

Repenser l'évaluation des apprentissages à l'ère de l'IA : une revue de la littérature anglophone

Reconsidering Learning Assessments in the Era of Artificial Intelligence: An English-language Literature Review

Cathia Papi, Amélie Beaudoin, Nilce Da Silva, Serge Gérin-Lajoie et Marie-Hélène Hébert

Université TÉLUQ

Pour citer cet article : Papi, C., Beaudoin, A., Da Silva, N., Gérin-Lajoie, S., & Hébert, M.-H. (2026). Repenser l'évaluation des apprentissages à l'ère de l'IA : Une revue de la littérature anglophone. *Distances et médiations des savoirs. Distance and Mediation of Knowledge*, (54). <https://doi.org/10.4000/16gfr>

Résumés

Français

Dans le cadre d'une revue systématique narrative de la littérature (Turnbull *et al.*, 2023), nous avons analysé 111 articles publiés dans des revues scientifiques, entre janvier 2020 et décembre 2024, afin de chercher à saisir l'impact de l'intelligence artificielle (IA) sur l'évaluation des apprentissages dans l'enseignement supérieur. Il ressort de ce corpus d'articles que l'IA générative (IAG) peut « réussir » certains types d'activités d'évaluation et que son usage est parfois difficile à détecter. Ceci conduit à la mise en place de systèmes de surveillance numérique ainsi qu'à la création d'activités d'évaluation pour lesquelles il est plus difficile ou moins pertinent de recourir à l'IA. De plus, l'IAG est également présentée comme un soutien potentiel tant pour les enseignants susceptibles de s'en servir pour concevoir ou corriger certaines activités d'évaluation des apprentissages, que pour les étudiants qui peuvent l'utiliser pour s'entraîner et bénéficier de rétroactions rapides en tout temps. Alors que les recours potentiels à l'IA dans l'évaluation des apprentissages sont nombreux, plusieurs auteurs avancent que les formations et politiques permettant d'encadrer ses usages demeurent à développer.

English

As part of a systematic-narrative literature review (Turnbull *et al.*, 2023), we analyzed 111 articles published in peer-reviewed journals between January 2020 and December 2024 to understand how artificial intelligence (AI) is influencing assessment in higher education. Through this process, we observed that generative AI (GenAI) can successfully replicate a human performance during certain assessment types and that GenAI misuse is sometimes challenging to detect. Thus, institutions and faculty members resort to AI-resistant examination designs and automated proctoring systems. Furthermore, GenAI is perceived as an opportunity for pedagogical support: for professors and instructors, it can assist with creating exams and grading; for students, it can provide immediate feedback and aid practice. While AI has several potential uses in assessments, many authors underscore the importance of developing training and policies.

Introduction

L'intelligence artificielle (IA) se développe de manière exponentielle, inspirant « enthousiasme et méfiance » (Roy *et al.*, 2025) aux acteurs de l'éducation. Le passage d'une intelligence artificielle plus élémentaire apportant certaines formes d'assistance notamment sous forme d'objets connectés, à l'intelligence artificielle générative (IAG) réalisant des productions textuelles, visuelles, sonores, voire programmatiques, marqué par la sortie de ChatGPT en novembre 2022, suscite de nombreux questionnements, notamment dans le milieu de l'enseignement supérieur (Audran, 2024).

Il est entre autres craint que l'IAG soit utilisée dans le cadre de travaux notés ou d'examens, car il est connu de longue date que la pression à réussir leurs études peut conduire les étudiants à plagier ou à tricher, voire à payer quelqu'un pour faire les travaux dont les résultats comptent pour la validation d'un cours ou d'une année (Callahan et Davidson, 2004; Hollis, 2018; Ison, 2012). Les étudiants font aussi parfois part de leur manque de connaissances de ce qui est autorisé ou non (Boudreau, 2020 ; Heckler et Forde, 2015). Plus généralement, Berdahl (2023) recense six motifs de fraude en milieu universitaire :

1. le manque d'intérêt pour le contenu ;
2. le manque de compétences des étudiants ;
3. le manque de compréhension des attentes ;
4. le manque d'intérêt envers le personnel enseignant ;
5. la tolérance de l'inconduite ;
6. le stress associé au manque de temps et à la performance attendue.

Ces divers comportements qui vont à l'encontre des normes, des règles et des principes éthiques (Denney *et al.*, 2021) tendent à être rassemblés dans l'expression « fraude académique » (Eckstein, 2003). Ce phénomène, certainement aussi vieux que l'existence d'évaluations notées, est expliqué par diverses théories (Wideman, 2008) telles que le triangle de la fraude de Cressey illustrant, depuis la fin des années 1960, le développement de la fraude entre opportunité, motivation et rationalisation (Pluchart, 2019). D'où les nombreuses inquiétudes ayant émergé lors du passage obligé par la formation à distance pendant la pandémie de COVID-19 (Detroz *et al.*, 2020 ; Papi *et al.*, 2020), en ce sens que la malhonnêteté académique peut sembler facilitée par l'absence de surveillance en présence (Burgason *et al.*, 2019) et est susceptible de remettre en cause la crédibilité ou la valeur des diplômes délivrés (Cizek, 1999 ; Fendler *et al.*, 2018). Ces préoccupations n'ont cessé de s'accroître au-delà de la pandémie, quel que soit le mode d'enseignement, en raison du développement de l'IAG.

De fait, les grands modèles de langage (*large language models*) constitutifs des outils d'IAG rendent ces derniers capables de générer des dissertations, de répondre à des questions, de

résumer des textes, etc. (Conseil supérieur de l'éducation et Commission de l'éthique en science et en technologie, 2024). Une enquête réalisée au Royaume-Uni concernant les usages malhonnêtes de l'IAG, recensés entre 2019 et 2024 dans 131 universités, présentée par le périodique *The Guardian*, fait par exemple ressortir une forte augmentation des cas de fraude réalisée avec l'IAG : 5,1 cas pour 1 000 étudiants en 2023-2024 contre 1,6 cas pour 1 000 une année auparavant (Goodier, 2025). Le plagiat traditionnel, consistant à recopier des passages de textes écrits par un autre être humain sans le reconnaître ou le référencer de manière appropriée (Yeo, 2007), tend ainsi à être remplacé par une forme plus sophistiquée combinant l'œuvre de l'IAG, qui crée à partir de contenus existants, et celle de l'étudiant, qui fait un copier-coller plus ou moins modifié de ce qu'a généré l'IAG sans mentionner cette dernière, ce que nous pourrions désigner de « plagIA » en écho au néologisme *AIgiarism* (Hern, 2022) anglophone.

Il est généralement estimé que l'IAG étant présente, il convient d'apprendre à composer avec, tout en encadrant ses usages (Conseil supérieur de l'éducation et Commission de l'éthique en science et en technologie, 2024). Pour autant, ce qui est mis en place sur le terrain et la manière dont l'IAG est acceptée ou non, intégrée ou non, dans les cours et activités d'évaluation semble varier selon les initiatives des enseignants, voire des établissements, et reste ainsi peu connu. Dès lors, dans quelle mesure l'IAG a-t-elle des répercussions sur les représentations, pratiques et politiques d'évaluation des apprentissages dans l'enseignement supérieur ?

Afin de faire le point sur la situation et de développer une réflexion sur les enjeux actuels et à venir, nous avons réalisé une revue systématique narrative de la littérature (Turnbull *et al.*, 2023) des articles publiés dans les revues scientifiques jusqu'à la fin de l'année 2024. Après avoir décrit la méthodologie, nous allons présenter les principaux résultats en nous intéressant d'une part aux risques que les recours à l'IA représentent pour l'intégrité académique, et d'autre part à ses potentialités et enjeux. Nous ouvrirons ensuite une discussion sur l'impact de l'IAG dans l'évaluation des apprentissages.

Méthodologie

Dans le but de réaliser une recherche exhaustive sur les répercussions de l'IA¹ dans l'évaluation dans l'enseignement supérieur, nous avons utilisé la méthode développée par le centre Evidence for Policy and Practice Information and Coordinating (EPPI) de l'University College de Londres (EPPI-Centre, 2010). La revue systématique narrative de la littérature (Turnbull *et al.*, 2023) menée entre octobre 2024 et mai 2025, visait plus précisément à documenter la manière dont les développements de l'IA amenaient ou non des changements concernant l'évaluation des apprentissages en prenant en compte les représentations et pratiques des étudiants et des enseignants ainsi que les politiques ou directives institutionnelles. Autrement dit, au niveau micro, dans quelle mesure la diffusion d'outils d'IA a-t-elle conduit étudiants et enseignants à

¹ Nous n'avons pas restreint à l'IAG, mais les articles recueillis évoquent presque exclusivement l'IAG.

modifier leurs regards ou pratiques concernant l'évaluation des apprentissages ? Du point de vue méso, dans quelle mesure les établissements d'enseignement supérieur encadrent-ils les usages de l'IA dans l'évaluation ou que font-ils pour garantir la validité des diplômes et certifications délivrés ?

Tel que présenté dans le rapport de processus de recension systématique (Papi *et al.*, 2025), afin de répondre à ces questions, nous avons défini différents groupes sémantiques caractérisant notre recherche et avons consulté les thésaurus des bases de données ERIC et Education Source² ainsi que diverses bases terminologiques comme la Vitrine linguistique (vitrinelinguistique.oqlf.gouv.qc.ca) et Termium Plus® (btb.termiumplus.gc.ca) pour obtenir un éventail de mots-clés le plus large possible³. Nous avons ensuite cherché leur fréquence dans les bases de données visées afin de sélectionner les plus pertinents pour chacun des groupes sémantiques⁴. Après avoir testé plusieurs agencements entre les groupes sémantiques, nous avons employé la requête qui produisait les meilleurs résultats :

Tableau 1.

Requête employée pour sonder les bases de données bibliographiques

Évaluation, plagiat ou tricherie	([<i>assess*</i> OR <i>evaluation</i> OR <i>test</i> OR <i>screening</i> OR <i>exam*</i>] OR [cheating OR plagiarism OR integrity OR fraud* OR dishonesty OR tricherie OR plagiat OR intégrité académique])
AND	
Intelligence artificielle	(ai OR artificial intelligence OR chatgpt OR chatbot OR conversational agent OR ia OR intelligence artificielle OR agent conversationnel)
AND	
Enseignement supérieur	(higher education OR college OR universit* OR post secondary OR postsecondary OR undergraduate OR graduate OR enseignement supérieur OR cégep)

Comme le font ressortir ces mots-clés, la recherche documentaire a été réalisée en français et anglais. Nous avons effectué des tentatives en espagnol et en portugais, mais ces dernières n'ont

² Comparables à des dictionnaires de synonymes, les thésaurus des bases de données ERIC et Education Source consistent en des répertoires lexicaux liés au domaine de l'éducation. Ils permettent de rechercher les mots-clés privilégiés au sein de leur moteur de recherche respectif et d'en repérer les termes voisins, qu'ils soient hyperonymes ou hyponymes.

³ Les bases terminologiques de la Vitrine Linguistique et de Termium Plus® sont des ressources traductionnelles consacrées au français et à l'anglais québécois et canadiens. Elles permettent de dégager des termes associés à des spécialités précises. Elles procurent des résultats plus fins que ceux générés par des dictionnaires généralistes qui, eux, pourraient ne pas fournir de traduction exacte ou ne pas saisir les nuances sémantiques propres à un domaine. Ainsi, en nous assurant une plus grande justesse dans nos choix terminologiques, ces deux types de ressources ont contribué à optimiser le jeu de mots-clés employé dans les requêtes.

⁴ Ceci a conduit à mettre de côté certains de nos mots-clés comme « grande école ».

pas abouti. Les bases de données sondées, à savoir ERIC, Education Source, Érudit et Cairn, ont été sélectionnées en raison de la richesse de leurs références concernant l'enseignement supérieur. Nous avons restreint la recherche documentaire aux articles publiés avant 2025 dans des revues scientifiques comprenant une évaluation en double aveugle.

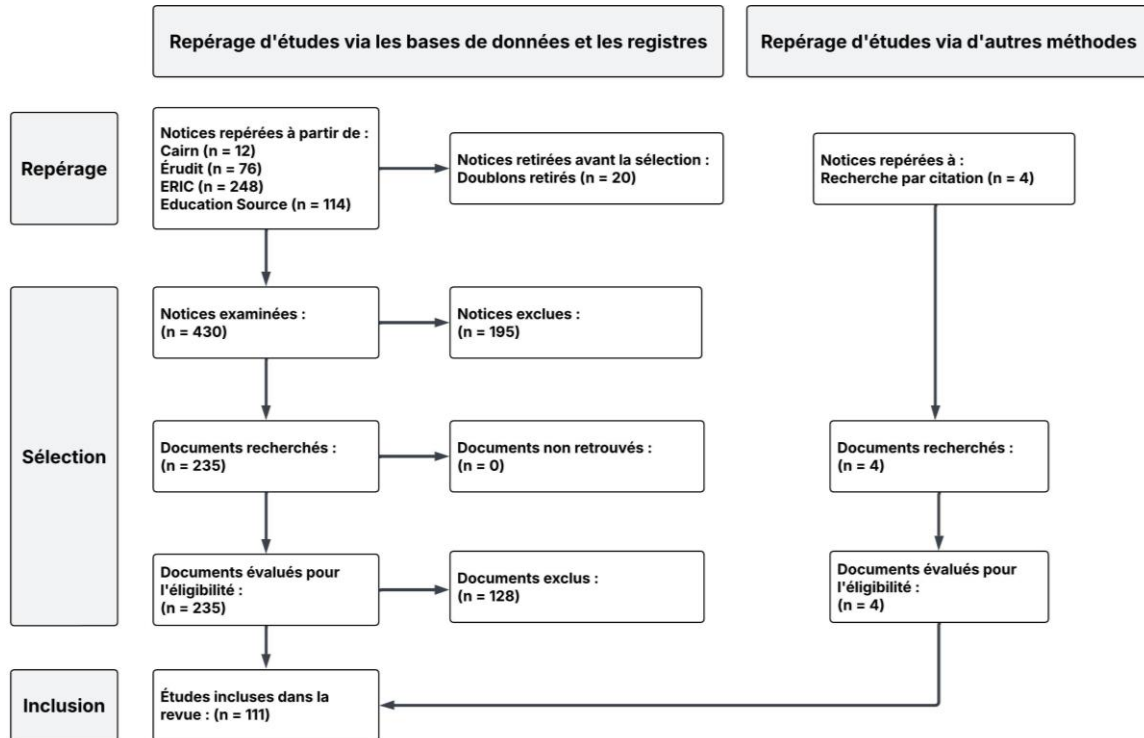
Ce sont ainsi 450 publications qui ont tout d'abord été retenues. Parmi ces dernières se trouvaient 20 doublons qui ont été retirés. L'ensemble des références et résumés des 430 articles ont été récupérés dans Zotero, puis reportés dans un fichier Excel dupliqué afin de faire une évaluation interjuges permettant de vérifier que les textes répondaient bien aux critères souhaités à savoir : porter sur l'IA et l'évaluation des apprentissages dans l'enseignement supérieur. Lorsque les résultats de l'évaluation divergeaient, des discussions entre les deux assistants de recherche ayant le rôle de juges ont eu lieu en présence de la chercheuse principale afin de trouver un consensus (Büchter *et al.*, 2020). Ceci a permis d'écarter 195 textes hors sujet (notamment ceux qui n'étaient pas centrés sur l'évaluation des apprentissages). Les fichiers PDF des 235 articles restants ont été récupérés. Une lecture rapide de ceux-ci dans leur intégralité a amené à mettre encore de côté 127 textes ne répondant d'aucune manière aux quatre questions de recherche. Ces dernières visaient à évaluer dans quelle mesure :

1. il est possible de relever ou non un accroissement de la fraude liée à l'usage de l'IA dans les travaux notés ou examens ;
2. l'IA a conduit à des modifications des activités d'évaluation des apprentissages, de la manière d'évaluer ou de rétroagir au sein du corps professoral ;
3. le recours à l'IA dans le cadre des évaluations est encadré au niveau des établissements d'enseignement supérieur ;
4. les usages relevés ou non de l'IA permettent de garantir la validité des diplômes et certifications délivrés par les établissements d'enseignement supérieur.

Enfin, quatre articles triés sur le volet n'apparaissant pas dans les résultats initiaux, mais correspondant au sujet selon les chercheurs ont aussi été ajoutés, comme l'illustre le schéma PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) ci-dessous :

Figure 1.

Schéma PRISMA - Recension systématique narrative sur l'évaluation et l'intelligence artificielle en enseignement supérieur



Adapté de Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, A. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews, *BMJ* 2021. 372 (71). <https://doi.org/10.1136/bmj.n71>. Traduction française par Emmanuelle Paquette Raynard, Martine Gagnon et Marianne Ruel, Bibliothèque de l'Université Laval.

In fine, ce sont ainsi 111 articles anglophones publiés entre janvier 2020 et décembre 2024 qui ont été retenus. Ces articles ont été analysés en saisissant des données dans une grille comprenant 27 colonnes. Cette dernière visait notamment à colliger des renseignements sur le cadre théorique, la méthodologie, le public de l'étude, la taille de l'échantillon, les principaux résultats de recherche et de discussion de chaque article ainsi que leurs limites. Dans un souci de rigueur, l'extraction des données s'est réalisée en deux temps. Un assistant de recherche a d'abord effectué l'analyse textuelle des articles retenus et a consigné l'information recherchée dans un tableur Excel. Puis, une seconde assistante de recherche, assistée par une IAG (ChatGPT 5.2 et 5.3), a parcouru sommairement ces mêmes références pour vérifier l'exactitude des éléments extraits par son homologue en comparant ce qui avait été relevé avec ce qui était mentionné par

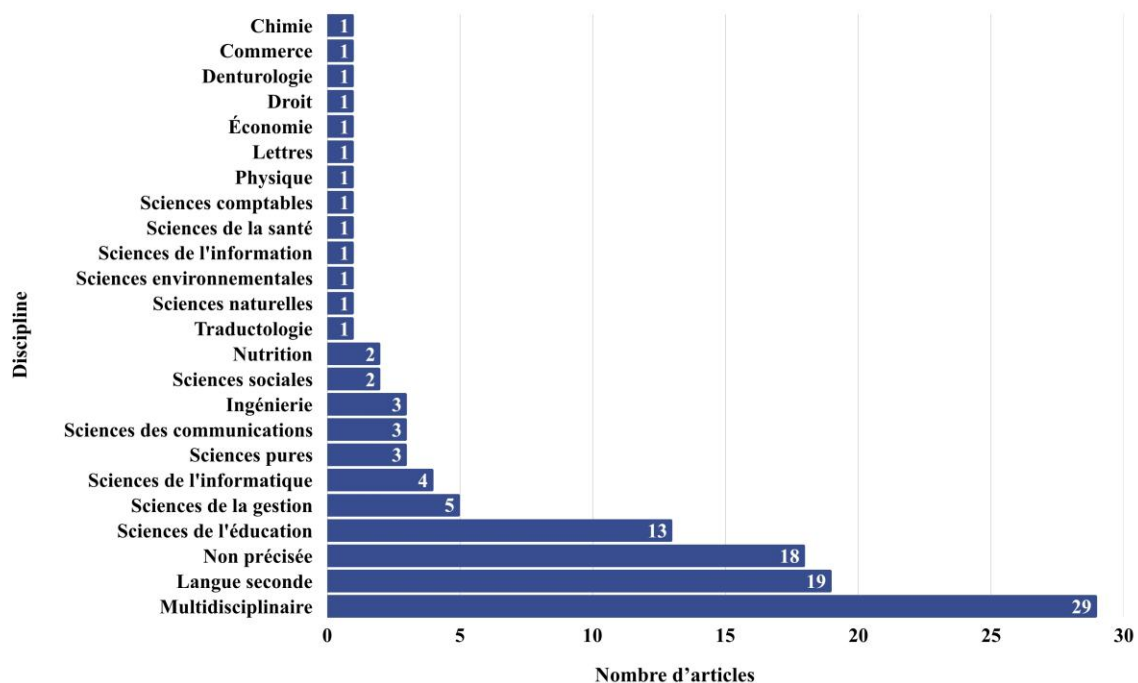
l'IAG afin de retourner lire l'article en cas de divergence et corriger ou enrichir, au besoin, certaines entrées dans le tableau.

En dépit de l'absence de délimitation temporelle (puisqu'il s'agissait de prendre en considération tout ce qui avait été publié antérieurement au moment de la recherche) et de la centration sur l'IA (sans préciser « générative »), force est de constater qu'à quatre exceptions près, les publications sont postérieures à la sortie de ChatGPT fin 2022 et semblent pour la plupart porter sur l'IAG⁵. Notre corpus est composé d'un article datant de 2020, de trois de 2021, de cinq de 2022, de 23 de 2023 et de 79 publiés en 2024. Une forte augmentation des publications est donc constatée depuis la diffusion grand public du premier outil d'OpenAI.

Les travaux du corpus ainsi constitué concernent diverses zones géographiques. La majorité d'entre eux ont été réalisés aux États-Unis (22 articles), au Royaume-Uni (15 articles), en Chine (13 articles), en Australie (11 articles) ou au Canada (9 articles) et 22 articles ont porté sur des enquêtes diffusées dans plusieurs pays, voire continents. De même, une grande partie des articles porte sur diverses disciplines, sur l'apprentissage de l'anglais comme langue étrangère ou sur les sciences de l'éducation comme le montre le graphique de la figure 2 ci-dessous :

Figure 2.

Répartition des disciplines visées par les articles du corpus



⁵ Il convient de noter que le terme « générative » n'est pas toujours clairement indiqué et que quelques articles sont davantage relatifs à des outils d'IA permettant la surveillance à distance.

Les méthodologies d'enquête sont de nature mixte dans 35 articles, qualitative dans 32 articles et quantitative dans 22 articles. Dix-sept articles emploient d'autres méthodologies relevant par exemple de l'analyse critique ou d'études pilotes et cinq articles détaillent peu l'approche méthodologique employée.

Résultats

Les articles rassemblés dans ce corpus mettent en évidence, d'une part, que les capacités de l'IAG constituent une menace à l'intégrité académique dès lors que des étudiants peuvent s'en servir pour réaliser leurs travaux, mais également un rempart potentiel en ce sens que certains outils peuvent être mis à contribution pour détecter les usages de l'IAG ou pour surveiller à distance les étudiants. D'autre part, il apparaît que l'IAG peut soutenir les enseignants dans la préparation et la correction d'activités d'évaluation ainsi que les étudiants dans leur apprentissage, si tant est qu'ils soient formés et que ces usages soient encadrés par des politiques institutionnelles adéquates.

L'IA comme menace ou protection de l'intégrité académique

Comme présenté en introduction, l'IAG est susceptible de favoriser différentes formes de fraude. Même si certains étudiants tendent encore à privilégier les ressources variées trouvées sur le web lors de la réalisation de devoirs et examens en ligne (Kortemeyer et Bauer, 2024), le recours à l'IAG est de plus en plus fréquent et les capacités de cette dernière sont en constant développement (Newton, 2024). Nous allons ainsi voir que, l'IAG permettant d'obtenir de bons résultats, certains étudiants y recourent dans le cadre d'activités d'évaluation des apprentissages, et ce, parfois en toute impunité, tant il peut être délicat d'identifier les cas de fraude ou d'imposer des systèmes de surveillance, ce qui conduit plusieurs auteurs à inviter les enseignants à modifier les activités d'évaluation sommative ou certificative.

Une IAG performante

Dix articles publiés en 2023 et 2024 se sont intéressés aux performances de différentes versions de ChatGPT dans le cadre d'évaluations sommatives ou certificatives de type QCM (Questionnaires à Choix Multiples), de rédactions de textes ou d'études de cas, et soulignent que, malgré quelques erreurs (Kortemeyer, 2023), ses capacités sont plutôt bonnes. Il est, par exemple, constaté que ChatGPT obtient un score d'environ 60 % à certains tests en physique (Kortemeyer, 2023) ou bien encore un score de 68,2 % (Nikolic *et al.*, 2023) à 90 % (Kortemeyer, 2023) à des examens de programmation. Dans le domaine des affaires, une étude conclut que les réponses données par ChatGPT atteignent une précision de 78,1 % aux QCM et 76,1 % aux exercices impliquant un peu de rédaction (Adilov *et al.*, 2024). Concernant plus particulièrement ChatGPT 3.5, une recherche conclut que l'IAG obtient de très bons résultats, mais que ceux-ci sont en deçà de ceux de la moyenne des étudiants en chirurgie dentaire, ces derniers ayant eux-mêmes

d'excellents résultats (Sallam *et al.*, 2023), tandis qu'une autre effectuée dans le cadre d'études en ressources humaines et management, conclut que les notes obtenues par l'IAG rivalisent ou dépassent celles des étudiants (Chaudhry *et al.*, 2023). En dépit de ces nuances, dans les deux cas, l'IAG est en mesure de réaliser de manière satisfaisante le travail attendu.

De même, dans un contexte d'études de premier et deuxième cycle d'ingénierie où les évaluations consistent en des QCM et des travaux écrits, il est remarqué que le recours à l'IAG permet de réussir les examens, mais que d'autres formes d'évaluation (orales, laboratoires, projets complexes) peuvent plus difficilement être confiées à l'IAG (Nikolic *et al.*, 2023). S'appuyant sur la typologie de Bloom, plusieurs auteurs (Nguyen Thanh *et al.*, 2023 ; Sallam *et al.*, 2023) expliquent d'ailleurs que plus on monte dans la pyramide, moins l'IAG est en mesure d'effectuer le travail de manière satisfaisante. Autrement dit, l'usage de l'IAG permet d'obtenir de meilleurs résultats dans le cadre de travaux faisant appel à la mémoire ou la compréhension qu'à l'application, l'analyse, l'évaluation ou la créativité (Sallam *et al.*, 2023). Ceci peut expliquer qu'une autre recherche mette en lumière que les travaux réalisés par ChatGPT obtiennent davantage la note de passage dans les évaluations de premier que de deuxième cycle (Richards *et al.*, 2024).

Une fraude plus ou moins assumée

Les activités d'évaluation des apprentissages apparaissent d'autant plus vulnérables aux usages potentiels de l'IAG qu'elles s'effectuent sans surveillance, comme cela est en particulier le cas dans le cadre des formations à distance asynchrones ou des travaux à réaliser en dehors des cours (Bhullar *et al.*, 2024). En 2024, quatre études quantitatives font effectivement ressortir un usage avoué de l'IAG lors d'évaluations sommatives. Ainsi, une recherche menée en Croatie auprès d'un échantillon de 350 étudiants et 12 professeurs indique que 50,6 % des étudiants utilisent ChatGPT pour réaliser leurs travaux à la maison et que les enseignants ne le détectent que dans 53,57 % des cas (Krecar *et al.*, 2024). Une autre recherche réalisée au Vietnam, met en évidence que parmi 1 386 étudiants de premier cycle, 9,6 % admettent avoir triché à l'aide de ChatGPT lorsque la question leur est posée directement, mais que l'usage d'un questionnaire indirect révèle une part plus importante (23,7 %) et fait ressortir que les étudiants plus avancés dans leurs parcours universitaires et ceux de sexe masculin tricheraient davantage que ceux en début de parcours et ceux de sexe féminin (Nguyen et Goto, 2024). Une enquête menée auprès de 337 étudiants de diverses disciplines en Australie met en avant que plus du tiers des étudiants ont utilisé un agent conversationnel pour réaliser des activités évaluées et ne perçoivent pas cela comme une violation de l'intégrité académique (Gruenhagen *et al.*, 2024).

Enfin, une étude menée auprès de 1 217 répondants (58,18 % d'étudiants, majoritairement de premier cycle, et 41,82 % d'enseignants), dans 76 pays, révèle que l'usage de l'IAG pour un travail scolaire est considéré comme de la fraude par 46,6 % des étudiants (tandis que 36,6 % ne considèrent pas cela comme une fraude et que 17,4 % ne savent pas) (Yusuf *et al.*, 2024). L'article

précise que 19,9 % des étudiants déclarent avoir utilisé l'IAG pour leurs travaux et avoir l'intention de poursuivre, tandis que 19,2 % évoquent un usage passé, qui ne se reproduira pas. Par ailleurs, une enquête menée auprès de 121 étudiants israéliens révèle que 50 % d'entre eux jugent acceptable d'être assisté par l'IA dans leurs activités d'évaluation pour faire face à la surcharge de travail et à la pression de réussir leurs études (Yavich et Davidovitch, 2024).

Un usage difficile à détecter

Huit articles mettent en exergue que l'IAG menace d'autant plus l'intégrité dite « académique » ou « intellectuelle » qu'il devient de plus en plus difficile de s'assurer que les travaux remis par les étudiants sont le fruit de leur travail (Wise *et al.*, 2024). Par exemple, une recherche dans le domaine de la gestion montre que ChatGPT réussit parfaitement différents types de travaux, notamment des études de cas, sans qu'aucun plagiat ne soit repéré par l'outil de détection du plagiat TurnItIn (Chaudhry *et al.*, 2023). Une autre étude montre quant à elle que TurnItIn et GPT-2 Output Detector tendent à bien identifier les textes écrits par une IAG, mais que le taux de faux positifs est élevé (jusqu'à 50 %) de telle sorte que cela ne permet pas de repérer la fraude de manière efficace (Richards *et al.*, 2024).

Bien que l'être humain semble plus à même de déceler l'œuvre de l'IAG, de nombreux usages restent difficiles à identifier. Dans le domaine de l'anglais comme langue étrangère, Saarna (2024) a soumis quatre travaux écrits par des étudiants japonais de premier cycle et quatre autres écrits par ChatGPT (ayant reçu des modèles d'erreurs communément faites par les étudiants) à 47 enseignants : 54,25 % de ces derniers ont su discerner correctement les travaux écrits par un humain de ceux produits par l'IAG. Dans celui de la psychologie cognitive et des sciences de l'apprentissage Waltzer *et al.* (2024) relèvent que les 140 enseignants participant à la recherche ont identifié les textes générés par ChatGPT dans 70 % des cas. La capacité à déceler l'usage de l'IAG est un peu plus faible chez les 145 étudiants participant à l'enquête, ces derniers effectuant un discernement correct dans 60 % des cas, et ChatGPT lui-même, repérant 63 % des cas. Distinguer l'écriture humaine de l'écriture générée par les grands modèles de langage s'avère d'autant plus ardu qu'il est possible de demander à l'IAG de s'inspirer de la façon de rédiger des étudiants. D'ailleurs, se fondant sur plusieurs études, Perkins (2023) souligne que plus l'IAG s'améliore, plus la capacité humaine à repérer son usage diminue. Les résultats de Krecar *et al.* (2024) dans une diversité de disciplines indiquent ainsi que les 12 enseignants participant à l'enquête ont distingué correctement ce qui était rédigé par un étudiant ou généré par une IAG dans 53,75 % des travaux évalués. En bref, les performances concernant la distinction entre le travail réalisé par l'humain et celui réalisé par l'IAG semblent toujours supérieures à 50 %, mais loin des 100 %, et le manque de fiabilité des outils de détection peut amener à accuser à tort des étudiants qui doivent alors prouver qu'ils ont effectué leur travail eux-mêmes (Gorichanaz, 2023).

Des systèmes de surveillance discutables

Percevant une grande différence de résultats entre les examens réalisés avec ou sans surveillance, quelques auteurs insistent sur l'importance de réaliser les examens les plus cruciaux sous surveillance pour préserver l'intégrité académique (Cotton *et al.*, 2024; Kortemeyer et Bauer, 2024). Trois articles s'intéressent ainsi aux systèmes de surveillance à distance (*intelligent online proctoring system*) mobilisant une diversité d'outils (de reconnaissance faciale des expressions, de détection des mouvements des yeux et de la bouche), parfois sophistiqués (comme un numériseur à balayage oculaire ou un clavier permettant de prendre le pouls) pour collecter des informations favorisant le repérage des cas de tricherie potentiels.

Alors que le coût des dispositifs existants est parfois élevé, Jia et He (2022) présentent quant à eux les résultats d'un projet pilote de conception et d'implémentation d'un système de surveillance des examens en ligne utilisant la reconnaissance faciale, la détection vocale et l'analyse multimodale, pendant la pandémie. Testé en Chine, auprès de 16 étudiants, le système n'a pas détecté de fraude et les étudiants semblent avoir vécu cette expérience de manière similaire à celle d'un examen surveillé en classe. Une étude longitudinale menée entre 2017 et 2019 auprès de 253 étudiants de premier cycle en Australie (Sefcik *et al.*, 2022) met en relief que les étudiants ont tendance à s'habituer au système de surveillance en ligne et que 79 % d'entre eux considèrent que cela permet de dissuader la fraude. Cependant, les problèmes techniques se sont avérés sources d'anxiété et les étudiants ont indiqué souhaiter davantage de précisions sur les comportements attendus et la confidentialité des données.

Se fondant sur les articles centrés sur la surveillance en ligne, Oravec (2022) souligne d'ailleurs que la fiabilité de ces dispositifs n'est pas garantie et que les logiques de suspicion intégrées à ces derniers risquent de générer un sentiment de surveillance permanent et d'anxiété chez les étudiants potentiellement amenés à devoir prouver qu'ils n'ont pas triché. De plus, l'auteur insiste sur le fait que l'usage de ces dispositifs soulève des questions éthiques en ce sens que les données biométriques et comportementales peuvent être collectées et stockées sur de longues périodes et utilisées pour constituer des profils ou des listes de « tricheurs potentiels » susceptibles d'avoir des répercussions sur la suite du parcours des étudiants.

Des activités d'évaluation à repenser

Une enquête menée auprès de 318 enseignants de différents niveaux, disciplines et continents permet de se rendre compte que seulement une minorité d'enseignants considère que les particularités ou complexités de leurs activités d'évaluation font en sorte qu'elles ne sont pas affectées par l'IAG, tandis que la majorité cherche des solutions pour promouvoir l'intégrité académique (Bower *et al.*, 2024). Ces auteurs incitent à modifier les évaluations pour y inclure plus de projets de groupes et d'observations en classe. Une recherche menée au Vietnam auprès de 150 enseignants de premier cycle relève une augmentation des présentations orales et

discussions en classe pour limiter le recours à l'IAG (Cong-Lem, 2024). Plus généralement, à l'issue de leur recension de 32 articles, Xia *et al.* (2024) indiquent que 29 % des publications suggèrent de modifier les activités d'évaluation.

Comme mentionné précédemment, les travaux de Nguyen Thanh *et al.* (2023) et Sallam *et al.* (2023) mettent en relief que l'IAG peut plus aisément réaliser des travaux faisant appel à des compétences à la base de la pyramide de Bloom comme la mémorisation, qu'à celles situées davantage à son sommet, en particulier la créativité, ce qui peut donner une orientation pour repenser les activités d'évaluation. D'ailleurs, plusieurs auteurs recommandent de privilégier les évaluations encourageant la créativité, la résolution de problème et le développement de la pensée critique, de promouvoir la participation active des étudiants et de trouver un équilibre entre l'enseignement centré sur l'être humain et l'assistance de l'IAG tout en jouant un rôle actif dans la vérification de la validité et de la fiabilité des matériaux générés par les modèles de langage (Xia *et al.*, 2024 ; Zirar, 2023). Kukol (2024) propose quant à lui de combiner évaluation ipsative et sociomatérielle⁶ pour renforcer l'engagement des étudiants, soutenir la rétention des connaissances et offrir une évaluation plus authentique des apprentissages scientifiques.

Une intégration progressive de l'IAG dans les pratiques et politiques

De plus en plus performante, l'IAG est certes, source de craintes, mais aussi de potentialités. Nous allons ainsi voir que des enseignants cherchent à l'intégrer dans leurs pratiques, potentiellement pour les alléger, mais aussi et surtout pour soutenir les étudiants dans leurs apprentissages par le biais d'activités d'évaluation formative créées ou accompagnées par des outils d'IAG, mais que les étudiants restent quelque peu attachés aux rétroactions réalisées par un être humain. Une formation aux usages éthiques des outils pertinents apparaît dès lors tout aussi nécessaire que l'avènement de politiques institutionnelles pour encadrer les pratiques.

Des outils d'IAG utiles pour créer et corriger des activités d'évaluation

Treize articles soutiennent l'idée que l'IAG peut alléger le travail enseignant grâce à l'automatisation partielle ou totale de différentes tâches allant de la planification des cours à l'évaluation des apprentissages (Favero, 2024). En matière d'évaluation, deux usages principaux de l'IAG peuvent être repérés, l'un concernant la création des activités d'évaluation et l'autre leur correction. Ainsi, Onal et Kulavuz-Onal (2024) remarquent que ChatGPT peut produire toute une diversité d'activités d'évaluation des apprentissages comme des QCM ou des études de cas, dans des disciplines aussi variées que le génie industriel et la linguistique appliquée. De plus, il

⁶ Selon Kukol (2024), la combinaison de l'évaluation ipsative et sociomatérielle représente une approche transformatrice pour l'enseignement des sciences au niveau universitaire. L'évaluation ipsative encourage les étudiants à mesurer leur progrès en fonction de leur performance antérieure et l'évaluation sociomatérielle prend en compte l'ensemble des dimensions cognitive, sociale et matérielle de l'apprentissage.

apparaît que les activités d'évaluation générées par l'IAG sont comparables à celles créées par des êtres humains, notamment en ce qui concerne leur niveau de difficulté (Kanik, 2024 ; O, 2024). Toutefois, Klyshbekova et Abbott (2024) signalent des hallucinations ainsi qu'un manque d'originalité et de profondeur dans les questions générées automatiquement. Onal et Kulavuz-Onal (2024) mettent d'ailleurs l'accent sur le fait que l'IAG ne remplace pas l'enseignant, mais vient soutenir et enrichir son travail et qu'il est important de valider les activités proposées, car il est entre autres constaté que certaines questions, bien que correctes en apparence, manquent de clarté ou de précision. Ceci entre quelque peu en écho avec les résultats d'une recherche menée auprès de 358 enseignants du Moyen-Orient et d'Afrique du Sud, qui montre que l'adoption de l'IAG dépend étroitement de la confiance des enseignants dans leurs propres capacités à encadrer les usages (Khlaif *et al.*, 2024).

Il en va de même en ce qui a trait à la correction qui n'est pas toujours fiable (Klyshbekova et Abbott, 2024) notamment lorsqu'il s'agit de textes dont l'évaluation implique une part de subjectivité. En effet, Geçkin *et al.* (2023), font ressortir que l'évaluation automatisée des productions écrites de 43 étudiants réalisée par ChatGPT 3.5 donne des résultats cohérents avec celle réalisée par les enseignants, mais non similaires, de telle sorte que son usage devrait être envisagé en complément plutôt qu'en substitution du jugement humain. Plus généralement, une étude dans le domaine de la traduction pose le constat que les résultats des corrections automatiques sont d'autant plus proches de ceux des évaluations humaines que la consigne est claire et fermée, les variations possibles limitées et que le contenu attendu est standardisé (Han et Lu, 2023). Une recherche réalisée aux États-Unis auprès de 24 étudiants de premier cycle fait aussi ressortir que ces derniers n'ont guère confiance en l'IAG et préfèrent être évalués par un enseignant seul ou en complément d'une IAG (Tossell *et al.*, 2024).

Des recours à l'IA dans l'évaluation formative

Avant même la sortie de ChatGPT, des travaux étaient réalisés pour chercher à comprendre dans quelle mesure divers outils pourraient favoriser la progression des étudiants et pendant la pandémie de COVID-19, l'usage d'agents conversationnels a parfois été considéré comme un moyen de compenser la baisse d'interactions humaines. Vingt-trois articles présentent ainsi les résultats de diverses études expérimentales consistant à proposer des usages de l'IA pour soutenir l'apprentissage⁷.

La recherche menée par Yang *et al.* (2021) dans le cadre d'un cours de comptabilité avance que 36 étudiants régulièrement amenés à effectuer des exercices de questions à trous générés automatiquement ont développé une meilleure compréhension conceptuelle et une plus grande rétention des informations lues que les 38 étudiants n'ayant pas fait usage de ces exercices. Dans

⁷ Un article partageant une expérience d'usage de l'IA prédictive visant à favoriser la réussite peut également être mentionné : dans un cours de biologie, une évaluation diagnostique est effectuée en début de cours et des évaluations ont lieu à la troisième et la sixième semaine, ce qui permet de mettre en place des interventions pédagogiques ciblées pour éviter les échecs (Bertolini *et al.*, 2021).

un cours de gestion d'une durée d'un semestre, des analyses de textes, des résolutions de dilemmes éthiques avec une rétroaction automatisée ont été proposées à 58 étudiants de premier cycle, ce qui semble avoir permis aux étudiants de mieux réussir leurs examens (Hyde *et al.*, 2024). Dans celui d'une formation chirurgicale (Trujillo *et al.*, 2024) impliquant trois chirurgiens, quatre résidents et deux experts, il est relevé que ChatGPT a été capable d'identifier et de nommer les erreurs chirurgicales avec un taux de détection comparable à celui des chirurgiens expérimentés et que ses rétroactions sont parfois perçues comme étant meilleures que celles des chirurgiens si bien que cet outil semble pouvoir soutenir l'apprentissage de manière pertinente.

Même si la tendance à inclure l'IAG dans l'évaluation formative est présente dans diverses disciplines, force est de constater que c'est dans le domaine de l'apprentissage de l'anglais comme langue étrangère que se trouve la majorité des articles de notre corpus. Kohnke (2023) présente ainsi le développement d'un agent conversationnel, créé à partir d'un outil proposé par Google, Dialogflow, et implanté dans un réseau social. L'auteur précise que 20 étudiants s'en sont servi comme partenaire d'apprentissage et ont dans l'ensemble émis des commentaires positifs relatifs, entre autres, à l'immédiateté et à la clarté des réponses. De même, les quatre enseignants convoqués dans cette recherche ont estimé que les étudiants avaient réalisé les tâches demandées de manière plus efficace tout en apprenant à mieux structurer leurs écrits. Une autre étude (Hsu *et al.*, 2023) a porté sur l'usage d'un assistant vocal intelligent, Alexa d'Amazon, mais cette fois-ci en classe. Il en ressort que l'usage de l'assistant à travers des jeux, des questions-réponses, des recherches lexicales et des échanges informels, crée un environnement d'apprentissage moins anxigène, plus favorable à la prise de parole, et que les 24 étudiants ayant bénéficié de sept séances de 50 minutes de travail avec cet assistant obtiennent de meilleurs résultats à la partie « écouter et parler » du TOEIC (*Test of English for International Communication*) que les 26 étudiants dépourvus de ce système (Hsu *et al.*, 2023).

La recherche de Xiao *et al.* (2024), portant sur 389 étudiants chinois, fait ressortir que les programmes d'IAG fournissent une rétroaction en temps réel, permettant aux étudiants de s'évaluer et d'améliorer continuellement leur apprentissage des langues en favorisant l'ajustement continu des stratégies d'apprentissage et l'autorégulation émotionnelle ainsi que le plaisir d'apprendre. De même, He (2024), qui s'intéresse aux outils d'évaluation automatisée de l'écriture, relève que grâce à ces derniers les 86 étudiants chinois faisant partie de son échantillon ont amélioré leurs textes et que leur motivation et leur plaisir d'écrire se sont accrus. La recherche de Gozali *et al.* (2024) menée auprès de 18 étudiants indonésiens conclut aussi que ces outils peuvent soutenir les étudiants en ce sens que les rétroactions de l'IAG permettent de développer les compétences rédactionnelles, mais rappelle que leur efficacité dépend de la capacité humaine à les utiliser de manière critique.

Des partenaires d'apprentissage qui ne font pas l'unanimité

Quelques recherches mentionnent que les performances des outils utilisés ne sont pas sans limites. Par exemple, Çakmak (2022), dont l'étude porte sur 90 étudiants turcs apprenant l'anglais, met en lumière que, bien que ces derniers soient conscients que leurs progrès linguistiques sont plus importants lorsqu'ils communiquent avec Replika qu'avec un camarade, leurs perceptions des interactions avec l'agent conversationnel se révèlent plutôt négatives. En effet, ils rapportent avoir éprouvé des difficultés à se faire comprendre, ce qui a engendré une augmentation de l'anxiété et de la frustration.

Dans une autre étude visant à tester quatre agents artificiels différents, Belda-Medina et Kokošková (2023) constatent que même si l'interaction avec l'IA tend à favoriser l'autonomie des étudiants, la satisfaction de ces derniers est somme toute modérée notamment en raison de certaines réponses prédéfinies et d'une reconnaissance vocale parfois déficiente. De même, ayant relevé que les gains d'apprentissage réalisés par des étudiants apprenant l'anglais lors d'interactions avec ChatGPT-4 ou avec un enseignant sont similaires, Escalante *et al.* (2023) remarquent que si la moitié des étudiants préfère la rétroaction apportée par l'agent conversationnel dont la clarté et la précision aident à améliorer la rédaction de textes, l'autre moitié préfère celle d'un humain, en présence, en raison de sa dimension affective. S'il ne remet pas en cause leur pertinence, l'aspect artificiel et imparfait des conversations avec les robots amène à chercher des formes de complémentarité entre soutien humain et agent conversationnel. Quelques auteurs comme Escalante *et al.* (2023) ou Liu *et al.* (2023) prônent ainsi une approche mixte où étudiants et enseignants peuvent se servir de l'IAG pour gagner en temps et précision de rétroaction tout en conservant des moments de discussion avec l'enseignant ou les pairs, afin de discuter des rétroactions générées par l'IAG et de soutenir l'engagement des étudiants.

Une formation nécessaire à la littératie de l'IA

Plusieurs auteurs insistent sur la nécessité de former les enseignants et les étudiants aux usages critiques et responsables de l'IA. Néanmoins, un seul décrit une formation de 35 enseignants, au Canada, ayant eu l'occasion de développer leur littératie de l'IA dans le cadre d'un cours en ligne de conception multimédia dans lequel ils ont eu l'occasion de corédiger un manuel collaboratif sur l'IAG. Cette approche expérientielle leur a permis d'apprendre à utiliser l'IAG de manière créative et responsable pour concevoir des environnements d'apprentissage (MacDowell *et al.*, 2024).

De même, seuls trois articles reviennent sur des exemples concrets, mais très limités, de formation des étudiants. Entre 2018 et 2020, MacDonald et Bézaire (2024) ont cherché à développer les compétences en recherche et en littératie informationnelle de trois cohortes d'étudiants de premier cycle en nutrition par la pratique de revues systématiques de la littérature. Ils expliquent que le fait de pratiquer chaque étape d'un processus de revue systématique a permis

aux étudiants de développer les compétences visées tout en réduisant les risques de plagiat puisque le processus de recherche est observable, ce qui pourrait désormais permettre d'éviter des usages frauduleux de l'IAG. En 2022-2023, Kirwan (2024) a souhaité développer une littératie critique de l'IAG chez 24 étudiants en leur demandant d'utiliser ChatGPT de sorte qu'ils se rendent compte des atouts et faiblesses de cet outil. L'article de Kong *et al.* (2023) porte sur la formation de 36 étudiants à la littératie de l'IA et rapporte également qu'après 30 heures de formation, les étudiants font preuve d'une meilleure compréhension des notions associées à l'IAG, d'un sentiment de maîtrise accru et d'une plus grande conscience des enjeux éthiques.

Ces formations peuvent sembler anecdotiques et moindres, cependant, elles tranchent avec les nombreux autres articles qui ne font qu'évoquer la nécessité de former les acteurs aux usages efficaces et éthiques de l'IA. La formation est effectivement avancée comme étant d'autant plus importante qu'il est estimé que l'usage de l'IAG entraîne un risque de dépendance, de réduction de la capacité à produire des travaux originaux, à développer sa créativité et à travailler en équipe (Bhullar *et al.*, 2024 ; Xia *et al.*, 2024 ; Zirar, 2023). Au moment où s'arrête la revue de la littérature (c'est-à-dire en décembre 2024, ce qui renvoie à des travaux réalisés antérieurement), la formation et l'accompagnement des étudiants et des enseignants aux usages pédagogiques et éthiques de l'IA semblent donc être encore à développer (Giray *et al.*, 2024 ; Luo, 2024 ; Xia *et al.*, 2024).

Des politiques et directives institutionnelles à adapter

Douze articles portent sur des politiques ou directives relatives à l'IA dans l'enseignement supérieur. Ayant analysé 142 politiques d'établissements d'enseignement supérieur ainsi que d'autres documents de divers pays dans le monde concernant l'intégrité académique, en novembre 2022 et mai 2023, Perkins et Roe (2024) remarquent que l'IA n'est alors que peu ou pas prise en compte dans les textes officiels et pointent les risques qui accompagnent ce flou réglementaire. Moorhouse *et al.* (2023) indiquent quant à eux que 23 des 50 universités auxquelles ils se sont intéressés en 2023 avaient des lignes directrices publiques relatives à l'IAG. Ayant analysé 102 politiques et guides d'usage de l'IA dans l'enseignement supérieur américain dans la deuxième moitié de 2023, Moore et Lookadoo (2024) décèlent que la majorité des textes adoptent un ton négatif ou restrictif, centré sur l'interdiction et la sanction, peu d'entre eux proposant des directives claires et des exemples concrets d'usages pédagogiques autorisés. Une étude des discours institutionnels canadiens sur l'IA dans 16 universités fait également ressortir une focalisation sur les enjeux de fraude et un manque de clarté concernant les usages pédagogiques, formatifs et responsables de l'IA (Marcel et Kang, 2024). Une autre recherche traitant de l'usage de l'IAG dans l'évaluation, à partir de l'analyse de 19 politiques institutionnelles dans divers pays, signale que l'originalité des travaux des étudiants est décrite comme menacée, sans être pour autant définie (Luo, 2024). Au Chili, l'analyse de 79 documents institutionnels met en relief que non seulement les politiques sont parfois difficiles à trouver, mais aussi que les termes juridiques employés sont complexes et donc potentiellement

incompréhensibles pour les étudiants (Moya et Eaton, 2024). L'analyse de 131 politiques institutionnelles de 11 pays présentée par Bannister *et al.* (2024a) met au jour que seulement quatre d'entre elles prennent en considération les besoins spécifiques des étudiants internationaux, notamment en matière d'outils de traduction souvent nécessaires pour ces derniers, comme le dénoncent également Vicary et Treffers-Daller (2024).

Certains auteurs mentionnent que les étudiants attendent des politiques nommant clairement ce qui est ou non autorisé et suggèrent de travailler avec eux à l'élaboration des directives institutionnelles (Johnston *et al.*, 2024). Xia *et al.* (2024) insistent quant à eux sur la nécessité d'une co-construction des attentes avec les étudiants et Bhullar *et al.* (2024) proposent de faire signer aux étudiants un gage d'authenticité du travail remis. Ceci semble cohérent avec l'observation selon laquelle les méthodes d'enseignement et d'évaluation influencent davantage les étudiants que ne le font les politiques institutionnelles (Ateeq *et al.*, 2024). Aussi, le rôle des enseignants reste-t-il central. De manière plus générale, Lentz (2024) énonce le besoin de développer des usages responsables et éthiques de l'IA, besoin auquel quelques travaux tendent à répondre. En effet, Firth *et al.* (2024) présentent un cadre pour utiliser l'IA de manière réfléchie et favoriser une meilleure compréhension des enjeux éthiques et pédagogiques de même qu'un renforcement du jugement critique des étudiants. Il en va de même de Shailendra *et al.* (2024) qui proposent un cadre d'adoption de l'IA dans les curriculums et universités ainsi qu'une définition des rôles des différents acteurs et de Bannister *et al.* (2024b) qui ont réuni 14 experts pour créer un cadre d'intégrité académique visant à partager de bonnes pratiques d'usage de l'IA dans l'enseignement supérieur.

Discussion

Cette revue systématique narrative de la littérature (Turnbull *et al.*, 2023) portant sur 111 articles publiés avant 2025 fait ressortir que les recherches sur les usages potentiels de l'IA dans l'évaluation en enseignement supérieur se sont développées depuis la pandémie de COVID-19 (Kohnke, 2023) et accélérées à partir de la sortie de ChatGPT puis de multiples outils d'IAG. À l'instar de ce qui est arrivé avec d'autres diffusions d'innovations technologiques telles que celle d'Internet, cette recension met en exergue que l'IAG peut être utilisée pour soutenir les activités d'apprentissage (Bertolini *et al.*, 2021 ; Hsu *et al.*, 2023 ; Kim, 2022 ; Yang *et al.*, 2021) et d'enseignement (Geçkin *et al.*, 2023 ; Khlaif *et al.*, 2024 ; Onal et Kulavuz-Onal, 2024) tout en soulevant un certain nombre de préoccupations.

En effet, l'évolution de l'IAG et ses appropriations croissantes font émerger des enjeux en matière d'évaluation sommative des apprentissages, tout particulièrement lorsque les tâches sont standardisées, les consignes précises et que l'évaluation est réalisée sans supervision (Adilov *et al.*, 2024 ; Bhullar *et al.*, 2024 ; Chaudhry *et al.*, 2023). Le tableau est clair : les étudiants tendent de plus en plus à se servir de l'IAG pour réaliser leurs travaux ou examens (Krecar *et al.*, 2024 ; Nguyen Thanh *et al.*, 2023 ; Yusuf *et al.*, 2024) et cela est de plus en plus difficile à détecter

puisque les outils et usages s'affinent constamment (Waltzer *et al.*, 2024; Wise *et al.*, 2024). Blondelle et Giorgini (2025) considèrent ainsi que l'exercice de détection de la fraude et des travaux générés par l'IAG est un défi perpétuel, voué à l'échec. L'intégrité académique semble donc vivement menacée (Bennett et Abusalem, 2024 ; Blackie, 2024 ; Cotton *et al.*, 2024; Zirar, 2023).

Afin de contrer la fraude académique assistée par l'IAG, venant s'ajouter aux multiples formes de plagiat ou tricherie déjà existants et recensés avant les récents développements de l'IA (Hollis, 2018; Ison, 2012 ; Kayisoglu et Temel, 2017 ; McCabe *et al.*, 2012), plusieurs stratégies semblent néanmoins pouvoir être mises en œuvre. Au-delà ou en plus des outils de détection du plagiat et du « plagIA » pour contrôler les travaux remis (Chaudhry *et al.*, 2023 ; Khlaif *et al.*, 2024), apparaissent effectivement deux tendances. D'un côté, un renforcement de la volonté de contrôle caractérisé par le choix de réaliser les examens sous surveillance en présence (Kortemeyer et Bauer, 2024) ou de mettre en place des dispositifs de télésurveillance dans le contexte d'examens à distance (Jia et He, 2022 ; Oravec, 2022) est observé. De fait, même si les systèmes informatisés de surveillance et contrôle ne sont pas toujours fiables et s'avèrent parfois sources d'anxiété pour les étudiants (Gorichanaz, 2023 ; Sefcik *et al.*, 2022), ils demeurent dignes d'intérêt en ce sens que leur présence est susceptible, à elle seule, de dissuader les étudiants de tricher. De l'autre côté, dans la lignée des démarches déjà anciennes d'intégration de l'audiovisuel, du multimédia, puis des technologies de l'information et de la communication dans l'enseignement, émerge une recherche plus pédagogique visant à intégrer l'IAG dans les activités d'évaluation aussi bien formative que sommative. Il s'agit alors d'amener les étudiants à l'utiliser dans des situations ou dispositifs pédagogiques précis, favorisant la réflexion et le développement de la pensée critique. Entre les deux, certains enseignants estiment que leurs activités d'évaluation sont peu à risque et les conservent ou les modifient modérément, sans intégrer l'IAG. Il s'agit alors de limiter les possibilités d'usage de l'IAG, en proposant, par exemple, davantage d'examens oraux (Kortemeyer, 2023 ; Richards *et al.*, 2024) ou d'enregistrements vidéo.

Une recension de la littérature prenant en compte 254 articles incite à repenser l'intégration de l'IA aux trois niveaux (micro, méso et macro) pour garantir le respect de normes éthiques (Castelló-Sirvent *et al.*, 2024). Selon les auteurs, afin de structurer l'implantation de l'IA de manière à maximiser ses bénéfices tout en minimisant les risques éthiques, il conviendrait, au niveau micro, de chercher à améliorer l'expérience d'apprentissage des étudiants tout en respectant les normes éthiques ; au niveau méso, d'optimiser les programmes et la surveillance des examens pour garantir l'intégrité académique ; au niveau macro, de définir des politiques éducatives incluant des formes de régulations éthiques et permettant de s'assurer que l'IA contribue positivement aux sociétés (Castelló-Sirvent *et al.*, 2024). Il importe d'ailleurs de noter que la perception de l'IA dans le contexte de l'évaluation n'est pas partout uniforme. L'enquête présentée par Yusuf *et al.* (2024) menée auprès de plus de mille répondants provenant de 76 pays révèle que les attitudes face à l'utilisation de l'IAG pour réaliser des travaux scolaires varient considérablement selon les régions, certains considérant cet usage comme une forme de

malhonnêteté académique, tandis que d'autres jugent un tel usage acceptable, voire bénéfique pour gérer la surcharge de travail et les pressions liées à l'objectif de réussite scolaire. De même, une étude comparative menée auprès d'étudiants sud-africains et canadiens tend à démontrer que les variations culturelles influencent non seulement les comportements face aux évaluations par simulation basées sur l'IAG, mais aussi les résultats obtenus, les styles de communication et les réactions des étudiants (Kotlyar *et al.*, 2024). Les différences observées en matière de schèmes communicationnels, de rapport à l'autorité, de construction de la confiance et de familiarité avec la technologie modifient tant les comportements que les perceptions concernant la validité et l'acceptabilité des dispositifs d'évaluation automatisés. Ces résultats invitent à considérer que les stratégies de mise en œuvre de l'IAG dans l'évaluation gagneraient à prendre en compte les particularités culturelles des contextes d'enseignement, même si les différences culturelles relatives à l'adoption de l'IA sont parfois moindres (Goto et Ramnarain, 2026).

De façon générale, une grande part des textes rassemblés dans le corpus incitent les enseignants et les institutions à dépasser la logique de détection de la fraude pour repenser l'évaluation et ses finalités en concevant davantage d'évaluations authentiques intégrant potentiellement, mais non nécessairement, l'IA (Bhullar *et al.*, 2024 ; Kortemeyer et Bauer, 2024). Si l'intérêt pour l'évaluation authentique impliquant des tâches complexes, signifiantes et contextualisées, proches de la vie quotidienne ou professionnelle (Gulikers *et al.*, 2004), n'est pas nouveau, les résultats préliminaires d'une autre recension systématique que nous venons de mener à ce sujet laissent entrevoir que celle-ci fait l'objet de nombreuses définitions, qui reflètent une diversité de compréhensions selon les contextes éducatifs et auteurs (Herrington et Herrington, 2006 ; Wiggins, 1990). Plus qu'un rempart contre l'IAG, cette approche peut ainsi inclure un usage raisonné de celle-ci en ce sens qu'elle vise à observer les processus cognitifs, la résolution de problèmes et la capacité d'adaptation favorisant le développement de compétences transférables et adaptées aux exigences des sociétés actuelles (Darling-Hammond *et al.*, 2014). D'ailleurs, au-delà de notre corpus, des publications plus récentes rappellent la pertinence des évaluations authentiques et de l'approche par compétence, permettant de dépasser l'évaluation du seul résultat (Kirsanov *et al.*, 2026 ; Kickbusch *et al.*, 2025). Ces derniers mois ont ainsi été publiés des travaux portant sur l'intégration de l'apprentissage automatique dans un jeu sérieux afin d'évaluer le savoir-être professionnel (*soft skills*) et le choix de programme d'études optimal pour des étudiants (Marengo *et al.*, 2025), sur l'usage de l'apprentissage profond pour évaluer les compétences collaboratives en présentiel (Valdes-Ramirez *et al.*, 2026) ou bien encore sur l'implémentation de l'IAG dans un dispositif de réalité virtuelle pour aider les apprenants à s'exercer à parler publiquement (Huang *et al.*, 2026). Alors que les dispositifs d'évaluation intégrant l'IAG se multiplient, il est également intéressant de signaler qu'un article propose une approche docimologique permettant d'apprécier l'usage de l'IAG dans les activités d'évaluation des apprentissages (Oliveira *et al.*, 2025).

Comme nous l'avons vu, quelques auteurs portent à notre attention l'absence de directives institutionnelles claires et le manque de formation des enseignants et étudiants qui créent un

certain flou et amènent à retrouver ce qui a toujours été relevé concernant l'intégration des TIC en éducation, à savoir une appropriation cloisonnée, avançant à rythmes variables selon les individus ou les groupes. Il ressort par exemple du corpus que certains enseignants recourent à l'IAG pour enrichir ou diversifier les activités d'évaluation et potentiellement pour gagner du temps de conception (Kanik, 2024 ; Onal et KulavuzOnal, 2024) et de correction (Geçkin *et al.*, 2023 ; Han et Lu, 2023). Les résultats de la correction automatisée s'avérant satisfaisants dans le cadre de certaines activités (Trujillo *et al.*, 2024 ; Xia *et al.*, 2024), il semble tentant de l'utiliser, notamment lorsque les travaux sont nombreux. Le gain de temps des enseignants peut aussi être favorable aux apprenants qui disposent dès lors plus rapidement de rétroactions. Toutefois, l'IAG ne paraît pas tant pouvoir remplacer que compléter le travail ou les interactions des étudiants et des enseignants puisque les échanges impliquent des capacités métacognitives ainsi qu'une intelligence émotionnelle particulières (Bhullar *et al.*, 2024 ; Kortemeyer, 2023 ; Hill, 2025).

De plus, l'usage de l'IAG soulève des enjeux majeurs concernant la collecte de données, enjeux rarement mentionnés dans les articles. Or, transférer l'évaluation formative ou la correction d'examen à des outils d'IAG confère à ces derniers la possibilité de recueillir et analyser de grandes quantités de données sur les étudiants, y compris sur leurs performances scolaires, leurs comportements d'apprentissage et leurs interactions en ligne. En outre, dans la mesure où certains outils peuvent engranger les contenus des travaux des étudiants et d'en partager des éléments lors de requêtes réalisées par d'autres personnes, il paraît crucial de garantir que toutes les données entrant dans le cadre d'un processus d'évaluation formative ou sommative sont protégées contre les accès et usages non autorisés ainsi que contre les violations de la confidentialité. Il serait donc opportun que les établissements d'enseignement garantissent que tout usage de l'IAG respecte les normes éthiques et implique le consentement des acteurs concernés. En ce sens, ils pourraient être encouragés à mettre en place des mesures de sécurité robustes pour protéger les informations personnelles des étudiants et l'ensemble des contenus issus du travail des étudiants et des enseignants.

Enfin, même si l'augmentation des risques de tricherie mine la confiance et la reconnaissance sociale envers les diplômes ou autres certifications et semble ainsi susceptible de provoquer leur perte de valeur symbolique (Bennett et Abusalem, 2024 ; Blackie, 2024 ; Cotton *et al.*, 2024 ; Zirar, 2023), aucune des recherches citées n'évalue concrètement l'impact actuel. Les articles recensés ne remettent d'ailleurs pas en cause ni la place conférée à l'évaluation sommative ou certificative des apprentissages dans les systèmes d'enseignement, ni l'importance accordée aux titres et certifications scolaires dans les processus de recrutement. Au mieux, un article suggère que l'évaluation formative serait désormais plus opportune que l'évaluation sommative, mais sans vraiment discuter la pertinence de cette dernière in fine (Farrokhnia *et al.*, 2024). Pourtant, il semble impératif de l'interroger, notamment au regard des travaux mettant en évidence de longue date la diversité des motifs et des formes de malhonnêteté académique (Berdahl, 2023) désormais facilitée par les outils d'IAG et les incertitudes entourant la légitimité de leurs usages. Tandis que la difficulté à cerner la frontière entre aide et triche, originalité ou délégation partielle

aux outils d'IAG (Murray, 2023 ; Perkins, 2023), engendre des situations délicates tant pour les étudiants que pour les enseignants, les publications de 2025 et 2026 que nous avons pu consulter continuent de soulever les questionnements, avantages, inconvénients et pistes de solutions relevés dans cette revue systématique narrative (Turnbull *et al.*, 2023) de la littérature.

Conclusion

Nous intéressants à l'évaluation dans l'enseignement supérieur, nous avons mené un processus de recension systématique narrative (Turnbull *et al.*, 2023) de la littérature au terme duquel nous avons sélectionné 111 articles parus dans des revues scientifiques. Dans la mesure où ils ont principalement été publiés entre 2022 et 2024, ces recherches portent souvent sur de premiers usages. Ainsi, bien que les résultats présentés soient généralement positifs, faisant ressortir une amélioration de l'apprentissage, voire de l'autonomie ou de la motivation des étudiants, ces derniers, tout comme les enseignants, n'étant pas toujours familiers de ces nouveaux outils, certaines réticences ou frustrations sont aussi mises en avant dans quelques publications. Ces dernières vont probablement diminuer avec l'intégration progressive de l'IAG dans les pratiques éducatives et sociales ainsi que l'amélioration rapide des outils d'IAG. Néanmoins, tandis que les étudiants tendent à percevoir positivement les interactions avec des outils d'IA (Marchal *et al.*, 2024), certains d'entre eux, affectés par des exercices qui ne fonctionnent pas et des rétroactions désincarnées, commencent à s'insurger contre l'usage de l'IAG fait par leurs enseignants (Hill, 2025). À cet égard, plusieurs études ayant fait le point plus récemment sur les performances de l'IAG concluent que les outils testés comportent toujours des lacunes importantes. Par exemple, l'IAG peinerait encore à évaluer les apprentissages dans la langue arabe (Ghazawi et Simpson, 2025) et les rétroactions générées automatiquement, bien que se rapprochant de celles faites par les enseignants, seraient plus répétitives, moins efficaces et de moins bonne qualité que celles offertes par les enseignants (Pecuchova *et al.*, 2026 ; Soyooof *et al.*, 2026).

Les divergences ou débats soulevés dans les articles recensés ne sont pas sans rappeler ceux ayant marqué de précédentes innovations technologiques, comme l'usage des moteurs de recherche ou la consultation de Wikipédia (Eijkman, 2010). Ce faisant, tandis que l'IAG tend à être présentée comme une opportunité, pour les étudiants et pour les enseignants, de diminuer le travail à réaliser, elle pourrait au contraire plutôt l'accroître en raison de la complexification des activités d'évaluation à proposer et à effectuer lorsque l'IAG est intégrée à ces dernières (Kortemeyer, 2023). Il semble dès lors important d'être au fait des différents essais technopédagogiques et de leurs limites documentées aussi bien par les chercheurs que les praticiens pour poursuivre la réflexion amorcée de longue date quant aux articulations entre humains et technologies dans l'enseignement supérieur. C'est en ce sens que plusieurs chercheurs concluent leurs articles en encourageant les établissements à former les acteurs de l'éducation aux usages éthiques de l'IAG et en incitant les enseignants à intégrer l'IAG à des activités authentiques d'évaluation permettant de développer une réflexion critique (Petit, 2022). Plus largement, ce sont des questions de

société qui sont ici soulevées, puisque, d'une part, tout usager a besoin d'être formé et tout citoyen requiert une protection contre les recours à l'IA se faisant à son insu et que, d'autre part, le système de formation-certification sur lequel reposent l'employabilité et la stratification sociale est quelque peu fragilisé.

Finalement, il semble important de revenir sur les limites d'une revue systématique narrative (Turnbull *et al.*, 2023) de la littérature. Tout d'abord, cette dernière intègre des recherches de différentes natures (résultats d'enquête, analyses de politiques, expérimentation, etc.), qui ne sont donc pas comparables, contrairement aux résultats d'une méta-analyse, mais permettent d'élargir le champ de réflexion. Ensuite, nous n'avons pas fait d'analyse systématique des biais, mais simplement recensé les limites perçues des publications, l'une d'entre elles étant notamment la moindre taille des populations étudiées dans plusieurs enquêtes. Enfin, le fait de n'avoir pris en considération que des articles publiés dans les revues scientifiques de quatre bases de données est une limite en soi, qui nous a amené à passer à côté de la littérature grise et de certains articles, d'autant plus qu'en réalisant ce travail en 2025, nous nous sommes arrêtés aux textes publiés fin 2024, ce qui explique que le principal outil d'IAG testé dans les recherches citées soit ChatGPT. L'émergence d'autres outils d'IAG et leur perfectionnement (Bousquet *et al.*, 2025 ; Roy *et al.*, 2025) ne remettent toutefois pas en cause la pertinence des enjeux identifiés dans cette revue, et semble au contraire renforcer la nécessité de repenser les stratégies d'évaluation au-delà de la simple détection de fraude. Alors que notre démarche ne nous a permis que de recueillir des textes en anglais, cet article et ce numéro de DMS dans son ensemble contribuent à enrichir les connaissances et réflexions francophones concernant l'impact de l'IAG dans l'éducation et la formation et il paraîtrait pertinent de refaire une telle revue de la littérature dans quelques années pour mieux cerner les évolutions en cours.

Bibliographie

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>.

Berdahl, L. (2023, 8 juin). Inconduite à l'université : Survol des causes et solutions potentielles. *Affaires universitaires*. <https://www.affairesuniversitaires.ca/career-advice-fr/inconduite-a-luniversite-survol-des-causes-et-solutions-potentielles/> (consulté le 12 juin 2026).

Blondelle, A. et Giorgini, P. (2025). *Intelligence artificielle et éducation : un double défi*. Hermann Éditeurs.

Boudreau, J.-P. (2020). Du copier-coller au créacollage numérique : Quelques jalons à poser pour contrer le plagiat. *Correspondance*, 25(5), 7.

Bousquet, M., Byrne, A., Forget, D., Gough, G., Rheault, L.-R., Roche, S. et Siaussat, D. (2025). National Context Impacts on SDG Mapping Needs and Approaches in Higher Education, a TriNational Comparison. *Sustainability*, 17(14), 6506. <https://doi.org/10.3390/su17146506>.

Büchter, R. B., Weise, A. et Pieper, D. (2020). Development, testing and use of data extraction forms in systematic reviews: A review of methodological guidance. *BMC Medical Research Methodology*, 20(1), 259. <https://doi.org/10.1186/s12874-020-01143-3>.

Burgason, K. A., Sefiha, O. et Briggs, L. (2019). Cheating Is in the Eye of the Beholder: An Evolving Understanding of Academic Misconduct. *Innovative Higher Education*, 44(3), 203-218.

Callahan, D. et Davidson, R. M. (2004). *The cheating culture: [Why more Americans are doing wrong to get ahead]*. Recorded Books.

Cizek, G. J. (1999). *Cheating on Tests: How To Do It, Detect It, and Prevent It*. Routledge. <https://doi.org/10.4324/9781410601520>.

Conseil supérieur de l'éducation et Commission de l'éthique en science et en technologie. (2024). *Intelligence artificielle générative en enseignement supérieur : Enjeux pédagogiques et éthiques*. Le Conseil et La Commission. <https://www.cse.gouv.qc.ca/publications/ia-enseignement-sup-50-0566/> (consulté le 12 juin 2026).

Darling-Hammond, L., Wilhoit, G. et Pittenger, L. (2014). Accountability for College and Career Readiness: Developing a New Paradigm. *Education Policy Analysis Archives*, 22, 86. <https://epaa.asu.edu/index.php/epaa/article/view/1724> (consulté le 12 juin 2026).

Denney, V., Dixon, Z., Gupta, A. et Hulphers, E. (2021). Exploring the Perceived Spectrum of Plagiarism: A Case Study of Online Learning. *Journal of Academic Ethics*, 19(2), 187-210. <https://doi.org/10.1007/s10805-020-09364-3>.

Detroz, P., Tessaro, W. et Younès, N. (2020). Évaluer en temps de pandémie. *e-JIREF*, 1-3. <https://journal.admee.org/ejiref/article/view/165/150> (consulté le 12 juin 2026).

Eckstein, M. A. (2003). *Combating academic fraud: Towards a culture of integrity*. International Institute for Educational Planning.

Eijkman, H. (2010). Academics and Wikipedia: Reframing Web 2.0+ as a Disruptor of Traditional Academic Power-Knowledge Arrangements. *Campus-Wide Information Systems*, 27(3), 173-185. <https://doi.org/10.1108/10650741011054474>.

EPPI-Centre. (2010). *EPPI-Centre Methods for Conducting Systematic Reviews*.

Fendler, R. J., Yates, M. C. et Godbey, J. M. (2018). Observing and Detering Social Cheating on College Exams. *International Journal for the Scholarship of Teaching and Learning*, 12(1). <https://doi.org/10.20429/ijstl.2018.120104>.

Gerstein, M. et Friedman, H. (2016). Rethinking higher education: focusing on skills and competencies. *Psychosociological Issues in Human Resource Management*, 4, 104-121.

Ghazawi, R. et Simpson, E. (2025). How well can LLMs grade essays in Arabic? *Computers and Education: Artificial Intelligence*, 9, 100449. <https://doi.org/10.1016/j.caeai.2025.100449>.

Goodier, M. (2025, 15 juin). Revealed: Thousands of UK University Students Caught Cheating Using AI. *The Guardian*. <https://www.theguardian.com/education/2025/jun/15/thousands-of-uk-university-studentscaught-cheating-using-ai-artificial-intelligence-survey> (consulté le 12 juin 2026).

Goto, J. et Ramnarain, U. (2026). Does culture matter in AI adoption? A predictive analysis of cultural dimensions, social influence, and personal innovativeness in the modified UTAUT model. *International Journal of Educational Technology in Higher Education*, 23(1), 13. <https://doi.org/10.1186/s41239-026-00588-8>.

Gulikers, J. T. M., Bastiaens, T. J. et Kirschner, P. A. (2004). A five-dimensional framework for authentic assessment. *Educational Technology Research and Development*, 52(3), 67-86. <https://doi.org/10.1007/BF02504676>.

Heckler, N. C. et Forde, D. R. (2015). The Role of Cultural Values in Plagiarism in Higher Education. *Journal of Academic Ethics*, 13(1), 61-75. <https://doi.org/10.1007/s10805-014-9221-3>.

Hern, A. (2022, 31 décembre). AI-assisted plagiarism? ChatGPT bot says it has an answer for that. *The Guardian*. <https://www.theguardian.com/technology/2022/dec/31/ai-assisted-plagiarism-chatgpt-bot-says-ithas-an-answer-for-that> (consulté le 12 juin 2026).

Herrington, J. A. et Herrington, A. J. (2006). *Authentic conditions for authentic assessment: Aligning task and assessment*.

Hill, K. (2025, 14 mai). The Professors Are Using ChatGPT, and Some Students Aren't Happy About It. *The New York Times*. <https://www.nytimes.com/2025/05/14/technology/chatgpt-college-professors.html> (consulté le 12 juin 2026).

Hollis, L. P. (2018). Ghost-Students and the New Wave of Online Cheating for Community College Students. *New Directions for Community Colleges*, 2018(183), 25-34. <https://doi.org/10.1002/cc.20314>.

Huang, P., Hwang, Y., Hsu, J. L., Peng, C. F., Tsai, C. H. et Wang, C. Y. (2026). The effectiveness of an AI-integrated VR oral training application in reducing public speaking anxiety and interview anxiety. *Computers and Education: Artificial Intelligence*, 10, 100514. <https://doi.org/10.1016/j.caeai.2025.100514>.

Ison, D. (2012). Plagiarism Among Dissertations: Prevalence at Online Institutions. *Journal of Academic Ethics*, 10(3), 227-236. <https://doi.org/10.1007/s10805-012-9165-4>.

Jeantheau, J. P. et Johnson, S. (2020). Assessment innovation and reform in France. *Assessment in Education: Principles, Policy & Practice*, 27(3), 290-308. <https://doi.org/10.1080/0969594X.2019.1577218>.

Kayisoglu, N. B. et Temel, C. (2017). An Examination of Attitudes towards Cheating in Exams by Physical Education and Sports High School Students. *Universal Journal of Educational Research*, 5(8), 1396-1402.

Kickbusch, S., Ashford-Rowe, K., Kemp, A., Boreland, J. et Huijser, H. (2025). Beyond Detection: Redesigning Authentic Assessment in an AI-Mediated World. *Education Sciences*, 15(11), 1537. <https://doi.org/10.3390/educsci15111537>.

Kirsanov, O., Kushwah, L. et Selvaretnam, G. (2025). Beyond Detection: How Students Use—and Hide—AI in Online Assessments and What Authentic Tasks Can Do About It. *Journal of Academic Ethics*, 24(1), 14. <https://doi.org/10.1007/s10805-025-09691-3>.

Marchal, P., Kumps, A., Floquet, C., Deruwé, O. et De Lièvre, B. (2024). Perceptions et usages d'un chatbot comme tuteur de cours en sciences de l'éducation. *Médiations & médiatisations : revue internationale sur le numérique en éducation et communication*, 18, 125-147. <https://doi.org/10.52358/mm.vi18.410>.

Marengo, A., Pagano, A. et Santamato, V. (2025). A machine learning framework for soft skills assessment: Leveraging serious games in higher education. *Computers and Education: Artificial Intelligence*, 9, 100469. <https://doi.org/10.1016/j.caeai.2025.100469>.

Maroy, C. et Pons, X. (2021). Re-conceptualising education policy trajectories in a globalised world: lessons from a multi-level comparison of accountability in France and Quebec. *Comparative Education*, 57(4), 560-578. <https://doi.org/10.1080/03050068.2021.1981714>.

McCabe, D. L., Butterfield, K. D. et Treviño, L. K. (2012). *Cheating in college: Why students do it and what educators can do about it*. The Johns Hopkins University Press. <http://www.scopus.com/inward/record.url?scp=84906003037&partnerID=8YFLogxK> (consulté le 12 juin 2026).

Moon, YI. (2007). Education reform and competency-based education. *Asia Pacific Educ. Rev.* 8, 337–341. <https://doi.org/10.1007/BF03029267>.

Murray, N. (2023). ChatGPT in the English Classroom: 18 Students Test it and Share their Insights. *Teaching and Learning Excellence through Scholarship*, 3(1). <https://doi.org/10.52938/tales.v3i1.2890>.

Nauffal, D. et Nasser, R. (2007). Differences Between Anglophone and Francophone Higher Educational Models: Students' Perceptions of The Quality of Performance Outcomes.

Oliveira, M., Zednik, C., Bombaerts, G., Sadowski, B. et Conijn, R. (2025). Assessing students' DRIVE: A framework to evaluate learning through interactions with generative AI. *Computers and Education: Artificial Intelligence*, 9, 100497. <https://doi.org/10.1016/j.caeai.2025.100497>.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, A. W., Mayo-Wilson, E., McDonald, S.,...Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews, *BMJ* 2021, 372(71). <https://doi.org/10.1136/bmj.n71>.

Papi, C., Gérin-Lajoie, S., Beaudoin, A., Verville, P. et Hébert, M.-H. (2025). *Processus de recension systématique sur l'intelligence artificielle et l'évaluation en enseignement supérieur 2024 et avant* (Version V2) [Jeu de données]. Borealis. <https://doi.org/10.5683/SP3/SIFEKS>.

Papi, C., Gérin-Lajoie, S. et Hébert, M.-H. (2020). Se rapprocher de l'évaluation à distance : dix pistes de réponse, *e-JIREF*, 201-206. <https://journal.admee.org/ejiref/article/view/186> (consulté le 12 juin 2026).

Pecuchova, J., Benko, L. et Drlik, M. (2026). Reimagining feedback through generative AI in engineering education. *Computers and Education: Artificial Intelligence*, 10, 100574. <https://doi.org/10.1016/j.caeai.2026.100574>.

Petit, L. (2022). La rencontre de l'Intelligence artificielle (IA) et de l'esprit critique (EC) : nouveaux enjeux ? nouvelle formation ? *Communication, technologies et développement*, 12. <https://doi.org/10.4000/ctd.8256>.

Pluchart, J.-J. (2019). Le triangle de la fraude. *Vie & sciences de l'entreprise*, 207(1), 83-97. <https://shs.cairn.info/revue-vie-et-sciences-de-l-entreprise-2019-1-page-83?lang=fr> (consulté le 12 juin 2026).

Roy, N., Proust-Androwkha, S., Gruslin, É., Vallerand, V. et Charles, É. (2025). L'intelligence artificielle au postsecondaire : Entre enthousiasme et méfiance – Introduction au numéro thématique. *Revue internationale des technologies en pédagogie universitaire*, 22(1). <https://doi.org/10.18162/ritpu-2025-v22n1-01>.

Soyoof, A., Reynolds, B. L., Rassaei, E., Kao, C.-W. et Van Ha, X. (2026). From teachers to chatbots: Scaffolded corrective feedback and student trust in online L2 English classrooms. *Computers and Education: Artificial Intelligence*, 10, 100530. <https://doi.org/10.1016/j.caeai.2025.100530>

Turnbull, D., Chugh, R. et Luck, J. (2023). Systematic-narrative Hybrid Literature Review: A Strategy for Integrating a Concise Methodology into a Manuscript. *Social Sciences & Humanities Open*, 7(1), 100381. <https://doi.org/10.1016/j.ssaho.2022.100381>.

Valdes-Ramirez, D., Beltran-Sanchez, J. A., Conant-Pablos, S. E., Ponce Lopez, R., Dominguez, A., Camacho-Zuñiga, C. et Zavala, G. (2026). A deep learning approach to estimating interaction levels in face-to-face lessons. *Computers and Education: Artificial Intelligence*, 10, 100528. <https://doi.org/10.1016/j.caeai.2025.100528>.

Wideman, M. A. (2008). Academic Dishonesty in Postsecondary Education: A literature review. *Transformative Dialogues: Teaching and Learning Journal*, 2(1), 1-15.

Wiggins, G. (1990). The Case for Authentic Assessment. *Practical Assessment, Research, and Evaluation*, 2(1), Article 1. <https://doi.org/10.7275/ffb1-mm19>.

Yeo, S. (2007). First-year university science and engineering students' understanding of plagiarism. *Higher Education Research & Development*, 26(2), 199-216. <https://doi.org/10.1080/07294360701310813>.

Annexe

Listes des références citées dans l'article venant du corpus

- Adilov, N., Cline, J. W., Hanke, H., Kauffman, K., Meneau, L., Resendez, E., Singh, S., Slaubaugh, M. et Suntornpithug, N. (2024). ChatGPT and the Course Vulnerability Index. *Journal of Education for Business*, 99(2), 125-132. <https://doi.org/10.1080/08832323.2023.2260929>.
- Ateeq, A., Alzoraiki, M., Milhem, M. et Ateeq, R. A. (2024). Artificial Intelligence in Education: Implications for Academic Integrity and the Shift Toward Holistic Assessment. *Frontiers in Education*, 01-11. <https://doi.org/10.3389/educ.2024.1470979>.
- Bannister, P., Peñalver, E. A. et Urbieta, A. S. (2024a). International Students and Generative Artificial Intelligence: A Cross-Cultural Exploratory Analysis of Higher Education Academic Integrity Policy. *Journal of International Students*, 14(3), 149-170.
- Bannister, P., Peñalver, E. A. et Urbieta, A. S. (2024b). Transnational Higher Education Cultures and Generative AI: A Nominal Group Study for Policy Development in English Medium Instruction. *Journal for Multicultural Education*, 18(1-2), 173-191. <https://doi.org/10.1108/JME-10-2023-0102>.
- Belda-Medina, J. et Kokošková, V. (2023). Integrating Chatbots in Education: Insights from the Chatbot-Human Interaction Satisfaction Model (CHISM). *International Journal of Educational Technology in Higher Education*, 20. <https://doi.org/10.1186/s41239-023-00432-3>.
- Bennett, L. et Abusalem, A. (2024). Artificial Intelligence (AI) and Its Potential Impact on the Future of Higher Education. *Athens Journal of Education*, 11(3), 195-212.
- Bertolini, R., Finch, S. J. et Nehm, R. H. (2021). Testing the Impact of Novel Assessment Sources and Machine Learning Methods on Predictive Outcome Modeling in Undergraduate Biology. *Journal of Science Education and Technology*, 30(2), 193-209. <https://doi.org/10.1007/s10956-020-09888-8>.
- Bhullar, P. S., Joshi, M. et Chugh, R. (2024). ChatGPT in Higher Education—A Synthesis of the Literature and a Future Research Agenda. *Education and Information Technologies*, 29(16), 21501-21522. <https://doi.org/10.1007/s10639-024-12723-x>.
- Blackie, M. A. L. (2024). ChatGPT is a game changer: Detection and eradication is not the way forward. *Teaching in Higher Education*, 29(4), 1109-1116. <https://doi.org/10.1080/13562517.2023.2300951>.
- Bower, M., Torrington, J., Lai, J. W. M., Petocz, P. et Alfano, M. (2024). How should we change teaching and assessment in response to increasingly powerful generative Artificial Intelligence?

Outcomes of the ChatGPT teacher survey. *Education and Information Technologies*, 29(12), 15403-15439. <https://doi.org/10.1007/s10639-023-12405-0>.

Çakmak, F. (2022). Chatbot-Human Interaction and Its Effects on EFL Students' L2 Speaking Performance and Anxiety. *Novitas-ROYAL (Research on Youth and Language)*, 16(2), 113-131.

Castelló-Sirvent, F., Roger-Monzó, V. et Gouveia-Rodrigues, R. (2024). Quo Vadis, University? A Roadmap for AI and Ethics in Higher Education. *Electronic Journal of E-Learning*, 22(6), 35-51.

Chaudhry, I. S., Sarwary, S. A. M., El Refae, G. A. et Chabchoub, H. (2023). Time to Revisit Existing Student's Performance Evaluation Approach in Higher Education Sector in a New Era of ChatGPT— A Case Study. *Cogent Education*, 10(1), 1-30. <https://doi.org/10.1080/2331186X.2023.2210461>.

Cong-Lem, N., Tran, T. N. et Nguyen, T. T. (2024). Academic Integrity in the Age of Generative AI: Perceptions and Responses of Vietnamese EFL Teachers. *Teaching English with Technology*, 24(1), 28-48.

Cotton, D. R. E., Cotton, P. A. et Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education & Teaching International*, 61(2), 228-239. <https://doi.org/10.1080/14703297.2023.2190148>.

Escalante, J., Pack, A. et Barrett, A. (2023). AI-Generated Feedback on Writing: Insights into Efficacy and ENL Student Preference. *International Journal of Educational Technology in Higher Education*, 20. <https://doi.org/10.1186/s41239-023-00425-2>.

Farrokhnia, M., Banihashem, S. K., Noroozi, O. et Wals, A. (2024). A SWOT Analysis of ChatGPT: Implications for Educational Practice and Research. *Innovations in Education and Teaching International*, 61(3), 460-474. <https://doi.org/10.1080/14703297.2023.2195846>.

Favero, T. G. (2024). Using artificial intelligence platforms to support student learning in physiology. *Advances in Physiology Education*, 48(2), 193-199. <https://doi.org/10.1152/advan.00213.2023>.

Firth, D. R., Derendinger, M. et Triche, J. (2024). Cheating Better with ChatGPT: A Framework for Teaching Students When to Use ChatGPT and Other Generative AI Bots. *Information Systems Education Journal*, 22(3), 47-60.

Geçkin, V., Kiziltas, E. et Çinar, Ç. (2023). Assessing Second-Language Academic Writing: AI vs. Human Raters. *Journal of Educational Technology and Online Learning*, 6(4), 1096-1108.

Giray, L., De Silos, P. Y., Adornado, A., Buelo, R. J. V., Galas, E., Reyes-Chua, E., Santiago, C. et Ulanday, Ma. L. (2024). Use and Impact of Artificial Intelligence in Philippine Higher

Education: Reflections from Instructors and Administrators. *Internet Reference Services Quarterly*, 28(3), 315-338. <https://doi.org/10.1080/10875301.2024.2352746>.

Gorichanaz, T. (2023). Accused: How Students Respond to Allegations of Using ChatGPT on Assessments. *Learning: Research and Practice*, 9(2), 183-196. <https://doi.org/10.1080/23735082.2023.2254787>.

Gozali, I., Wijaya, A. R. T., Lie, A., Cahyono, B. Y. et Suryati, N. (2024). Leveraging the potential of ChatGPT as an automated writing evaluation (AWE) tool: Students' feedback literacy development and AWE tools integration framework. *The JALT CALL Journal*, 20(1), 1-22. <https://doi.org/10.29140/jaltcall.v20n1.1200>.

Gruenhagen, J. H., Sinclair, P. M., Carroll, J.-A., Baker, P. R. A., Wilson, A. et Demant, D. (2024). The rapid rise of generative AI and its implications for academic integrity: Students' perceptions and use of chatbots for assistance with assessments. *Computers and Education: Artificial Intelligence*, 7, 100273. <https://doi.org/10.1016/j.caeai.2024.100273>.

Han, C. et Lu, X. (2023). Can Automated Machine Translation Evaluation Metrics Be Used to Assess Students' Interpretation in the Language Learning Classroom? *Computer Assisted Language Learning*, 36(5-6), 1064-1087. <https://doi.org/10.1080/09588221.2021.1968915>.

He, Y. (2024). The Metaphor of AI in Writing in English: A Reflection on EFL Learners' Motivation to Write, Enjoyment of Writing, Academic Buoyancy, and Academic Success in Writing. *International Review of Research in Open and Distributed Learning*, 25(3), 271-286. <https://doi.org/10.19173/irrodl.v25i3.7769>.

Hsu, H.-L., Chen, H. H.-J. et Todd, A. G. (2023). Investigating the Impact of the Amazon Alexa on the Development of L2 Listening and Speaking Skills. *Interactive Learning Environments*, 31(9), 5732-5745. <https://doi.org/10.1080/10494820.2021.2016864>.

Hyde, S. J., Busby, A. et Bonner, R. L. (2024). Tools or Fools: Are We Educating Managers or Creating Tool-Dependent Robots? *Journal of Management Education*, 48(4), 708-734. <https://doi.org/10.1177/10525629241230357>.

Jia, J. et He, Y. (2022). The Design, Implementation and Pilot Application of an Intelligent Online Proctoring System for Online Exams. *Interactive Technology and Smart Education*, 19(1), 112-120. <https://doi.org/10.1108/ITSE-12-2020-0246>.

Johnston, H., Wells, R. F., Shanks, E. M., Boey, T. et Parsons, B. N. (2024). Student Perspectives on the Use of Generative Artificial Intelligence Technologies in Higher Education. *International Journal for Educational Integrity*, 20. <https://doi.org/10.1007/s40979-024-00149-4>.

Kanik, M. (2024). The Use of ChatGPT in Assessment. *International Journal of Assessment Tools in Education*, 11(3), 608-621.

Khlaif, Z. N., Ayyoub, A., Hamamra, B., Bensalem, E., Mitwally, M. A. A., Ayyoub, A., Hattab, M. K. et Shadid, F. (2024). University Teachers' Views on the Adoption and Integration of Generative AI Tools for Student Assessment in Higher Education. *Education Sciences*, 14(10), 1090. <https://doi.org/10.3390/educsci14101090>.

Kim, N.-Y. (2022). AI-Integrated Mobile-Assisted Language Learning: Is It an Effective Way of Preparing for the TOEIC Test in Classroom Environments? *English Teaching*, 77(3), 79-102.

Kirwan, A. (2024). ChatGPT and university teaching, learning and assessment: Some initial reflections on teaching academic integrity in the age of Large Language Models. *Irish Educational Studies*, 43(4), 1389-1406. <https://doi.org/10.1080/03323315.2023.2284901>.

Klyshbekova, M. et Abbott, P. (2024). ChatGPT and Assessment in Higher Education: A Magic Wand or a Disruptor? *Electronic Journal of e-Learning*, 22(2), 30-45. <https://doi.org/10.34190/ejel.21.5.3114>.

Kohnke, L. (2023). A Pedagogical Chatbot: A Supplemental Language Learning Tool. *RELC Journal: A Journal of Language Teaching and Research*, 54(3), 828-838. <https://doi.org/10.1177/00336882211067054>.

Kong, S.-C., Cheung, W. M.-Y. et Zhang, G. (2023). Evaluating an Artificial Intelligence Literacy Programme for Developing University Students' Conceptual Understanding, Literacy, Empowerment and Ethical Awareness. *Educational Technology & Society*, 26(1), 16-30.

Kortemeyer, G. (2023). Could an Artificial-Intelligence Agent Pass an Introductory Physics Course? *Physical Review Physics Education Research*, 19(1). <https://doi.org/10.1103/PhysRevPhysEducRes.19.010132>.

Kortemeyer, G. et Bauer, W. (2024). Cheat Sites and Artificial Intelligence Usage in Online Introductory Physics Courses: What Is the Extent and What Effect Does It Have on Assessments? *Physical Review Physics Education Research*, 20(1). <https://doi.org/10.1103/PhysRevPhysEducRes.20.010145>.

Kotlyar, I., Pearse, N. J. et Krasman, J. (2024). Understanding Cross-Country Differences in Assessment Simulations: Insights from South African and Canadian Students. *Discover Education*, 3. <https://doi.org/10.1007/s44217-024-00125-7>.

Krecar, I. M., Kolega, M. et Jurcec, L. (2024). Perception of ChatGPT Usage for Homework Assignments: Students' and Professors' Perspectives. *IAFOR Journal of Education*, 12(2), 33-60.

Kukol, A. (2024). Combination of Ipsative and Sociomaterial Assessment Methodologies within University-Level Science Education. *Higher Education for the Future*, 11(2), 220-234. <https://doi.org/10.1177/23476311241268970>.

Lentz, P. (2024). Ethical Authorship and Moral Motivation: The Key to Ethical AI Use. *Business and Professional Communication Quarterly*, 87(4), 592-609. <https://doi.org/10.1177/23294906241259153>.

Liu, C.-C., Liu, S.-J., Hwang, G.-J., Tu, Y.-F., Wang, Y. et Wang, N. (2023). Engaging EFL Students' Critical Thinking Tendency and In-Depth Reflection in Technology-Based Writing Contexts: A Peer Assessment-Incorporated Automatic Evaluation Approach. *Education and Information Technologies*, 28(10), 13027-13052. <https://doi.org/10.1007/s10639-023-11697-6>.

Luo, J. (2024). A Critical Review of GenAI Policies in Higher Education Assessment: A Call to Reconsider the "Originality" of Students' Work. *Assessment & Evaluation in Higher Education*, 49(5), 651-664. <https://doi.org/10.1080/02602938.2024.2309963>.

MacDonald, H. et Bézaire, V. (2024). Leveraging Systematic Review Practice for Research Skill Development in an Undergraduate Science Course: A Case Study. *Advances in Physiology Education*, 48(3), 518-526. <https://doi.org/10.1152/advan.00171.2023>.

MacDowell, P., Moskalyk, K., Korchinski, K. et Morrison, D. (2024). Preparing Educators to Teach and Create With Generative Artificial Intelligence. *Canadian Journal of Learning and Technology / Revue Canadienne de l'apprentissage et de La Technologie*, 50(4), 1-23. <https://doi.org/10.21432/cjlt28606>.

Marcel, F. et Kang, P. (2024). Examining AI Guidelines in Canadian Universities: Implications on Academic Integrity in Academic Writing. *Discourse and Writing/Rédactologie*, 34, 93-126. <https://doi.org/10.31468/dwr.1051>.

Moore, S. et Lookadoo, K. (2024). Communicating Clear Guidance: Advice for Generative AI Policy Development in Higher Education. *Business and Professional Communication Quarterly*, 87(4), 610-629. <https://doi.org/10.1177/23294906241254786>.

Moorhouse, B. L., Yeo, M. A. et Wan, Y. (2023). Generative AI tools and assessment: Guidelines of the world's top-ranking universities. *Computers and Education Open*, 5, 100151. <https://doi.org/10.1016/j.caeo.2023.100151>.

Moya, B. A. et Eaton, S. E. (2024). Academic Integrity Policy Analysis of Chilean Universities. *Journal of Academic Ethics*, 22(4), 639-663. <https://doi.org/10.1007/s10805-024-09515-w>.

Newton, P. et Xiromeriti, M. (2024). ChatGPT performance on multiple choice question examinations in higher education. A pragmatic scoping review. *Assessment & Evaluation in Higher Education*, 49(6), 781-798. <https://doi.org/10.1080/02602938.2023.2299059>.

Nguyen, H. M. et Goto, D. (2024). Unmasking academic cheating behavior in the artificial intelligence era: Evidence from Vietnamese undergraduates. *Education & Information Technologies*, 29(12), 15999-16025. <https://doi.org/10.1007/s10639-024-12495-4>.

Nguyen Thanh, B., Vo, D. T. H., Nguyen Nhat, M., Pham, T. T. T., Thai Trung, H. et Ha Xuan, S. (2023). Race With the Machines: Assessing the Capability of Generative AI in Solving Authentic Assessments. *Australasian Journal of Educational Technology*, 39(5), 59-81. <https://doi.org/10.14742/ajet.8902>.

Nikolic, S., Daniel, S., Haque, R., Belkina, M., Hassan, G. M., Grundy, S., Lyden, S., Neal, P. et Sandison, C. (2023). ChatGPT versus engineering education assessment: A multidisciplinary and multi-institutional benchmarking and analysis of this generative artificial intelligence tool to investigate assessment integrity. *European Journal of Engineering Education*, 48(4), 559-614. <https://doi.org/10.1080/03043797.2023.2213169>.

O, K.-M. (2024). A comparative study of AI-human-made and human-made test forms for a university TESOL theory course. *Language Testing in Asia*, 14, 1-17. <https://doi.org/10.1186/s40468-024-00291-3>.

Onal, S. et Kulavuz-Onal, D. (2024). A Cross-Disciplinary Examination of the Instructional Uses of ChatGPT in Higher Education. *Journal of Educational Technology Systems*, 52(3), 301-324. <https://doi.org/10.1177/00472395231196532>.

Oravec, J. A. (2022). AI, Biometric Analysis, and Emerging Cheating Detection Systems: The Engineering of Academic Integrity?: IA, análisis biométrico y sistemas emergentes de detección de trampas: ¿La ingeniería de la integridad académica? *Education Policy Analysis Archives / Archivos Analíticos de Políticas Educativas / Arquivos Analíticos de Políticas Educativas*, 30(175-177), 1-18. <https://doi.org/10.14507/epaa.30.5765>.

Perkins, M. (2023). Academic Integrity considerations of AI Large Language Models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching & Learning Practice*, 20(2), 1-24. <https://doi.org/10.53761/1.20.02.07>.

Perkins, M. et Roe, J. (2024). Decoding Academic Integrity Policies: A Corpus Linguistics Investigation of AI and Other Technological Threats. *Higher Education Policy*, 37(3), 633-653. <https://doi.org/10.1057/s41307-023-00323-2>.

Richards, M., Waugh, K., Slaymaker, M. A., Petre, M., Woodthorpe, J. et Gooch, D. (2024). Bob or Bot: Exploring ChatGPT's Answers to University Computer Science Assessment. *ACM Transactions on Computing Education*, 24(1). <https://doi.org/10.1145/3633287>.

Saarna, C. (2024). Identifying Whether a Short Essay Was Written by a University Student or ChatGPT. *International Journal of Technology in Education*, 7(3), 611-633.

Sallam, M. et Al-Salahat, K. (2023). Below average ChatGPT performance in medical microbiology exam compared to university students. *Frontiers in Education*, 111. <https://doi.org/10.3389/educ.2023.1333415>.

Sefcik, L., Veeran-Colton, T., Baird, M., Price, C. et Steyn, S. (2022). An Examination of Student User Experience (UX) and Perceptions of Remote Invigilation during Online Assessment. *Australasian Journal of Educational Technology*, 38(2), 49-69. <https://doi.org/10.14742/ajet.6871>.

Shailendra, S., Kadel, R. et Sharma, A. (2024). Framework for Adoption of Generative Artificial Intelligence (GenAI) in Education. *IEEE Transactions on Education*, 67(5), 777-785. <https://doi.org/10.1109/TE.2024.3432101>.

Tossell, C. C., Tenhundfeld, N. L., Momen, A., Cooley, K. et Visser, E. J. de. (2024). Student Perceptions of ChatGPT Use in a College Essay Assignment: Implications for Learning, Grading, and Trust in Artificial Intelligence. *IEEE Transactions on Learning Technologies*, 17, 1069-1081. <https://doi.org/10.1109/TLT.2024.3355015>.

Trujillo, J. C., Vela Ulloa, J., Escalona Vivas, G., Grasset Escobar, E., Villagrán Gutiérrez, I., Achurra Tirado, P. et Varas Cohen, J. (2024). Surgeons vs ChatGPT: Assessment and Feedback Performance Based on Real Surgical Scenarios. *Journal of Surgical Education*, 81(7), 960-966. <https://doi.org/10.1016/j.jsurg.2024.03.012>

Vicary, A. et Treffers-Daller, J. (2024). Reading against All Odds: Creative Coping Strategies of International Students in HE. *International Journal of Bilingual Education and Bilingualism*, 27(8), 1128-1141. <https://doi.org/10.1080/13670050.2024.2345698>.

Waltzer, T., Pilegard, C. et Heyman, G. D. (2024). Can you spot the bot? Identifying AI-generated writing in college essays. *International Journal for Educational Integrity*, 20(1), 1-18. <https://doi.org/10.1007/s40979-024-00158-3>.

Wise, B., Emerson, L., Luyn, A. V., Dyson, B., Bjork, C. et Thomas, S. E. (2024). A Scholarly Dialogue: Writing Scholarship, Authorship, Academic Integrity and the Challenges of AI. *Higher Education Research and Development*, 43(3), 578-590. <https://doi.org/10.1080/07294360.2023.2280195>.

Xia, Q., Weng, X., Ouyang, F., Lin, T. J. et Chiu, T. K. F. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education*, 21(1), 1-22. <https://doi.org/10.1186/s41239-024-00468-z>.

Xiao, T., Yi, S. et Akhter, S. (2024). AI-Supported Online Language Learning: Learners' Self-Esteem, Cognitive-Emotion Regulation, Academic Enjoyment, and Language Success. *International Review of Research in Open and Distributed Learning*, 25(3), 77-96. <https://doi.org/10.19173/irrodl.v25i3.7666>.

Yang, A. C. M., Chen, I. Y. L., Flanagan, B. et Ogata, H. (2021). Automatic Generation of Cloze Items for Repeated Testing to Improve Reading Comprehension. *Educational Technology & Society*, 24 (3), 147-158. <https://www.jstor.org/stable/27032862>.

Yavich, R. et Davidovitch, N. (2024). Plagiarism among Higher Education Students. *Education Sciences*, 14(8), 908. <https://doi.org/10.3390/educsci14080908>.

Yusuf, A., Pervin, N. et Román-González, M. (2024). Generative AI and the future of higher education: A threat to academic integrity or reformation? Evidence from multicultural perspectives. *International Journal of Educational Technology in Higher Education*, 21(1), 1-29. <https://doi.org/10.1186/s41239-024-00453-6>.

Zirar, A. (2023). Exploring the Impact of Language Models, Such as ChatGPT, on Student Learning and Assessment. *Review of Education*, 11(3). <https://doi.org/10.1002/rev3.3433>.

Remerciements

Ce projet de recherche a été rendu possible grâce au financement de l'Université TÉLUQ (FAR1) et de celui du GRIIPTIC.

