nécessaire, dans le cadre d'une activité similaire, d'encourager une analyse critique des réponses fournies par les IA, et de promouvoir les capacités d'autoréflexion et d'autoévaluation chez les étudiants (Lee *et al.*, 2025). À la fin de l'expérience, il serait souhaitable d'organiser une session de retour d'expérience où les étudiants pourront partager ce qu'ils ont appris de l'utilisation des IA génératives dans ce contexte, ainsi que les difficultés rencontrées.

Dans le cadre de notre étude, les accès à la version payante de ChatGPT ont été fournis aux étudiants. Pour mettre en œuvre un dispositif reposant sur le *feedback* généré par l'IA dans un autre contexte, il sera important d'être attentif aux problématiques de disparités d'accès à ces outils afin de garantir l'équité entre les étudiants.

Pour que ce processus de *feedback* avec les IA génératives soit durable, il est crucial de prendre en compte ces considérations de mise en œuvre lors d'autres activités. Cela permettra aux étudiants de s'approprier ces pratiques et de les intégrer de façon autonome et efficace dans leur apprentissage.

En conclusion, l'utilisation des IA pour la génération de *feedbacks* implique une vigilance sur l'égalité d'accès à ces outils, la formation des étudiants à leur usage et un accompagnement des enseignants sur la rédaction de prompts. La question de la sécurité des données et de l'éthique interroge l'utilisation d'outils commerciaux internationaux au profit d'outils institutionnels spécifiques.

Partie 2

Évaluer à l'échelle de l'institution

Infrastructures, stratégies
et transitions numériques

Hiba Asri & Abdelali Rochdi

Intégration des outils numériques dans l'évaluation des apprentissages
Étude de cas: Université Cadi Ayyad – Maroc

5

Résumé

L'intégration des technologies numériques dans l'évaluation des apprentissages des étudiants a profondément transformé les pratiques pédagogiques, en apportant des solutions innovantes visant à améliorer la qualité et l'efficacité des évaluations. À l'Université Cadi Ayyad (UCA), l'augmentation rapide du nombre d'étudiants, notamment dans les établissements à accès ouvert, a rendu la gestion des examens plus complexe et le processus traditionnel d'évaluation de plus en plus laborieux. Face à ces défis, l'UCA a adopté progressivement deux méthodes d'évaluation : l'évaluation sur ordinateur via la plateforme Moodle et l'évaluation papier à l'aide de l'outil *Auto Multiple Choice* (AMC). Elles ont permis de créer des évaluations satisfaisantes utilisant plusieurs types de questions, y compris la possibilité d'introduire des exercices ou des problèmes à résoudre pour une évaluation plus approfondie des connaissances et du savoir-faire de l'étudiant. L'objectif est d'accélérer le processus d'évaluation et de remise des notes grâce à la correction automatique des rendus, d'alléger la charge de travail des enseignants tout en maintenant la qualité et le niveau des évaluations. Ces deux méthodes sont appréciées par la plupart des enseignants et des responsables de l'Université, favorisant ainsi leur adoption au sein de l'UCA. L'automatisation et l'objectivité offertes par

ces outils ont non seulement permis de répondre aux exigences croissantes de l'enseignement supérieur dans un contexte de massification des effectifs des étudiants, mais également d'assurer une évaluation plus équitable, transparente et rassurante pour nos étudiants en éliminant les erreurs liées à la correction manuelle des copies d'examen.

Introduction

L'intégration des outils numériques a révolutionné les pratiques pédagogiques, offrant des solutions innovantes pour améliorer la qualité de l'évaluation des apprentissages des étudiants. L'utilisation de ces outils permet de faciliter la gestion des examens et d'améliorer l'expérience globale d'apprentissage à travers les divers types d'évaluation : diagnostique, formative et sommative. L'usage du numérique dans l'évaluation a également permis d'assurer une grande rapidité et une meilleure fiabilité dans la correction des rendus, en proposant des moyens appréciés par tous pour mesurer les connaissances et les compétences des étudiants. Selon un rapport de l'UNESCO (Organisation des Nations unies pour l'éducation, la science et la culture), l'adoption du numérique dans l'évaluation contribue à une évaluation plus équitable et transparente, permettant un *feedback* immédiat aux étudiants et une analyse plus claire des données de performance. De plus, l'automatisation des corrections réduit significativement la charge de travail des enseignants et minimise les erreurs humaines (UNESCO, 2020).

L'Université Cadi Ayyad (UCA) fait face à un défi majeur qui impacte la qualité des enseignements et la gestion des évaluations, à savoir la massification des étudiants[15], en particulier dans les établissements à accès ouvert[16]. Avec des cohortes d'environ 60 000 étudiants en première année (toutes filières confondues), la charge de travail demandé aux professeurs est fastidieuse. Ils doivent assurer les cours magistraux, les travaux dirigés, les travaux pratiques durant tout le semestre, puis évaluer juste après les apprentissages de leurs étudiants dans un délai de quelques jours.

Depuis plusieurs années, l'UCA a commencé progressivement à utiliser le numérique dans les pratiques pédagogiques. Avec l'introduction de la plateforme e-learning Moodle, les certificats Internet et

[15] Nombre d'étudiants à l'UCA en 2024 : environ 120 000.
[16] Accès ouvert : établissement universitaire ouvert à tous les bacheliers sans sélection.

Informatique C2i niveau 1 et C2i niveau 2 « Métiers de l'enseignant » en 2010, puis le certificat TIC-Maroc[17] en 2014, l'UCA a fait un grand pas vers l'usage du numérique dans les pratiques pédagogiques. Le nombre d'établissements, de filières, de modules qui passent au numérique ne cesse d'augmenter d'année en année, et c'est en 2019, avec l'arrivée de la pandémie de Covid-19, que tous les établissements, toutes les filières et tous les modules ont adopté l'utilisation de la plateforme Moodle comme outil principal d'enseignement en ligne.

La mise en place de la plateforme eCampus[18] par l'Université a permis d'atténuer en partie le problème de massification sur le volet enseignement. Cependant, le volet évaluation présente encore des défis à relever pour les professeurs comme :

- la résistance à l'usage des QCM (questionnaires à choix multiples) dans l'évaluation des apprentissages ;
- la difficulté à produire assez de questions pour évaluer de grandes cohortes ;
- le manque de maîtrise des outils numériques pour certains enseignants.

Le Centre d'innovation pédagogique de l'UCA (CIP)[19], en collaboration avec le pôle digitalisation ont contribué à surmonter ces défis en mettant en place les outils numériques et les formations nécessaires pour une évaluation simple et efficace des apprentissages des étudiants. Deux types d'évaluation fondée sur l'usage du numérique ont été utilisés :

- évaluation numérique avec la plateforme eCampus Moodle[20], en exploitant tous les types de questions fournis par la plateforme dont principalement les QCM/QCU (question à choix unique) pour une évaluation rapide et objective des connaissances des étudiants d'une part et, d'autre part, les questions « Cloze »[21] permettant d'intégrer des exercices et des problèmes à résoudre, favorisant ainsi une évaluation plus approfondie, qui correspond aux niveaux « appliquer » et « analyser » de la taxonomie de Bloom ;

17 Produit dans le cadre du projet Tempus « P@lmes » entre 2013 et 2015.
18 La plateforme eCampus regroupe toutes les plateformes e-learning (Moodle 4.2) des établissements de l'UCA
19 https://cip.uca.ma/ (consulté le 14.01.2026).
20 https://moodle.org/ (consulté le 14.01.2026).
21 Voir :https://docs.moodle.org/4x/fr/Question_cloze_%C3%A0_r%C3%A9ponses_int%C3%A9gr%C3%A9es (consulté le 14.01.2026).

- évaluation papier avec l'outil AMC[22] (*Auto Multiple Choice*) qui permet, d'une part, d'utiliser des questions de type QCM/QCU pour une évaluation efficace des connaissances des étudiants et, d'autre part, de regrouper des questions pour élaborer un exercice ou un problème à résoudre.

Ce chapitre commence par une présentation de l'état de l'art relatif à la formation et à l'évaluation des apprentissages des étudiants dans un contexte de massification. Nous aborderons ensuite les spécificités des deux méthodes adoptées par l'UCA : l'évaluation numérique via Moodle et l'évaluation papier à l'aide de l'outil AMC. Nous procéderons par la suite à une analyse de la mise en œuvre de ces méthodes, en examinant les aspects des infrastructures matérielles, logicielles et réseau, ainsi que la formation et l'accompagnement des enseignants. Enfin, nous présenterons les résultats obtenus et l'impact de cette expérimentation sur les étudiants, les enseignants, et les établissements de l'UCA.

Contexte et problématique

L'évaluation des apprentissages dans l'enseignement supérieur constitue un défi majeur pour les équipes pédagogiques en raison de sa complexité. Il est important de garantir un alignement rigoureux entre les objectifs, le contenu et les méthodes d'évaluation. De plus, il faut gérer efficacement la charge de travail et la concentration nécessaires pour évaluer rapidement et de manière équitable les copies des étudiants. Ces contraintes deviennent encore plus importantes pour les grandes cohortes d'étudiants.

L'usage du numérique dans les pratiques pédagogiques, et particulièrement dans l'évaluation des apprentissages, est devenu une obligation pour tous les professeurs de l'enseignement supérieur. Plusieurs études ont été conduites afin d'examiner l'impact de ces outils sur les processus d'évaluation.

L'étude menée par l'Institut Symbiosis de management des opérations (Symbiosis Institute of Operations Management, IIC) souligne les atouts de Moodle en matière de flexibilité, de personnalisation et de sécurité pour l'enseignement, l'apprentissage, et l'évaluation (Barge et Londhe, 2014). Cette étude, fondée sur des entretiens avec des

[22] https://www.auto-multiple-choice.net/index.fr (consulté le 14.01.2026).

enseignants et des étudiants, révèle que l'utilisation de Moodle a permis une réduction significative du temps consacré aux évaluations avec un taux de satisfaction de 78 % à 86 % parmi les utilisateurs. Une autre étude (Maqableh et Alia, 2021) a montré qu'une analyse comparative a été menée entre des groupes d'étudiants utilisant Moodle et d'autres ne recourant pas aux outils numériques. Les résultats de cette étude révèlent que les étudiants ayant accès à Moodle obtiennent des performances académiques supérieures et expriment une satisfaction accrue vis-à-vis de leur processus d'apprentissage. Cette amélioration est attribuée à la participation à diverses activités interactives proposées sur Moodle, telles que les quiz, les forums et les ateliers.

L'amélioration de l'apprentissage des étudiants en troisième cycle grâce à l'intégration de Moodle a été mise en évidence par une comparaison de son utilisation dans les cours liés aux technologies de l'information et de la communication (TIC) avec d'autres disciplines. Une analyse portant sur 104 cours, complétée par un questionnaire administré à 393 étudiants, révèle que ces derniers ont exprimé une grande satisfaction sur le rôle de Moodle dans l'amélioration de leur apprentissage (Costa *et al.*, 2012).

En comparaison avec d'autres plateformes de *Learning Management System* (LMS), et en évaluant spécifiquement l'utilité des quiz Moodle pour l'évaluation formative, il apparaît que ces quiz offrent une grande variété d'options et de fonctionnalités de *feedback*. Cependant, ils présentent des limites, notamment l'absence de révision par les pairs et un manque d'interactions directes entre étudiants et tuteurs (Fernando, 2020).

À l'Université d'Aveiro au Portugal et à l'Université Politehnica de Bucarest (Gamage *et al.*, 2022 ; Nyagope, 2024), Moodle est principalement utilisé comme dépôt de matériel pédagogique. Une analyse de contenu et un questionnaire auprès de 278 étudiants d'Aveiro et de 52 étudiants de Politehnica ont révélé une insatisfaction générale concernant cette utilisation limitée de Moodle. Malgré cette insatisfaction, les étudiants reconnaissent l'importance d'exploiter plus de fonctionnalités afin de renforcer l'enseignement et l'apprentissage.

Les défis liés à la massification dans l'enseignement supérieur, tels que le taux d'encadrement trop faible, la surcharge liée à la correction des copies d'examen, et la diminution de la qualité des évaluations, ont été largement relevés dans plusieurs études (Maringe et Sing, 2014 ; Olmos *et al.*, 2015). Les chercheurs proposent l'intégration de nouvelles

technologies, ainsi que la formation continue des enseignants, comme solution pouvant améliorer de manière significative la qualité des évaluations et faciliter les retours pédagogiques. Ils recommandent aussi d'impliquer les étudiants dans le processus d'évaluation à travers des activités de révision par les pairs.

De nombreuses études (Hornsby et Osman, 2014 ; Varghese, 2015 ; Nyagope, 2024) dans la littérature ont exploré les défis liés à l'évaluation formative et sommative dans les classes de grande taille. Elles proposent l'adoption d'outils tels que les plateformes d'apprentissage en ligne, les logiciels de vidéoconférence et les forums de discussion, qui offrent des avantages significatifs en matière d'accessibilité et de flexibilité. Cependant, ces solutions sont également confrontées à des défis, notamment des problèmes techniques, un manque de compétences numériques chez certains enseignants et étudiants, ainsi que des difficultés à maintenir l'engagement des étudiants. En Afrique du Sud, la massification de l'enseignement a conduit à des classes surpeuplées, ce qui a considérablement réduit l'interaction entre enseignants et étudiants et compromis la qualité de l'évaluation. Les études soulignent que l'intégration des plateformes numériques apparaît comme une solution pertinente face à des problèmes tels la massification et le manque de ressources humaines (Maringe et Sing, 2014 ; Nyagope, 2024). Cependant, ces outils ne permettent pas de répondre à toutes les situations, notamment celles impliquant une évaluation complexe des compétences.

Le tableau 5.1 présente un récapitulatif comparant les différentes études citées auparavant, en déterminant leurs objectifs et leurs résultats pertinents.

Les travaux recensés dans la littérature confirment que l'UCA a adopté une approche cohérente et efficace pour répondre aux défis posés par une évaluation rapide, efficace et équitable dans un contexte de massification. Cette approche repose sur l'intégration avancée des technologies numériques, notamment la plateforme Moodle et l'application AMC. Grâce à ces deux outils, il est possible de mettre en place des examens incluant des QCM et des exercices tout en respectant l'alignement pédagogique, et particulièrement les niveaux cognitifs visés (comprendre, appliquer, voire analyser). Pour cela, les professeurs sont amenés à utiliser des tables de spécifications pour s'assurer que les questions couvrent tous les points à évaluer du programme, ce qui permet une évaluation systématique des apprentissages des étudiants.

TABLEAU 5.1 État de l'art : comparatif des études.

Référence	Objectifs	Résultats
Barge et Londhe, 2014	Étudier l'utilisation de Moodle à l'Institut Symbiosis de management des opérations	• Satisfaction élevée (86 % des professeurs, 78 % des étudiants) • Amélioration de l'engagement
Fernando, 2020	Évaluer l'utilité des quiz Moodle pour l'évaluation formative	• Grande variété d'options et de *feedback* • Limitations dans la révision par les pairs
Olmos *et al.*, 2015	Comparer l'apprentissage dans les cours liés aux TIC et autres	• Cours TIC plus interactifs et efficaces • Satisfaction élevée des étudiants
Costa *et al.*, 2012	Analyser l'utilisation de Moodle comme dépôt de matériel à l'Université d'Aveiro	• Insatisfaction générale • Importance d'exploiter d'autres fonctionnalités
Gamage *et al.*, 2022	Analyser l'utilisation de Moodle comme dépôt de matériel à l'Université Politehnica de Bucarest	• Insatisfaction générale • Importance d'exploiter d'autres fonctionnalités
Maqableh et Alia, 2021	Améliorer les résultats d'apprentissage avec Moodle dans l'enseignement supérieur	• Meilleurs résultats académiques • Grande satisfaction des étudiants
Lynch et Pappas, 2017	Analyser les défis de l'évaluation dans l'enseignement supérieur	• Défis majeurs dans l'évaluation • Importance des technologies et formation continue
Maringe et Sing, 2014	Examiner les défis et opportunités des outils numériques pour l'évaluation des grandes classes	• Accessibilité et flexibilité accrues • Défis techniques et d'engagement
Gilles, 2016	Analyser les défis de l'évaluation dans l'enseignement supérieur	• Défis majeurs dans l'évaluation • Importance des technologies et formation continue
Nyagope, 2024	Étudier les défis de la massification en Afrique du Sud et les solutions numériques	• La massification entraîne des classes surpeuplées • Adoption des plateformes numériques recommandée

Le tableau 5.2 illustre une table de spécification avec trois niveaux cognitifs : mémoriser, comprendre et appliquer. Chaque ligne correspond à un chapitre ou une section, tandis que chaque colonne représente un type de question pour un niveau cognitif spécifique. Les valeurs dans le tableau indiquent le nombre de questions « n11, n12… », à générer par chapitre, par niveau cognitif et par type de question. La répartition de ces nombres de questions dépend à la fois de l'importance relative de chaque chapitre dans le cours et de la durée totale de l'épreuve. Elle tient également compte du temps estimé (t) nécessaire

pour répondre à chaque type de question. L'enseignant doit ajuster cette répartition de manière à ce que la somme des produits du nombre de questions par leur temps de réponse respectif corresponde à la durée prévue de l'épreuve.

TABLEAU 5.2 Exemple d'une table de spécification.

	Niveau cognitif 1 : mémoriser				Niveau cognitif 2 : comprendre				Niveau cognitif 3 : appliquer				Nombre total de questions
	QCM	QCU	Correspondance	V/F	QCM	QCU	Correspondance	V/F	QCM	QCU	Correspondance	V/F	
Chap. 1	n11[23]	n12	n13	n14	n21	n22	n23	n24	n31	n32	n33	n34	N[24]
Chap. 2	m11	m12	m13	m14	m21	m22	m23	m24	m31	m32	m33	m34	M
...													
Chap. 10	x11	x12	x13	x14	x21	x22	x23	x24	x31	x32	x33	x34	X
Nombre total de questions	z11[25]	z12	z13	z14	z21	z22	z23	z24	z31	z32	z33	z34	Z[26]

Cadre théorique

L'UCA cherche à créer un environnement d'évaluation adapté aux besoins de tous, qui permet d'alléger la charge de travail des enseignants sur le plan pédagogique pour qu'ils puissent continuer à développer leurs travaux de recherche et à garantir des évaluations équitables et transparentes pour tous les étudiants.

L'évaluation des apprentissages des étudiants représente une composante fondamentale du processus éducatif, permettant de mesurer la compréhension des étudiants et leur maîtrise des acquis. À l'UCA, nous faisons face à des défis majeurs en matière d'évaluation en raison de plusieurs facteurs :

1. le premier défi est la massification. Tous les étudiants marocains ayant un baccalauréat ont le droit de suivre leurs études à l'université. L'inscription dans les établissements à accès ouvert est garantie par la loi. De ce fait, le nombre d'étudiants inscrits en première

[23] Nombre total de questions à choix multiples relatives au chapitre 1.
[24] Nombre total de toutes les catégories de questions relatives au chapitre 1.
[25] Nombre total de questions à choix multiples relatives à tous les chapitres (niveau cognitif spécifique).
[26] Nombre total de questions.

et en deuxième année au sein des établissements à accès ouvert[27] (facultés) est considérablement élevé (voir tableau 5.3) et le nombre de copies à corriger par professeur varie de 150 à 300. Cette tâche devient presque irréalisable pour les épreuves qui utilisent exclusivement des questions ouvertes ;
2. le planning des examens est fortement contraint et les enseignants ont un délai de cinq jours pour effectuer la correction, la saisie des notes, la délibération et l'affichage des résultats ;
3. la lutte contre la fraude impose de limiter l'occupation des salles à un maximum de 40 % des places assises, ce qui entraîne une augmentation significative du nombre de salles utilisées et du personnel de surveillance requis ;
4. l'absentéisme des étudiants aux examens atteint 30 %, mobilisant inutilement de nombreuses salles, des ressources humaines et matérielles.

TABLEAU 5.3 Vue d'ensemble de l'effectif des étudiants à l'UCA (du semestre 1 au semestre 4 de l'année universitaire 2024-2025).

Institution	Semestre 1	Semestre 2	Semestre 3	Semestre 4
Faculté des sciences juridiques, économiques et sociales, Marrakech (FSJES)	28 951	28 062	11 898	11 761
Faculté des lettres et sciences humaines, Marrakech (FLSH)	11 533	11 460	5781	5592
Faculté des sciences Semlalia, Marrakech (FSSM)	5778	5195	3050	2799
Faculté des sciences juridiques, économiques et sociales, Kelâa des Sraghna (FSJES)	4031	3995	1923	1839
Faculté polydisciplinaire, Safi (FPS)	12 824	12 257	3112	3092
Faculté de langue arabe, Marrakech (FLAM)	1575	1540	914	902

[27] Tout bachelier ayant obtenu 10 ou plus dans la moyenne générale du baccalauréat a le droit d'accéder aux facultés.

Pour répondre à ces défis, l'UCA a mis en place les moyens qu'il faut pour évaluer les étudiants dans de meilleures conditions tout en veillant à l'égalité des chances. Deux dispositifs d'évaluation ont été adoptés, l'évaluation numérique via Moodle et l'évaluation papier avec l'application AMC (voir figure 5.1).

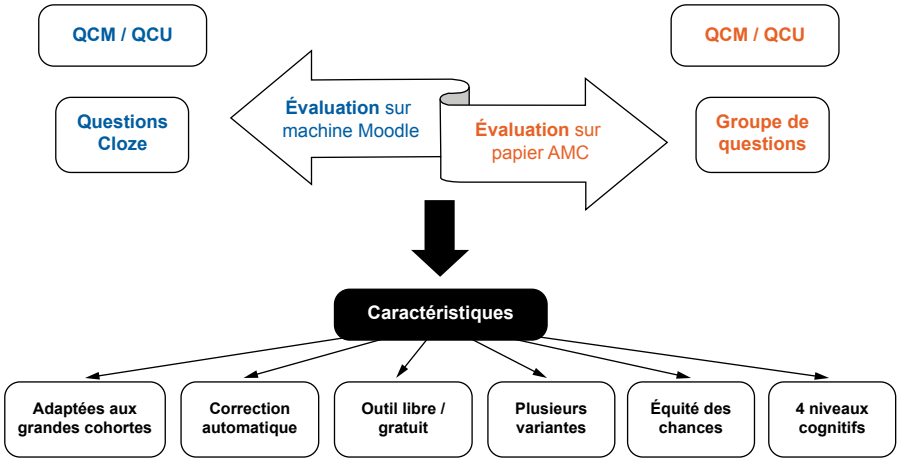


FIGURE 5.1 Approches d'évaluation à l'UCA.

Dispositif et scénario pédagogique

Dispositif 1 : Évaluation numérique via la plateforme Moodle

L'évaluation des apprentissages des étudiants sur Moodle peut s'effectuer à travers une variété d'activités d'évaluation, telles que les tests, les devoirs, ou les ateliers. Ces activités permettent de réaliser des évaluations formatives tout au long du semestre, qui aident les étudiants à apprendre de leurs erreurs grâce aux *feedback*s intégrés dans les QCM ou à une évaluation par les pairs dans les ateliers. Ces outils sont également utilisés pour des évaluations sommatives, notamment dans le cadre des contrôles continus ou du contrôle de mi-semestre, où l'équipe pédagogique dispose généralement d'un à deux mois pour corriger les copies et communiquer les notes. En revanche, pour les contrôles de fin de semestre ou les examens de rattrapage, l'évaluation repose principalement sur l'activité « test » de Moodle, en raison de sa possibilité de correction automatique et de sa rapidité.

Ce premier dispositif présente les caractéristiques suivantes :

- correction automatique : les copies des étudiants sont corrigées automatiquement et en temps réel, ce qui permet d'afficher les notes et les résultats le jour même de l'examen. De plus, la correction automatique des copies élimine les erreurs humaines, augmentant ainsi l'équité des chances pour tous les étudiants ;
- diversité des formats de questions : Moodle propose plusieurs types de questions. Le professeur choisit le nombre et le type de questions à utiliser en fonction de sa table de spécification (Alade et Omoruyi, 2014). Parmi les types de questions les plus couramment utilisés, on trouve :
  - QCM/QCU, Associations, Vrai/Faux : ces types de questions sont utilisés pour évaluer les connaissances des étudiants au niveau de la compréhension et de la mémorisation,
  - réponses courtes, calculées : idéales pour de petits exercices d'application, elles ciblent les niveaux de la compréhension et d'application,
  - questions Cloze : utilisées pour créer des exercices d'application ou des problèmes à résoudre, visant les niveaux d'application et d'analyse ;
- taille de la banque de questions : les professeurs ont la possibilité de créer une vaste banque de questions en organisant celles-ci par catégories selon les séquences de cours, le niveau cognitif, et le type de question. Cette approche permet de générer plusieurs variantes d'un même examen, contribuant ainsi à réduire les risques de tricherie ;
- accessibilité et sécurité : tous les examens à l'UCA se déroulent en mode présentiel. Étant donné que la plateforme Moodle est accessible via un navigateur, un dispositif de sécurité a été instauré pour éviter toute tentative de fraude. En effet, l'UCA a mis en place des plateformes Moodle dédiées exclusivement aux évaluations, accessibles uniquement depuis les salles d'examen par un contrôle d'adresses IP ;
- simplification logistique : il n'y a quasiment plus d'impression papier et les ordinateurs ne sont allumés que lorsqu'un étudiant est effectivement présent et passe son examen ;
- ressources humaines et charge horaire : la taille de la banque de questions et le paramétrage des « tests » sur Moodle permettent de créer une multitude de variantes dans une même salle d'examen, ce qui réduit le nombre de surveillants de plus de 50 % par rapport aux examens antérieurs ;

- étudiants avec besoins spécifiques : une attention particulière est accordée aux étudiants à besoins spécifiques, soit en utilisant la fonctionnalité d'accessibilité sur Moodle, soit par un autre biais (papier, entretien oral, etc.) selon les besoins exprimés par les étudiants en amont de l'examen ;
- suivi et analyse des performances : Moodle permet de suivre les performances des étudiants en temps réel et de générer des rapports détaillés, facilitant l'analyse des résultats et des statistiques.

Dispositif 2 : Évaluation papier avec l'outil AMC

Depuis la pandémie de Covid-19, l'utilisation de QCM papier avec une correction automatique des copies scannées des étudiants est devenue courante. À l'UCA, nous avons opté pour une solution libre et gratuite, à savoir l'outil *Auto Multiple Choice* (AMC). Cet outil nous a également permis de relever les défis identifiés pour les établissements ne disposant pas d'assez d'ordinateurs pour des évaluations numériques. De plus, il a permis à des enseignants ayant des lacunes en compétences TICE (technologies de l'information et de la communication pour l'enseignement) d'évaluer facilement leurs étudiants.

Le choix de l'outil AMC dans les évaluations de fin de semestre a été justifié par :

- outil libre et gratuit : comme l'UCA comporte 15 établissements avec 120 000 étudiants, le choix d'un outil libre et gratuit est une nécessité absolue ;
- correction automatique : comme tous les outils QCM papier, AMC permet une correction rapide et précise des copies d'examen[28]. Chaque établissement dispose de plusieurs scanners professionnels permettant ainsi de réduire considérablement le temps de correction et la charge de travail des enseignants ;
- types de questions : l'outil AMC en propose trois. Le professeur choisit le nombre et le type de questions à utiliser en fonction de sa table de spécification[29] :

[28] En première année, à la FLSH et à la FSJES, le nombre de copies d'une épreuve peut atteindre 12 000.

[29] La table de spécification aligne les objectifs d'apprentissage avec le contenu du cours, les types de questions et leur poids, garantissant ainsi une évaluation cohérente et équitable.

- QCM/QCU : ces types de questions sont choisis pour une évaluation de niveau cognitif : connaître ou comprendre,
- question ouverte : utilisée par les professeurs ayant des cohortes raisonnables en troisième année ou en master pour une évaluation de niveau cognitif « Appliquer » ou « Analyser »,
- groupe de questions : les professeurs avec de grandes cohortes utilisent les groupes de questions pour créer des exercices ou des problèmes et ainsi préparer des épreuves avec des niveaux cognitifs « Appliquer » ou « Analyser » ;
- variantes multiples : AMC permet de générer plusieurs variantes pour une même banque de questions, ce qui réduit considérablement les risques de fraudes en proposant l'évaluation la plus équitable possible pour tous les étudiants ;
- consultation des copies et fiabilité des résultats : l'outil AMC facilite l'annotation des copies scannées des étudiants et la génération d'un dossier contenant toutes les copies annotées au format PDF. Cette fonctionnalité permet aux professeurs de gérer aisément la phase de consultation des copies après l'affichage des résultats, renforçant ainsi la transparence et la fiabilité vis-à-vis des étudiants.

En combinant l'évaluation numérique via Moodle et l'évaluation papier à l'aide d'AMC, nous avons pu relever les défis posés par la massification des étudiants et les exigences administratives liées à la gestion des calendriers d'examens et à la qualité de l'évaluation.

Certes, les deux dispositifs présentent beaucoup d'avantages, mais aussi quelques inconvénients. L'évaluation numérique via Moodle peut être interrompue par des problèmes techniques comme une coupure internet ou d'électricité et devient lourde avec la gestion du grand nombre d'étudiants. L'évaluation papier avec AMC ne résout pas l'ensemble des défis auxquels nos établissements doivent faire face.

Le tableau 5.4 présente une comparaison des contributions des deux méthodes à la résolution des défis identifiés.

TABLEAU 5.4 Comparatif des méthodes proposées.

Défi identifié	Solution	Moodle	AMC
Massification	Correction automatique	✓	✓
Alignement pédagogique	Table de spécification et types de questions	✓	✓
Équité et lutte contre la fraude	Large banque de questions Plusieurs variantes d'un même examen	✓ ✓	✓ ✓
Accessibilité et sécurité	Serveurs dédiés Contrôle d'accès IP Contrôle d'identité	✓ ✓ ✓	✓ ✓
Logistique d'impression papier	Outils numériques	✓	
Étudiants à besoin spécifique	Accessibilité sur Moodle Épreuves adaptables aux étudiants	✓ ✓	 ✓
Consultation des copies	Copies corrigées et annotées automatiquement	✓	✓
Coût des licences	Outil libre et gratuit	✓	✓

Méthodologie

L'UCA regroupe 15 établissements universitaires, comprenant des facultés, des écoles et des instituts. Parmi ces établissements, les sept à accès ouvert enregistrent une augmentation continue du nombre d'étudiants chaque année. Cette dynamique de massification engendre des défis communs à tous ces établissements, notamment le respect des délais fixés par les calendriers d'examens, l'alignement pédagogique, la garantie de la qualité et de l'équité des évaluations, ainsi qu'une logistique complexe.

Depuis 2014, l'UCA a investi dans l'infrastructure matérielle et logicielle pour faciliter l'usage du numérique dans les pratiques pédagogiques des enseignants. Cette infrastructure a permis à quelques établissements de l'UCA comme la faculté des sciences Semlalia (FSSM) de se lancer dans la production de cours en ligne sur la plateforme Moodle. L'évaluation numérique à la FSSM a débuté avec une capacité de 200 ordinateurs. L'intérêt qu'a reçu ce mode d'évaluation a encouragé l'UCA à acquérir un *datacenter* (centre de données)[30] et plus d'ordinateurs pour permettre à tous les établissements à accès ouvert d'avoir un minimum de 500 postes pour gérer les évaluations numériques.

[30] Nutanix, 5 nœuds, 80 To de stockage et 3 To de RAM.

L'acquisition du *datacenter* et l'implémentation de la virtualisation des serveurs nous ont permis non seulement de mettre en ligne une plateforme Moodle d'enseignement pour chaque établissement, mais aussi de déployer des plateformes Moodle dédiées exclusivement aux évaluations[31] pour certains établissements.

Concernant l'outil AMC, nous avons opté pour la création de comptes « Bureau à distance Linux » au niveau du *datacenter*, ce qui évite l'installation individuelle par les professeurs d'un tel environnement. Chacun d'entre eux dispose ainsi d'un espace privé sécurisé pour mener à bien ses évaluations, tout en bénéficiant des performances de stockage et de RAM (*Random Access-memory*) du *datacenter*. Cela rend la correction des copies scannées plus fluide et plus rapide.

Le CIP de l'UCA a mis en place un ensemble de formations adaptées aux besoins des enseignants, organisées en présentiel ou en mode hybride dans l'ensemble des établissements. Ces formations visent à familiariser les enseignants avec les technologies éducatives, notamment celles liées à l'évaluation. Elles abordent l'utilisation de la plateforme Moodle pour la gestion des cours et l'évaluation des apprentissages des étudiants, ainsi que l'outil AMC pour la création et la correction des examens QCM au format papier. Le CIP assure également un accompagnement personnalisé aux professeurs de l'UCA, particulièrement pendant les périodes d'examens de fin de semestre.

Dans ce qui suit, nous nous concentrerons sur la FSSM et nous analyserons également l'état d'avancement de l'utilisation des outils numériques pour l'évaluation des apprentissages dans les autres établissements de l'UCA.

La faculté des sciences Semlalia (FSSM) accueille 12 000 étudiants répartis sur plus de 25 filières académiques couvrant des domaines variés tels que l'informatique, les mathématiques, la biologie, la physique, la géologie, et la chimie : c'est un établissement expérimental de l'UCA, dans lequel sont mises en œuvre de nouvelles pratiques éducatives.

L'évaluation des apprentissages des étudiants se déroule exclusivement en présentiel, en utilisant divers types d'épreuves : des questions ouvertes, des évaluations numériques via Moodle, des évaluations papier avec AMC ou sous forme de travaux à réaliser tout au long du module (devoirs, mini-projets, études de cas…). Le choix du type d'évaluation est influencé par plusieurs facteurs, tels que le niveau cognitif

[31] https://exam.uca.ma

visé, le nombre d'étudiants inscrits, ainsi que les compétences TICE des membres de l'équipe pédagogique.

Les informations clés sur les effectifs étudiants par semestre ainsi que sur les ressources humaines et matérielles disponibles, présentées dans le tableau 5.5 et explicitées ensuite, offrent un aperçu de l'environnement de travail et des défis rencontrés dans la gestion des examens.

TABLEAU 5.5 Vue d'ensemble de l'expérience Moodle et AMC à la FSSM.

Nombres d'étudiants par semestre (toutes filières confondues)					
S1 : 5778	S2 : 5195	S3 : 3050	S4 : 2799	S5 : 1983	S6 : 1896

Évaluation numérique (Moodle)			Évaluation papier (AMC)		
Salle numérique (SN)	Ordinateurs	Surveillants*	Locaux	Effectif[32]	Surveillants*
SN1	200	5	Amphithéâtres	540	30
SN2	180	4	Salles TD	1772	100
SN3	80	2	Salles TP	240	13
SN4	80	2	Bibliothèque	150	8
SN5	80	2	Autre	130	7
Total	620	15	Total	2830	158

* Nombre de professeurs et d'administratifs.

Évaluation numérique : plateforme Moodle

La mise en œuvre de cette pratique a été limitée par le ratio entre le nombre d'ordinateurs disponibles (environ 600) et le nombre d'étudiants inscrits par module. Cette contrainte nous a obligés à programmer une série d'épreuves successives tout au long de la journée pour les grandes cohortes d'étudiants.

La répartition des étudiants par groupe et par salle numérique nous permet de mieux gérer les effectifs. La présence sur place est indispensable pour tous les étudiants, car l'accès aux épreuves (tests sur

[32] L'effectif correspond au nombre global d'étudiants pouvant passer l'examen dans un type de local (amphithéâtre, salle TD, salle TP…). Chaque local dispose d'une capacité d'accueil déterminée, qui fixe le nombre maximal d'étudiants pouvant y passer l'examen.

Moodle) est strictement lié à l'adresse IP des salles numériques. De plus, pour renforcer la sécurité, une seule instance de Moodle peut être ouverte à la fois, afin d'empêcher un étudiant d'accéder à la session d'un autre pendant l'examen.

Notre approche offre aux enseignants une solution solide pour améliorer la qualité des évaluations. En effet, nos enseignants souhaitaient, en plus des questionnaires QCM/QCU habituels, pouvoir proposer des exercices d'application, voire des problèmes à résoudre, qui soient autocorrigés, notamment pour des matières scientifiques comme l'informatique, les mathématiques, la chimie, la physique, etc.

Pour répondre à ce besoin, nous avons utilisé les questions Cloze de Moodle pour la conception des exercices tout en respectant les règles de la docimologie (Aouatif et Awatif, 2023).

La structure d'un exercice/problème comporte :

- un énoncé de l'exercice/problème et les données à utiliser ;
- les consignes importantes pour répondre aux questions ;
- pour chaque question de l'exercice/problème :
  - énoncé de la question,
  - les options de réponses possibles s'il s'agit d'un QCM/QCU,
  - une case pour saisir la réponse s'il s'agit d'une question de type « réponse courte ».

Dans le cas de questions de type QCM/QCU, les distracteurs (les réponses incorrectes) doivent être plausibles. Cela signifie qu'ils doivent correspondre à des résultats que l'on pourrait obtenir en cas d'erreur lors de la résolution de l'exercice/problème. Cette rigueur dans la rédaction des questions rend la tâche de production de la base de questions plus délicate.

La figure 5.2 montre un exemple de question Cloze dans le cadre d'un examen du module « Excel avancé ». L'étudiant dispose des fichiers de données Excel nécessaires pour répondre aux questions en mobilisant l'ensemble de ses connaissances et de ses savoir-faire, avant de saisir la réponse dans la case appropriée.

Il est important de rappeler que l'export simplifié des notes permet de communiquer rapidement les résultats à l'administration pour affichage et que le recours au numérique facilite considérablement la logistique d'organisation, la conservation des copies, et la réponse aux demandes des étudiants de consulter leurs copies (voir figure 5.3).

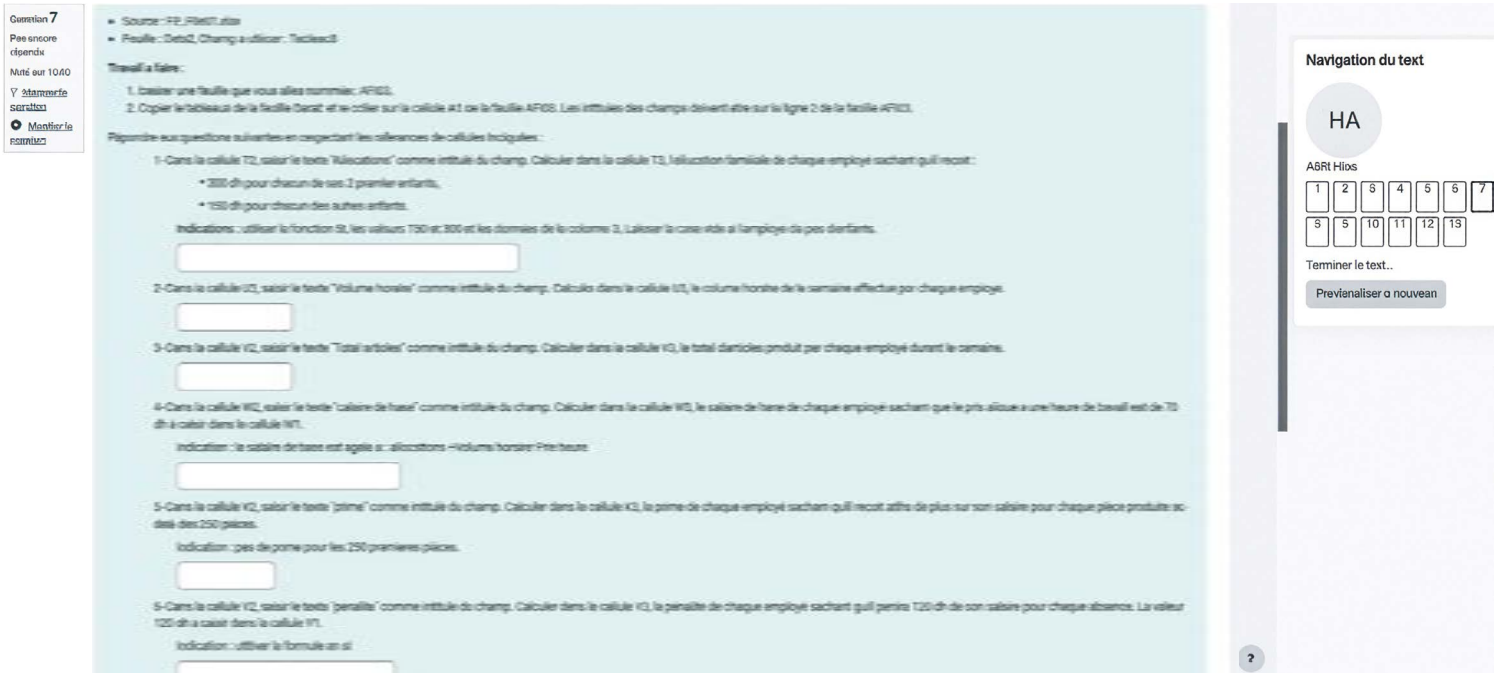


FIGURE 5.2 Question Cloze dans un examen final du module « Excel avancé ».

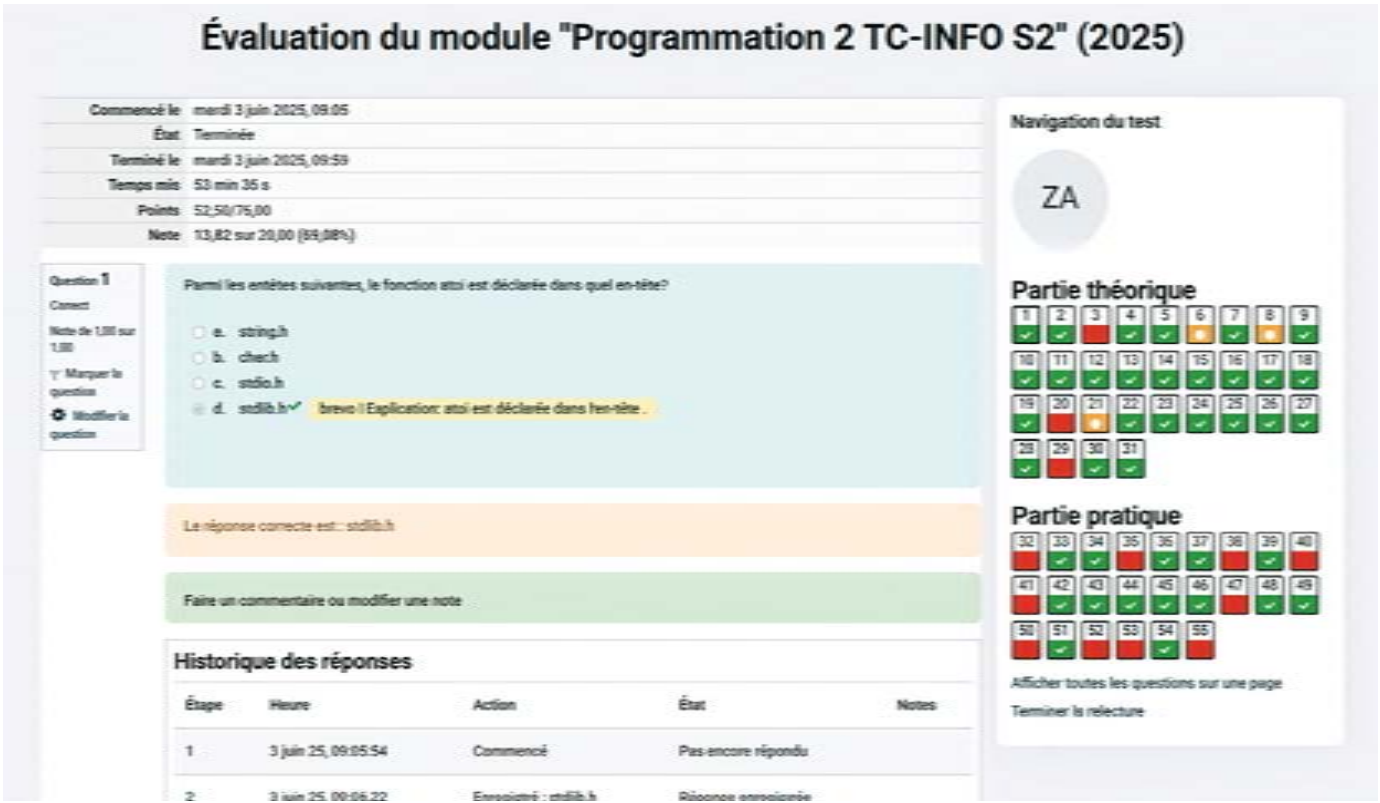


FIGURE 5.3 Interface de la copie corrigée et annotée d'un étudiant.

Évaluation papier avec l'outil AMC

Le QCM papier est un outil d'évaluation employé depuis plusieurs années dans nos établissements, avec une correction manuelle ou à l'aide d'applications sélectionnées par les professeurs. Cette méthode a été appliquée principalement pour des classes de petite taille.

Nous avons cherché à uniformiser les pratiques d'évaluation en mettant en place un système simple d'utilisation, fiable et à la portée de tous. Le choix a été de recourir à l'outil AMC pour produire et corriger les épreuves de type QCM papier. Cette pratique a été mise en place pour permettre aux enseignants ne souhaitant pas adopter l'évaluation numérique de trouver une alternative satisfaisante.

Son objectif principal est de rationaliser les aspects infrastructurels et ceux liés à la formation et à l'accompagnement. Elle repose, d'une part, sur l'utilisation des ressources du *datacenter* pour créer des environnements de bureau à distance Linux, accessibles à l'ensemble des enseignants qui souhaitent pratiquer l'évaluation sur papier avec AMC. D'autre part, elle repose sur la mise en place de formations sous forme d'ateliers pratiques et d'un support tout au long de la période d'évaluation. Plusieurs enseignants ne maîtrisent ni l'environnement Linux ni le langage LaTeX[33] (l'outil AMC nécessite la rédaction des épreuves en format LaTeX ou TXT), nos ateliers ont donc été conçus en se focalisant particulièrement sur la rédaction des épreuves en format TXT.

Comme pour l'évaluation numérique, nos enseignants ont souhaité inclure des exercices et des problèmes autocorrigés, en plus des questions habituelles de type QCM/QCU. Cette demande traduit leur volonté de diversifier les modes d'évaluation pour mieux refléter l'alignement entre objectifs de formation et nature des évaluations. Certains ont également fait part d'un sentiment de frustration face aux limites des questions fermées, estimant qu'elles ne permettaient pas toujours d'évaluer de manière exhaustive les acquis des étudiants. Pour répondre à ce besoin, nous avons utilisé les « groupes de questions » de AMC pour concevoir des exercices et des problèmes, tout en tenant compte des principes de la docimologie (Aouatif et Awatif, 2023). Nos enseignants ont conscience que les distracteurs proposés doivent correspondre à des résultats pouvant être obtenus en cas d'erreur lors de la résolution des exercices et des problèmes.

[33] LaTeX est un langage informatique qui permet de formater et de générer des documents scientifiques. Il est connu pour sa qualité de rendus au format DPF. Source : https://orbilu.uni.lu/bitstream/10993/16701/1/Tutorial_Latex.pdf (consulté le 15.01.2026).

La structure adoptée pour une épreuve avec l'outil AMC est la suivante :

- en-tête de l'épreuve ;
- questions sous forme de groupes :
  - groupe 1 : questions de cours sous forme de QCM et QCU indépendantes,
  - groupe 2 : exercice d'application :
    - énoncé de l'exercice et consignes,
    - pour chaque question :
      - énoncé de la question avec consigne,
      - options de réponses :
  - groupe 3 : exercice d'application 2,
  - groupe 4 : problème à résoudre,
  - etc.

AMC permet de générer automatiquement plusieurs variantes de l'épreuve en réorganisant de manière aléatoire l'ordre des questions et les choix de réponses. Cependant, afin de préserver la cohérence logique des exercices où les questions doivent suivre un enchaînement spécifique, l'option de mélange des questions a été désactivée pour les groupes qui proposent des exercices ou des problèmes (voir figure 5.4).

Les épreuves sont préparées en PDF en utilisant deux formats de mise en page : A4 ou A3. Les enseignants privilégient généralement le format A3 recto verso, sans grille séparée de réponse. Cette méthode permet à chaque étudiant de recevoir une seule feuille A3 contenant toutes les questions, ce qui évite les erreurs humaines lors de la distribution des variantes et la collecte des copies.

La création d'une banque de questions suffisamment riche pour permettre de nombreuses variantes d'un même examen est une tâche laborieuse. Il est également important de valider rigoureusement ces questions afin de garantir qu'elles respectent les critères de qualité docimologiques. Cette richesse de la banque de questions (catalogue AMC) assure aussi une transparence renforcée et une égalité des chances pour tous les étudiants.

Résultats et discussion

Après les résultats positifs des expérimentations à la FSSM, l'UCA a généralisé le déploiement des outils numériques pour l'évaluation des apprentissages dans tous les autres établissements. Le passage à

Question 14 Les propriétés magnétiques des trois molécules précédentes sont :

- ☐ P_2 : paramagnétique ; P_2^+ : diamagnétique ; P_2^- : paramagnétique
- ☐ P_2 : diamagnétique ; P_2^+ : paramagnétique ; P_2^- : diamagnétique
- ☐ P_2 : diamagnétique; P_2^+ : diamagnétique ; P_2^-: paramagnétique
- ☐ P_2 : paramagnétique ; P_2^+ : paramagnétique ; P_2^- : paramagnétique
- ■ P_2 : diamagnétique ; P_2^+ : paramagnétique ; P_2^- : paramagnétique

Fin du Sujet II

Sujet III

On dispose d'un récipient fermé, de volume V contenant, de l'acide hypochloreux pur sous une pression $P_{HClO} = P_i(\text{initiale}) = 0{,}060$ bar. A 298K, il s'établit l'équilibre de déshydratation de l'acide hypochloreux, de constante d'équilibre Kp = 11,1, selon la réaction :

$$2HClO_{(g)} \overset{1}{\rightleftarrows} \frac{1}{2} Cl_2O_{(g)} + H_2O_{(g)}$$

Données : Le gaz est supposé parafait ; R = 8,31J/K.mol; 1 bar = 10^5 Pa ;
$\Delta_f H^\circ_{298K}(Cl_2O, g)$=80,3 kJ/mol ; $\Delta_f H^\circ_{298K}(H_2O, g)$=-241,3 kJ/mol ;
$S^\circ_{298K}(Cl_2O, g)$=266,1 J/K.mol ; $S^\circ_{298K}(H_2O, g)$=188,7 J/K.mol ; $S^\circ_{298K}(HClO, g)$=236,6 J/K.mol;

Question 15 Calculer, à 298 K, l'enthalpie libre standard de la décomposition de HClO. Commenter le résultat obtenu.

- ☐ 5,96 kJ/mol. Réaction endothermique.
- ☐ 5,60 kJ/mol. Réaction exothermique.
- ☐ - 5,60 kJ/mol. Réaction non spontanée.
- ■ - 5,96 kJ/mol. Réaction spontanée.

Question 16 Calculer, à 298 K, la valeur de l'enthalpie standard de la décomposition de HClO. Commenter le résultat obtenu.

- ☐ 11,80 kJ/mol. Réaction endothermique.
- ■ - 11,44 kJ/mol. Réaction exothermique.
- ☐ 11,44 kJ/mol. Réaction endothermique.
- ☐ - 11,80 kJ/mol. Réaction endothermique.

Question 19 Etablir les expressions des pressions partielles des gaz qui composent le système à l'équilibre sachant que $P_t(\text{totale}) = P_i(\text{initiale})$. Calculer les valeurs de ces pressions.

- ☐ $P_{HClO} = 0{,}060(1+\alpha) = 0{,}08$Pa ; $P_{Cl_2O} = P_{H_2O} = 0{,}060(\alpha) = 0{,}26$Pa
- ■ $P_{HClO} = 0{,}060(1-\alpha) = 0{,}008$bar ; $P_{Cl_2O} = P_{H_2O} = 0{,}060(\frac{\alpha}{2}) = 0{,}026$bar
- ☐ $P_{HClO} = 0{,}060(1-\alpha) = 0{,}008$bar ; $P_{Cl_2O} = P_{H_2O} = 0{,}060(2\alpha) = 2{,}6$bar
- ☐ $P_{HClO} = 0{,}060(1-\alpha) = 0{,}008$bar ; $P_{Cl_2O} = P_{H_2O} = 0{,}060(\alpha) = 0{,}26$bar

Question 20 Quelle est l'influence d'une augmentation de la pression totale, à température constante, sur l'équilibre de déshydratation de l'acide hypochloreux ?

- ☐ L'augmentation de la pression totale entraîne l'évolution dans le sens directe (sens 1).
- ☐ L'augmentation de la pression totale entraîne l'évolution dans le sens indirecte (sens 2).
- ■ La pression totale n'a pas d'influence sur l'équilibre.

Question 21 Quelle est l'influence d'une augmentation de la température, à pression totale constante, sur l'équilibre de déshydratation de l'acide hypochloreux ?

- ☐ Il n'y a pas d'effet de la température sur l'équilibre.
- ■ L'augmentation de la température entraîne l'évolution dans le sens indirecte (sens 2).
- ☐ L'augmentation de la température entraîne l'évolution dans le sens directe (sens 1).

Fin du Sujet III

FIGURE 5.4 Exemple d'exercice sur AMC.

grande échelle est progressif : dans un premier temps, seuls les professeurs ayant une maîtrise des TICE ont été contactés. Par la suite, tous les enseignants de l'UCA ont eu la possibilité d'utiliser les deux dispositifs dans leurs évaluations. Dans les établissements ne disposant pas assez de salles numériques équipées avec un nombre suffisant d'ordinateurs, seul le dispositif QCM papier avec AMC a été mis en place.

La réforme des filières de licence, lancée en septembre 2023 dans tous les établissements à accès ouvert, se distingue par l'introduction de plusieurs modules transversaux[34] communs à l'ensemble des filières. Les cohortes inscrites à ces modules varient entre 5000 et 12 000 étudiants. Pour relever le défi de l'évaluation de ce large effectif, le CIP, en collaboration avec les enseignants responsables de chaque module, a instauré une banque de questions partagée avec tous les intervenants afin de simplifier les évaluations de fin de semestre. Ce dispositif innovant a facilité la génération de multiples versions d'épreuves à partir de cette banque de questions, que ce soit pour les évaluations numériques via Moodle ou celles sur papier à l'aide du système AMC.

Analyse des résultats

Pour évaluer notre approche, un questionnaire a été proposé aux équipes pédagogiques de l'UCA ayant utilisé l'un des dispositifs ou les deux pour leurs évaluations.

L'analyse des 88 réponses obtenues (figures 5.5, 5.6 et 5.7) révèle que toutes les disciplines sont représentées, avec une prédominance des sciences techniques, regroupant plus de 50 % des participants, et une majorité d'enseignants provenant de la FSSM. Il a également été observé que plus de 60 % des enseignants avaient à gérer des cohortes allant de 500 à plus de 3000 étudiants, ce qui explique leur intérêt pour ces deux dispositifs.

Ce troisième graphique (figure 5.7) permet d'illustrer la répartition des cohortes étudiantes et d'identifier les effectifs les plus représentés à l'université. Il montre, par exemple, que 26,4 % des cohortes comptent entre 100 et 500 étudiants.

[34] MTU (unité de transmission maximale), *soft skills* (compétences douces), *digital skills* (compétences numériques), langues étrangères, etc.

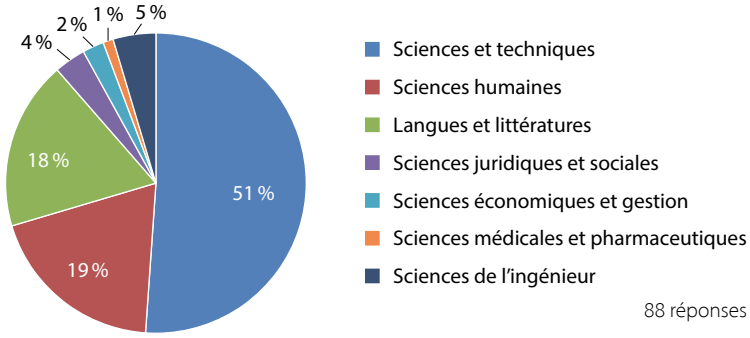


FIGURE 5.5 Champs disciplinaires utilisant le numérique pour évaluer.

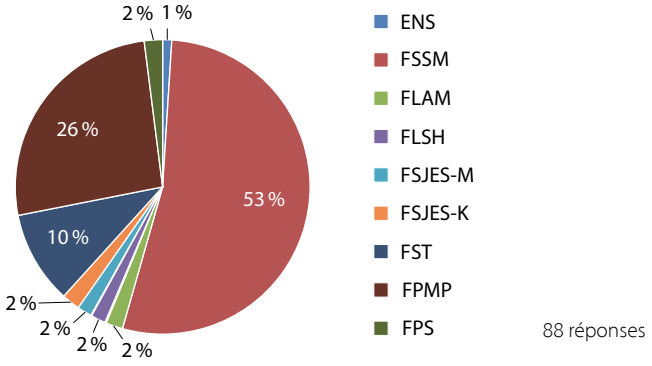


FIGURE 5.6 Établissements utilisant le numérique pour évaluer.

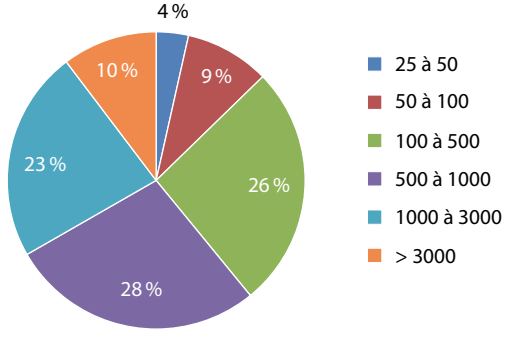


FIGURE 5.7 Nombre d'étudiants par cohorte.

Enfin, le graphique ci-dessous (figure 5.8) montre l'avis des enseignants quant à l'utilisation des deux méthodes. Il en ressort qu'elles sont perçues comme utiles et permettent un gain de temps. Cependant, certains enseignants se disent insatisfaits de la qualité de l'évaluation, ce qui pourrait être lié au choix du type de questions et à la qualité de leur rédaction. Ces éléments suggèrent un besoin de renforcer le volet formation sur les outils numériques en générale et sur la docimologie en particulier.

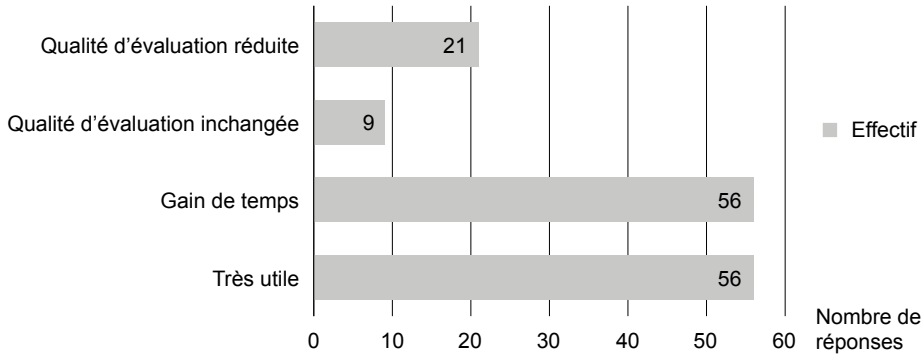


FIGURE 5.8 Avis des professeurs sur l'utilisation des outils numériques pour l'évaluation des apprentissages des étudiants[35].

Pour compléter notre étude, nous présentons, dans le tableau 5.6, les points forts et les points faibles sous forme de synthèse réalisée par le CIP. Nous proposerons également quelques pistes d'amélioration.

Afin de pallier la complexité de production d'une large base de questions, nous avons déjà entamé le développement d'une application conçue avec l'intelligence artificielle, destinée à générer des questions à partir des cours dispensés par les professeurs. Cette application est actuellement en phase de test, et les premiers résultats s'avèrent très prometteurs. Par ailleurs, l'UCA est en train de mettre en place un système d'inscription préalable aux examens afin d'optimiser les ressources nécessaires aux évaluations numériques (salles, ordinateurs et surveillants) pour l'organisation des examens.

[35] L'axe des abscisses représente le nombre de réponses recueillies. Par exemple, 56 étudiants ont jugé l'utilisation des outils numériques « très utile ».

TABLEAU 5.6 Points forts et points faibles des dispositifs proposés.

Dispositif 1	**Évaluation numérique – Moodle**
Points forts	• Possibilité de produire des questions QCM/QCU ou autres pour une évaluation des connaissances et de la compréhension • Possibilité de produire des **exercices ou problèmes** via les questions Cloze • Possibilité d'intégrer des **images en couleur, de l'audio et de la vidéo** dans les questions • Réduction du temps d'évaluation • Diminution de la charge de travail des enseignants • Consultation facile des copies par les étudiants • Export facile et rapide des notes • Amélioration de la qualité des enseignements • Amélioration de l'infrastructure de l'établissement • Stockage des copies • Réduction des cas de fraude • Pas de consommation de papier
Points faibles	• Évaluation prolongée dans le temps (plusieurs groupes) • Production d'une large banque de questions • Ressources humaines du CIP insuffisantes
Dispositif 2	**Évaluation papier – AMC**
Points forts	• Possibilité de produire des questions QCM/QCU pour une évaluation des connaissances et compréhension • Possibilité de produire des **exercices ou problèmes** via les groupes de questions • Réduction du temps d'évaluation • Consultation facile des copies par les étudiants • Export facile et rapide des notes • Amélioration de l'infrastructure de l'établissement • Stockage des copies • Réduction des cas de fraude
Points faibles	• Logistique complexe d'impression et de distribution des différentes variantes des sujets d'examen

Impact de l'utilisation des dispositifs proposés

La mise en place des deux dispositifs d'évaluation des apprentissages a eu un impact sur les étudiants, les professeurs et les établissements. Nous allons dans ce qui suit détailler quelques impacts identifiés.

Impact sur les étudiants

Dans la plupart des modules, les étudiants sont mal informés des objectifs, de la méthode d'apprentissage, et des modalités d'évaluation. Depuis la mise en place de notre dispositif, les enseignants clarifient ces éléments lors de la première séance en présence. L'adoption de l'évaluation numérique ou papier incite les enseignants à enrichir leurs espaces de cours dans la plateforme pédagogique Moodle

avec plus d'activités interactives, notamment des QCM autocorrectifs, ce qui a aussi pour effet de préparer les étudiants aux contrôles. Nous n'avons pas conduit d'enquête auprès des étudiants pour mesurer finement leur avis, mais ces deux outils contribuent à réduire le délai d'affichage des résultats, ce qui diminue la période de stress liée à leur attente, et ont tendance à limiter la fraude grâce aux multiples variantes d'examens générées automatiquement.

Impact sur les enseignants

L'utilisation des deux outils proposés a considérablement réduit le temps consacré à la correction des copies, permettant ainsi aux enseignants de consacrer plus de temps à d'autres tâches académiques, telles que la recherche scientifique ou la préparation des cours. La formation offerte aux enseignants sur l'utilisation des outils numériques, comme Moodle et AMC, a renforcé leurs compétences en TIC, les rendant plus autonomes et capables, à leur tour, de former ou d'accompagner d'autres collègues.

Impact sur les établissements

La mise en place des deux dispositifs d'évaluation s'inscrit dans le cadre de la digitalisation de l'UCA. Ils ont eu un impact positif sur les établissements de l'UCA à trois niveaux :

- infrastructure : l'adoption des évaluations numériques via Moodle et des évaluations papier à l'aide de l'outil AMC a apporté des avantages considérables aux établissements, tant sur le plan de l'équipement matériel (*datacenter*, ordinateurs, scanners, etc.) que sur celui de l'infrastructure réseau, notamment grâce à la mise en place de la fibre optique reliant les différents établissements ;
- finances : le fait que les deux dispositifs reposent sur des outils libres et gratuits permet de réaliser des économies en évitant l'acquisition de solutions similaires payantes ;
- ressources humaines : la surveillance des examens représente une tâche gourmande en ressources humaines. L'adoption de l'évaluation numérique via Moodle permet de diminuer considérablement le besoin en surveillants (voir tableau 5.5), réduisant ainsi le nombre de créneaux de surveillance pour l'ensemble des enseignants de l'établissement.

Conseils de mise en œuvre pratique

Pour mettre en œuvre les deux dispositifs d'évaluation adoptés par notre Université, nous recommandons de :

- implanter l'infrastructure matérielle et réseau nécessaire pour une évaluation numérique ou papier ;
- produire des exercices, en utilisant les types de questions proposées par Moodle (questions Cloze) ou AMC (groupe de questions) ;
- former les enseignants, pour une utilisation optimale :
  - à la création d'une table de spécification pour leurs examens,
  - aux critères de qualité de rédaction des questions (docimologie),
  - à la création et à la gestion des banques de questions et des paramétrages des activités « test »,
  - à l'outil AMC, surtout concernant la préparation des questions ;
- privilégier un travail en groupe pour diminuer la charge de travail de production des questions ;
- faire valider les questions par un comité de lecture ;
- mettre en place une équipe support pour l'accompagnement des enseignants en difficulté.

Conclusion et travaux futurs

L'intégration des outils numériques dans l'évaluation des apprentissages à l'Université Cadi Ayyad est devenue une nécessité pour répondre aux grands défis rencontrés dans un contexte de massification du nombre des étudiants, surtout dans les établissements à accès ouvert. Les investissements de l'UCA sur le plan des infrastructures matérielle et réseau couplés aux formations des enseignants dans le domaine des TIC ont permis d'établir deux approches d'évaluations fondées sur l'usage du numérique : l'évaluation numérique via Moodle et l'évaluation papier sur l'outil AMC.

Les exigences de nos enseignants en matière de qualité des évaluations nous ont conduits à opter pour les questions de type Cloze sur Moodle, ainsi que pour les groupes de questions sur AMC. Ces choix permettent de générer des épreuves autocorrectives, qui, en plus des traditionnelles questions de type QCM/QCU, incluent des exercices d'application ou d'analyse. Cela garantit une évaluation complète des connaissances, de la compréhension, du savoir-faire, et des capacités d'analyse de nos étudiants. Les résultats issus de la mise en œuvre de

ces deux approches révèlent un impact positif sur la charge horaire des enseignants, ainsi que sur l'infrastructure des établissements. Toutefois, il demeure essentiel de poursuivre les efforts en matière de formation continue, afin de toucher un plus grand nombre de professeurs et, surtout, de les sensibiliser aux bonnes pratiques de rédaction des QCM, conformément aux principes de la docimologie.

Dans la continuité de ce travail, l'UCA prévoit, sur le plan infrastructurel, de renforcer ses équipements en procédant à l'acquisition de nouveaux ordinateurs et scanners. Cette initiative vise à optimiser le temps consacré aux examens, permettant ainsi de respecter les calendriers pédagogiques et administratifs annuels des établissements. Sur le plan pédagogique, une application conçue avec l'intelligence artificielle est en cours de finalisation, dans le but de faciliter la création de banques de questions pour chaque module. Enfin, l'UCA compte perpétuer la formation continue pour l'ensemble des enseignants, avec une attention particulière portée aux nouvelles recrues.

Laurent Moccozet, Omar Benkacem,
André Camacho, Allan Bowman
& Patrick Roth

6 Évolution des modalités et infrastructures des examens en ligne

Résumé

Pendant le confinement dû au Covid-19, les universités ont radicalement transformé l'enseignement et les examens, passant à des méthodes entièrement en ligne. Même les enseignants réticents ont dû rapidement s'adapter. Cette transition urgente a souvent impliqué l'adaptation empirique des méthodes traditionnelles aux outils en ligne. Après cette période, il est intéressant d'examiner l'évolution des pratiques : retour aux méthodes traditionnelles, adoption massive des modalités en ligne, ou phénomène de post-numérisation de l'éducation. L'Université de Genève avait déjà une infrastructure pour les examens en ligne, renforcée pendant la pandémie. Notre étude analyse l'évolution de cette infrastructure et des modalités d'examen avant, pendant et après le Covid-19, ainsi que les pratiques innovantes actuelles des enseignants.

Introduction

L'e-évaluation ou évaluation numérique consiste à administrer un examen à l'aide d'un dispositif numérique (Viberg *et al.*, 2024) soit à distance, soit dans une salle de classe. Elle peut être appliquée dans un contexte à la fois formatif et sommatif. Pendant la pandémie de Covid-19 et les périodes de confinement, l'évaluation numérique a été

largement acceptée comme le seul moyen de maintenir l'évaluation dans un contexte d'apprentissage à distance (St-Onge *et al.*, 2022). Cela a conduit à un déploiement de l'évaluation numérique à une rapidité inédite. Mais l'évaluation numérique n'est pas seulement un moyen pratique et technique de passer des examens. Elle peut fournir de nouvelles approches pédagogiques pour l'évaluation des compétences et, en particulier, pour l'évaluation active et authentique (Hillier, 2023). Au niveau d'un établissement d'enseignement supérieur, le processus d'encouragement de l'utilisation de l'évaluation numérique par l'ensemble du personnel enseignant nécessite la mise en place d'une stratégie appropriée. Il s'agit de savoir comment établir les capacités qui permettront aux enseignants de réaliser des examens de ce type ; comment les aider à s'approprier les outils et à les intégrer dans leurs pratiques d'évaluation pour les faire évoluer ; et, enfin, comment les motiver à adopter cette modalité tout en respectant les bonnes pratiques.

Le contexte et la problématique

Dès 2014 (Moccozet *et al.*, 2017), l'Université de Genève a commencé ses premiers essais pour développer une stratégie de soutien à l'évaluation numérique fondée sur la mise à disposition d'un environnement numérique flexible permettant de réaliser des évaluations selon une variété de scénarios. Cette stratégie a pu être largement implémentée avec la décision de l'institution de recourir systématiquement à l'évaluation en ligne pour les examens 2019-2020. D'une façon générale, la pratique des technologies numériques dans l'éducation a été redéfinie à de nombreux égards par le premier confinement qui a accompagné la pandémie de Covid-19. Il a promu au premier rang des enjeux éducatifs la qualité de l'enseignement hybride et la maîtrise des outils et des compétences numériques (Siddiq *et al.*, 2024). L'évaluation numérique est alors devenue la norme pour l'ensemble de la population académique durant cette courte période. L'expérience acquise reposant sur un environnement résilient a permis une utilisation massive de l'évaluation numérique. Dans la période post-Covid, il est intéressant de se pencher sur la question des évaluations numériques après que tous les enseignants et étudiants y ont été confrontés. À quels phénomènes assiste-t-on, entre adoption massive et rejet ? Et si elles sont adoptées, dans quelle direction évoluent-elles, de la simple automatisation à des méthodes innovantes axées sur les compétences ?

Ce chapitre propose d'étudier, d'analyser et d'illustrer les changements dans le déploiement du soutien institutionnel et l'utilisation des évaluations numériques par les enseignants, avant, pendant et après le Covid-19 à l'Université de Genève. Il se terminera par un état des lieux et des perspectives quant aux méthodes d'évaluation numérique (Souppez *et al.*, 2023 ; Viberg *et al.*, 2024) en relation avec le contexte post-numérique vers lequel évolue l'enseignement supérieur (Fawns, 2019).

Le cadre théorique

La pandémie de Covid-19, avec la période de confinement qui lui a été largement imposée, a eu l'effet d'un tremblement de terre sur le secteur de l'éducation. Dans l'enseignement supérieur, le confinement a rapidement nécessité l'introduction de ce que l'on appelle aujourd'hui l'enseignement à distance d'urgence (*Emergency Remote Teaching*, ERT) et, dans le cas des examens, l'évaluation à distance d'urgence (*Emergency Remote Assessment*, ERA). Hodges *et al.* (2020) définissent l'ERT comme suit :

> Contrairement aux expériences planifiées dès le départ et conçues pour être en ligne, l'ERT est un transfert temporaire de l'enseignement vers un autre mode de transmission en raison de circonstances de crise. Il implique l'utilisation de solutions d'enseignement entièrement à distance pour l'instruction ou l'éducation qui seraient autrement dispensées en face-à-face ou sous forme de cours hybrides (c.-à-d. combinant des activités présentielles et à distance en modes synchrone et asynchrone) et qui reviendront à ce format une fois la crise ou l'urgence passée. L'objectif principal dans ces circonstances n'est pas de recréer un écosystème éducatif robuste, mais plutôt de fournir un accès temporaire à l'enseignement et aux supports pédagogiques d'une manière qui est rapide à mettre en place et qui est disponible de façon fiable pendant une urgence ou une crise.

Comme le soulignent les auteurs, l'ERT se caractérise par une situation d'urgence, qui n'attend pas de résultats durables, et qui n'est que temporaire avant un retour à la situation antérieure à la période de crise. Cependant, compte tenu de l'ampleur de cette crise, la question se pose de savoir si la période post-Covid s'est effectivement limitée à un retour aux pratiques habituelles, notamment en matière d'examens.

La littérature établit clairement que le confinement a immédiatement conduit à la décision générale de faire passer les examens à distance dans l'enseignement supérieur (Montenegro-Rueda *et al.*, 2021). Il s'agit de connaître les modalités de surveillance d'une institution à l'autre, car le passage à l'enseignement à distance fait craindre un recours massif à la tricherie. Certains établissements ont choisi d'utiliser des outils de surveillance, d'autres non, pour ne pas augmenter le stress des étudiants (Chouhan, 2024 ; Kharbat et Abu Daabes, 2021). Dans la plupart des cas, les enseignants transposent simplement leurs pratiques habituelles dans un format à distance, avec le désir de revenir le plus rapidement possible à la situation initiale (Hancock *et al.*, 2023 ; Papanastasiou et Solomonidou, 2023 ; Slade *et al.*, 2022). Parmi les raisons invoquées pour ne pas recourir aux examens à distance, on note notamment la tricherie, jugée trop facile pour les élèves et trop compliquée à contrer pour les enseignants, ainsi que le fait que les nouvelles méthodes d'évaluation ont été introduites trop rapidement et dans l'urgence, sans capacité d'analyse des incidences sur la qualité et la significativité des résultats obtenus (Papanastasiou et Solomonidou, 2023). Cependant, une prise de conscience est également en train d'émerger. Elle a été provoquée par la transition obligatoire induite par le changement des méthodes d'examen. Elle concerne les implications des procédures d'examen pour les étudiants, l'alignement et l'articulation des examens avec l'enseignement, l'authenticité des procédures d'examen par rapport aux compétences évaluées et, enfin, les ressources disponibles pour la réalisation des examens (St-Onge *et al.*, 2022). La situation est même parfois perçue comme une opportunité d'innover et de revoir les pratiques et les procédures (Broadbent *et al.*, 2023). Certains résultats sont perçus comme positifs et pouvant être maintenus par la suite (Patael *et al.*, 2022). Mottiar *et al.* (2024) fournissent une analyse plus détaillée du comportement des enseignants en situation de crise, en présentant une typologie des réponses d'évaluation au Covid-19 (modèle TARC, *Typology of Assessment Responses to COVID-19*). Cette typologie identifie quatre catégories d'enseignants, dont le comportement permet de prédire les effets à long terme de la pandémie sur leurs pratiques d'évaluation et l'utilisation de l'évaluation numérique. Ils sont positionnés dans un espace à deux dimensions qui identifie l'ampleur des changements apportés aux méthodes d'examen et leur inscription dans un plan à court terme (réaction) ou à long terme (stratégie). Les auteurs identifient ainsi : (1) les « réactifs », qui sont des enseignants

ayant simplement transféré leurs évaluations en ligne ; (2) les « réactifs adaptatifs », qui ont légèrement modifié les évaluations pour les adapter à l'environnement en ligne ; (3) les « innovateurs opportunistes », qui ont profité de la pandémie pour mettre en œuvre des stratégies qu'ils avaient envisagées, mais qu'ils avaient repoussées par manque de temps et de motivation ; et enfin, (4) les « innovateurs engagés », qui sont engagés dans l'innovation des méthodes d'enseignement et d'évaluation de manière continue et qui ont continué à faire ce qu'ils ont toujours fait. Les enseignants qui appartiennent aux deux dernières catégories sont les plus susceptibles d'intégrer et d'adapter les changements introduits dans les examens pour les reproduire dans le temps.

Pour finir, plusieurs constats se dégagent pour la période post-Covid :

- tout d'abord, il y a eu des développements qui sont considérés comme positifs concernant la qualité des méthodes d'examen (Broadbent *et al.*, 2023) ;
- les voies à suivre et les orientations à prendre une fois la période de crise terminée sont décrites afin de tirer parti de l'expérience acquise, même si elle l'a été dans la précipitation (Hartle, 2023 ; Kim, 2022) ;
- on craint notamment que les mesures adéquates ne soient pas maintenues et étendues (Broadbent *et al.*, 2023).

L'étude menée par Birrada *et al.* (2024) indique que les méthodes d'interaction et d'évaluation en ligne continueront à être utilisées dans les universités durant la période post-Covid. Cependant, l'efficacité globale de cette expérience d'apprentissage et d'évaluation en ligne « forcée » reste perçue comme moindre que dans le cas de l'enseignement en face-à-face. Les auteurs posent la question de savoir si les formes d'examen traditionnelles seront vraiment remplacées par des formes d'évaluation plus contemporaines. Ils rappellent le constat fait par Wareing (2022), pour qui l'apprentissage et l'enseignement ont été et seront facilités par la technologie plutôt que transformés par elle. Il prédit par ailleurs que l'enseignement supérieur avec un enseignement en présentiel, complété par le numérique, restera le modèle à la fois apprécié et privilégié. Selon Douglas et Chapman (2023), les enseignants du monde entier ont soudainement dû s'adapter pour enseigner en mode d'apprentissage mixte, en démultipliant les activités en ligne, et tout porterait à croire que cette modalité continuera à occuper une place prépondérante post-Covid et deviendrait la « nouvelle norme » en matière d'enseignement supérieur.

Le dispositif et le scénario pédagogique

En 2014, l'Université de Genève a mis en place une infrastructure institutionnelle pour les examens numériques, composée de trois éléments principaux : le *Learning Management System* (LMS) Moodle pour les tests, le navigateur sécurisé *open source* SEB (*Safe Exam Browser*) pour contrôler l'accès aux ressources, et une infrastructure de bureau virtuel. Cette configuration permet différents scénarios d'examens, que ce soit sur les ordinateurs de l'Université ou sur ceux des étudiants.

Une équipe technique assure le développement et la maintenance de cet environnement numérique. Elle propose des formations, développe des ressources pour les enseignants et peut fournir une assistance pendant les examens. Des scénarios types ont été documentés pour faciliter l'adoption du système. La promotion repose sur le partage d'expériences des « innovateurs engagés » (selon le modèle TARC) et sur des présentations directes aux facultés par l'équipe de soutien. L'organisation de l'environnement numérique d'examen est illustrée à la figure 6.1.

Lors de la pandémie, l'Université a basculé l'ensemble des examens en ligne. Pour les examens écrits, l'infrastructure existante a été renforcée, tandis que Zoom a été utilisé pour les examens oraux. L'Université a choisi de ne pas recourir aux technologies de surveillance pour les examens à distance. Le renforcement de l'infrastructure a nécessité l'installation d'instances Moodle supplémentaires sur des serveurs optimisés dédiés aux examens. Le rectorat et les facultés se sont fortement impliqués dans le déploiement des ressources et la coordination, notamment en désignant du personnel formé pour faire le lien entre l'équipe de support général et les enseignants. Ces personnels sont soit des techniciens qui ont des compétences dans la mise en œuvre des outils numériques ou des profils de technopédagogues qui ont des compétences dans les deux domaines.

Après le retour à l'enseignement en présentiel, l'infrastructure numérique d'examens mise en place pendant la pandémie a été maintenue. Certaines facultés ont conservé une personne-ressource pour soutenir leurs enseignants et assurer la liaison avec l'équipe de support global.



FIGURE 6.1 Organisation de l'environnement institutionnel d'examens numériques.

La méthodologie

Pour analyser l'évolution de l'usage des e-évaluations à l'université, nous nous sommes appuyés sur trois sources de données :

1. les statistiques par année académique des typologies d'examens réalisés sur l'infrastructure d'e-évaluation institutionnelle répertoriées par l'équipe de support de l'Université. Cette équipe support assure le déploiement et le maintien de toutes les plateformes dédiées à l'e-évaluation et a accès à l'intégralité des données qui y sont présentes. Par ailleurs, elle répertorie toutes les évaluations qui s'y déroulent et monitore le déroulement de toutes les évaluations. Ces données sont disponibles depuis l'année 2013-2014. Nous les complétons par les statistiques des activités d'évaluation sur les serveurs Moodle dédiés aux examens en ligne. Ces données sont disponibles depuis l'année 2020-2021 ;
2. une étude des pratiques post-Covid concernant les examens numériques a été menée auprès d'une population d'enseignants de l'Université de Genève. L'étude, anonyme, a consisté en un questionnaire et des entretiens individuels semi-directifs. Le questionnaire a été envoyé à 36 enseignants et 24 réponses ont été reçues. Toutes les méthodes d'examen numérique sont représentées dans cet échantillon. Il couvre des classes allant de 30 à plus de 400 étudiants. Le questionnaire a été suivi de 11 entretiens individuels ;

3. une analyse des pratiques d'évaluation à partir du portail des innovations pédagogiques de l'Université de Genève[36]. Ce portail répertorie 155 exemples pédagogiques sous forme de fiches structurées en quatre parties : situation de départ, mise en place et déroulement, retour et conseils et, enfin, avis des étudiants. Nous avons identifié et étudié 20 exemples qui mettent en avant le recours à une évaluation numérique. Ces exemples correspondent aux pratiques des « innovateurs engagés » selon le modèle TARC dont nous pouvons considérer qu'ils ouvrent la voie à des pratiques plus innovantes qui serviront de modèle pour le reste des enseignants. Les descriptifs complets des cours permettent de comprendre comment les enseignants recourent au numérique pour l'évaluation des étudiants et l'intègrent dans leur démarche pédagogique.

Les résultats et leur analyse

Analyse des statistiques des examens numériques

La figure 6.2 montre l'évolution du nombre de passations d'examens au cours des années académiques selon les trois principales méthodes soutenues par l'infrastructure institutionnelle :

- *Bring Your Own Device* (BYOD, en bleu) : sur l'équipement des étudiants sur le campus de l'Université ;
- sur les ordinateurs de l'Université (INSTIT, en rouge) ;
- à distance sur l'équipement des étudiants (DISTANCE, en vert).

L'analyse du graphique révèle clairement que la période post-Covid n'a pas marqué un retour aux pratiques initiales en matière d'examens numériques. Au contraire, l'expérience forcée de l'évaluation numérique pendant la pandémie a eu un impact durable, transformant significativement les pratiques d'évaluation au sein de l'institution. Cette évolution est encore plus remarquable si l'on considère une nouvelle méthode d'examen introduite après le Covid-19 en réponse à une demande des enseignants, à la suite de leurs expériences pendant la période de confinement. Il s'agit d'un système qui permet de passer des tests sur Moodle, d'imprimer divers documents pour effectuer ces tests en classe, puis de les corriger automatiquement sur Moodle après avoir numérisé les feuilles de réponses des

[36] https://www.unige.ch/innovations-pedagogiques/catalogue (consulté le 15.01.2026).

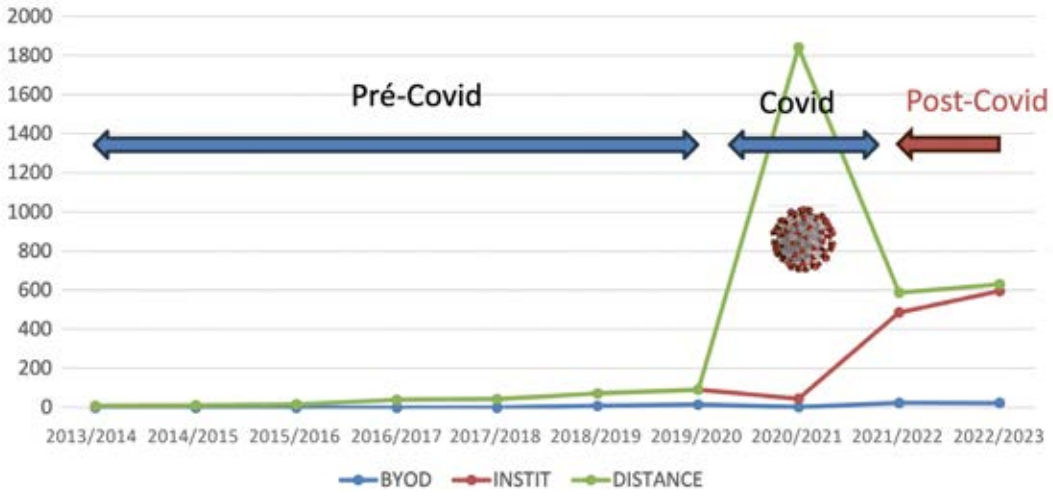


FIGURE 6.2 Progression du nombre d'examens passés au cours des années académiques selon les trois modalités principales (BYOD en bleu : équipement des étudiants sur le campus de l'Université ; INSTIT en rouge : ordinateurs de l'Université ; DISTANCE en vert : à distance, équipement des étudiants).

élèves. Cette méthode est représentée en jaune dans le schéma de la figure 6.3 ci-dessous et est appelée « Off-Quiz ». Avec cette méthode supplémentaire, l'infrastructure institutionnelle supporte actuellement quatre modes d'évaluation électronique différents : BYOD, à distance, institutionnel (examens réalisés sur le campus avec l'équipement numérique de l'Université) et Off-Quiz ou quiz hors ligne (désormais nommé *offline*).

Si l'on considère les trois périodes, pré-Covid, Covid et post-Covid, on constate que les tendances à l'intérieur de chacune d'entre elles sont cohérentes. Pré-Covid, les modalités sont très peu utilisées, ce qui correspond principalement à une utilisation par des enseignants motivés. On observe le pic de la modalité « distance » et la disparition des autres modalités pendant la période Covid. Sur la période post-Covid, on remarque une diminution globale du nombre d'examens numériques, mais une augmentation significative par rapport à la période pré-Covid, avec un glissement principalement vers les modalités institutionnelles et quiz *offline*. Parallèlement, les modalités BYOD et à distance restent très faibles, et l'on constate même qu'elles ont tendance à diminuer au fil du temps.

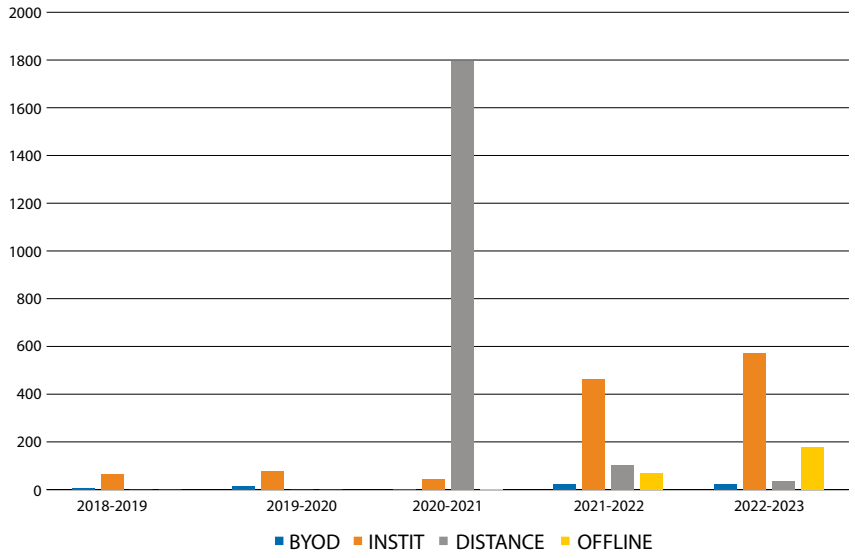


FIGURE 6.3 Évolution du nombre d'examens par méthode au cours des années académiques.

En ce qui concerne la modalité institutionnelle, il convient de noter les points suivants :

- 70 % des examens numériques sont réalisés de cette manière ;
- 100 % des examens de la faculté de traduction et d'interprétation sont passés selon cette méthode ;
- les facultés de traduction et d'interprétation et d'économie et de gestion ont des référents facultaires[37]. Elles organisent leurs examens numériques de manière indépendante, avec le soutien occasionnel de l'équipe d'assistance de l'Université ;
- les facultés de sciences, de sciences sociales, de psychologie et sciences de l'éducation, de sciences humaines et de médecine organisent leurs examens avec la participation active de l'équipe d'évaluation en ligne ;
- les postes de travail informatiques ont été aménagés pour les personnes ayant des besoins particuliers. L'Université appelle « étudiant à besoins particuliers » un étudiant porteur d'un trouble

[37] Les référents facultaires sont des personnels qui sont dédiés par les facultés à aider sous une forme ou sous une autre à l'organisation des examens sur ordinateur et font l'intermédiaire avec l'équipe d'assistance centrale de l'Université.

diagnostiqué et/ou d'un handicap, qui rencontre des obstacles dans son parcours d'études ;
- 500 machines virtuelles (VDI) ont été déployées pour compléter l'environnement numérique institutionnel.

Nous avons répertorié les différents types d'activités d'évaluation disponibles sur Moodle sur les cinq serveurs dédiés aux examens répartis entre les facultés. Les activités d'évaluation sont au nombre de trois : les devoirs (*assignments*), les tests (quiz) et les tests hors ligne (quiz *offline*) :
- devoir : il permet à l'enseignant de fournir aux étudiants des tâches, de récupérer les travaux et de leur transmettre des *feedbacks* et des notes ;
- test : il permet à l'enseignant de créer des évaluations comportant des questions à choix multiple, vrai-faux, d'appariement, à réponses courtes ou calculées. La notation des réponses est automatisée ;
- test hors ligne : il permet à l'enseignant de rédiger des QCMs papiers. Ces QCMs peuvent être téléchargés sous différents formats et imprimés. Les étudiants donnent les réponses sur une grille à cocher. Ces grilles sont ensuite scannées, déposées dans le système, puis les réponses sont évaluées et notées.

Chaque type d'activité d'évaluation offre aux enseignants la possibilité d'intégrer dans sa pratique le support numérique. Les tests hors ligne représentent une modalité d'évaluation (les QCMs) standard à laquelle il est possible de recourir sans le numérique et qui facilite la correction des examens pour l'enseignant avec une prise de risque technologique et pédagogique minimale. Les devoirs laissent potentiellement la possibilité de rendus variés qui vont de l'essai écrit à du code de programme. Ils ne peuvent pas être corrigés automatiquement et leur recours peut donc faire appel soit à un confort de lecture de travaux écrits ou à des scénarios d'évaluation authentique. Les tests se placent en intermédiaire et leur contexte dépend du type de questions qui y sont intégrées.

La synthèse de leur utilisation cumulée pour l'ensemble de l'Université est présentée dans la figure 6.4. Pour chaque année académique, chaque serveur dédié aux examens cumule toutes les e-évaluations des sessions des deux semestres d'automne et de printemps ainsi que la session additionnelle de rattrapage. Nous présentons l'évolution de

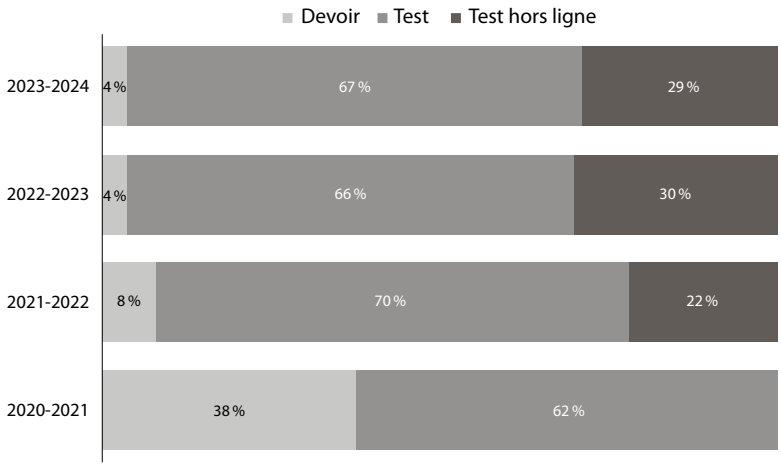


FIGURE 6.4 Répartition des types d'activités d'évaluation numérique pour l'ensemble de l'Université par année académique.

la répartition entre les trois types d'évaluation. L'observation de l'évolution entre 2020-2021 et 2023-2024 montre une forte tendance à un recours massif aux tests (96 % en cumulant les tests et tests hors ligne sur les deux dernières années) et une quasi-disparition des devoirs. On constate aussi que la répartition entre tests et tests hors ligne tend vers une répartition qui se stabilise aux environs de 70 % en ligne contre 30 % hors ligne. Cette évolution traduit un penchant important pour des scénarios d'évaluation choisis.

Pour approfondir l'analyse, nous avons aussi répertorié les types de questions posées dans les tests en ligne et les tests hors ligne. Les serveurs Moodle dédiés aux examens offrent une vingtaine de types de questions différentes pour les tests. Nous les avons regroupés selon cinq catégories : (1) questions à choix ; (2) questions ouvertes ; (3) questions de calculs ou numériques ; (4) questions d'appariement et (5) questions hybrides qui mélangent celles à choix, ouvertes et numériques. La figure 6.5 présente l'évolution de la répartition des usages de ces catégories de questions pour l'ensemble de l'Université au cours des années académiques. On constate qu'en période de post-Covid, il y a une tendance forte à l'augmentation du recours aux questions à choix qui finissent par représenter environ 40 %. Viennent ensuite les questions numériques et de calculs, qui représentent environ 30 %. Les questions d'appariement se stabilisent aux environs de 20 % et les questions ouvertes autour de 10 %.

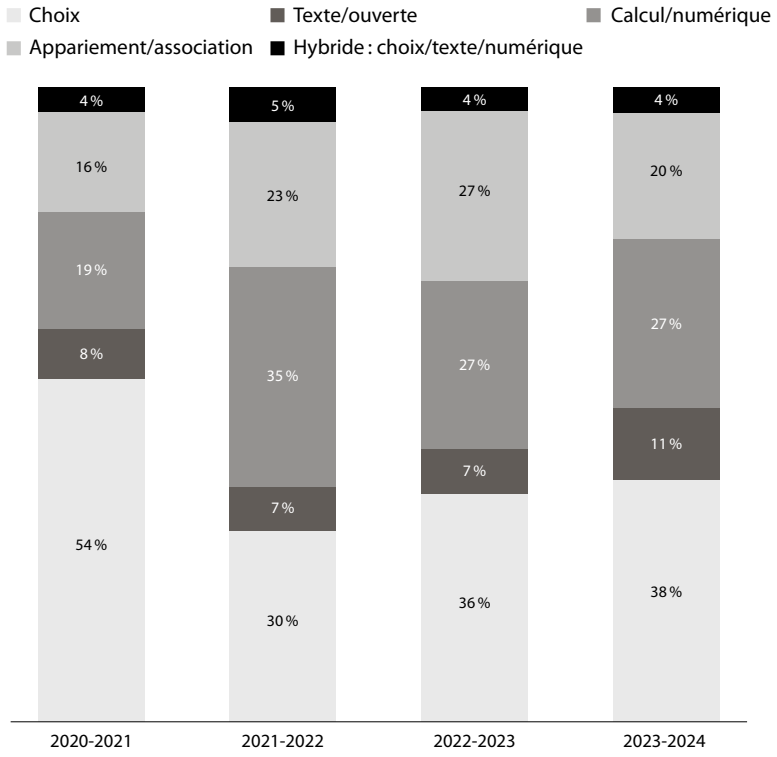


FIGURE 6.5 Répartition des types de questions pour les évaluations au format test et test hors ligne pour l'ensemble de l'Université par année académique.

On constate que les questions de type hybride sont très peu utilisées, ce qui peut principalement s'expliquer par la complexité pratique de leur rédaction qui nécessite de recourir à un langage spécifique à intégrer directement dans le texte des consignes.

L'analyse de ces tendances laisse penser que les nouveaux utilisateurs de l'e-évaluation post-Covid y ont sans doute plus recours pour les facilités offertes à l'enseignant (correction automatique, lecture du texte…) que pour développer des scénarios d'évaluation innovants. Toutefois, le taux relativement élevé de questions de type numérique et de calculs montre probablement un recours à l'évaluation des compétences intégrant celles de l'utilisation de logiciels, type R pour les statistiques, qui peut être considéré comme un aspect positif par la dimension de mise en situation ou mise en contexte authentique qu'il induit.

Analyse de l'étude post-Covid auprès des enseignants

La figure 6.6 montre les résultats (en pourcentage de réponses) concernant les obstacles identifiés par les enseignants qui les empêchent de passer un examen en mode numérique. L'analyse de ces résultats révèle un schéma intéressant dans les obstacles perçus par les enseignants. La période Covid semble avoir eu un impact déterminant sur les barrières technologiques initiales. En effet, ni le manque de familiarité avec l'informatique (zone 1, vert foncé), ni l'utilisation de Moodle pour la production, la réalisation et la correction des examens (zone 2, vert clair) ne constituent des obstacles significatifs. Cette absence de réticence technologique s'explique probablement par l'expérience forcée pendant la pandémie, qui a contraint même les plus réfractaires à maîtriser ces outils. En revanche, des blocages persistent concernant des aspects plus techniques de l'exécution des examens sur le campus (zone 3, rouge clair), suggérant un besoin de soutien spécifique pour ces aspects pratiques. Enfin, la question de la tricherie (zone 4, rouge foncé) reste une préoccupation majeure, probablement renforcée par l'expérience des examens à distance pendant le Covid.

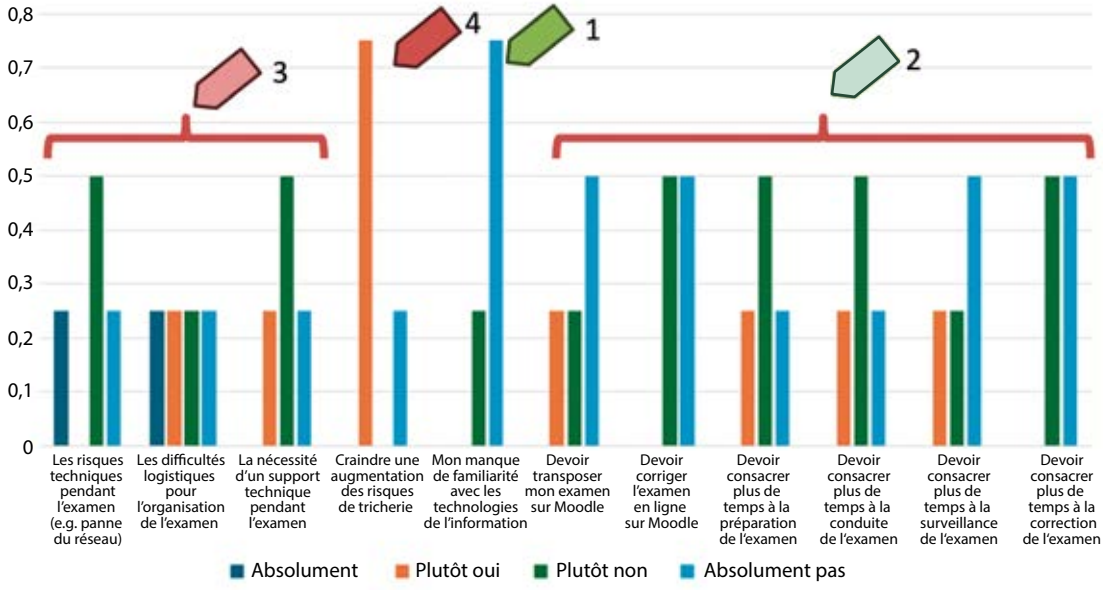


FIGURE 6.6 Réponses des enseignants concernant les obstacles à la transition des examens vers le mode numérique.

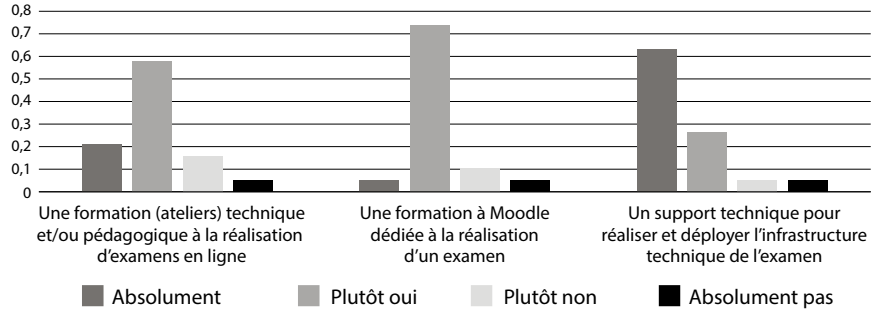


FIGURE 6.7 Réponses des enseignants concernant les besoins complémentaires à l'infrastructure existante qui les inciteraient à passer au mode numérique.

La figure 6.7 illustre les réponses en pourcentage des enseignants aux questions sur leurs besoins pour les accompagner vers le mode numérique. Les propositions faites en matière de formation et de soutien technique sont plutôt souhaitables, mais c'est la proposition de soutien technique pour le déploiement et la réalisation de l'examen qui est la plus attendue. Cela fait écho à la zone 3 de la figure 6.6 concernant les obstacles liés à la réalisation de l'examen, à savoir un appui technique pour rassurer les enseignants et leur apporter une assistance technique.

En ce qui concerne les raisons de passer au mode numérique, il y a un groupe pour lequel la somme des réponses « Absolument » et « Plutôt oui » est de loin majoritaire et par ordre de préférence :

- adapter les méthodes d'évaluation à l'évolution numérique de la discipline et à la réalité du terrain ;
- supprimer les copies papier ;
- réponses dactylographiées lisibles ;
- automatiser le marquage ;
- correction à distance ;
- correction par plusieurs personnes ;
- faciliter la vérification du plagiat.

Il y a ensuite un deuxième groupe de motivations pour lesquelles la majorité n'est pas exprimée, dans l'ordre de la plus exprimée à la moins exprimée :

- fournir un accès aux ressources en ligne ;
- anonymiser les copies ;
- utiliser des contenus multimédias dans les questions.

Cette division en deux groupes suggère que l'intérêt des enseignants pour les examens numériques reste dans les schémas traditionnels, ou dans un contexte qui facilite la réalisation de l'examen. Les utilisations qui pourraient conduire à des méthodes innovantes, telles que les ressources en ligne et le contenu numérique, sont moins prises en compte.

Lors des entretiens individuels, une synthèse des avantages mis en avant par les enseignants concernant la méthode institutionnelle (examens réalisés sur le campus avec l'équipement numérique de l'Université) permet de dégager les points suivants :

- normalisation de l'interface et des logiciels utilisés par les étudiants ;
- le logiciel utilisé peut être contrôlé et sécurisé à l'aide de SEB ;
- des statistiques d'examen peuvent être obtenues ;
- logistique simplifiée grâce à la salle dédiée (cette méthode dépend toutefois de la disponibilité des salles équipées) ;
- permet d'évaluer les compétences des étudiants sur le logiciel enseigné en classe.

En ce qui concerne le BYOD (sur le campus avec l'équipement numérique des étudiants), les avantages sont les suivants :

- les étudiants sont plus à l'aise avec leur propre matériel ;
- les examens peuvent être organisés au sein de l'Université, avec un plus grand choix de salles ;
- les étudiants ont accès à leurs documents pour les examens à livre ouvert ;
- pour le mode à distance, les enseignants citent :
  - une plus grande flexibilité pour les étudiants et la possibilité de répondre à certains problèmes d'accessibilité et de handicap,
  - un avantage géographique, puisqu'il n'est pas nécessaire de faire venir les étudiants sur le campus pour passer les examens.

Enfin, la méthode du quiz hors ligne (un test préparé et corrigé à l'aide de Moodle, mais réalisé sur papier par les étudiants) présente un intérêt particulier pour les enseignants :

- correction rapide ;
- correction automatisée ;
- permet d'administrer rapidement un examen à un grand nombre d'étudiants.

On constate que les répondants sont convaincus des bénéfices de l'utilisation de l'e-évaluation et de son potentiel. Mais il y a aussi des besoins identifiés et orientés vers une infrastructure dédiée, soulignant les défis logistiques à relever. L'e-évaluation est perçue comme une opportunité, favorisant la réflexion sur les méthodes d'évaluation et incitant à repenser les pratiques pédagogiques, même si l'on constate que ce n'est sans doute pas l'objectif premier de l'utilisation du numérique.

Le questionnaire (voir figure 6.7) révèle les besoins de soutien et les entretiens soulèvent également des obstacles :

- nécessité d'une infrastructure dédiée et disponible – salles, équipements, logiciels, etc. ;
- nécessité de lignes directrices spécifiques en matière d'évaluation électronique ;
- renforcement des ressources existantes, en particulier du personnel spécialisé.

En ce qui concerne la nécessité d'une infrastructure dédiée, on assiste actuellement à une augmentation de l'utilisation des examens institutionnels malgré le manque de salles équipées disponibles. Plus cette méthode se popularise, plus elle nécessite des investissements lourds, notamment en salles équipées. Face à cette situation, la faculté de droit a décidé de promouvoir l'utilisation du BYOD, qui permet aux étudiants d'avoir accès à des salles non équipées. Mais le BYOD est beaucoup plus délicat à mettre en place, car il nécessite des expérimentations poussées pour optimiser la configuration de ce mode d'évaluation et le rendre acceptable par les enseignants.

Analyse des pratiques innovantes

Les analyses statistiques et l'étude auprès des enseignants révèlent une adoption croissante, mais prudente de l'e-évaluation, avec une préférence pour des modalités qui minimisent les risques techniques tout en offrant des avantages pratiques comme la correction automatisée. Cependant, certains enseignants, correspondant aux « innovateurs engagés » du modèle TARC, développent des pratiques explorant plus pleinement le potentiel pédagogique du numérique. L'analyse de 20 scénarios de cours innovants, répertoriés dans le portail des innovations pédagogiques de l'Université, permet de compléter notre compréhension de l'évolution des pratiques en illustrant comment

l'e-évaluation peut transformer en profondeur les approches pédagogiques. L'analyse de ces scénarios innovants montre que ces enseignants dépassent largement le cadre des tests automatisés qui dominent les statistiques globales.

Les dispositifs d'évaluation numérique se déclinent sous plusieurs formes complémentaires. Les quiz automatisés sur Moodle constituent une pratique largement utilisée tant pour l'évaluation formative que sommative, souvent enrichie par des systèmes de bonus pour encourager un travail régulier. L'évaluation par les pairs s'est également développée, s'appuyant sur des grilles critériées en ligne et des systèmes d'évaluation anonyme. Le travail collaboratif occupe une place prépondérante à travers différents formats : création collective de contenus pédagogiques (vidéos, podcasts) en petits groupes, développement de wikis collaboratifs, et de projets interdisciplinaires. Ces activités sont généralement accompagnées d'évaluations intermédiaires et de séances de coaching. Les productions numériques créatives, qu'elles soient individuelles ou collectives, permettent d'évaluer la compréhension approfondie des concepts. Les portfolios électroniques servent à documenter la progression des apprentissages, tandis que les outils de vote interactif permettent des évaluations formatives en temps réel.

Selon les retours des enseignants qui les ont mises en place, ces pratiques d'évaluation numérique présentent des avantages pédagogiques significatifs. Ils considèrent qu'elles favorisent un apprentissage actif et la construction des connaissances à travers un engagement plus constant des étudiants. La production de contenus numériques permet à ces derniers de démontrer leur compréhension tout en développant des compétences transversales. L'évaluation par les pairs et l'autoévaluation renforcent les capacités réflexives et critiques des étudiants, améliorant leur autonomie et leur jugement évaluatif. Le caractère régulier et formatif des évaluations facilite un suivi précis de la progression et permet des interventions pédagogiques ciblées.

Cependant, ces pratiques soulèvent plusieurs enjeux importants. La question de la validité et de la fiabilité est centrale. L'évaluation des productions créatives et collaboratives pose des défis spécifiques en matière de critérisation, nécessitant des grilles d'évaluation détaillées et une explicitation claire des attentes. Les portfolios électroniques, bien qu'ils offrent une évaluation plus authentique, requièrent des outils robustes pour garantir la validité des jugements.

La protection des données personnelles constitue un autre enjeu majeur, en ce qui concerne notamment les traces d'apprentissage, les productions des étudiants et les données d'interaction. L'équité et l'accessibilité doivent également être prises en compte, tant au niveau de l'accès aux outils numériques que des compétences technologiques requises. La propriété intellectuelle des productions étudiantes et les questions de confidentialité dans l'évaluation par les pairs nécessitent un encadrement clair.

Conseils de mise en œuvre pratique

L'analyse de l'évolution de l'utilisation des examens numériques à l'Université de Genève avant, pendant et après la pandémie de Covid-19 montre clairement que la crise a eu un impact important dans la période post-pandémique et qu'il n'est pas question de revenir à la situation antérieure. En complétant cette analyse par une étude auprès d'un échantillon d'enseignants, il est possible de mieux comprendre l'impact et l'effet de ce qui s'est passé pendant le Covid et d'éclairer les usages actuels.

Nous pouvons supposer que la confrontation obligatoire de tous les enseignants avec les examens en ligne pendant la pandémie, même ceux qui étaient le plus opposés à cette modalité, a provoqué un changement dans leur attitude et leur pratique à l'égard des examens en ligne. Cela leur a permis d'expérimenter, d'acquérir de l'expérience et d'analyser les points positifs et négatifs. Si nous comparons ce résultat avec le modèle TARC (Mottiar *et al.*, 2024), nous pouvons expliquer le changement dans le recours aux évaluations numériques entre l'avant et l'après-Covid par le fait que les « innovateurs engagés » ont été rejoints par des « innovateurs opportunistes » et peut-être aussi par quelques « réactifs adaptatifs ».

Les résultats de la figure 6.4 montrent pourquoi les méthodes institutionnelles et *offline* sont les plus performantes dans la période post-Covid. Ce sont les deux méthodes qui offrent le plus d'avantages et le moins d'inconvénients redoutés par les enseignants. Par exemple, la méthode *offline* évite les risques qui correspondent aux raisons de la zone 3 de la figure 6.6, notamment les problèmes techniques qui peuvent survenir lors de l'examen. Cette méthode permet également de limiter les restrictions d'accès aux salles, puisque la modalité *offline* peut être réalisée dans n'importe quelle salle, quel que soit son

équipement. Elle correspond probablement à des enseignants qui sont passés de la modalité test à la modalité QCM (population dite « réactifs adaptatifs » selon la typologie du modèle TARC : « La peur du plagiat a rendu les questions d'examen de type essai problématiques, de sorte que les QCM à temps contrôlé ont été privilégiés » [Mottiar *et al.*, 2024, p. 107]) pendant le Covid et pour qui l'expérience, en particulier la rapidité de la correction, a été suffisamment positive pour adopter ce dispositif par la suite, mais sans les risques. La modalité institutionnelle représente un niveau plus risqué, mais atténue encore les risques et les bénéfices de l'e-évaluation pour une population d'enseignants qui ont été favorablement motivés par la période « forcée » du Covid (probablement parmi les « innovateurs opportunistes »). Dans la pratique, on constate que le succès croissant du mode institutionnel pose un problème : il atteint les limites des ressources disponibles concernant les salles équipées. Cela impliquera probablement un changement à l'avenir, la faculté de droit ayant décidé par exemple d'expérimenter le BYOD en 2025 afin de pouvoir l'offrir en complément du mode institutionnel pour accompagner avec succès le développement de l'usage de l'e-évaluation. Les résultats des réponses aux questions présentées dans la figure 6.5 confirment que l'institution devrait continuer à déployer des moyens et des ressources si elle veut soutenir et encourager le développement de la pratique de l'e-évaluation qui a émergé à la suite de la période Covid.

Enfin, selon les réponses aux questions sur les motivations des enseignants, les aspects innovants (accès aux ressources en ligne, utilisation de contenus multimédias, etc.) restent secondaires, mais la question de l'authenticité des examens et de leur adéquation avec la réalité du terrain est une considération essentielle pour les enseignants. C'est pourquoi, à côté des équipements et infrastructures nécessaires, il est indispensable de produire des cadres pour orienter les choix pédagogiques des enseignants. Cette demande est également ressortie lors des entretiens individuels. Des cadres de ce type existent, comme celui proposé par Pauli et Ferrell (2020) qui définissent cinq principes pour les examens futurs : authentique, accessible, automatisé de manière appropriée, continu et sécurisé :

- l'évaluation authentique : les pratiques innovantes illustrent ce principe à travers la création collective de contenus pédagogiques et les projets interdisciplinaires qui préparent les apprenants à des scénarios du monde réel ;

- l'évaluation accessible : les portfolios électroniques et la diversité des formats d'évaluation (vidéos, podcasts, wikis) permettent de s'adapter aux différents profils d'étudiants ;
- l'évaluation automatisée appropriée : au-delà de la simple correction automatique des QCM, les pratiques innovantes montrent comment combiner l'automatisation (quiz formatifs, systèmes de vote) avec des évaluations plus complexes (grilles critériées en ligne pour l'évaluation par les pairs) ;
- le contrôle continu : les innovateurs engagés démontrent son efficacité à travers les évaluations intermédiaires, le coaching régulier et le suivi via les portfolios ;
- l'évaluation sécurisée : les pratiques innovantes apportent une nouvelle perspective en privilégiant des formats d'évaluation qui rendent la fraude moins pertinente, comme les productions créatives personnalisées.

Par ailleurs, on constate au travers des pratiques innovantes en particulier que la démultiplication des productions de traces et de contenus numériques pose la question d'un sixième principe, la définition d'un cadre éthique élargi qui protège les étudiants non seulement par rapport à l'accès, mais aussi à la réutilisation de leurs productions numériques. On peut penser qu'un tel cadre faciliterait la transition des populations enseignantes de chacune des trois premières catégories du modèle TARC vers des catégories plus ambitieuses.

En conclusion, nous avons présenté dans ce chapitre l'évolution du dispositif institutionnel de soutien aux examens numériques mis en place à l'Université de Genève, ainsi que celle de l'utilisation de cette méthode d'examen par les enseignants au niveau de l'ensemble de l'Université et en analysant plus finement les innovations pédagogiques les plus récentes et significatives. En y associant une étude auprès d'un échantillon d'enseignants, nous avons pu apporter un éclairage sur les évolutions des pratiques d'examen en lien avec la pandémie de Covid-19. Nous avons également pu esquisser quelques tendances globales et innovantes et identifier des besoins pour soutenir le développement de l'e-évaluation au niveau institutionnel.

Evi Belsack & Patricia Mercier

7 Évaluations en ligne post-Covid : quels accompagnements et quel alignement pédagogique ?

Résumé

Depuis la pandémie de Covid-19 en 2020, l'Université libre de Bruxelles (ULB) a vu son rapport au numérique transformé. La plateforme Moodle, également connue sous le nom d'« Université virtuelle », a été fortement utilisée pour diffuser et évaluer les cours pendant les périodes de confinement. Cinq ans après, malgré un retour à la normale, de nombreux et nombreuses enseignants et enseignantes continuent d'utiliser cette plateforme pour les évaluations certificatives, soulevant des questions sur le soutien technique et pédagogique actuel. Pendant la pandémie, le Centre d'appui pédagogique de l'ULB a redoublé d'efforts pour former les enseignants et adapter la plateforme à une utilisation massive. Aujourd'hui, bien que les examens en ligne persistent, l'accompagnement a diminué, suggérant une meilleure maîtrise des outils par les enseignants. Cependant, des questions subsistent sur la pertinence de maintenir ces pratiques après la pandémie et sur leur impact sur l'apprentissage. Nous proposons une analyse quantitative des données des derniers examens pour examiner ces questions et les mettre en débat.

Introduction

La pandémie de Covid-19 a profondément transformé (Jorro, 2021 ; Parent *et al.*, 2023) les pratiques pédagogiques et les modalités d'évaluation dans les établissements d'enseignement supérieur à travers le monde. À l'Université libre de Bruxelles (ULB), cette période a marqué un tournant décisif dans l'utilisation des outils numériques pour l'enseignement et l'évaluation : l'ensemble des enseignants ont dû utiliser ces derniers, même si l'empan de cette utilisation est assez large (de l'utilisation du courriel à des dispositifs en ligne intégrés). La plateforme Moodle, également connue sous le nom d'« Université virtuelle » à l'ULB, est devenue un pilier central pour la diffusion et l'évaluation des acquis d'apprentissage durant les confinements. Cinq ans après, alors que les contraintes sanitaires ont disparu, de nombreuses et nombreux enseignantes et enseignants continuent de privilégier les évaluations écrites en ligne, soit par le biais de dépôt de documents (« devoir » Moodle), soit sous forme de questionnaire (« test » Moodle). Cette persistance soulève des questions importantes sur la mise en place adéquate de ces activités et l'accompagnement technique et pédagogique qui la soutient, ainsi que sur l'impact de ces pratiques sur l'apprentissage des étudiants.

Ce chapitre propose une analyse des données récentes concernant les examens en ligne à l'ULB. Il explore les aspects techniques, technopédagogiques, éthiques et psychologiques de ces pratiques, en s'appuyant sur une méthodologie tenant compte à la fois des données disponibles et de données récoltées auprès des enseignants.

Nous avons examiné les données disponibles concernant le nombre moyen d'étudiants ventilés selon les facultés et les avons mises en regard de l'utilisation des tests et devoirs dans l'Université virtuelle (UV). Cette analyse a permis de mesurer l'évolution de l'utilisation de ces activités avant et après la pandémie.

Un questionnaire court (10 questions) a été soumis aux enseignants utilisant les activités « test » ou « devoir » pendant les sessions d'examen. Ce questionnaire comprenait des questions fermées (choix multiple ou réponse multiple) et des questions ouvertes, permettant de recueillir des données qualitatives sur les raisons de l'utilisation persistante des évaluations en ligne.

En examinant les raisons pour lesquelles ce type d'évaluation perdure, malgré le retour à l'enseignement en présentiel, ce texte vise à

éclairer les enjeux et les défis liés à l'alignement pédagogique et à l'accompagnement des enseignants et des étudiants dans ce nouveau contexte post-pandémique.

Contexte et problématique

Cinq ans après le début de la pandémie, la vie a repris sur les campus. Bien que les contraintes sanitaires aient disparu, certains et certaines enseignants et enseignantes ont maintenu les évaluations certificatives en ligne, principalement au travers des devoirs et tests dans Moodle. L'objet de ce texte est de comprendre pourquoi et dans quelles conditions.

Avant cela, il nous semble nécessaire d'exposer la situation actuelle selon les axes développés dans la littérature pour décrire les dispositifs existants ou ayant été instaurés pendant le Covid, soit les aspects techniques, technopédagogiques, pédagogiques, éthiques et psychologiques en situation d'urgence.

La mise en place en urgence de différents palliatifs a permis de répondre, pour une grande part en tout cas, aux problèmes rencontrés en raison du confinement et de l'obligation de travailler à distance.

Actuellement, bien que la demande de formation soit moins intensive – il n'existe d'ailleurs plus de formation centrée exclusivement sur les devoirs et les tests Moodle –, et que les déclarations d'examens en ligne soient moins fréquentes, le Centre d'appui pédagogique de l'ULB (CAP) continue, sur demande d'enseignants, de vérifier le paramétrage d'activités d'évaluation certificative dans Moodle pour assurer leur bon déroulement technique et d'apporter un support au moment de ou après l'épreuve.

Du point de vue de l'infrastructure, Moodle est hébergé sur des serveurs externes performants. Qui plus est, ceux-ci sont renforcés lors de chaque session d'examen.

Cependant, durant les périodes destinées aux examens, nous observons aussi l'utilisation de ces activités hors de tout support ou accompagnement de la part du CAP, et ce, pour un nombre important de cas.

Par ailleurs, d'autres évolutions liées au numérique, comme l'arrivée massive de l'intelligence artificielle générative (IAg), ont renforcé les problématiques associées au contrôle du plagiat, que ce soit sur les évaluations à questions ouvertes, telles qu'organisables dans le « devoir » Moodle, ou à questions fermées gérées par le « test » Moodle.

Comprendre la persistance des examens à distance

Nous cherchons à comprendre pourquoi les examens en ligne continuent d'être utilisés malgré le retour à l'enseignement en présentiel et la diminution ou disparition d'une partie des dispositifs d'accompagnement technopédagogiques (comme les formations ciblées), alors que le questionnement lié à l'alignement entre objectifs, méthodes d'enseignement/d'apprentissage et méthodes d'évaluation semble être absent des préoccupations des enseignants et enseignantes dont ils et elles nous font part.

Nous émettons trois hypothèses quant aux raisons pour lesquelles les examens en ligne subsistent :

1. une solution à de nouvelles contraintes : l'augmentation du nombre d'étudiants et les difficultés logistiques, telles que le manque d'auditoire ou de salles suffisamment grandes pour accueillir tous et toutes les étudiants et étudiantes, ont rendu les examens en ligne plus pratiques ;
2. la montée en compétence : les enseignants et les étudiants ont acquis ou renforcé certaines compétences numériques, entraînant la persistance de formes d'évaluation désormais considérées comme maîtrisées, une sorte d'effet rebond du Covid ;
3. la découverte des possibilités et plus-values des modalités hybrides d'enseignement, notamment des outils d'évaluation Moodle ou, par le biais de ceux-ci, des gains du numérique dans l'enseignement : organisation logistique facilitée (centralisation des fichiers, communication des résultats), apprentissage différent et différencié, etc.

La pandémie de Covid-19 a bouleversé les modes d'enseignement et d'évaluation à l'Université libre de Bruxelles (ULB), comme dans toutes les universités à travers le monde. La littérature qui a émergé pendant et après le Covid a mis en avant différents axes qu'il était essentiel de prendre en compte : les aspects techniques, pédagogiques et affectifs liés au passage de l'enseignement à distance et l'urgence qui a contextualisé tous ces éléments. Nous les présentons ici avant de décrire leur issue pour l'ULB.

La transition du présentiel au distanciel n'a pu se faire que par le biais technologique avec des conséquences aussi bien pédagogiques que psychologiques pour les étudiants comme pour les enseignants.

D'un point de vue technologique, même si des plateformes de type *Learning Management System* (LMS, par exemple Moodle) existaient déjà depuis longtemps et qu'une partie des enseignants s'en servaient déjà aussi (Champoux et Hilman, 2022).

Le cadre théorique

La pandémie de Covid-19 a provoqué un bouleversement sans précédent dans l'enseignement supérieur, forçant une transition rapide vers l'enseignement à distance et nécessitant une adaptation majeure des pratiques d'évaluation, mais également des outils d'évaluation. Différentes nouvelles préoccupations ont surgi à ce moment.

Du point de vue des outils d'évaluation à distance, bien que plusieurs outils, tels que les LMS, aient déjà été intégrés dans le monde de l'enseignement universitaire (Yerly et Issaïeva, 2021), tous les enseignants et toutes les enseignantes n'y avaient pas recours. Le confinement a impliqué leur formation à l'utilisation de ces outils.

Des questions se sont posées concernant les étudiants et leur rapport au numérique, et également la fraude, remettant en question la fidélité de l'évaluation. La littérature a montré des différences significatives entre les étudiants pour l'accès au numérique, qu'il s'agisse d'ordinateurs ou d'accès fiable au réseau internet (OCDE, 2020 ; Duroisin *et al.*, 2021). Il n'est par ailleurs pas évident que tous maîtrisent les outils numériques, la chimère du *digital native* (Roy *et al.*, 2018) s'est révélée avec force.

Enfin, l'environnement relativement calme qui entoure les épreuves écrites à l'université n'a pu être garanti lors d'examens à distance, parfois dans des lieux exigus partagés, pouvant déconcentrer les étudiants (OCDE, 2020).

Tous ces éléments tendent à un accroissement des inégalités dans l'enseignement supérieur (Raaper et Brown, 2020 ; OCDE, 2020 ; Duroisin *et al.*, 2021).

La situation particulière à l'ULB selon trois dimensions

Aspects techniques

Pendant la pandémie, la mise en place des examens à distance s'est effectuée sur deux outils principaux, choisis selon la modalité écrite ou orale de l'examen. Alors que Moodle a été investi pour les examens

écrits, Teams (Microsoft) a été choisi pour les examens oraux. Dans le cadre de ce texte, nous nous sommes concentrées sur les examens écrits, donc sur les éléments mis en place sur la plateforme Moodle.

Du point de vue technique, la pandémie a imposé le passage de tous les cours de l'Université à l'enseignement à distance, ce qui a nécessité une infrastructure informatique solide. Les serveurs ont dû être en mesure de supporter une charge accrue, ce qui a nécessité de recourir dans l'urgence à un prestataire externe. Pour les deux premières sessions, toutes à distance (juin et septembre 2020), une seconde plateforme « spécial examen » a été dupliquée sur la base de celle où s'étaient déroulés les cours durant le semestre. Sur cette Université virtuelle « bis », les cours étaient vides de contenu, permettant aux enseignants de créer plus librement leur évaluation. Par la suite, et jusqu'à maintenant encore, l'ULB a opté pour un renforcement des serveurs où est hébergée la plateforme Moodle principale, uniquement durant les périodes officielles de session. Ces dispositifs ont été et sont indispensables pour soutenir une charge importante d'utilisation simultanée rencontrée lors des tests ou devoirs organisés en modalité synchrone.

Dans Moodle, l'activité « test » offre différentes possibilités : la création, la catégorisation et la centralisation de questions dans une banque. On compte environ 18 formats de question distincts. Bien qu'il soit possible de créer une question ouverte, ce sont les questions fermées qui sont majoritairement utilisées. Elles peuvent se présenter sous diverses formes, comme des questions à réponses multiples, des correspondances, des affirmations vraies ou fausses, des textes à trous avec ou sans options de réponse, des questions à réponses numériques, des questions avec calcul (avec ou sans choix de réponse et d'unités aléatoires), des glisser-déposer sur un texte ou une image, etc.

Du point de vue du paramétrage, il est possible de configurer la durée d'accès et de passage au test (temps imparti), la proposition aléatoire de questions et/ou des propositions de réponses, des restrictions d'accès de différents ordres. La correction est automatisée, puisque, au moment de créer les questions et les réponses, l'enseignant renseigne la ou les bonnes réponses. Différents barèmes peuvent être configurés.

De multiples combinaisons de ces options ont été mises en œuvre durant la pandémie dans l'objectif de ralentir la transmission de réponses, ou la collaboration, alors qu'un travail individuel était

requis. Une série de recommandations pour assurer la sécurisation des épreuves et une ergonomie maximale des activités pour les étudiants ont été proposées par le CAP. Par exemple, bien que les étudiants puissent écrire directement leur réponse dans l'interface de la question ouverte dans texte (texte en ligne), nous avons rapidement conseillé le dépôt de documents dans une activité devoir pour éviter la perte du texte en cas de défaut de connexion au réseau internet.

Il est possible également de fournir, pendant ou après l'achèvement du test, des commentaires sur les questions, fonctionnalité largement utilisée en prévision des visites de copies distancielles, et sur les réponses, ce qui permet une rétroaction plus détaillée et plus personnalisée.

De son côté, l'activité « devoir » permet à l'enseignant de donner une consigne à un moment convenu et aux étudiants de déposer un ou plusieurs fichier(s) de réponses, sous formats courants (texte, image, présentation) ou spécifiques, dans un laps de temps défini. L'enseignant corrige manuellement sous forme classique (valeur) ou sous forme de grille critériée (créée par ses soins) permettant de calculer automatiquement la note en fonction de la sélection des niveaux de critères atteints.

Enfin, au moment des épreuves, le personnel du CAP, tous métiers confondus, se relayait pour faire face aux incidents techniques ou administratifs qui pouvaient survenir. De leur côté, les étudiants pouvaient s'adresser à une ligne téléphonique pour les urgences (pendant un examen) et à une adresse électronique pour les demandes moins pressantes. De manière générale, il était fortement conseillé aux enseignants de proposer des épreuves « blanches » (non cotées) pour que les étudiants s'approprient le dispositif technique et s'y confrontent avant l'examen final.

Parallèlement, l'Université s'est inquiétée de mettre à disposition des lieux proposant un réseau internet correct et un environnement propice à l'étude et l'évaluation des apprentissages, voire fournir le matériel adéquat.

Aspects technopédagogiques

Sur le plan technopédagogique, l'Université a organisé des séances de formation en groupe afin de familiariser l'ensemble de la communauté universitaire aux plateformes Moodle (qui existait depuis plus de dix ans à ce moment-là) et Teams (nouvellement installé). Elle a proposé,

par l'intermédiaire de CAP, des ateliers spécifiques et des séances d'accompagnement individuel aux enseignants pour l'utilisation des devoirs et tests dans Moodle (de mars à août 2020 : 50 ateliers – de septembre 2020 à juin 2021 : 65 ateliers) avec, pour contenu : prise en main, configuration des paramètres, correction, consultation des résultats et prévention des incidents.

De plus, ces dispositifs spécifiques mettaient en évidence les implications pédagogiques sous-jacentes à ces activités, notamment l'alignement pédagogique (Biggs, 1996), c'est-à-dire la cohérence entre les objectifs des enseignants en matière de compétences et de contenus pour les étudiants, les méthodes utilisées pour faire acquérir ces compétences et la manière dont elles sont évaluées.

Ainsi, l'équipe du CAP a non seulement familiarisé les enseignants aux aspects technopédagogiques portant sur l'utilisation de l'outil Moodle, mais elle a aussi sensibilisé aux implications pédagogiques que peut avoir l'utilisation de ces outils. Que l'enseignante sollicite l'accompagnement individuel des conseillers technopédagogiques (CTP) pour vérifier la cohérence et l'intérêt pédagogique des questions ou pour s'assurer des différents paramètres techniques, l'un permettait d'aborder l'autre.

Schématiquement, l'usage des questions fermées, principalement utilisées dans l'activité « test », ne permet pas une évaluation de compétences aussi poussée que celle permise par les questions ouvertes, telles que les essais, les dissertations, les travaux de recherche ou les devoirs sur table, mieux prises en charge dans l'activité « devoir ».

Parallèlement, parmi les enseignants qui ont opté pour les évaluations écrites dans Moodle à l'occasion de cette utilisation intensive et contrainte (partiellement, d'autres options comme Teams ou le courriel restaient possibles), une part importante a non seulement découvert, mais aussi s'est réellement approprié ces formes d'évaluation en ligne. Plus largement, ce sont les possibles modalités hybrides d'enseignement et leurs plus-values (Céci, 2022) qui ont été mises au jour, comme le gain de temps (par exemple, correction automatisée), la possibilité de proposer un rythme d'apprentissage et un *feedback* plus personnalisé malgré un nombre important d'étudiants, ainsi que des façons différentes d'apprendre (grâce, par exemple, à des types de questions moins classiques).

Aspects éthiques et psychologiques

Ces bouleversements dans les outils et les modalités d'examen, sur une période très courte, ont nécessité et justifié un soutien accru des enseignants. De notre côté, nous avons offert aux étudiants une interface et un paramétrage les plus rassurants possibles, tout en tenant compte des enjeux éthiques d'intégrité et d'originalité des travaux remis ou des tests effectués. En effet, le contrôle du travail des étudiants n'était plus exercé de manière habituelle. Au fil des sessions, en regard des résultats des évaluations certificatives à distance, certains et certaines enseignants et enseignantes ont fait preuve de beaucoup de créativité pour déployer un système de paramétrages complexes des activités afin de contrer les tentatives de tricherie et de maintenir un contrôle de l'évaluation.

À l'époque, l'institution s'est positionnée sur certains dispositifs en tenant notamment à préserver la vie privée de ses étudiants : les enseignants ne pouvaient pas exiger des étudiants qu'ils branchent leur caméra pendant la réalisation d'un examen écrit (se déroulant généralement à la maison).

La méthodologie

Afin de tester nos hypothèses, nous avons examiné les données disponibles concernant le nombre moyen d'étudiants ventilés selon les facultés et nous les avons mises en regard de l'utilisation des tests et des devoirs dans Moodle.

Nous avons également soumis un questionnaire en ligne aux enseignants qui, selon les données accessibles aux administrateurs de la plateforme, utilisent l'activité « test » et/ou « devoir » pendant une session.

À ce stade, il est nécessaire de préciser qu'en plus des possibilités offertes par Moodle (examen à distance ou en auditoire), les enseignants ont l'opportunité, s'ils le souhaitent, de recourir à un système classique de questions à choix multiples sur grille papier (avec correction automatisée).

Nous nous sommes concentrées sur les données d'avant Covid, soit l'année académique 2018-2019, et sur les données post-Covid complètes les plus récentes, soit celles de l'année académique 2023-2024.

À la lecture de la figure 7.1, on peut observer un accroissement général de la population estudiantine de 18 %, soit près d'un cinquième.

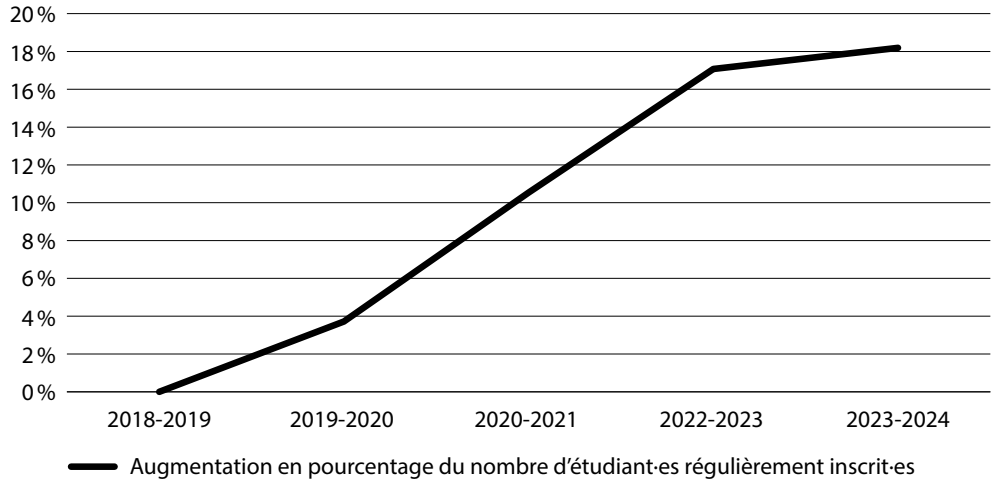


FIGURE 7.1 Évolution des inscriptions étudiantes.

La figure 7.2 montre à la fois l'évolution de l'utilisation des deux activités Moodle visées ainsi que sa répartition dans les différentes facultés. On observe que, dans certaines facultés, l'utilisation des activités était inexistante avant la pandémie et que, l'année dernière, toutes utilisaient, dans différentes mesures, l'activité « test » et/ou « devoir ».

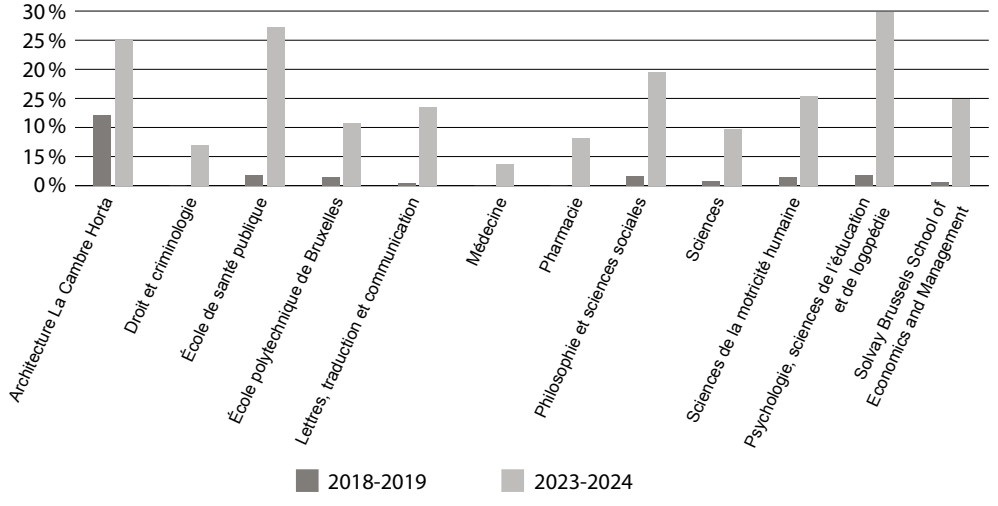


FIGURE 7.2 Évolution de l'utilisation des activités « devoir » et/ou « test » en 2018-2019 et en 2023-2024.

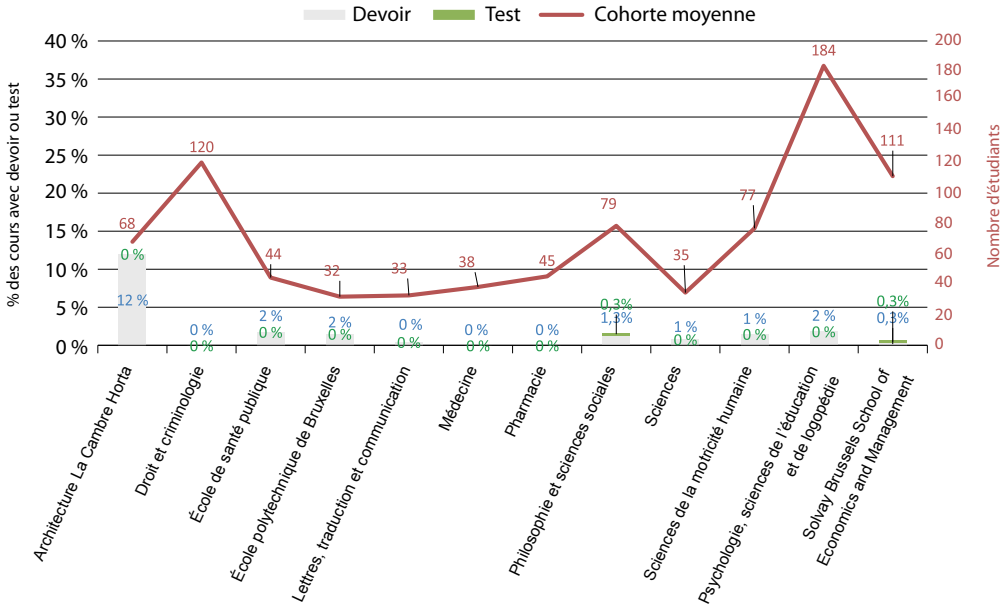


FIGURE 7.3 Cohorte moyenne et pourcentage d'utilisation de « test » et de « devoir » en 2018-2019.

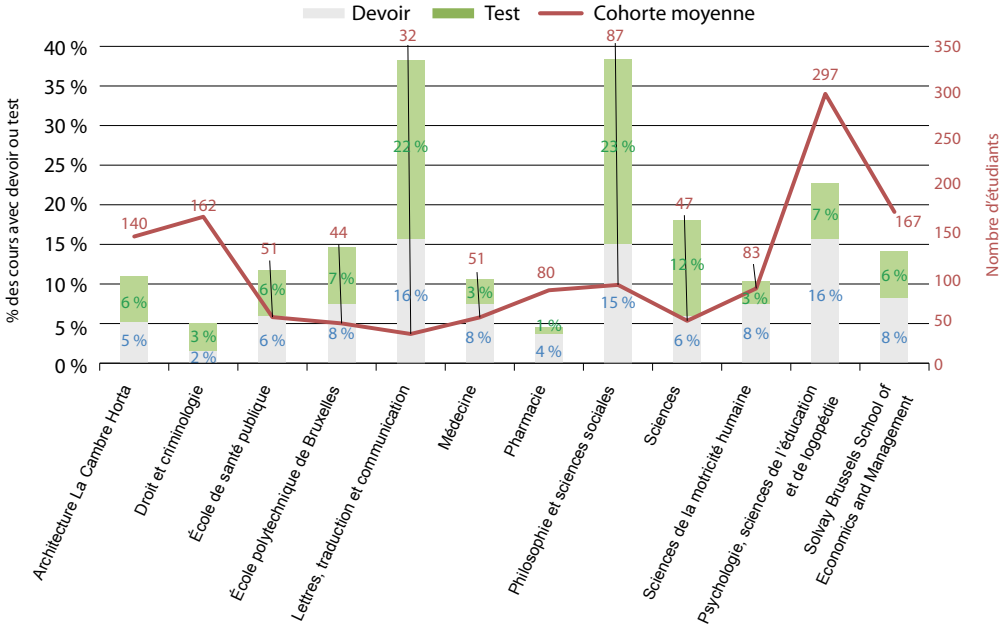


FIGURE 7.4 Cohorte moyenne et pourcentage d'utilisation de « test » et de « devoir » en 2023-2024.

Avant le Covid (figure 7.3), c'était la faculté d'architecture qui regroupait le plus d'activités « devoir » ou « test », avec une cohorte moyenne de 68 étudiants. Les facultés comptant une cohorte moyenne particulièrement élevée, telles que les facultés de sciences psychologiques et de l'éducation (PSE), de droit et criminologie (DC) et de philosophie et sciences sociales ainsi que la Solvay Brussels School (SBS) ne comptaient pas ou peu d'utilisation de ces mêmes activités.

Après le Covid (figure 7.4), toutes les facultés comptent des activités « test » et « devoir », mais, contre-intuitivement, ce ne sont pas les facultés qui ont vu leur population augmenter qui ont le plus développé leur utilisation des activités « test » et « devoir » : la faculté de lettres, traduction et communication totalise 16 % de cours utilisant les devoirs et 22 % utilisant les tests. C'est le score le plus élevé avec la faculté de philosophie et sciences sociales qui compte 15 % de devoirs et 23 % de tests, pour une cohorte moyenne de 87 étudiants alors qu'elle ne figure plus parmi les quatre facultés avec le plus d'étudiants en moyenne (dans l'ordre décroissant : PSE 297 ; SBS 167 ; DC 162 ; Archi 140).

La figure 7.5 montre bien à quel point notre hypothèse 1 ne se vérifie pas en fonction de la répartition du nombre moyen d'étudiants. Ce ne seraient pas les enseignants qui travaillent dans les facultés comptant les cohortes moyennes les plus élevées qui utilisent le plus les activités « test » et « devoir » pendant les sessions d'examen. Pourtant, certains et certaines enseignants et enseignantes témoignent, dans les réponses aux questions ouvertes, que le nombre d'étudiants justifie l'utilisation des activités « test » et « devoir ». Cette contradiction impose une réflexion plus approfondie sur le sujet.

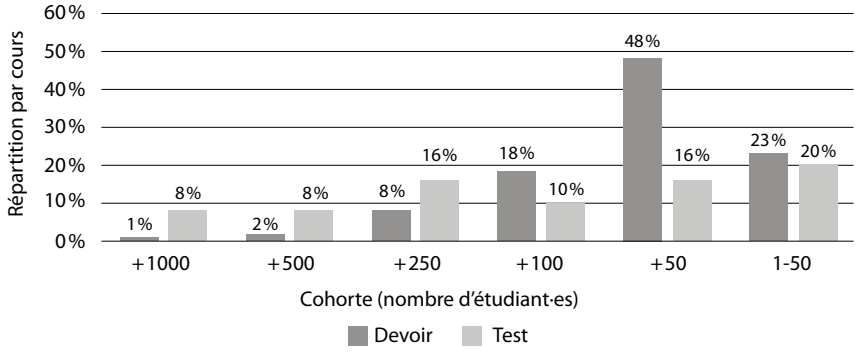


FIGURE 7.5 Répartition des cours avec « test » et « devoir » par taille de cohorte en période de session 2023-2024.

Toutefois, les données disponibles ne permettent pas de distinguer clairement les évaluations certificatives des évaluations continues. Ces dernières peuvent se répartir sur toute l'année, y compris durant les périodes de session. A contrario, il est vraisemblable qu'une partie résiduelle des évaluations certificatives se déroulent hors session.

En revanche, en ce qui concerne l'hypothèse selon laquelle le Covid a permis de révéler aux enseignants les possibilités des outils d'évaluation en ligne, les données après le Covid montrent qu'un changement s'est opéré.

Enquête en ligne

Afin d'approfondir cette hypothèse et de voir si elle se confirme, nous avons créé une enquête en ligne, puis nous l'avons soumise aux 624 enseignants qui ont recouru à un devoir ou à un test parmi les activités relevées sur le Moodle ULB pendant une session d'examen.

Nous avons opté pour un questionnaire court (10 questions), dont six questions fermées (choix multiple ou réponse multiple) et quatre questions ouvertes, espérant ainsi que les enseignants répondent au questionnaire sans se sentir trop sollicités, dans l'espoir qu'ils et elles soient nombreux et nombreuses à répondre.

Comme dans toute recherche, des biais ont été identifiés, et ce, malgré le soin que nous avons apporté à la construction de notre méthodologie. Les principaux sont les suivants : l'usage d'une enquête en ligne a peut-être empêché les enseignants moins à l'aise avec l'outil informatique de répondre. De ce fait, les enseignants et les enseignantes qui ont répondu pourraient avoir des opinions ou des expériences différentes de ceux et celles qui n'ont pas répondu.

Par ailleurs, des limites sont également à signaler. L'échantillon est non représentatif : avec seulement 10 % de taux de réponse, cet échantillon peut ne pas être représentatif de l'ensemble des enseignants de l'ULB. Cela limite la généralisation des résultats à l'ensemble de la population enseignante.

De plus, les informations recueillies par le biais du questionnaire sont autorapportées, ce qui peut entraîner des biais de mémoire ou de désirabilité sociale, où les répondants et les répondantes peuvent fournir des réponses qu'ils et elles pensent être attendues ou socialement acceptables.

Enfin, bien que des questions ouvertes aient été incluses dans le questionnaire, une analyse plus approfondie, sous forme d'entretiens individuels ou de groupes de discussion, pourrait fournir des *insights* plus riches et nuancés.

Les résultats et leur analyse

Malgré les biais et les limites évoqués plus haut, les informations que nous avons pu recueillir nous ont tout de même permis de faire quelques observations qui induisent un questionnement à propos de l'usage et surtout des raisons de l'usage des activités « test » et « devoir » dans Moodle durant les sessions d'examens. Comme nous l'avons mentionné plus haut, malheureusement, seuls 10 % des enseignants ont répondu à ce questionnaire, ce qui est sans doute une des conséquences de la surcharge de travail des enseignants au sein de notre institution, voire de l'enseignement supérieur en Belgique francophone. Bien que le peu de réponses ne nous permette pas d'effectuer des analyses statistiques poussées et généralisables, nous pouvons néanmoins avoir une idée plus précise des raisons de leur choix d'activités au travers des réponses aux questions ouvertes, ainsi que de l'aide dont ils ont bénéficié ou non, en fonction de leur réponse à une question fermée.

Au moins un enseignant ou une enseignante de chaque faculté a répondu au questionnaire, mais la répartition est inégale et ne correspond pas non plus à l'intensité de l'usage de Moodle pour les tests et les devoirs. Seule la faculté de philosophie et sciences sociales compte à la fois un grand nombre de répondants et une utilisation intensive de l'UV pendant les sessions d'examen.

De manière cohérente avec nos observations des données disponibles, par le biais de l'administration Moodle, ce ne sont pas les enseignants des cohortes les plus importantes qui ont répondu et déclarent utiliser cette ou ces activités (figure 7.6).

La majorité des enseignants (52 %) utilisent le « devoir » seul, soit la possibilité de remettre un travail sur différents supports, plus réfléchis et faisant appel à des compétences de haut niveau, telles que la critique, la comparaison (Bloom, 1956), contre un quart qui utilise le test ou la combinaison « devoir » et « test ».

L'une des questions ouvertes portait sur la raison pour laquelle les enseignants recouraient à cette modalité (le test) d'évaluation malgré le retour à l'enseignement sur le campus. Nous avons catégorisé

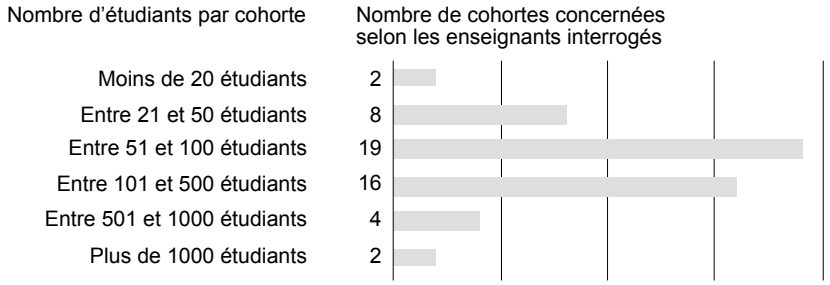


FIGURE 7.6 Nombre d'étudiants dans les cohortes visées par les activités « test » et « devoir » (2023-2024).

ces réponses dans une approche réduite de la théorisation ancrée telle qu'elle a été conçue par Glaser et Strauss (1995), puis reprise par Paillé et Muchelli (2021) : les raisons logistiques et les raisons pédagogiques. Il est intéressant de noter que seule une catégorie est en général repérable dans une réponse, et ce, malgré la complexité du problème de l'évaluation d'un cours et des choix qui sont posés pour celle-ci. Ainsi, en ce qui concerne les raisons logistiques, les enseignants évoquent le manque de places dans les auditoires, la taille des cohortes et le gain de temps de correction ainsi que la possibilité d'inclure des étudiants Erasmus retournés dans leur pays d'origine lors de la session.

Pour ce qui est de la catégorie « raisons pédagogiques », les enseignants évoquent la cohérence : lorsque le sujet porte sur l'enseignement ou la recherche d'éléments sur le Web, il paraît logique à l'enseignant de proposer un examen en ligne. L'idée de favorisation de l'évaluation continue est évoquée : la correction automatique permet de nombreuses évaluations sans pour autant surcharger les enseignants de corrections et favorise un *feedback* (disponible par réponse, par question ou globalement).

Enfin, elle permet l'entraînement des étudiants et des étudiantes aux types de questions auxquelles ils et elles seront confrontés à l'examen similaire sur papier, comme l'expriment ces enseignants :

> Cela permet d'établir une méthode d'évaluation homogène et cohérente pour des étudiants provenant de différents programmes de bachelier [...] réfléchir en amont à des questions à correction automatique critériées selon les objectifs pédagogiques du cours et sur base des apprentissages envisagés me permet de gagner beaucoup de temps à la correction. Les

questions à réponse automatique sont combinées avec des questions ouvertes permettant d'évaluer l'analyse critique sur base des apprentissages du cours.

Les étudiants sont évalués sur leurs capacités à utiliser des outils de recherche documentaire en ligne. Il est donc logique que le test se fasse en ligne.

Pour ce qui est des devoirs, les catégories sont du même ordre. D'un point de vue logistique, l'usage de Moodle permet de rassembler tous les travaux, d'avoir une vision claire de qui a rendu quoi et à quel moment, et d'effectuer une correction, mais il permet également de contourner la difficulté de rassembler les étudiants de différentes facultés pour un examen transversal dans un seul auditoire et dans le même créneau horaire. La possibilité de remettre des travaux de différents formats ainsi que celle de configurer les moments d'ouverture et de fermeture de l'activité, voire une facilité à lire en imprimé plutôt que de déchiffrer les écritures manuscrites, sont des aspects cités par les répondants.

Voici ce qu'ils et elles en disent :

[j'utilise l'activité devoir] :

- pour poster le support d'une présentation orale (PowerPoint, lien Canva) ;
- [parce que] le dépôt sur l'UV permet de scanner le travail remis par le logiciel de détection de plagiat ;
- [parce qu'elle] permet… de donner/faire les corrections et de rendre transparent[e] la grille d'évaluation ;
- [pour] traiter divers fichiers ;
- [pour] la clarté de la communication avec les étudiants : énoncé, date limite de remise.

D'un point de vue pédagogique, les enseignants évoquent des compétences différentes à évaluer : rédaction, critique, synthèse…

Les tests et devoirs Moodle proposent de nombreuses possibilités, raison pour laquelle les enseignants y font appel. L'hypothèse selon laquelle il y a eu prise de conscience sur l'utilité de ces outils d'évaluation est confortée par les 49 % des répondants à l'enquête déclarant ne pas utiliser ces outils avant la pandémie, mais également par certains commentaires :

- [La pandémie] m'a obligée à surmonter ma paresse et mes barrières naturelles et à découvrir les activités en ligne.
- Prise de conscience d'un intérêt de la part des étudiants pour des activités différentes.

- Je pense que, comme pour tout le monde, elle a remis en cause nos pratiques pédagogiques en permettant de mieux intégrer les technologies informatiques dans nos enseignements [...].

Notons bien que le test est plus difficile à paramétrer et organisé en mode synchrone dans le cadre des examens, ce qui engendre plus de stress chez l'enseignant que le devoir. Cette différence pourrait expliquer un besoin plus important d'accompagnement.

Pourtant, les constats sont clairs : le nombre de demandes d'accompagnement sur les outils d'évaluations, de vérification des épreuves et de soutien lors de celles-ci ont drastiquement diminué ces dernières années et il est interpellant de constater que l'accompagnement au choix de l'activité et l'accompagnement pédagogique, même rassemblés, restent nettement inférieurs à la demande de vérification de paramètres ou de résolution de problèmes d'ordre purement technique. Par ailleurs, 44 % (29, voir figure 7.7) des répondants à l'enquête déclarent ne pas avoir été accompagnés pour les activités devoir et test.

Dès lors, pourquoi les enseignants ne demandent-ils pas plus d'accompagnement, et particulièrement d'accompagnement pédagogique (conception, analyse des résultats) ou technopédagogique (choix, paramétrage) ? Sont-ils réellement montés en compétence au point de devenir autonomes ? Font-ils appel à d'autres ressources que celles du CAP ? Interrogent-ils encore l'utilité et les conséquences pédagogiques ainsi que l'ergonomie des activités disponibles dans Moodle ?

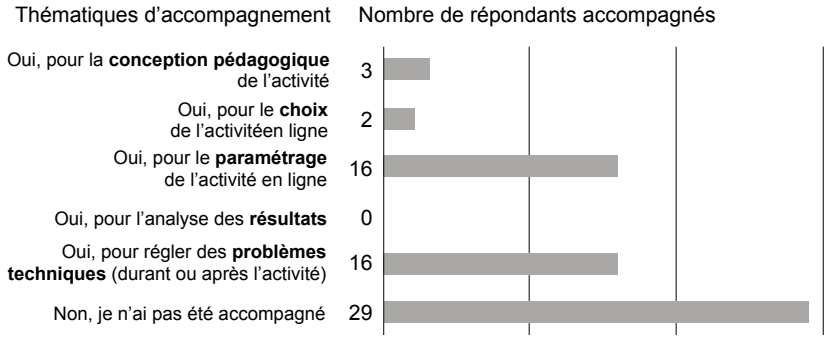


FIGURE 7.7 Avez-vous été accompagné pour les activités « devoir » et « test » sur Moodle ?

La montée en compétences techniques des enseignants au point de devenir autonomes peut être mise en doute par le nombre de répondants qui ont fait appel au CAP pour régler des problèmes techniques, durant ou après l'activité.

Néanmoins, nous pouvons postuler que la situation d'urgence leur a effectivement permis d'accroître leurs compétences à des degrés divers ou leur sentiment de maîtrise des activités en ligne, comme le montrent les verbatim suivants :

- J'ai appris à utiliser les activités « test » et « devoir ».
- J'ai généralisé leur utilisation et je me sens plus confiant dans leur usage, car j'ai testé de nombreuses fonctionnalités pendant la période Covid.

D'une part, notre enquête révèle qu'une grande proportion (59 %) des enseignants utilisent les tutoriels en ligne (ULB ou autre) pour mettre en œuvre ces activités. Bien que principalement orientée vers les aspects techniques (paramétrage, extraction des résultats, correction, etc.), notre documentation ULB tend également à attirer l'attention sur certains des savoirs pédagogiques en lien avec l'activité proposée. En effet, alors que, au CAP, nous considérons les aspects techniques et pédagogiques comme intimement intriqués, nous avons constaté qu'ils se distinguent chez les enseignants. Nous postulons que l'urgence de la pandémie a permis d'occulter cette intrication. Nous considérons également que le paramétrage technique constitue une bonne porte d'entrée pour accéder aux préoccupations pédagogiques, dont les enseignants ignorent parfois l'existence. Dès lors, le contenu proposé par les tutoriels suffit-il à une partie de notre public ? Le recours important à ce type de documentation à disponibilité facile peut aussi traduire une volonté d'autonomie logistique, en réaction à la difficulté d'organiser des examens en présentiel pour de grandes cohortes (résultant du manque de locaux et de surveillants), soit des raisons éloignées de considérations pédagogiques.

D'autre part, la diminution des demandes d'accompagnement interroge la visibilité et l'utilité ressentie de certains de nos services : pourrait-elle être altérée malgré la diversité des canaux de contact ? Nous disposons de pages internet dédiées à nos missions sur le portail du personnel de l'ULB. Néanmoins, ces informations ne sont accessibles qu'après authentification (non accessibles lors d'une recherche sur le

Web). Par ailleurs, les enseignants peuvent rapidement recevoir un support ou entamer un accompagnement (techno)pédagogique par de multiples moyens : deux adresses électroniques généralistes sont disponibles tandis que l'ensemble des conseillers CAP restent joignables directement. De plus, un canal Teams dédié à ces questions fonctionne sur le mode « entraide entre enseignants », même si les conseillers y interagissent également.

Conclusion et conseils de mise en œuvre pratique

Même si les données ne sont pas généralisables, elles permettent de lever le voile sur une partie des pratiques post-Covid en matière d'évaluation en ligne, particulièrement sur la plateforme Moodle, et sur les raisons qui poussent les enseignants à les faire perdurer ou évoluer.

Les réponses des enseignants ont ouvert d'autres pistes que nous n'avions pas imaginées : l'entraînement à des QCM en ligne, alors que l'examen se fait sur papier, la facilitation de l'évaluation continue grâce à la correction automatique et aux *feedbacks* intégrables, l'ouverture à de nouvelles pratiques pédagogiques intégrant le numérique malgré une mise au bain forcée et ensuite, une levée de cette contrainte.

Des raisons que nous avions anticipées ont également été rapportées lors de l'enquête et sont apparues lors de la catégorisation des réponses ouvertes des enseignants : dans un certain nombre de cas, la facilité logistique de quelques activités en ligne comme palliatif d'une massification intense de l'accès à l'Université et des difficultés qui en découlent telles que le manque de places dans les auditoires, les difficultés à trouver des créneaux horaires et des surveillants pour les examens lorsqu'ils se déroulent dans plusieurs auditoires ; une prise en main d'outils numériques qui n'aurait pas eu lieu dans des circonstances classiques, mais qui a contribué à une diffusion de ces activités.

Toutes ces raisons poussent à une réflexion pédagogique sur la manière dont on envisage l'évaluation. En effet, en début de chapitre, nous avons parlé de l'alignement constructiviste de Biggs (1996) qui a subi différentes mutations. Biggs, dans son article de 1996, affirme que, si les objectifs et les méthodes (le curriculum) sont alignés avec les manières dont l'étudiant est évalué, les résultats s'améliorent. Cependant, lors de la réflexion et de l'élaboration d'un cursus ou d'un cours, on réfléchit souvent aux objectifs, tel un idéal à atteindre, sans tenir suffisamment compte des contraintes qui pèsent sur les

enseignants, qu'elles soient liées aux conditions de travail ou à l'attitude de certains et certaines étudiants et étudiantes, soumis et soumises, eux et elles aussi, à différents aléas et difficultés. Or, la vie académique est pétrie de contraintes techniques (prise en main d'outils numériques), mais aussi bien souvent logistiques, telles que la taille des lieux d'examen, le nombre d'étudiants à évaluer. Ces contraintes pèsent lourd, particulièrement lors des premières années. Compte tenu de ces constats, nous posons la question de savoir s'il ne serait pas plus raisonnable et rationnel de commencer par circonscrire les contraintes d'évaluation et d'adapter les objectifs de cours en fonction de celles-ci. En effet, l'articulation dont parle Biggs postule bien une cohérence entre les trois éléments centraux du dispositif, mais elle ne contraint aucunement à commencer par l'un ou par l'autre. Il appelle à réfléchir et à concevoir les trois éléments de manière symbiotique, mais également, et c'est le propos de Romainville (2023) et de Lanarès *et al.* (2023), à intégrer ce système à un contexte qui lui donnera sa pertinence.

Par ailleurs, une autre grande crainte des enseignants est celle de la tricherie, ce qui remettrait en question l'authenticité des travaux réalisés par les étudiants et, par là, l'évaluation des compétences visées. Nous pensons qu'une solution possible consisterait à instaurer une évaluation coconstruite par les étudiants en utilisant les outils tels que l'IA, par exemple. Cette méthode favoriserait et cadrerait la collaboration, permettant de démystifier l'utilisation de l'intelligence artificielle générative (IAg) et d'en appréhender les limites, en répondant notamment à des questions d'exercices critiques.

Les résultats de nos investigations nous poussent à maintenir une vigilance constante à propos de la réflexion sur l'usage de la technologie et de l'envisager conjointement avec des préoccupations pédagogiques de base.

Bernadette Charlier
& Jean-Luc Gilles

Conclusion
Évaluer à l'ère numérique, entre potentialités, risques et responsabilités

À l'issue de cet ouvrage, il apparaît clairement que les transformations à l'œuvre dans les pratiques d'évaluation des apprentissages, portées par les innovations pédagogiques et numériques, ouvrent des perspectives stimulantes tout en soulevant de nouveaux défis. C'est dans cette tension entre potentialités, risques et responsabilités que s'inscrit l'analyse finale proposée ici, à la fois synthèse, mise en perspective et appel à une approche qualité en ingénierie des évaluations.

Tant l'introduction de cet ouvrage que ses contributions permettent d'appréhender les transformations en cours aujourd'hui dans les pratiques d'évaluation des institutions d'enseignement supérieur, des enseignants et des étudiants. En effet, comme le soulignent bien Emmanuel Sylvestre et Jean-Michel Jullien dans leur introduction, des transformations sont à l'œuvre dans toutes les composantes de l'évaluation, qu'elles soient pédagogique, docimologique, technologique ou éthique. Si une approche analytique permet d'envisager ces transformations pour chacune des dimensions, nous vous invitons à en appréhender les effets par une analyse transversale à la lumière des contributions de l'ouvrage : comment les transformations pédagogiques dans les pratiques d'évaluation interrogent-elles la docimologie ? Dans quelle mesure les critères de qualité docimologiques sont-ils appliqués, appelés à évoluer ou mis en brèche par les évolutions technologiques ? Dans quelle mesure le regard sur la performance, fenêtre

ouverte sur les compétences et connaissances, est-il transformé avec l'émergence des IA ? Quelles questions éthiques sont soulevées ? Sur cette base, les enjeux pour les pratiques des étudiants, des enseignants et des institutions pourront être envisagés.

Mais avant de procéder à cette analyse des contributions de l'ouvrage, rappelons quelques critères de qualité qui pourront servir de repères (Gilles et Charlier, 2020) :

1. validité : les évaluations doivent mesurer précisément ce que l'enseignant souhaite évaluer, en étant alignées avec les objectifs et le contenu enseigné, tout en permettant des inférences solides ;
2. fidélité : les résultats doivent être objectifs, avec un contrôle de la subjectivité des correcteurs, notamment grâce à des grilles d'évaluation critériées ;
3. sensibilité : les évaluations doivent refléter des phénomènes subtils et fournir des mesures précises des apprentissages ;
4. diagnosticité : les *feedbacks* doivent permettre un diagnostic précis des points forts et des points à améliorer, facilitant la régulation des apprentissages ;
5. équité : les apprenants doivent être traités de manière juste et sans discrimination ;
6. praticabilité : les évaluations doivent être réalisables dans des délais raisonnables avec les ressources disponibles ;
7. transparence : les processus et objectifs de l'évaluation doivent être clairement communiqués et compris par les apprenants ;
8. authenticité : les tâches proposées doivent avoir du sens pour les apprenants et être pertinentes par rapport à leur contexte ;
9. autoévaluation : l'autoévaluation est encouragée afin de permettre aux apprenants de prendre conscience de leurs progrès et des points à améliorer.

Si, comme le rappellent des contributeurs de cet ouvrage, il est nécessaire de prendre en compte les contraintes (techniques, humaines, financières) pour la mise en œuvre de ces critères de qualité, sur le plan institutionnel, les services d'appui à la pédagogie peuvent jouer un rôle central pour faciliter leur application en invitant l'ensemble des acteurs (y compris administratifs) à mettre en pratique les neuf étapes du processus attendues pour une gestion de la qualité des évaluations telles que décrites par Gilles et Leclercq (1995) et Gilles (2002, 2023).

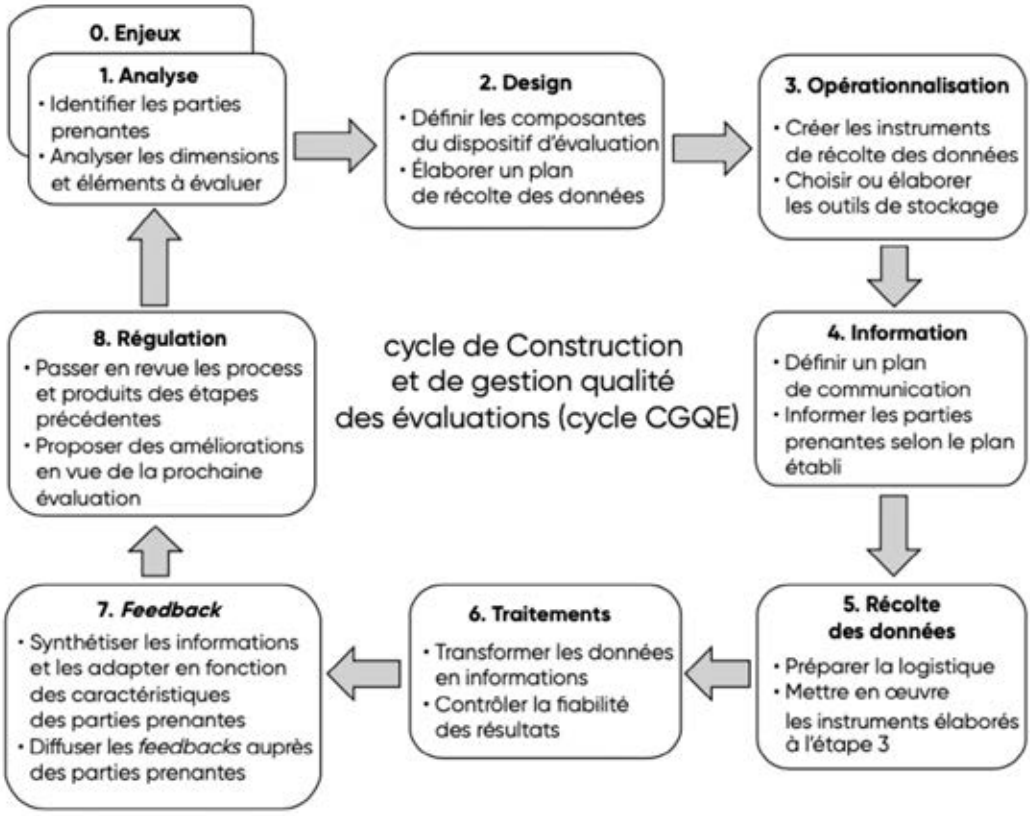


FIGURE 1 Schéma du cycle de Construction et de gestion qualité des évaluations en formation. Source : Gilles (2023).

En premier, les enjeux de l'évaluation peuvent être interrogés. Au-delà de la certification, il s'agit d'examiner comment l'évaluation des apprentissages et ses modalités servent le projet de l'établissement. Sa conception de sa fonction sociale ? Sa vision des individus qu'il forme ? Selon l'institution, l'évaluation des apprentissages relève-t-elle de la seule responsabilité de l'enseignant ? S'agit-il d'une activité réalisée par un collectif d'enseignants ? En quoi l'institution est-elle responsable ? En effet, l'existence de services d'appui à l'évaluation des apprentissages est conditionnée par l'adoption de réponses spécifiques. Dans de nombreux cas, bien que des services soient dispensés pour ce qui concerne la formation à l'enseignement ou son évaluation, il n'en va pas de même pour l'évaluation des apprentissages laissée entièrement à la responsabilité de l'enseignant. Une démarche plus collective s'avère pourtant

déterminante, en particulier pour ce qui concerne l'e-évaluation, pour garantir la qualité de celle-ci. Car, comme cela est évoqué dans l'ouvrage, des questions d'accessibilité à des services de qualité pour tous se posent, tout comme des questions de protection des données personnelles qui peuvent difficilement trouver des réponses adéquates au niveau des pratiques d'enseignants isolés, même de bonne volonté. Pour la suite, il est essentiel de souligner que les huit étapes suivantes du processus d'analyse à la régulation des évaluations pourraient être soutenues par des procédures, des instruments et des accompagnements adaptés, aussi bien pour les étudiants que pour les enseignants. Une telle démarche, accompagnée par des centres d'appui compétents, favoriserait également la capitalisation des bonnes pratiques et leur diffusion au sein et à l'extérieur de l'institution.

Avant d'en analyser les effets, examinons à présent les transformations à l'œuvre à la lumière des contributions du présent ouvrage.

Quelles transformations ?

Les contributions à cet ouvrage donnent à voir les transformations que connaissent ces dernières années les pratiques d'évaluation des apprentissages, en particulier à la suite de la crise du Covid-19.

Parmi les principales transformations pédagogiques identifiées, on observe :

- un élargissement des objets, lieux et temporalités d'évaluation : ouverture accentuée vers de nouveaux objets d'évaluation comme les compétences, ouverture et étendue à des espaces variés pour la réalisation des performances, notamment en lien avec le développement du BYOD (*Bring Your Own Device*), et la prise en compte de temporalités plus longues pour les apprentissages en développant des évaluations à différents moments du processus ;
- l'ouverture à de nouvelles sources de données comme l'évaluation de l'oral, pour l'apprentissage des langues en mode asynchrone ;
- une attention plus grande portée aux apprentissages émergents en y intégrant des formes d'apprentissage non formalisées par les curricula traditionnels ;
- un plus grand déploiement institutionnel de l'évaluation numérique, impulsé et amplifié par la pandémie de Covid-19 et ses impacts positifs pour la formation et l'accompagnement des enseignants et des surveillants ;

- au-delà de cette crise, le maintien de pratiques d'évaluations à distance (synchrone et asynchrone), même après le retour au présentiel ;
- une évolution des pratiques de *feedback*, favorisant un retour plus explicite et immédiat pour les domaines bien définis, et ouvrant la voie à l'utilisation d'outils d'IA comme ChatGPT pour soutenir l'autoévaluation ;
- le déploiement de dispositifs d'évaluation ouverts et continus tels que l'e-portfolio.

Ces mutations interrogent profondément les fondements de la docimologie et appellent à repenser les critères traditionnels de qualité, comme nous allons le voir dans la section suivante.

Quels effets ?

À la lumière des transformations identifiées, il importe d'en analyser les principaux effets. Nous nous intéresserons tout d'abord aux effets sur la qualité des évaluations, en mobilisant les critères fondamentaux rappelés précédemment, avant d'examiner ceux relatifs au développement des compétences des étudiants.

Effets sur la qualité des évaluations

Validité

Avec les transformations récentes, on ne se limite plus seulement à évaluer des savoirs académiques traditionnels (comme des connaissances théoriques), mais on élargit aux compétences, aux pratiques, aux attitudes, etc. Cet élargissement des objets évalués, notamment vers les compétences, ouvre la possibilité d'une meilleure validité des évaluations, en assurant un alignement plus étroit avec les finalités pédagogiques (Dupré et Giraudin, 2025). Toutefois, cette amélioration reste conditionnée à une définition explicite des objectifs et à un alignement rigoureux des modalités d'évaluation avec les objectifs. Dans le cas contraire, si l'élargissement des objets d'évaluation n'est pas rigoureusement aligné avec les objectifs pédagogiques, il risque de fragiliser la validité en introduisant des écarts entre ce qui est enseigné et ce qui est évalué. En d'autres termes, sans un alignement explicite, l'extension des objets évalués peut compromettre la validité en évaluant des dimensions peu ou mal reliées aux objectifs.

À cet égard, il faut, comme cela est fait dans certaines contributions, interroger le rôle spécifique des usages du numérique dans la transformation des performances des étudiants (Asri, 2024). Ainsi, l'interface d'une application de e-portfolio guide l'étudiant (ou le contraint) dans sa réflexivité ou, encore, la vidéo utilisée pour évaluer les compétences de production orale dans une langue étrangère évalue – dans certains cas – uniquement la compétence à lire à voix haute un texte écrit (Belgheddouche, 2025).

Fidélité

La diffusion de grilles d'évaluation critériées et l'automatisation de certaines corrections peuvent accroître la fidélité des évaluations. Cependant, la dépendance accrue à des dispositifs technologiques expose à de nouveaux biais, notamment ceux liés aux algorithmes (Audran, 2024). En effet, dans les dispositifs d'évaluation automatisée (par exemple via l'intelligence artificielle, le *machine learning*, ou d'autres algorithmes), il peut apparaître des biais spécifiques qui n'existaient pas, ou qui étaient moins présents, dans l'évaluation traditionnelle humaine. Trois exemples :

- biais dans les données d'entraînement : si un algorithme est entraîné sur des données qui reflètent des inégalités ou des stéréotypes (par exemple, linguistiques, culturels, socio-économiques), il risque de reproduire, et même d'amplifier ces biais dans ses évaluations ;
- biais d'interprétation : certains algorithmes d'évaluation automatique (par exemple, correction automatique d'essais ou d'oraux) peuvent privilégier certains types de formulations, de syntaxe ou de contenus qui ne correspondent pas nécessairement à la diversité des styles d'expression des étudiants ;
- manque de prise en compte du contexte : les algorithmes évaluent souvent sur la base de critères « objectifs » prédéfinis, mais ils peuvent passer à côté d'éléments contextuels, créatifs ou pertinents pour l'apprenant, que seul un évaluateur humain expérimenté serait capable de valoriser.

Par ailleurs, l'ouverture à des contextes variés pour l'évaluation des performances suscite des questions, notamment quant à l'équité : dans quelle mesure tous les étudiants sont-ils bien préparés à des situations qui leur sont étrangères ?

Sensibilité

L'introduction d'outils permettant de capter plus finement l'évolution des apprentissages, grâce à des évaluations multipliées et diversifiées, améliore la sensibilité des dispositifs. Cette évolution invite néanmoins à rester attentif à la pertinence des données collectées, pour éviter une accumulation d'informations peu exploitables (Romero, 2019). Un dispositif numérique peut enregistrer une multitude d'indicateurs pendant une activité d'apprentissage : le temps passé sur chaque question, le nombre de clics, les déplacements de la souris, les mots saisis, les changements de fenêtre, etc. Mais tous ces éléments n'ont pas la même pertinence pour évaluer l'acquisition d'une compétence ciblée. Le temps passé pourrait être trompeur (par exemple, un étudiant peut rester longtemps bloqué sans progresser). Les clics peuvent refléter des essais aléatoires sans réflexion… De telles données comportementales, si elles ne sont pas contextualisées, risquent d'induire des interprétations erronées et de détourner l'attention des indicateurs véritablement pertinents. Ainsi, collecter massivement des traces d'activité sans avoir défini ce qui est réellement informatif risque de saturer les enseignants ou les étudiants d'informations inutiles, de masquer les signaux pertinents et de diluer la capacité d'analyse.

Diagnosticité

La multiplication de *feedbacks* automatisés et personnalisés soutient le développement de dispositifs plus diagnostiques, capables d'orienter efficacement les apprentissages. Ces *feedbacks*, générés par des outils numériques ou des intelligences artificielles, peuvent offrir aux étudiants un retour immédiat sur leurs erreurs, leurs réussites, ou encore sur les stratégies mobilisées, ce qui facilite la régulation en temps réel de leurs apprentissages. Par exemple, un système peut identifier les étapes mal maîtrisées d'un raisonnement mathématique ou signaler des points de confusion récurrents dans un texte argumentatif.

Cependant, pour que ces dispositifs remplissent pleinement leur fonction diagnostique, encore faut-il que les *feedbacks* soient pertinents, clairs et adaptés au niveau de compréhension de l'étudiant (Calone et Lafontaine, 2018). Des retours trop génériques, trop techniques ou mal formulés risquent de perdre l'étudiant, voire de le décourager. De plus, lorsqu'ils sont générés par des systèmes automatisés, il devient crucial de soutenir leur interprétation : l'étudiant doit être

capable de comprendre ce que le *feedback* signifie, comment il a été produit, et surtout comment l'utiliser pour progresser.

Dans cette perspective, il est également utile de prévoir des dispositifs permettant aux étudiants de donner leur avis sur la pertinence et l'utilité des *feedback*s qu'ils reçoivent. Ce retour d'expérience contribue non seulement à soutenir l'engagement des apprenants, mais aussi à alimenter un processus d'amélioration continue des systèmes de *feedback*, en permettant à l'humain comme à l'IA d'ajuster, d'affiner et de fiabiliser progressivement les mécanismes d'évaluation automatisée.

Équité

L'accès facilité aux dispositifs d'évaluation par le biais des technologies numériques peut renforcer l'équité en réduisant certaines barrières.

Un dispositif numérique permet aux étudiants de passer des évaluations sans être physiquement présents dans un lieu spécifique. Cela réduit les inégalités liées à la distance, au coût et à la disponibilité des infrastructures (Laferrière *et al.*, 2022). Par exemple, un étudiant vivant loin de l'université, dans une zone rurale ou à l'étranger, peut accéder aux mêmes évaluations que ses pairs sans devoir se déplacer.

Certaines évaluations en ligne asynchrones offrent aux étudiants la possibilité de passer l'épreuve à un moment qui leur convient dans une plage horaire donnée. Cela bénéficie particulièrement aux étudiants ayant des contraintes spécifiques, comme un emploi, des responsabilités familiales ou des besoins particuliers liés à la santé.

Les outils numériques permettent également d'adapter plus facilement l'évaluation aux besoins particuliers :

- agrandir les textes pour les malvoyants ;
- proposer des énoncés audio pour les malentendants ;
- ajuster les temps impartis ;
- utiliser des logiciels d'assistance (magnificateurs d'écran, sous-titrage automatique, reconnaissance vocale…).

Ce type d'aménagements, essentiels pour l'équité, est plus facilement réalisable en environnement numérique. Cependant, de nouveaux risques d'iniquités apparaissent aussi (en particulier dans les contextes sud) comme cela est rappelé dans certains cas présentés dans cet ouvrage : accès inégal à des outils performants ; qualités variables des connexions internet ; déficit de compétences numériques.

Praticabilité

Le recours aux outils numériques simplifie certaines procédures et réduit la charge de correction pour les enseignants. Par exemple, la correction automatisée de questionnaires à choix multiples ou l'exploitation de plateformes de dépôt de travaux permettent de gagner du temps et de standardiser certaines tâches administratives. De plus, les outils numériques facilitent l'archivage des productions, le suivi des parcours étudiants et la production de statistiques, soutenant ainsi la gestion des évaluations à grande échelle.

Cependant, cette simplification locale s'accompagne d'une complexification globale de l'organisation des évaluations. La multiplication des interfaces, des formats de fichiers, des plateformes d'évaluation ou des outils complémentaires (*proctoring*, antiplagiat, e-portfolios, etc.) exige des enseignants une montée en compétences numériques constante et une capacité accrue à gérer des environnements techniques diversifiés. À cela s'ajoutent les risques techniques qu'il convient d'anticiper : instabilité des plateformes, problèmes d'accès, compatibilités entre systèmes, sécurité des données, ou encore cybersécurité en cas de fraudes. Ainsi, si les outils numériques peuvent alléger certaines tâches individuelles, ils imposent également une vigilance accrue en matière d'organisation, de support technique et de formation continue des acteurs (Lab'UA, 2023). On assiste dès lors à un déplacement de la charge liée à certaines tâches spécifiques chronophages (correction manuelle, gestion administrative) vers d'autres, car de nouvelles exigences apparaissent liées à la maîtrise des outils, à la sécurisation des dispositifs, à la gestion des incidents, etc. L'effort n'est donc pas vraiment supprimé, mais déplacé.

Transparence

La possibilité de rendre visibles les critères, les attentes et les processus d'évaluation s'en trouve accrue grâce aux technologies numériques. Les plateformes permettent en effet de partager facilement avec les étudiants des grilles critériées, des barèmes détaillés, des exemples de productions attendues ou des explications sur le fonctionnement des dispositifs d'évaluation. Cette visibilité renforce la compréhension des objectifs, la prévisibilité des attentes et l'engagement des étudiants dans leur apprentissage (Martínez *et al.*, 2015).

Toutefois, cette transparence peut être fragilisée lorsque les processus d'évaluation reposent sur des systèmes automatisés utilisant des algorithmes complexes. Dans ces cas, le mode de fonctionnement du système, les critères réellement mobilisés pour générer une note ou un *feedback*, ainsi que les raisons précises de certaines décisions restent souvent difficiles à expliciter, même pour les concepteurs du système. Cette opacité algorithmique peut créer une rupture dans la confiance des étudiants envers l'évaluation, limiter leur capacité à comprendre les résultats obtenus et, par conséquent, réduire les effets formatifs potentiels du dispositif.

Ainsi, si les technologies numériques offrent des opportunités importantes pour renforcer la transparence, elles appellent également à une vigilance accrue quant à la capacité des dispositifs automatisés à rendre leurs mécanismes compréhensibles, justifiables et ouverts à la discussion.

Authenticité

Les dispositifs favorisant le BYOD (*Bring Your Own Device*, c'est-à-dire l'utilisation par les étudiants de leurs propres appareils numériques comme outils d'apprentissage et d'évaluation) et les évaluations réalisées en contexte réel ou simulé offrent l'opportunité de proposer des tâches plus authentiques, ancrées dans des situations significatives. En mobilisant les outils que les étudiants utilisent dans leur quotidien ou leur futur environnement professionnel, ces dispositifs permettent de concevoir des évaluations plus proches de la réalité des pratiques et plus pertinentes pour les apprenants (Kozanitis, 2021).

Ces approches soutiennent une forme d'appropriation active des savoirs, en sollicitant non seulement les connaissances, mais aussi les compétences de mobilisation, d'adaptation et de jugement dans des familles de situations complexes qui ont du sens pour les apprenants. Toutefois, le défi demeure de concevoir des évaluations véritablement authentiques, qui ne se contentent pas d'imiter la réalité de manière superficielle. Une formalisation excessive des consignes, des contextes trop scénarisés ou un encadrement trop rigide, défauts pouvant être renforcés par certaines plateformes, peuvent conduire à une artificialisation des situations, vidant les tâches de leur potentiel formatif. Il s'agit donc de trouver un équilibre entre cadrage méthodologique et ouverture à la diversité des contextes et des réponses possibles, afin de préserver la richesse des apprentissages situés, et ce, sans perdre de vue l'importance d'apprendre avec les autres dans une communauté de pratique.

Autoévaluation

L'émergence d'outils soutenant l'autoévaluation, tels que les e-portfolios ou les solutions fondées sur l'intelligence artificielle, encourage le développement de l'autorégulation chez les apprenants. Ces dispositifs leur permettent de suivre l'évolution de leurs acquis, de documenter leurs progrès et de recevoir un retour individualisé sur leurs performances. Utilisés de manière judicieuse, ils favorisent une prise de conscience des points forts et des axes d'amélioration, condition essentielle à une démarche réflexive et autonome.

Pour que cette dynamique soit pleinement effective, elle nécessite toutefois un accompagnement explicite. Tous les étudiants ne disposent pas spontanément des compétences métacognitives nécessaires pour tirer parti de ces dispositifs. Il est donc essentiel de leur fournir des repères pour interpréter les *feedbacks*, structurer leur réflexion et orienter leurs actions correctives.

Dans cette perspective, certaines techniques pédagogiques peuvent jouer un rôle central pour renforcer l'autoévaluation, notamment l'usage de la technique des degrés de certitude (Leclercq, 1983) dans des dispositifs d'évaluations standardisées telles que des quiz en ligne. En invitant les étudiants à indiquer, pour chaque réponse donnée, leur niveau de confiance, on les amène à s'interroger systématiquement et de manière réflexive sur leurs connaissances. Cette pratique permet de croiser la justesse des réponses avec la conscience qu'ils en ont, révélant ainsi des zones de méconnaissance ignorées ou au contraire des connaissances assurées. Elle procure aussi à l'enseignant un matériau précieux pour identifier les réponses correctes, mais incertaines et les réponses erronées, mais données avec assurance (ce qui peut s'avérer dangereux, par exemple dans le secteur de la santé), facilitant un accompagnement plus ciblé. Cette approche permet également d'identifier chez les apprenants des tendances récurrentes à la surestimation (assurance excessive malgré des erreurs) ou à la sous-estimation (doute malgré des réponses correctes), offrant ainsi des informations précieuses pour développer des stratégies de remédiation adaptées aux profils métacognitifs de chacun. Encourager une autoévaluation active et outillée, en intégrant des indicateurs comme les degrés de certitude, contribue à faire évoluer les étudiants vers une posture d'acteurs de leur apprentissage, tout en enrichissant la fonction formative des évaluations.

Effets sur le développement des compétences des étudiants

À côté de leurs effets sur la qualité des évaluations, les transformations récentes influencent également de manière significative le développement des compétences des étudiants. Nous analysons ici les principaux effets repérables à travers quatre dimensions : les compétences transversales, les compétences numériques, les compétences métacognitives, ainsi que les risques d'accentuation des inégalités.

Renforcement des compétences transversales

La diversification des formes d'évaluation, rendue possible par les outils numériques, contribue au développement de compétences transversales essentielles pour la réussite académique et professionnelle (Redecker, 2017 ; SBFI, 2018). En mobilisant différentes modalités d'évaluation (projets collaboratifs, évaluations en contexte réel, autoévaluations, etc.), les dispositifs actuels sollicitent des compétences telles que l'autonomie, la capacité à s'autoévaluer, l'adaptabilité à des environnements variés et la gestion du temps et des priorités. La possibilité d'évoluer dans des situations ouvertes, parfois partiellement imprévisibles, amène les étudiants à renforcer leur capacité à mobiliser leurs acquis de manière souple et pertinente dans des contextes diversifiés.

Maîtrise des compétences numériques

L'intégration systématique d'outils numériques dans les dispositifs d'évaluation favorise l'acquisition et la consolidation de compétences techniques devenues incontournables dans de nombreux secteurs professionnels (OCDE, 2023). L'usage de plateformes de dépôt, de logiciels collaboratifs, d'outils de présentation numérique, de traitements de données ou encore d'espaces de travail en ligne contribue à développer chez les apprenants une aisance numérique et une capacité à utiliser efficacement les technologies au service de leurs apprentissages. Ces compétences techniques ne se limitent pas à un savoir-faire instrumental : elles incluent également des capacités d'adaptation à de nouveaux outils, de sélection critique des ressources numériques disponibles et de gestion responsable des données.

Développement des compétences métacognitives et empowerment

L'explicitation des attentes, la place croissante donnée à l'autoévaluation et à l'autorégulation participent au développement des compétences

métacognitives des étudiants (OCDE, 2021). En les amenant à mieux comprendre les critères de réussite, à évaluer leur propre niveau de performance et à ajuster leurs stratégies d'apprentissage, les dispositifs actuels renforcent leur *empowerment*, c'est-à-dire leur pouvoir d'agir sur leurs apprentissages et leurs choix de régulation. De surcroît, des pratiques telles que l'usage des degrés de certitude dans les évaluations leur permettent de prendre conscience de leurs marges d'incertitude et de leur rapport à la connaissance, deux dimensions cruciales pour développer une approche réflexive, autonome et critique vis-à-vis de leurs apprentissages.

Risques de fracture

En contrepartie, ces évolutions peuvent accentuer les écarts entre apprenants, selon leur familiarité préalable avec les outils numériques, leur environnement technologique et leur capacité d'autoorganisation (Castaño-Muñoz *et al.*, 2018). Les apprenants disposant d'un accès limité aux technologies performantes, d'une faible qualité de connexion internet, ou présentant une moindre maîtrise des compétences numériques se trouvent en situation de désavantage par rapport à leurs pairs mieux équipés et plus autonomes (Aissaoui et Ben Lassen, 2016). De même, les exigences accrues en matière d'autorégulation et de gestion du travail à distance peuvent pénaliser certains profils d'apprenants moins expérimentés dans ces démarches. La prise en compte proactive de ces risques est indispensable pour éviter que les dispositifs d'évaluation innovants ne deviennent un facteur supplémentaire de reproduction des inégalités. Si les transformations récentes offrent des opportunités importantes pour améliorer la qualité des évaluations et soutenir le développement de compétences clés chez les étudiants, elles ne sont pas exemptes de risques et de défis. L'ambivalence des effets constatés rappelle la nécessité d'une réflexion approfondie sur les enjeux que ces évolutions soulèvent, tant pour la recherche que pour la qualité des pratiques évaluatives. C'est à cette analyse des enjeux que nous consacrons à présent la suite de notre réflexion.

Quels enjeux ?

Les effets mis en évidence à travers les transformations récentes de l'évaluation des apprentissages dans l'enseignement supérieur ne sont pas univoques. S'ils ouvrent des perspectives intéressantes en matière

de qualité, de développement des compétences et d'individualisation des parcours, ils soulèvent également un ensemble d'enjeux qui méritent d'être analysés avec attention. Ces enjeux concernent à la fois la recherche en évaluation, la qualité et l'éthique des pratiques, ainsi que les risques émergents qu'il convient de réguler pour garantir des dispositifs justes, efficaces et durables.

Enjeux pour la recherche

La recherche en évaluation reste malheureusement un domaine sous-développé en sciences de l'éducation. En témoigne le petit nombre de postes académiques ouverts dans ce domaine. Pourtant, les questions ouvertes sont extrêmement nombreuses comme nous venons de le voir en examinant les effets potentiels des usages du numérique ; il s'agit de les décrire d'abord pour ensuite en comprendre les conditions, et en mesurer les effets. Les récits de pratiques ou les études de cas si elles sont inspirantes ne suffisent pas. Il devient dès lors nécessaire de renforcer les cadres méthodologiques, de favoriser la mobilisation de modèles d'analyse éprouvés, comme ceux issus des approches qualité (Pottiez et Gilles, 2025) et de soutenir la construction d'outils conceptuels adaptés aux spécificités des environnements numériques d'apprentissage (Educa.ch, 2021). Par ailleurs, la recherche devrait accorder une attention particulière aux effets systémiques des dispositifs numériques d'évaluation, en étudiant par exemple comment ils influencent l'agencement des curricula, les dynamiques institutionnelles, ou encore les inégalités d'apprentissage.

Enjeux pour la qualité et l'éthique des pratiques évaluatives

Les transformations pédagogiques et technologiques génèrent de nombreux enjeux éthiques. Nous avons déjà évoqué le respect de l'égalité d'accès par une lutte résolue contre les inégalités numériques entre étudiants. Cela passe par des formations, un soutien financier et une écoute des besoins exprimés, comme la mise à disposition d'espaces et d'outils de travail adéquats.

La question de l'authenticité des productions renvoyant à la question souvent posée : « comment s'assurer que le travail évalué est bien celui de l'apprenant ? » suppose à notre avis un recadrage. Une réponse simple serait d'empêcher tout usage du numérique pendant une épreuve.

Si pour une question de validité, cette réponse peut s'avérer adéquate, dans d'autres cas, il s'agit de revoir la compétence évaluée en prenant en compte le travail distribué par l'étudiant entre lui et le logiciel. Dans ce cas, l'application d'intelligence artificielle, par exemple, est un partenaire en cognition dont il s'agit de décrire et de justifier le rôle, et éventuellement de le critiquer. Ainsi, l'étudiant devra démontrer comment sa performance est transformée et enrichie par l'IA. La responsabilité éthique des enseignants et des institutions est donc engagée : il s'agit de garantir des pratiques évaluatives justes, transparentes et respectueuses des droits fondamentaux.

Enjeux liés aux opportunités et aux risques émergents

Le développement des usages du numérique dans l'évaluation des apprentissages offre l'opportunité aux institutions de s'impliquer davantage dans ces processus et de prendre au sérieux leurs responsabilités à cet égard. Ce sont elles qui délivrent les diplômes et pourtant les pratiques d'évaluation mises en œuvre en leur sein restent souvent des points aveugles. Les règlements, s'ils protègent des plus graves dérives en matière de plagiat ou de tricherie ou encore de non-transparence, ne considèrent bien souvent aucune des étapes d'un cycle de qualité tel que présenté plus haut. Les services d'appui à la pédagogie universitaire se limitent dans beaucoup de cas à l'enseignement et à l'apprentissage sans appréhender l'évaluation qui reste une chasse gardée. Quelques contributions de cet ouvrage montrent cependant que des actions d'accompagnement et de formation sont possibles et déjà en œuvre ici et là.

Dans cette perspective, une seconde opportunité à ne pas manquer serait de permettre à l'étudiant d'intervenir davantage dans ce processus d'évaluation en étant formé à l'évaluation, à l'art du *feedback*, à la critique des situations d'évaluation qui lui sont proposées, à la maîtrise des instruments, à la compréhension de leurs valeurs ajoutées et de leurs biais. Ces connaissances ne sont pas seulement réservées aux spécialistes. Les processus d'évaluation des connaissances et des compétences faisant partie du travail et de toute activité, même de loisirs, elles sont utiles et nécessaires pour tous, tout au long de leur vie.

L'essor des dispositifs numériques d'évaluation soulève aussi plusieurs risques qui exigent une régulation active pour préserver l'équité et la légitimité des pratiques. Un premier risque tient à l'opacité des

algorithmes, qui rend difficiles la compréhension et la contestation des décisions automatisées. Sans transparence ni contrôle humain, la confiance dans l'évaluation est fragilisée. Un second enjeu est la normalisation des comportements : les outils numériques valorisent ce qui est mesurable, au détriment, parfois, d'apprentissages plus profonds ou singuliers. Cela peut biaiser la validité et brider la diversité des approches d'apprentissage. La fracture numérique va également au-delà de l'accès aux outils : elle concerne aussi les inégalités dans les compétences numériques, renforçant le risque d'injustice cognitive pour certains publics. Enfin, les dispositifs de télésurveillance (*proctoring*) peuvent menacer la vie privée, générer du stress ou reproduire des biais discriminatoires.

Face à ces dérives possibles, il est indispensable de poser des garde-fous éthiques et méthodologiques clairs, en mobilisant des critères de qualité tels que ceux que nous avons évoqués plus haut, et en associant les parties prenantes à la gouvernance des dispositifs.

Bibliographie

Abou Haidar, L. (2021). L'oral à l'ère du numérique : enseigner et apprendre autrement ? *Alsic*, *24*(2). https://doi.org/10.4000/alsic.5739

Aissaoui, R. et Ben Hassen, T. (2016). Diffusion technologique et inégalités numériques : une exploration de la fracture numérique dans l'espace MENA. *STATECO*, *110*, 105-121. https://www.insee.fr/fr/statistiques/fichier/2576743/stec110g.pdf (consulté le 19.01.2026).

Alade, O. M. et Omoruyi, I. V. (2014). Table of specification and its relevance in educational development assessment. *European Journal of Educational and Development Psychology*, *2*(1), 1-17.

Alam, M. I., Malone, L., Nadolny, L., Brown, M. et Cervato, C. (2023). Investigating the impact of a gamified learning analytics dashboard : Student experiences and academic achievement. *Journal of Computer Assisted Learning*, *39*(5), 1436-1449. https://doi.org/10.1111/jcal.12853

Altet, M. (2013). Chapitre 2. Formes de résistance des pratiques de formation d'enseignants à la pratique réflexive et conditions de développement de la réflexivité. Dans M. Altet, J. Desjardins, R. Étienne, L. Paquay et P. Perrenoud (dir.), *Former des enseignants réflexifs : Obstacles et résistances* (p. 39-60). De Boeck Supérieur. https://doi.org/10.3917/dbu.altet.2013.01.0039

Aouatif, K. et Awatif, B. (2023). The e-docimology in the context of corona virus : Through trial and error. *African Scientific Journal*, *3*(16), 1-21. https://doi.org/10.5281/zenodo.7556691

Arias-Calderón, M., Castro, J. et Gayol, S. (2022). Serious Games as a Method for Enhancing Learning Engagement : Student Perception on Online Higher Education During COVID-19. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.889975

Asri, H. (2024). *L'intégration des outils numériques dans l'évaluation des apprentissages : étude de cas à l'Université Cadi Ayyad – Maroc* [communication orale]. Journée scientifique du 26.08.2025 du Projet P8 Future University soutenu par swissuniversities, Lausanne, Suisse. https://tube.switch.ch/videos/XmY5suqUds (consulté le 19.01.2026).

Astuti, F. S. et Retnawati, H. (2022). Online Learning Assessment : Solution in Industrial Revolution 4.0. Series : Advances in Social Science, Education and Humanities Research. *Proceedings of the 5th International Conference on Current Issues in Education (ICCIE 2021)*. https://doi.org/10.2991/assehr.k.220129.064

Audran, J. (2024). Cinq enjeux d'évaluation face à l'émergence des IA génératives en éducation. *Mesure et évaluation en éducation*, *47*(1), 6-26. https://doi.org/10.7202/1114564ar

Bailly, B., Demougeot-Lebel, J. et Lison, C. (2015). La formation d'enseignants universitaires nouvellement recrutés : quelles retombées ? *Revue internationale de pédagogie de l'enseignement supérieur*, *31*(3). http://ripes.revues.org/990 (consulté le 02.01.2026).

Barana, A., Conte, A., Fissore, C., Marchisio, M. et Rabellino, S. (2019). Learning analytics to improve formative assessment strategies. *Journal of E-Learning and Knowledge Society*, *15*(3), 75-88. https://dx.doi.org/10.20368/1971-8829/1135057

Barge, P. et Londhe, B. R. (2014). From Teaching, Learning to Assessment : MOODLE Experience at B'School in India. *Procedia Economics and Finance*, *11*, 857-865. https://doi.org/10.1016/S2212-5671(14)00249-4

Barras, H. (2020). Évaluer dans l'urgence, en repensant sa planification à l'aide des principes issus de la gestion de crises. *Évaluer. Journal international de recherche en éducation et formation*, Numéro hors-série, *1*, 17-24.

Barras, H. (2023). Impact de la formation à l'enseignement, analyse de bilans de compétences. *Évaluer. Journal international de recherche en éducation et formation*, *9*(3), 3-30. https://doi.org/10.48782/e-jiref-9-3-3

Béchard, N. (2023). Repenser ses pratiques d'évaluation : dix étapes faciles. *Apprendre et enseigner aujourd'hui*, *13*(1), 52-56. https://doi.org/10.7202/1107544ar

Belgheddouche, A. (2025). *La vidéo comme outil d'évaluation de la production orale : cas d'un groupe d'étudiantes de première année en licence de français à l'ENSB* [communication orale]. Journée scientifique du 26.08.2025 du Projet P8 Future University soutenu par swissuniversities, Lausanne, Suisse. https://tube.switch.ch/videos/XmY5suqUds (consulté le 19.01.2026).

Bidarra, J., Rocio, V., Sousa, N. et Coutinho-Rodrigues, J. (2024). Problems and prospects of hybrid learning in higher education. *Open Learning : The Journal of Open, Distance and e-Learning*, *40*(3), 304-323. https://doi.org/10.1080/02680513.2024.2404036

Biémar, S., Daele, A., Malengrez, D. et Oger, L. (2015). Le « Scholarship of Teaching and Learning » (SoTL). Proposition d'un cadre pour l'accompagnement des enseignants par les conseillers pédagogiques. *Revue internationale de pédagogie de l'enseignement supérieur*, *31*(2). https://doi.org/10.4000/ripes.966

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, *32*(3), 347-364. https://doi.org/10.1007/BF00138871

Biggs, J. (2003). *Teaching for quality learning at university* (2e éd.). Open University Press/ Society for Research into Higher Education.

Blandin, B. (2004). La relation pédagogique à distance : que nous apprend Goffman ? *Distances et savoirs*, *2*(2-3), 357-381. https://shs.cairn.info/revue-distances-et-savoirs-2004-2-page-357?lang=fr# (consulté le 19.01.2026).

Bloom, B. S. (1956). *Taxonomy of educational objectives : The classification of educational goals* (1re éd.). Longman Group.

Bosc-Miné, C. (2014). Caractéristiques et fonctions des feed-back dans les apprentissages. *Année psychologique*, *114*, 315-353. https://doi.org/10.4074/S000350331400205X

Broadbent, J., Ajjawi, R., Bearman, M., Boud, D. et Dawson, P. (2023). Beyond emergency remote teaching : Did the pandemic lead to lasting change in university courses ? *International Journal of Educational Technology in Higher Education*, *20*(58). https://doi.org/10.1186/s41239-023-00428-z

Calone, A. et Lafontaine, D. (2018). Feedback normatif vs feedback élaboré : quel impact sur la performance et le sentiment de contrôlabilité des élèves ? *Évaluer. Journal international de recherche en éducation et formation*, *4*(2), 47-76. https://journal.admee.org/index.php/ejiref/article/view/103 (consulté le 02.01.2026).

Cao, S. et Zhong, L. (2023). Exploring the effectiveness of ChatGPT-based feedback compared with teacher feedback and self-feedback : Evidence from chinese to english translation. *ArXiv*. https://doi.org/10.48550/arXiv.2309.01645

Carless, D. (2013). Sustainable feedback and the development of student self-evaluative capacities. Dans S. Merry, M. Price, D. Carless et M. Taras (dir.), *Reconceptualising feedback in higher education : Developing dialogue with students* (p. 113-122). Routledge.

Carless, D. et Boud, D. (2018). The development of student feedback literacy : Enabling uptake of feedback. *Assessment and Evaluation in Higher Education*, *43*(8), 1315-1325. https://doi.org/10.1080/02602938.2018.1463354

Castaño-Muñoz, J., Colucci, E. et Smidt, H. (2018). Free digital learning for inclusion of migrants and refugees in Europe : A qualitative analysis of three types of learning purposes. *The International Review of Research in Open and Distributed Learning*, *19*(2), 1-21. https://doi.org/10.19173/irrodl.v19i2.3382

Céci, J.-F. (2022, octobre). *Innovation pédagogique et hybridations, au prisme du numérique : une taxonomie de l'innovation pédagogique instrumentée* [communication orale]. Colloque

International TICEMED 13, Athènes, Grèce. https://hal.science/hal-04124604/document (consulté le 02.01.2026).

Cellule IPN (2022). Enquête : Évaluation des pratiques de feedback au sein de l'EPF. EPF École d'ingénieurs. https://ilearn.epf.fr/recherche/etudes/Evaluation-des-Pratiques-de-Feedback.pdf (consulté le 02.01.2026).

Cellule IPN (2024). Enquête : Analyse des usages des IA génératives au sein de l'EPF. EPF École d'ingénieurs. https://ilearn.epf.fr/recherche/etudes/Analyse-usages-IAg.pdf (consulté le 02.01.2026).

Champoux, M. et Hillman, A. (2022). *Évaluer à distance en temps de pandémie : le cas d'un programme de maîtrise professionnelle en environnement*. https://hal.science/hal-03779551 (consulté le 15.01.2026).

Chan, C. K. Y. et Hu, W. (2023). Students' voices on generative AI : Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, *20*(43). https://doi.org/10.1186/s41239-023-00411-8

Choi-Lundberg, D. L., Butler-Henderson, K., Harman, K. et Crawford, J. (2023). A systematic review of digital innovations in technology-enhanced learning designs in higher education. *Australasian Journal of Educational Technology*, *39*(3), 133-162. https://doi.org/10.14742/ajet.7615

Chouhan, R. (2024). Strategies for maintaining academic integrity in remote unproctored and proctored online assessments for engineering courses. *Learning : Research and Practice*, *10*(1), 75-92. https://doi.org/10.1080/23735082.2023.2216198

Costa, C., Alvelos, H. et Teixeira, L. (2012). The Use of Moodle e-learning Platform : A Study in a Portuguese University. *Procedia Technology*, *5*, 334-343. https://doi.org/10.1016/j.protcy.2012.09.037

Coulet, J.-C. (2011). La notion de compétence : un modèle pour décrire, évaluer et développer les compétences. *Le travail humain*, *74*(1), 1-30. https://doi.org/10.3917/th.741.0001

Coulon, A. (1997). *Le métier d'étudiant. L'entrée dans la vie universitaire*. Presses universitaires de France.

Daele, A. et Berthiaume, D. (2009). *L'identification et la rédaction des objectifs pédagogiques*. Université de Lausanne. https://enseigner.unil.ch/ressources/document/lidentification-et-la-redaction-des-objectifs-pedagogiques/ (consulté le 02.01.2026).

Dawson, P. (2024). Remote proctoring : Understanding the debate. Dans S. E. Eaton (dir.), *Second handbook of academic integrity* (p. 1511-1526). Springer Nature Switzerland.

De Vecchi, G. (2021). *Évaluer sans dévaluer*. Hachette Éducation.

Demougeot-Lebel J. (2017) « Le « bon cours » et le « mauvais cours » : analyse descriptive et essai de typologie des conceptions de jeunes enseignants-chercheurs en France ». *Les dossiers des Sciences de l'éducation*, *38*, 131-149. https://doi.org/10.4000/dse.1548

Demougeot-Lebel, J. et Perret, C. (2011). Qu'attendent les enseignants universitaires français en termes de formation et d'accompagnement pédagogiques ? *Revue internationale de pédagogie de l'enseignement supérieur*, *27*(1). https://doi.org/10.4000/ripes.456

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*, Numéro hors-série, *1*, 111-119.

Dias-Chiaruttini, A. et Dolz, J. (2023). Nouveaux objets et nouveaux contextes d'enseignement de l'oral. *Repères*, 68. https://doi.org/10.4000/reperes.5963

Digital Education Council (2024). *Digital Education Council Global AI Student Survey 2024*. https://www.digitaleducationcouncil.com/post/digital-education-council-global-ai-student-survey-2024 (consulté le 02.01.2026).

Dolz, J. et Schneuwly, B. (1998). *Pour un enseignement de l'oral. Initiation aux genres formels à l'école*. ESF.

Douglas, T. et Chapman, J. (2023). Supporting Blended Learners in the New Normal. Dans D. Kember, R. A. Ellis, S. Fan et A. Trimble (dir.), *Adapting to Online and Blended Learning in Higher Education : Supporting the Retention and Success of the Expanded and Diversified Intake* (p. 351-376). Springer Nature. https://doi.org/10.1007/978-981-99-0898-1_16

Dumais, C. (2010). *L'évaluation de l'oral et la progression : la grille d'observation progressive* [communication orale]. XI[e] Colloque de l'Association internationale pour la recherche en didactique du français (AIRDF). Liège, Belgique.

Dumais, C., Messier, G., Viola, S. et Meunier, H. (2013). Des grilles d'autoévaluation et de coévaluation pour améliorer la compétence orale. *Québec français*, *170*, 64-68. https://id.erudit.org/iderudit/70512ac (consulté le 02.01.2026).

Dupré, A., et Giraudin, É. (2025). *Le e-portfolio comme outil pour l'évaluation des compétences : analyse d'usage dans une formation de bachelor universitaire de technologie* [communication orale]. Journée scientifique du 26.08.2025 du Projet P8 Future University soutenu par swissuniversities, Lausanne, Suisse. https://tube.switch.ch/videos/XmY5suqUds (consulté le 19.01.2026).

Duroisin, N., Beauset, R. et Tanghe, C. (2021). Education and digital inequalities during COVID-19 confinement : From the perspective of teachers in the French speaking Community of Belgium. *European Journal of Education*, *56*(4), 515-535. https://doi.org/10.1111/ejed.12475

Educa.ch. (2021). *La numérisation dans l'éducation : enjeux, perspectives et recommandations*. Educa.ch. https://www.educa.ch/sites/default/files/2021-11/La_numerisation_dans_l_education.pdf (consulté le 02.01.2026).

El Asame, M., Wakrim, M. et Battou, A. (2022). Designing e-assessment activities appropriate to learner's competency levels : Hybrid pedagogical framework and authoring tool. *Education and Information Technologies*, *27*, 2543-2567. https://doi.org/10.1007/s10639-021-10607-y

El Kamoun, N., Akhasbi, H., Lakrami, F., Gilles, J.-L., Rigo, J.-M. et Aliss, E. (2025, janvier). *L'intelligence artificielle générative : une opportunité pour améliorer l'évaluation formative dans un contexte universitaire massif* [communication orale]. 36[e] Colloque de l'AD-MEE-Europe, Belval, Luxembourg. http://hdl.handle.net/20.500.12162/8096 (consulté le 02.01.2026).

Endrizzi, L. (2012, octobre). *Les technologies numériques dans l'enseignement supérieur, entre défis et opportunités*. Dossier d'actualité Veille et analyses, 78. http://veille-et-analyses.ens-lyon.fr/DA/detailsDossier.php?parent=accu%20eil&dossier=*78&lang=fr* (consulté le 02.01.2026).

Erard, S. et Schneuwly, B. (2005). La didactique de l'oral : savoirs ou compétences ? Dans J.-P. Bronckart, E. Bulea et M. Pouliot (dir.). *Repenser l'enseignement des langues* (p. 69-97). Presses universitaires du Septentrion. https://doi.org/10.4000/books.septentrion.14817

Escotet, M. Á. (2024). The optimistic future of Artificial Intelligence in higher education. *Prospects*, *54*(3), 531-540. https://doi.org/10.1007/s11125-023-09642-z

Espasa, A., Guasch, T., Mayordomo, R. M., Martínez-Melo, M. et Carless, D. (2018). A Dialogic Feedback Index measuring key aspects of feedback processes in online learning environments. *Higher Education Research & Development*, *37*(3), 499-513. https://doi.org/10.1080/07294360.2018.1430125

Fawns, T. (2019). Postdigital Education in Design and Practice. *Postdigital Science and Education*, *1*(1), 132-145. https://doi.org/10.1007/s42438-018-0021-8

Fernando, W. (2020). Moodle quizzes and their usability for formative assessment of academic writing. *Assessing Writing*, *46*. https://doi.org/10.1016/j.asw.2020.100485

Frenay, M., Saroyan, A., Taylor, K. L., Bédard, D., Clement, M., Colet, N. R., Paul, J.-J. et Kolmos, A. (2010). Accompagner le développement pédagogique des enseignants

universitaires à l'aide d'un cadre conceptuel original. *Revue française de pédagogie. Recherches en éducation*, 172, 63-76. https://doi.org/10.4000/rfp.2253

Gamage, D., Staubitz, T. et Whiting, M. (2021). Peer assessment in MOOCs : Systematic literature review. *Distance Education*, *42*(2), 268-289. https://doi.org/10.1080/01587919.2021.1911626

Gamage, S. H. P. W., Ayres, J. R. et Behrend, M. B. (2022). A systematic review on trends in using Moodle for teaching and learning. *International Journal of STEM Education*, *9*(1), 9. https://doi.org/10.1186/s40594-021-00323-x

Garcia-Debanc, C. (1999). Évaluer l'oral. *Pratiques : linguistique, littérature, didactique*, *103-104*, 193-212. https://doi.org/10.3406/prati.1999.1867

Garcia-Debanc, C. et Delcambre, I. (2001). Enseigner l'oral ? *Repères, recherches en didactique du français langue maternelle*, *24-25*, 3-21. https://doi.org/10.3406/reper.2001.2367

Georges, F. et Poumay, M. (2020). *Créer des SAE. Guide de soutien à la création de situations d'apprentissage et d'évaluation (SAE) en contexte d'APC*. Université de Liège. https://hdl.handle.net/2268/254034 (consulté le 02.01.2026).

Gilles, J.-L. (2016). *Ingénierie des évaluations à l'ère du Web 2.0 – Working Paper (octobre 2016)*. Lausanne : Haute école pédagogique du canton de Vaud. http://tinyurl.com/ingenierieevaluations (consulté le 15.01.2026).

Gilles, J.-L. (2002). *Qualité spectrale des tests standardisés universitaires : Mise au point d'indices édumétriques d'analyse de la qualité spectrale des évaluations des acquis des étudiants universitaires et application aux épreuves MOHICAN check up '99* [Thèse de doctorat, Université de Liège]. http://hdl.handle.net/20.500.12162/824 (consulté le 02.01.2026).

Gilles, J.-L. (2023). *Introduction à la docimologie et aux questions à correction automatique* [communication orale]. Conférence-atelier du Service Écoles-Médias (SEM) du Département de l'instruction publique (DIP) de l'État de Genève, Genève, Suisse. http://hdl.handle.net/20.500.12162/7015 (consulté le 02.01.2026).

Gilles, J.-L. et Charlier, B. (2020). Dispositifs d'évaluation à distance à correction automatisée versus non automatisée : analyse comparative de deux formes emblématiques. *Évaluer. Journal international de recherche en éducation et formation*, Numéro hors-série, *1*, 143-154. http://hdl.handle.net/20.500.12162/3865 (consulté le 02.01.2026).

Gilles, J.-L. et Leclercq, D. (1995). *Procédures d'évaluation adaptées à des grands groupes d'étudiants universitaires : Enjeux et solutions pratiquées à la FAPSE-ULG* [communication orale]. Symposium international sur la Rénovation didactique en biologie, Tunis, Tunisie. http://hdl.handle.net/20.500.12162/2709 (consulté le 02.01.2026).

Glaser, B. G. et Strauss, A. L. (1995). La production de la théorie à partir des données (traduit par J.-L. Fabiani). *Enquête. Cahiers du CERCOM*, *1*, 183-195. https://doi.org/10.4000/enquete.282

Guàrdia, L., Crisp, G. et Alsina, I. (2017). Trends and challenges of e-assessment to enhance student learning in higher education. *Innovative practices for higher education assessment and measurement*, 36-56. https://doi.org/10.4018/978-1-5225-0531-0.ch003

Guasch, T. et Espasa, A. (2015). Collaborative Writing Online : Unravelling the Feedback Process. Dans M. Deane et T. Guasch (dir.), *Learning and Teaching Writing Online : Strategies for Success* (p. 13-30). Brill. https://doi.org/10.1163/9789004290846_003

Guasch, T., Espasa, A., Alvarez, I. M. et Kirschner, P. A. (2013). Effects of feedback on collaborative writing in an online learning environment. *Distance Education*, *34*(3), 324-338. https://doi.org/10.1080/01587919.2013.835772

Guo, B., Zhang, X., Wang, Z., Jiang, M., Nie, J., Ding, Y., Yue, J. et Wu, Y. (2023). How Close is ChatGPT to Human Experts ? Comparison Corpus, Evaluation, and Detection. *ArXiv*. https://doi.org/10.48550/arXiv.2301.07597

Halté, J.-F. et Rispail, M. (dir.). (2005). *L'oral dans la classe : Compétences, enseignement, activités*. L'Harmattan.

Hancock, P., Birt, J., De Lange, P., Fowler, C., Kavanagh, M., Mitrione, L., Rankin, M., Slaughter, G. et Williams, A. (2023). Integrity of assessments in challenging times. *Accounting Education*, *32*(5), 501-522. https://doi.org/10.1080/09639284.2022.2137818

Hartle, S. (2023). Integrating the Formative and Summative Post Covid-19 in a Blended Learning Approach to Higher Educational Assessment : Challenges and Opportunities. *Iperstoria*, 21, 190-216. https://doi.org/10.13136/2281-4582/2023.i21.1191

Hattie, J. et Timperley, H. (2007). The Power of Feedback. *Review of Educational Researh*, *77*(1), 81-112. https://doi.org/10.3102/003465430298487

Hébert, M.-H., Gérin-Lajoie, S., Papi, C., Fournier Dubé, N., Beauparlant, R. et Beaudoin, A. (2024, novembre). *Le plagiat et la tricherie en formation à distance à l'enseignement supérieur : qu'en sait-on ?* [communication orale]. 45e session d'étude de l'ADMEE-Canada, Sherbrooke, Canada.

Heil, J. et Ifenthaler, D. (2023). Online Assessment in Higher Education : A Systematic Review. *Online Learning*, *27*(1), 187-218. https://doi.org/10.24059/olj.v27i1.3398

Hettiarachchi, E., Mor, E., Huertas, M. A. et Guerrero-Roldán, A. E. (2015). Introducing a Formative E-Assessment System to Improve Online Learning Experience and Performance. *Journal of Universal Computer Science*, *21*(8), 1001-1021.

Hillier, M. (2023). Making Online Assessment Active and Authentic. Dans M. D. Sankey, H. Huijser et R. Fitzgerald (dir.), *Technology-Enhanced Learning and the Virtual University* (p. 1-46). Springer. https://doi.org/10.1007/978-981-19-9438-8_17-1

Hodges, C., Moore, S., Lockee, B., Trust, T. et Bond, A. (2020). The Difference Between Emergency Remote Teaching and Online Learning. *EDUCAUSE Review*. https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning

Holmes, N. (2014). Student perceptions of their learning and engagement in response to the use of a continuous e-assessment in an undergraduate module. *Assessment & Evaluation in Higher Education*, *40*(1), 1-14. https://doi.org/10.1080/02602938.2014.881978

Hornsby, D. J. et Osman, R. (2014). Massification in higher education : Large classes and student learning. *Higher Education*, *67*(6), 711-719. https://doi.org/10.1007/s10734-014-9733-1

Horwitz, E. K., Horwitz. M. B. et Cope, J. (1986). Foreign Language Classroom Anxiety. *Modern Language Journal*, *70*(2), 125-132.

Ifenthaler, D. et Greiff, S. (2021). Leveraging learning analytics for assessment and feedback. Dans J. Liebowitz (dir.), *Online learning analytics* (p. 1-18). Auerbach Publications. https://doi.org/10.1201/9781003194620

Jorro, A. (2021). Épreuves, mises en tension, inventivité des étudiants en situation de confinement. *Le sujet dans la cité*, *11*, 13-19. https://doi.org/10.3917/lsdlc.011.0013

Kaddouri, M. et Loiola, F. (2022). Quelques enjeux et paradoxes de l'approche par compétences : une introduction. *Revue des sciences de l'éducation*, *48*(2). https://doi.org/10.7202/1098336ar

Kamalov, F., Calonge, D. S. et Gurrib, I. (2023). New Era of Artificial Intelligence in Education : Towards a Sustainable Multifaceted Revolution. *Sustainability*, *15*(16). https://doi.org/10.3390/su151612451

Kharbat, F. F. et Abu Daabes, A. S. (2021). E-proctored exams during the COVID-19 pandemic : A close understanding. *Education and Information Technologies*, *26*(6), 6589-6605. https://doi.org/10.1007/s10639-021-10458-7

Kim, K.-J. (2022). Moving forward : Embracing challenges as opportunities to improve medical education in the post-COVID era. *Humanities and Social Sciences Communications*, *9*, 1-4. https://doi.org/10.1057/s41599-022-01451-7

Kozanitis, A. (2021). Évaluation des apprentissages en formation à distance : les situations authentiques à la rescousse. *Revue internationale des technologies en pédagogie universitaire*, *18*(1), 276-290. https://doi.org/10.18162/ritpu-2021-v18n1-24

Kugel, P. (1993). How professors develop as teachers. *Studies in Higher Education*, *18*(3), 315-328. https://doi.org/10.1080/03075079312331382241

Lab'UA (2023). *L'évaluation en ligne, optimisation ou dégradation ?* Université d'Angers. https://labua.univ-angers.fr/levaluation-en-ligne-optimisation-ou-degradation/ (consulté le 02.01.2026).

Lac, M., Mias, C., Labbé, S., Bataille, M. et Wittorski, R. (2010). Les représentations professionnelles et l'implication professionnelle comme modèles d'intelligibilité des processus de professionnalisation. *Les dossiers des sciences de l'éducation*, 24, 133-145. https://journals.openedition.org/dse/960 (consulté le 02.01.2026).

Laferrière, T., Baron, G.-L., Allaire, S., Nadeau-Tremblay, S., Tremblay, M., Boisjoly, F., Nicole, M.-C. et Beaudoin, J. (2022). Numérique, éducation et forme scolaire : enjeux d'équité, *Diversité*, 200. http://journals.openedition.org/diversite/1625 (consulté le 02.01.2026).

Lafleur, F., Nolla, J.-M. et Samson, G. (dir.) (2021). *Évaluation des apprentissages en formation à distance. Enjeux, modalités et opportunités de formation en enseignement supérieur*. Presses de l'Université du Québec.

Lafontaine, L. et Messier, G. (2009). Représentations de l'enseignement et de l'évaluation de l'oral chez des enseignants et des élèves du secondaire en français langue d'enseignement. *Revue du Nouvel-Ontario*, *34*, 119-144. https://doi.org/10.7202/038722ar

Lafortune L. et Saint Pierre L. (1998). *Affectivité et métacognition dans la classe : Des idées et des applications concrètes pour l'enseignant*. De Boeck Université.

Lameul, G. (2016). Postures et activité du sujet en formation : de l'intention au geste professionnel. *Revue internationale de pédagogie de l'enseignement supérieur*, *32*(3). https://doi.org/10.4000/ripes.1160

Lanarès, J., Laperrouza, M. et Sylvestre, E. (2023). *Design pédagogique*. Éditions Épistémé.

Lanarès, J. et Poteaux, N. (2013). Comment répondre aux défis actuels de l'enseignement supérieur ? Dans D. Berthiaume et N. Rege Colet (dir.), *La pédagogie de l'enseignement supérieur : repères théoriques et applications pratiques. Tome 1 : Enseigner au supérieur* (p. 9-24). Peter Lang.

Laparra, M. (2008). L'oral, un enseignement impossible ? *Pratiques*, *137-138*, 117-134. https://doi.org/10.4000/pratiques.1155

Lavoie, C. et Bouchard, É. (2017). Formation universitaire à l'évaluation de l'oral : regard sur la capacité d'autoévaluation de futurs enseignants. Dans J.-F. de Pietro, C. Fisher et R. Gagnon (dir.), *L'oral aujourd'hui : Perspectives didactiques* (p. 259-274). Presses universitaires de Namur. https://books.openedition.org/pun/6687 (consulté le 02.01.2026).

Le Boterf, G. (1994). *De la compétence : Essai sur un attracteur étrange*. Éditions d'organisation.

Le Boterf, G. (2002). De quel concept de compétence avons-nous besoin ? *Soins cadres*, *41*. http://www.guyleboterf-conseil.com/images/Soins%20cadres.PDF (consulté le 14.01.2026).

Leclercq, D. (1983). Confidence marking, its use in testing. *Evaluation in Education : An International Review Series*, *6*(2), 163-287. https://orbi.uliege.be/bitstream/2268/9482/1/Leclercq_Confidence_Marking.pdf (consulté le 02.01.2026).

Lee, H., Sarkar, A., Tankelevitch., L., Drosos, I., Rintel, S., Banks, R. et Wilson, N. (2025). The Impact of Generative AI on Critical Thinking : Self-Reported Reductions in Cognitive Effort and Confidence Effects From a Survey of Knowledge Workers. *CHI'25 : Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, *1121*, 1-22. https://doi.org/10.1145/3706598.3713778

Lee, V. W. Y., Lam, P. L. C., Lo, J. T. S., Lee, J. L. F. et Li, J. T. S. (2022). Rethinking online assessment from university students' perspective in COVID-19 pandemic. *Cogent Education*, *9*(1), 2082079. https://doi.org/10.1080/2331186X.2022.2082079

Lemieux, M.-M. (2021). Inégalités, compétences et conditions numériques. *Revue internationale des technologies en pédagogie universitaire*, *18*(1), 157-169. https://doi.org/10.18162/ritpu-2021-v18n1-14

Lepage, I., Leduc, D. et Stockless, A. (2019). E-évaluation dynamique et engagement cognitif en contexte de grand groupe à l'université. *Évaluer. Journal international de recherche en éducation et formation*, *5*(3), 9-33.

Lison, C. et Paquelin, D. (2019). La formation du supérieur : un levier de transformation des universités québécoises. *Revue internationale d'éducation de Sèvres*, 80. https://doi.org/10.4000/ries.8184

Lynch, R. P. et Pappas, E. (2017). A Model for Teaching Large Classes : Facilitating a "Small Class Feel". *International Journal of Higher Education*, *6*(2), 199-212. https://doi.org/10.5430/ijhe.v6n2p199

Malglaive, G., Migne, J., Bazile, J., Bonnet, J., Chaix, M.-L., Darré, J.-P., Pastré, P., Perrot, J.-L. et Weil-Barais, A. (1994). Représentations et apprentissage chez les adultes, *Éducation permanente*, *119*, 5-175.

Many, H., Shvetsova, M. et Forestier, G. (2024). Transformation numérique : comment enseigner (avec) l'IA générative dans l'enseignement supérieur ? *Études & Pédagogies*. https://doi.org/10.20870/eep.2024.8100

Maqableh, M. et Alia, M. (2021). Evaluation online learning of undergraduate students under lockdown amidst COVID-19 Pandemic : The online learning experience and students' satisfaction. *Children and Youth Services Review*, *128*, 106160. https://doi.org/10.1016/j.childyouth.2021.106160

Maringe, F. et Sing, N. (2014). Teaching large classes in an increasingly internationalising higher education environment : Pedagogical, quality and equity issues. *Higher Education*, *67*(6), 761-782. https://doi.org/10.1007/s10734-013-9710-0

Martínez, M. L., Álvarez Valdivia, I. M. et Ortiz, A. R. (2015). Making learning more visible through e-assessment : Implications for feedback. *Journal of Computing in Higher Education*, *27*(1), 10-27. https://doi.org/10.1007/s12528-015-9091-8

Marty, S. et Thomas Vasquez, K. (2021). Dispositifs pédagogiques numériques et représentations technophiles/technophobes : le cas du Projet Up, dispositif pédagogique hybride innovant en contexte de crise sanitaire. *Interfaces numériques*, *10*(2). https://doi.org/10.25965/interfaces-numeriques.4624

Masson, S. (2020). *Activer ses neurones pour mieux apprendre et enseigner.* Odile Jacob.

Mate, K. et Weidenhofer, J. (2021). Considerations and strategies for effective online assessment with a focus on the biomedical sciences. *FASEB bioAdvances*, *4*(1), 9-21. https://doi.org/10.1096/fba.2021-00075

Maurer, B. (2001). *Une didactique de l'oral, du primaire au lycée*. Bertrand-Lacoste.

MESRI (2013). Livre blanc « La démarche ePortfolio dans l'enseignement supérieur français ». *Cahier n° 1 Enjeux et recommandations*.

MESRI (2019, juin). *Repères pour l'exercice du métier d'enseignant-chercheur.* https://www.enseignementsup-recherche.gouv.fr/fr/reperes-pour-l-exercice-du-metier-d-enseignant-chercheur-47766 (consulté le 02.01.2026).

Migne, J. (1994). Pédagogie et représentations. *Éducation permanente*, *119*, 11-31.

Moccozet, L., Benkacem, O. et Burgi, P.-Y. (2017). Towards a Technology-Enhanced Assessment Service in Higher Education. *Interactive Collaborative Learning*, 453-467. https://doi.org/10.1007/978-3-319-50340-0_40

Mollick, E. et Mollick, L. (2023). Student use cases for AI : Part 1 : AI as feedback generator. Harnessing the power of instant input. *Harvard Business School*. https://hbsp.harvard.edu/inspiring-minds/ai-as-feedback-generator (consulté le 02.01.2026).

Montenegro-Rueda, M., Luque-de la Rosa, A., Sarasola Sánchez-Serrano, J. L. et Fernández-Cerero, J. (2021). Assessment in Higher Education during the COVID-19

Pandemic : A Systematic Review. *Sustainability*, *13*(19). https://doi.org/10.3390/su131910509

Moscovici, S. (1961). *La psychanalyse, son image et son public*. PUF. http://doi.org/10.3917/puf.mosco.2004.01

Mottiar, Z., Byrne, G., Gorham, G. et Robinson, E. (2024). An examination of the impact of COVID-19 on assessment practices in higher education. *European Journal of Higher Education*, *14*(1), 101-121. https://doi.org/10.1080/21568235.2022.2125422

Nicol, D. (2010). From monologue to dialogue : Improving written feedback processes in mass higher education. *Assessment & Evaluation in Higher Education*, *35*(5), 501-517. https://doi.org/10.1080/02602931003786559

Noguera, I., Guerrero-Roldán, A. E., Peytcheva-Forsyth, R. et Yovkova, B. (2018). Perceptions of students with special educational needs and disabilities towards the use of e-assessment in online and blended education : barrier or aid ? *Proceedings 12th International Technology, Education and Development Conference* (p. 817-828). https://doi.org/10.21125/inted.2018

Nolla, J.-M., Ntebutse, J.-G. et Leroux, J. L. (2022). L'évaluation des apprentissages en formation à distance : l'expérience des professeur·es universitaires dans la régulation des émotions. *La Revue LEeE*, *5*. https://doi.org/10.48325/rleee.005.06

Nonnon, E. (1999). L'enseignement de l'oral et les interactions verbales en classe : champs de référence et problématiques. *Revue française de pédagogie*, 129, 87-131.

Nonnon, E. (2011). L'histoire de la didactique de l'oral, un observatoire de questions vives de la didactique du français. *Pratiques*, *149-150*, 184-206. https://doi.org/10.4000/pratiques.1739

Nonnon, E. (2016). 40 ans de discours sur l'enseignement de l'oral : la didactique face à ses questions. *Pratiques*, *169-170*. https://doi.org/10.4000/pratiques.3115

Nyagope, T. S. (2024). Massification at Higher Education Institutions : Challenges Associated with Teaching Large Classes and How it Impacts the Quality of Teaching and Learning in South Africa. *International Journal of Environmental, Sustainability, and Social Science*, *5*(1). https://doi.org/10.38142/ijesss.v5i1.852

OCDE (2021). *Perspectives de l'OCDE sur les compétences 2021 : Se former pour la vie*. Éditions OCDE. https://doi.org/10.1787/fc97e6d3-fr

OCDE (2023). *Perspectives de l'OCDE sur l'éducation numérique 2023*. Éditions OCDE. https://doi.org/10.1787/78afb124-fr

OECD (2020). Remote online exams in higher education during the COVID-19 crisis. *OECD Education Policy Perspectives*, *6*. https://doi.org/10.1787/f53e2177-en

Olmos, S., Mena, J., Torrecilla, E. M. et Iglesias, A. I. (2015). Improving graduate students learning through the use of Moodle. *Educational Research and Reviews*, *10*(5), 604-614.

Paillé, P. et Mucchielli, A. (2021). *L'analyse qualitative en sciences humaines et sociales*. Armand Colin.

Papanastasiou, E. C. et Solomonidou, G. (2023). Evaluating Emergency Remote Assessment Adaptations in Higher Education due to COVID-19 : Faculty Insights and Challenges. *Education Sciences*, *13*(2). https://doi.org/10.3390/educsci13020184

Parent, S., Depover, C. et Poellhuber, B. (2023). La formation à distance et l'apprentissage à l'aide du numérique : une perspective postpandémique – introduction au numéro thématique. *Revue internationale des technologies en pédagogie universitaire / International Journal of Technologies in Higher Education*, *20*(2), 1-4. https://doi.org/10.18162/ritpu-2023-v20n2-01

Patael, S., Shamir, J., Soffer, T., Livne, E., Fogel-Grinvald, H. et Kishon-Rabin, L. (2022). Remote proctoring : Lessons learned from the COVID-19 pandemic effect on the large scale on-line assessment at Tel Aviv University. *Journal of Computer Assisted Learning*, *38*(6), 1554-1573. https://doi.org/10.1111/jcal.12746

Pauli, M. et Ferrell, G. (2020). *The future of assessment : Five principles, five targets for 2025*. Jisc Bristol, UK. https://www.voced.edu.au/content/ngv :87328 (consulté le 02.01.2026).

Polyportis, A. et Pahos, N. (2024). Understanding students' adoption of the ChatGPT chatbot in higher education : The role of anthropomorphism, trust, design novelty and institutional policy. *Behaviour & Information Technology*, *44*(2), 315-336. https://doi.org/10.1080/0144929X.2024.2317364

Pottiez, J. et Gilles, J.-L. (2025). *L'évaluation de la formation* (4e éd.). Dunod.

Poumay, M. et Georges, F. (2022). *Comment mettre en œuvre une approche par compétences dans le supérieur ?* De Boeck Supérieur.

Poumay, M., Tardif, J. et Georges, F. (dir.) (2017). *Organiser la formation à partir des compétences : Un pari gagnant pour l'apprentissage dans le supérieur.* De Boeck Supérieur.

Poumay, M. (2017). Chapitre 9. Séminaires et portfolios de traces pour soutenir et évaluer le développement de la compétence. Dans M. Poumay, J. Tardif et F. Georges (dir.), *Organiser la formation à partir des compétences* (p. 189-212). De Boeck Supérieur. https://doi.org/10.3917/dbu.pouma.2017.01.0189

Prokofieva, V., Brandt-Pomares, P., Velay, J.-L., Hérold J.-F. et Kostromina, S. (2017). Stress de l'évaluation scolaire : un nouveau regard sur un problème ancien. *Recherches & éducations*, 18. https://doi.org/10.4000/rechercheseducations.4657

Qadir, J. (2023). Engineering Education in the Era of ChatGPT : Promise and Pitfalls of Generative AI for Education. *2023 IEEE Global Engineering Education Conference (EDUCON)*, 1-9. https://doi.org/10.1109/EDUCON54358.2023.10125121

Qvortrup, A. et Keiding, T. B. (2014). Portfolio assessment : production and reduction of complexity. *Assessment & Evaluation in Higher Education*, *40*(3), 407-419. https://doi.org/10.1080/02602938.2014.918087

Raaper, R. et Brown, C. (2020). The Covid-19 pandemic and the dissolution of the university campus : Implications for student support practice. *Journal of Professional Capital and Community*, *5*(3-4), 343-349. https://doi.org/10.1108/JPCC-06-2020-0032

Redecker, C. (2017). *European framework for the digital competence of educators : DigCompEdu*. Publications Office of the European Union. https://doi.org/10.2760/159770

Rege Colet, N. et Berthiaume, D. (dir.) (2015). *La pédagogie de l'enseignement supérieur : repères théoriques et applications pratiques. Tome 2 : Se développer au titre d'enseignant*. Peter Lang.

Rege Colet, N., McAlpine, L., Fanghanel, J. et Weston, C. (2011). Le concept de *Scholarship of Teaching and Learning*. *Recherche et formation*, 67, 91-104. https://doi.org/10.4000/rechercheformation.1412

Richardot, S. (2022). Chapitre 7. Étudier les représentations sociales en contexte : l'emploi de scénarios. Dans B. Albero et J. Thievenaz (dir.), *Enquêter dans les métiers de l'humain : Traité de méthodologie de la recherche en sciences de l'éducation et de la formation. Tome II* (2e éd., p. 379-395). Éditions Raison et Passions. https://doi.org/10.3917/rp.alber.2022.02.0379

Ridgway, J., McCusker, S. et Pead, D. (2004). *Literature review of e-assessment*. Futurelab. https://durham-repository.worktribe.com/output/1681299 (consulté le 02.01.2026).

Roegiers, X. (2000). *Une pédagogie de l'intégration. Compétences et intégration des acquis dans l'enseignement*. De Boeck Université.

Romainville, M. (2023). L'« alignement pédagogique » : un arbre qui cache la forêt ? *Revue française de pédagogie. Recherches en éducation*, *221*, 109-116. https://doi.org/10.4000/11w0e

Romainville, M. et Fischer, L. (2020). Dilemmes relatifs à l'usage du portfolio en enseignement supérieur. *Évaluer. Journal international de recherche en éducation et formation*, *6*(1), 79-90.

Romero, M. (2019). Analyser les apprentissages à partir des traces. *Distances et médiations des savoirs*, *26*. https://doi.org/10.4000/dms.3754

Roy, N., Gareau, A. et Poellhuber, B. (2018). Les natifs du numérique aux études : enjeux et pratiques. *La revue canadienne de l'apprentissage et de la technologie*, *44*(1), 1-24. https://doi.org/10.21432/cjlt27558

Sahin, M. et Ifenthaler, D. (dir.) (2024). *Assessment Analytics in Education : Designs, Methods and Solutions*. Springer.

Saussez, F. et Allal, L. (2007). Réfléchir sur sa pratique : le rôle de l'autoévaluation ? *Mesure et évaluation en éducation*, *30*(1), 97-124. https://doi.org/10.7202/1086174ar

Scallon, G. (2015). *Des savoirs aux compétences. Exploration en évaluation des apprentissages*. De Boeck Supérieur.

Secrétariat d'État à la formation, à la recherche et à l'innovation (SBFI) (2018). *Compétences transversales : Cadre de référence pour la formation professionnelle initiale*. SBFI. https://www.sbfi.admin.ch/dam/sbfi/fr/dokumente/webshop/2018/transversale-kompetenzen.pdf.download.pdf/Transversale_Kompetenzen_BB2030_f.pdf (consulté le 02.01.2026).

Seufert, S., Guggemos, J. et Sonderegger, S. (2019). *Learning Analytics in Higher Education using Peer-feedback and Self-assessment : Use Case of an Academic Writing Course* [communication orale]. 11th International Conference on Computer Supported Education, Héraklion, Grèce.

Siddiq, F., Olofsson, A. D., Lindberg, J. O. et Tomczyk, L. (2024). Special issue : What will be the new normal ? Digital competence and 21st-century skills : critical and emergent issues in education. *Education and Information Technologies*, *29*(6), 7697-7705. https://doi.org/10.1007/s10639-023-12067-y

Slade, C., Lawrie, G., Taptamat, N., Browne, E., Sheppard, K. et Matthews, K. E. (2022). Insights into how academics reframed their assessment during a pandemic : Disciplinary variation and assessment as afterthought. *Assessment & Evaluation in Higher Education*, *47*(4), 588-605. https://doi.org/10.1080/02602938.2021.1933379

Souppez, J.-B. R. G., Goswami, D. et Yuen, J. (2023). *Assessment and Feedback in the Generative AI Era : Transformative Opportunities, Novel Assessment Strategies and Policies in Higher Education* [communication orale]. International Federation of National Teaching Fellows Symposathon 2023. https://publications.aston.ac.uk/id/eprint/45763/ (consulté le 02.01.2026).

St-Onge, C., Ouellet, K., Lakhal, S., Dubé, T. et Marceau, M. (2022). COVID-19 as the tipping point for integrating e-assessment in higher education practices. *British Journal of Educational Technology*, *53*(2), 349-366. https://doi.org/10.1111/bjet.13169

Stes, A. et Van Petegem, P. (2011). La formation pédagogique des professeurs dans l'enseignement supérieur : une étude d'impact. *Recherche & formation*, 67, 15-30. https://doi.org/10.4000/rechercheformation.1360

Sung, G., Guillain, L. et Schneider, B. (2023). Can AI help teachers write higher quality feedback ? Lessons learned from using the GPT-3 engine in a makerspace course. Dans P. Blikstein, J. Van Aalst, R. Kizito et K. Brennan (dir.), Proceedings of the 17th International Conference of the Learning Sciences – ICLS 2023 (p. 2093-2094). International Society of the Learning Sciences. https://doi.org/10.22318/icls2023.904961

Sylvestre, E, van de Poël, J.-F., Carbonel, H. et Jullien, J.-M. (2020, avril). Vade-mecum pour l'évaluation à distance des étudiant·e·s. https://enseigner.unil.ch/ressources/document/vade-mecum-pour-levaluation-a-distance-des-etudiants/ (consulté le 02.01.2026).

Tardif, J. (2006). *L'évaluation des compétences : Documenter le parcours de développement*. Chenelière Éducation.

Tardif, J. (2017). Chapitre 1. Des repères conceptuels à propos de la notion de compétence, de son développement et de son évaluation. Dans M. Poumay, J. Tardif et F. Georges (dir.), *Organiser la formation à partir des compétences* (p. 15-37). De Boeck Supérieur. https://doi.org/10.3917/dbu.pouma.2017.01.0015

UNESCO (2020). *Global education monitoring report, 2020 : Inclusion and education : all means all*. https://doi.org/10.54676/JJNK6989 (consulté le 19.01.2026).

Varghese, N. V. (2015). *Challenges of massification of higher education in India* [document inédit]. https://doi.org/10.13140/RG.2.1.4843.3681

Viau, R. (1995). L'état des recherches sur l'anxiété en contexte scolaire. *Cahiers de la recherche en éducation*, *2*(2), 375-398. https://doi.org/10.7202/1018209ar

Viberg, O., Mutimukwe, C., Hrastinski, S., Cerratto-Pargman, T. et Lilliesköld, J. (2024). Exploring teachers' (future) digital assessment practices in higher education : Instrument and model development. *British Journal of Educational Technology*, *55*(6), 2597-2616. https://doi.org/10.1111/bjet.13462

Wahas, Y. M. A. et Syed, A. J. A. (2024). E-assessment challenges during e-learning in higher education : A case study. *Education and Information Technologies*, *29*, 14431-14450. https://doi.org/10.1007/s10639-023-12421-0

Wang, L., Chen, X., Wang, C., Xu, L., Shadiev, R. et Li, Y. (2024). ChatGPT's capabilities in providing feedback on undergraduate students' argumentation : A case study. *Thinking Skills and Creativity*, *51*. https://doi.org/10.1016/j.tsc.2023.101440

Wareing, S. (2022). Chapter 29. Digital education : less change and more change than predicted. Dans R. Sharpe, S. Bennett et T. Varga-Atkins (dir.), *Handbook of Digital Higher Education* (p. 363-373). Elgaronline. https://doi.org/10.4337/9781800888494.00040

Winstone, N. E. (2022). Characterising feedback cultures in higher education : an analysis of strategy documents from 134 UK universities. *Higher Education*, *84*, 1107-1125. https://doi.org/10.1007/s10734-022-00818-8

Winstone, N. E., Nash, R. A., Parker, M. et Rowntree, J. (2017). Supporting Learners' Agentic Engagement With Feedback : A Systematic Review and a Taxonomy of Recipience Processes. *Educational Psychologist*, *52*(1), 17-37. https://doi.org/10.1080/00461520.2016.1207538

Yale-Soulière, G., Turgeon, L., Brière, F. N., Alawie, F., Imbeault, A., Morizot, J. et Janosz, M. (2023). Impact de l'anxiété d'évaluation sur le fonctionnement scolaire et psychologique des adolescents. *Revue de psychoéducation*, *52*(1), 136-157. https://doi.org/10.7202/1099291ar

Yang, M. et Carless, D. (2012). The feedback triangle and the enhancement of dialogic feedback processes. *Teaching in Higher Education*, *18*(3), 285-297. https://doi.org/10.1080/13562517.2012.719154

Yerly, G. et Issaieva, E. (2021). (Re)penser l'évaluation des apprentissages au postsecondaire en temps de crise : défis à relever et occasions à saisir en période de COVID-19. *Revue internationale des technologies en pédagogie universitaire*, *18*(1), 89-101. https://doi.org/10.18162/ritpu-2021-v18n1-09

Young, Dolly J. (1990). An Investigation of Students' Perspectives on Anxiety and Speaking. *Foreign Language Annals*, *23*(6), 539-553.

Présentation des auteurs et autrices

Hiba Asri est professeure associée au département d'informatique de la faculté des sciences Semlalia et membre du Centre d'innovation pédagogique à l'Université Cadi Ayyad de Marrakech. Elle travaille principalement sur le Big Data, l'intelligence artificielle et l'Internet des objets, avec des applications dans la santé, l'éducation et l'agriculture. Lauréate du prix d'excellence en recherche de l'Université en 2024 et du prix UNESCO-ICHEI en 2022, elle est aussi instructrice certifiée Huawei ICT Academy.

Assia Belgheddouche est professeure à l'École normale supérieure de Bouzaréah (Alger) et cheffe d'une équipe de recherche en pédagogie universitaire au laboratoire LISODIP. Elle est titulaire d'un doctorat en sciences du langage à l'Université Paul Valéry (Montpellier). Elle s'intéresse aux questions en lien avec la pédagogie universitaire, l'enseignement/apprentissage des langues étrangères, l'analyse des pratiques professionnelles et la didactique du plurilinguisme. Elle a été directrice du laboratoire LISODIP et fait partie du comité de rédaction de la revue *Socles*. Elle a également participé à des projets de coopération internationale.

Evi Belsack est conseillère pédagogique au Centre d'appui pédagogique de l'Université libre de Bruxelles depuis 2016. Elle y accompagne les enseignants et enseignantes dans la création ou la modification de leurs dispositifs d'enseignement, avec ou sans outils technologiques. Titulaire d'un doctorat en sciences de l'éducation de l'Université libre de Bruxelles et de Nantes Université, elle s'intéresse particulièrement aux pédagogies actives ainsi qu'à l'évaluation des apprentissages. Elle collabore régulièrement avec Patricia Mercier sur la réalisation de tutoriels et de dispositifs impliquant l'utilisation de Moodle.

Omar Benkacem est chef de projet informatique et spécialiste en e-learning, responsable du service e-assessment à l'Université de Genève. Titulaire d'un master en informatique de l'Université de Genève et d'un master en physique obtenu au Maroc, il a été CTO du

projet CoseLearn (2003-2010), soutenu par la DDC, visant à promouvoir l'e-learning en Afrique francophone et arabophone. Depuis 2010, à l'Université de Genève, il a piloté des projets stratégiques, notamment l'institutionnalisation et le déploiement à grande échelle du e-assessment, la certification des diplômes via la *blockchain*, le développement d'un environnement d'apprentissage personnalisé, ainsi que la gestion de l'identité numérique académique. Il est également actif dans les communautés SWITCH Eduhub et copréside le groupe SIG-eassessment en Suisse.

Allan Bowman est analyste-programmeur au département ELearning – eAssessment de l'Université de Genève. Titulaire d'un bachelor en informatique de gestion de la HES-SO Genève, il poursuit actuellement un Master of Science en Business Information Systems à la FHNW. Fort d'une expérience diversifiée dans le développement web, le design interactif et l'analyse de données, Allan Bowman se spécialise dans la gestion de projets numériques innovants.

André Camacho est analyste-programmeur à l'Université de Genève. Diplômé en sciences de l'éducation et titulaire d'un Master of Science in Learning and Teaching Technologies à l'Université de Genève, il s'est spécialisé dans les technologies éducatives, la programmation et le développement. Parallèlement à son métier et à ses recherches, il travaille en tant que consultant et développeur indépendant. Actuellement employé à l'Université de Genève dans l'équipe e-assessment et e-learning, il participe à divers projets d'innovation et coopère sur la notion d'examens sur ordinateur avec diverses universités et hautes écoles suisses.

Bernadette Charlier est professeure ordinaire émérite à l'Université de Fribourg (Suisse) depuis août 2022. Docteure en sciences de l'éducation (Université catholique de Louvain, 1996) et titulaire d'une licence en sciences de l'éducation (Université de Liège, 1979), elle est une spécialiste internationalement reconnue en technologie de l'éducation et de la pédagogie universitaire. Après avoir commencé sa carrière académique à l'Université de Namur (Belgique), elle a rejoint l'Université de Fribourg en 2002, où elle a dirigé en tant que professeure le Centre de didactique universitaire. Ses travaux de recherche portent notamment sur les dispositifs de formation hybrides, le développement et l'évaluation de dispositifs de formation, et les environnements personnels d'apprentissage. Très active au sein de la communauté scientifique, elle a coordonné de nombreux projets de recherche européens et collaboré

étroitement avec des instances d'évaluation de l'enseignement supérieur. Depuis son éméritat, elle poursuit ses activités d'enseignement à l'UCLouvain (Belgique) et ses activités de recherche, d'expertise et de consultance dans le cadre de l'entreprise YllYl en Suisse.

Aurélie Dupré est directrice d'appui à la pédagogie et à l'innovation et chercheuse au laboratoire CIREL à l'Université de Lille. Docteure en sciences de l'éducation et de la formation, ses travaux s'inscrivent dans le champ de la pédagogie universitaire. Elle s'intéresse au développement professionnel en pédagogie des enseignants et à l'accompagnement des innovations curriculaires et pédagogiques dans l'enseignement supérieur (accompagnement des équipes pédagogiques dans la mise en œuvre de l'approche par compétences et la flexibilisation des parcours de formation). Aurélie Dupré est responsable du DU de pédagogie universitaire de l'Université de Lille et coordonnatrice pédagogique nationale du projet Avenir(s)-volet ESR.

Étienne Gibaud est enseignant-chercheur permanent à l'EPF École d'ingénieurs sur le campus de Montpellier. Docteur en mathématiques appliquées et modélisation, il se spécialise en physique numérique dans ses activités d'enseignement. Passionné par l'éducation, il est impliqué dans la Cellule d'innovation pédagogique et numérique de l'EPF, notamment en tant que pilote du groupe de travail Recherche et valorisation. Intéressé par les thématiques de la réflexivité, de la métacognition et de l'apprentissage, mais aussi par la formation des enseignants par compagnonnage, il participe à diverses initiatives de développement pédagogiques internes et de nombreux efforts de communications en colloques et séminaires.

Jean-Luc Gilles est professeur honoraire à la Haute école pédagogique du canton de Vaud (HEP Vaud, Lausanne, Suisse) depuis août 2025. Docteur en sciences de l'éducation (Université de Liège, 2002) et titulaire d'un master en sciences psycho-pédagogiques (Université de Mons, 1992), il s'est spécialisé en ingénierie des évaluations et en approche qualité des dispositifs de formation. Après avoir dirigé le CAFEIM puis le SMART à l'Université de Liège, il a rejoint la HEP Vaud en 2010, où il a contribué à la recherche et à l'enseignement du master MASPE. Cofondateur et président de la spin-off Assess-Group S.A. (de 2006 à 2016), il est également cofondateur du Groupe international de recherche et expertises en ingénierie des évaluations en formation (GIRIEF) ainsi que de l'association

EvalU. Depuis 2025, il intervient comme expert dans les audits institutionnels de l'enseignement supérieur et propose des activités de consultance en évaluation et formation.

Éric Giraudin, affecté à l'Université Savoie Mont-Blanc et membre du réseau de recherche ERECO, est directeur du volet ESR du programme Avenir(s) – France 2030. Depuis plus de dix ans, il pilote des projets de transformation pédagogique et numérique dans l'enseignement supérieur, avec une expertise particulière dans la mise en œuvre de la démarche portfolio pour l'APC et le projet de vie des apprenants. Il collabore avec les acteurs de la pédagogie, du numérique et de la recherche ainsi qu'avec des acteurs du monde socio-économique pour développer des solutions innovantes, notamment autour : des portfolios numériques, de la professionnalisation des étudiants, et de l'accompagnement des équipes pour l'évolution des pratiques pédagogiques.

Diana Griffoulières occupe le poste de responsable de l'ingénierie pédagogique numérique à l'EPF École d'ingénieurs en France. Elle est titulaire d'un doctorat en sciences de l'éducation et e-learning, obtenu à l'Université ouverte de Catalogne (UOC) en Espagne. Son expérience et ses travaux se concentrent sur la conception et la mise en œuvre de dispositifs pédagogiques, visant à promouvoir le développement professionnel des enseignants et la réussite des étudiants. Elle s'intéresse également à la réflexion sur l'intégration efficace du numérique dans les pratiques pédagogiques. De plus, elle est engagée dans la recherche appliquée en pédagogie de l'enseignement supérieur, contribuant ainsi à l'évolution des méthodes d'enseignement et à l'amélioration des expériences d'apprentissage pour les étudiants.

Anne-Céline Grolleau est chargée de développement pédagogique depuis 2007. Après avoir monté, piloté et animé un dispositif de développement professionnel pédagogique en pédagogie numérique et de l'enseignement supérieur pour les grandes écoles de la région Pays de la Loire, elle a rejoint l'Université de Nantes en 2013 où elle a accompagné les acteurs de l'Université dans leurs projets pédagogiques avant de coordonner les accompagnements à l'intégration des logiques compétences dans les formations à partir de 2019, dans le cadre du projet NeptUNe. Depuis 2024, elle est responsable du service d'appui au développement professionnel en pédagogie au sein de la direction d'appui à la pédagogie et à l'innovation de l'Université de Lille.

Jean-Michel Jullien est responsable du service Educational Development Unit for Distance Learning (EDUDL+) d'UniDistance Suisse. Ses centres d'intérêt sont l'enseignement et l'apprentissage avec le digital, le développement professionnel des enseignants et la qualité dans l'enseignement à distance. Ses travaux de recherche-développement récents visent à accompagner les gouvernances des institutions à l'aide de matrices qualitatives : matrice de maturité institutionnelle du déploiement des compétences digitales (2022), matrice de qualité des espaces d'enseignement et d'apprentissage (2023), d'évaluation des stratégies d'e-assessment (2025).

Patricia Mercier est conseillère technopédagogique au Centre d'appui pédagogique (CAP) de l'Université libre de Bruxelles (ULB) depuis 2017 pour l'accompagnement des enseignants et des enseignantes à l'utilisation des TICE. Elle est actuellement cheffe de projet pour l'Université virtuelle, la plateforme d'apprentissage de l'ULB. Formatrice, rédactrice de tutoriels et gestionnaire d'espaces d'aide aux utilisateurs et utilisatrices, elle s'est peu à peu spécialisée dans l'enseignement hybride, notamment sur Moodle, un des systèmes de gestion de l'apprentissage (ou LMS) open source le plus utilisé au monde.

Laurent Moccozet est maître d'enseignement et de recherche au Centre universitaire d'informatique de l'Université de Genève. Il est titulaire d'un doctorat en systèmes d'information obtenu à l'Université de Genève. Il enseigne dans divers domaines des systèmes d'information et codirige le bachelor en systèmes d'information. Il développe une partie de ses activités de recherche dans les applications des technologies numériques pour soutenir les activités d'enseignement et l'apprentissage : les environnements personnels d'apprentissage, l'évaluation des connaissances et des compétences, etc. Il s'intéresse aussi aux espaces de fabrications, les FabLabs, comme lieux d'innovation et d'expérimentation pour l'apprentissage à l'université. Il collabore régulièrement avec l'équipe e-learning de l'Université de Genève.

Abdelali Rochdi est directeur du Centre d'innovation pédagogique à l'Université Cadi Ayyad de Marrakech. Docteur en chimie bio-organique, il s'est progressivement intéressé à des thématiques transversales telles que les compétences numériques et l'enseignement à distance. En qualité d'enseignant-chercheur, ses recherches portent sur l'innovation pédagogique, en particulier concernant l'apport de l'intelligence artificielle dans le domaine éducatif.

Patrick Roth est responsable du pôle e-learning à l'Université de Genève. Il est en charge du développement et de l'exploitation des technologies éducatives. Docteur en sciences informatiques, il a obtenu son doctorat à l'UNIGE en 2002 dans le domaine de l'interaction homme-machine (IHM). Lauréat du prix LATSIS en 2004, il a ensuite effectué un post-doctorat à l'Université Humboldt. Parallèlement à ses responsabilités au pôle e-learning, il enseigne au Centre universitaire d'informatique de l'UNIGE et supervise des étudiants de bachelor et de master. Ses domaines d'intérêt portent notamment sur les plateformes d'apprentissage en ligne (LMS, streaming, etc.), l'e-portfolio, les compétences numériques et les IHMs. Il est également membre des communautés LERU DIGE, Switch eduhub, ainsi que de plusieurs groupes d'intérêt suisses tels que SIG-Portfolio, Digital Skills et SIG-AI.

François Stephan est directeur du campus de Montpellier de l'EPF École d'ingénieurs. Docteur en mécanique, il se passionne pour la pédagogie et assure en école d'ingénieurs des postes de directeur des Études. Il crée à l'EPF une « cellule » d'innovation pédagogique en 2017 avec les objectifs de mener avec les enseignants des réflexions sur l'acte d'enseigner et d'expérimenter de nouvelles pratiques. Il transforme son campus en un vaste champ d'expérimentation tant sur l'aménagement des espaces que sur les méthodes d'enseignement, l'évaluation et le rapport à l'étudiant.

Emmanuel Sylvestre est professeur titulaire et directeur du Centre de soutien à l'enseignement à l'Université de Lausanne. Dans ses fonctions, il accompagne les enseignants et les équipes pédagogiques dans leurs démarches réflexives, et contribue à l'élaboration de la stratégie institutionnelle en matière d'enseignement et d'apprentissage. Ses recherches conduites au sein de l'Institut des sciences sociales et de l'Observatoire de l'éducation et de la formation de l'Université de Lausanne se concentrent principalement sur la pédagogie universitaire, avec un intérêt spécifique pour l'évaluation des dispositifs pédagogiques, le design pédagogique et la pédagogie inclusive. En outre, il assume le rôle de rédacteur en chef de la revue e-Jiref de l'Association pour le développement des méthodologies d'évaluation en éducation (ADMEE-Europe).

Table des matières

Sommaire 5

Acronymes et sigles 7

Introduction 9
Emmanuel Sylvestre & Jean-Michel Jullien

Dimension pédagogique 10
Dimension docimologique 11
Dimension technologique 12
Dimension éthique 14
Que propose cet ouvrage ? 16

Partie 1
Évaluer à l'échelle de l'enseignement
Innovations pédagogiques et dispositifs en contexte de classe 19

1 Le e-portfolio comme outil pour l'évaluation des compétences: analyse d'usage dans une formation de bachelor universitaire de technologie 21
Aurélie Dupré & Éric Giraudin

Résumé 21
Introduction 21
Le contexte et la problématique 22
Éléments de cadrage théorique 24
L'approche par compétences 24
L'évaluation des compétences et la démarche e-portfolio 25
Présentation du dispositif 26
La méthodologie 27

Variables d'analyse des productions étudiantes 30
Variables d'analyse des autopositionnements des étudiants 31
Variables d'analyse des évaluations portées par les enseignants 32
Présentation des résultats 32
Les productions des étudiants dans le e-portfolio 33
L'évaluation du niveau de compétence visé 37
La preuve du développement de la compétence 40
Conclusion et conseils de mise en œuvre pratique 42

2 La vidéo comme outil d'évaluation de la production orale: cas d'un groupe d'étudiantes de première année en licence de français à l'ENSB 45
Assia Belgheddouche

Résumé 45
Introduction 45
Le contexte et la problématique 46
Le cadre théorique 49
L'enseignement et l'évaluation de l'oral 49
L'anxiété d'évaluation 50
L'évaluation de l'oral à l'ère du numérique 51
Le dispositif et le scénario pédagogique 53
La méthodologie 54
Les résultats et leur analyse 56
Ce que le questionnaire a révélé 56
Les modalités de réalisation de l'activité de production orale 56
Conseils de mise en œuvre pratique 65

3 Pratiques évaluatives d'enseignants du supérieur 67
Aurélie Dupré & Anne-Céline Grolleau

Résumé 67
Introduction 67
Contexte et problématique 68
Éléments de cadrage théorique 70
Le développement professionnel en pédagogie des enseignants du supérieur 70

Représentations et numérique 71
Méthodologie 72
Analyse des représentations des enseignants 72
Analyse des dispositifs d'évaluation 73
Résultats et analyse 74
Représentations des enseignants sur les usages du numérique pour l'évaluation 74
Représentations sur l'évaluation 77
Représentations sur le numérique 79
Représentations sur l'usage du numérique pour l'évaluation 79
Usages des outils numériques dans les pratiques d'évaluation des enseignants 80
Des outils numériques pour la mise en place de QCM lors d'évaluations formatives ou diagnostiques 80
Des pratiques d'évaluation certificatives sans usage du numérique 81
L'usage du numérique comme outil de dépôt de productions support à l'évaluation 82
Le e-portfolio, un usage émergent du numérique dans une visée certificative 84
Conclusion et conseils de mise en œuvre pratique 86

4 *Feedbacks* par intelligence artificielle en école d'ingénieurs: un premier regard 89
Étienne Gibaud, François Stephan & Diana Griffoulières

Résumé 89
Introduction 90
Le contexte et la problématique 90
Le cadre théorique 91
Le dispositif et le scénario pédagogique 94
La méthodologie 95
Matrice d'évaluation 96
Les résultats et leur analyse 97
Discussion et conclusions 101
Conseils de mise en œuvre pratique 102

Partie 2

Évaluer à l'échelle de l'institution

Infrastructures, stratégies et transitions numériques

105

5 Intégration des outils numériques dans l'évaluation des apprentissages
Étude de cas: Université Cadi Ayyad – Maroc 107
Hiba Asri & Abdelali Rochdi

Résumé 107
Introduction 108
Contexte et problématique 110
Cadre théorique 114
Dispositif et scénario pédagogique 116
Dispositif 1 : Évaluation numérique via la plateforme Moodle 116
Dispositif 2 : Évaluation papier avec l'outil AMC 118
Méthodologie 120
Évaluation numérique : plateforme Moodle 122
Évaluation papier avec l'outil AMC 126
Résultats et discussion 127
Analyse des résultats 129
Impact de l'utilisation des dispositifs proposés 132
Conseils de mise en œuvre pratique 134
Conclusion et travaux futurs 134

6 Évolution des modalités et infrastructures des examens en ligne 137
Laurent Moccozet, Omar Benkacem, André Camacho, Allan Bowman & Patrick Roth

Résumé 137
Introduction 137
Le contexte et la problématique 138
Le cadre théorique 139
Le dispositif et le scénario pédagogique 142
La méthodologie 143
Les résultats et leur analyse 144

Analyse des statistiques des examens numériques 144
Analyse de l'étude post-Covid auprès des enseignants 150
Analyse des pratiques innovantes 153
Conseils de mise en œuvre pratique 155

7 Évaluations en ligne post-Covid: quels accompagnements et quel alignement pédagogique ? 159
Evi Belsack & Patricia Mercier

Résumé 159
Introduction 160
Contexte et problématique 161
Comprendre la persistance des examens à distance 162
Le cadre théorique 163
La situation particulière à l'ULB selon trois dimensions 163
La méthodologie 167
Enquête en ligne 171
Les résultats et leur analyse 172
Conclusion et conseils de mise en œuvre pratique 177

Conclusion
Évaluer à l'ère numérique, entre potentialités, risques et responsabilités 179
Bernadette Charlier & Jean-Luc Gilles

Quelles transformations ? 182
Quels effets ? 183
Effets sur la qualité des évaluations 183
Effets sur le développement des compétences des étudiants 190
Quels enjeux ? 191
Enjeux pour la recherche 192
Enjeux pour la qualité et l'éthique des pratiques évaluatives 192
Enjeux liés aux opportunités et aux risques émergents 193

Bibliographie 195

Présentation des auteurs et autrices 207