



Recension des usages d'intelligences artificielles génératives (IAg) pour offrir de la rétroaction en enseignement supérieur

Use of Generative Artificial Intelligence (GenAI) to Provide Feedback in Higher Education: A review

<https://doi.org/10.18162/ritpu-2025-v22n1-11>

Chantal TREMBLAY ✉  Université du Québec à Montréal, Canada

Sylvain MIKLOHOUN ✉ Université de Montréal, Canada

Bruno POELLHUBER ✉  Université de Montréal, Canada

Mis en ligne : 3 juin 2025

Résumé

L'intelligence artificielle générative (IAg) connaît actuellement un essor sans précédent en éducation. Parmi les potentialités de l'IAg pour favoriser l'apprentissage étudiant en enseignement supérieur, on retrouve sa capacité à fournir une rétroaction personnalisée. Pour comprendre l'état de la recherche actuelle, cet examen de la portée synthétise les recherches traitant de l'usage d'IAg pour fournir de la rétroaction sur une production étudiante. Les résultats montrent une variété d'IAg utilisées et les caractéristiques des rétroactions. La discussion souligne le besoin de mieux documenter les approches conceptuelles mobilisées pour concevoir les IAg afin de favoriser la compréhension des études et le transfert des connaissances.

Mots-clés

Intelligence artificielle générative, rétroaction, étayage, évaluation des apprentissages, enseignement supérieur, enseignement postsecondaire, enseignement universitaire

Abstract

Generative artificial intelligence (GenAI) is currently undergoing unprecedented growth in education. Among the ways that GenAI could potentially enhance learning in higher education is its ability to provide personalized feedback. To understand the current state of research, this scoping review synthesizes studies focusing on the use of GenAI to provide feedback on student work. The results reveal a variety of GenAI tools used and the characteristics of the feedback provided. The discussion emphasizes the need for better documentation of the methodological design processes used for GenAI, in order to promote understanding of the studies and the transfer of knowledge.



Keywords

Generative artificial intelligence, feedback, scaffolding, learning assessment, higher education, postsecondary education, university students

Introduction

L'intelligence artificielle générative (IAg) connaît actuellement un essor sans précédent, notamment depuis l'accessibilité au grand public de l'agent conversationnel ChatGPT, suivie de l'apparition rapide de nombreux autres services fondés sur de grands modèles de langage et l'IAg (Gemini, Llama, Claude, etc.). Les systèmes fonctionnant à partir de l'IAg sont préentraînés avec de grands ensembles de données et recourent à des systèmes de traitement du langage naturel. Bien qu'ils soient fondés sur des modèles probabilistes plutôt que sur une véritable compréhension du langage naturel, ils sont capables de répondre textuellement, oralement ou visuellement à des instructions ou requêtes (*prompts*) qui sont écrites sous la forme d'un texte continu ou d'une conversation.

Les potentialités de l'IAg pour favoriser l'apprentissage étudiant sont nombreuses et variées, mais le sujet de l'IAg en enseignement supérieur fait l'objet de vives discussions, notamment sur les plans du plagiat et de la probité intellectuelle. Sa capacité à biaiser ou à inventer des données engendre un risque important de désinformation (Reporters sans frontières, 2023). Plusieurs autres risques éthiques ou légaux potentiels ont été signalés dans la littérature : confidentialité des renseignements personnels, confabulations, biais, manque de transparence, empreinte carbone, etc.

Malgré cela, les potentialités de l'IAg en enseignement supérieur sont de plus en plus reconnues. Plusieurs d'entre elles sont relevées par le Conseil de l'innovation du Québec (2024) et une liste systématique a été publiée par Yan *et al.* (2024). La recherche de Lepage (2024) montre que les personnes enseignantes sont particulièrement disposées à adopter les IAg qui ont trait à la préparation de matériel pédagogique et didactique. La personnalisation de l'apprentissage constitue l'une des potentialités les plus importantes de l'IAg, grâce à un apprentissage adaptatif et personnalisé adapté au niveau de chaque personne apprenante. Or, les résultats des recherches récentes sur les tuteurs destinés aux étudiants et étudiantes rapportent des résultats contradictoires. Pardos et Bhandari (2024) rapportent des résultats positifs en mathématiques, tandis que Bastani *et al.* (2024) observent plutôt le contraire. Les résultats de leur étude montrent que le groupe expérimental ayant utilisé ChatGPT-4 pour s'exercer à résoudre des problèmes mathématiques a obtenu des résultats significativement plus faibles que le groupe témoin n'ayant pas utilisé d'IAg lors d'un examen où toute forme d'aide était interdite.

S'il ne semble techniquement pas trop difficile de développer des systèmes qui répondent aux questions des étudiants et étudiantes et leur donnent une rétroaction simple, développer un véritable système de tutorat intelligent adaptatif est une entreprise autrement difficile. La conception d'un système de tutorat intelligent adaptatif serait idéalement fondée sur une capacité à étayer l'apprentissage en fournissant une rétroaction personnalisée à chaque étudiante ou étudiant en fonction de son niveau précis dans la maîtrise de la matière, et de manière à lui fournir les indices lui permettant de progresser plutôt qu'en lui donnant des réponses toutes faites. L'IAg pourrait donc substituer ou compléter le rôle du tuteur ou de la tutrice tel que défini dans la théorie de l'étayage de Bruner (1983). Pour cela, il serait nécessaire d'utiliser des techniques

avancées pour configurer les systèmes fondés sur l'IAg afin qu'ils évaluent différentes productions étudiantes et fournissent une rétroaction adaptée selon certains paramètres.

Le potentiel qu'a l'IAg de transformer la manière de développer des tuteurs intelligents adaptatifs est ce qui nous a amenés à entreprendre cette recension d'écrits. Ainsi, cet article présente les résultats d'un examen de la portée (*scoping review*) (Arksey et O'Malley, 2005; Sabiston *et al.*, 2022) qui vise à synthétiser les recherches traitant des usages d'IAg pour fournir de la rétroaction (*feedback*) sur une production étudiante. Le cadre conceptuel expose les concepts d'étayage et d'IAg afin d'expliquer les critères de sélection des articles de cette recension. Les cinq étapes de la méthodologie employée dans cette étude sont ensuite décrites. Les résultats présentent une synthèse des études, en faisant ressortir leurs ressemblances et différences en ce qui concerne la rétroaction fournie par l'IAg à l'étudiant ou l'étudiante. La discussion met en évidence certaines limites observées dans ces études, ce qui permet en conclusion de proposer des pistes de recherche futures.

1. La rétroaction : un levier pour soutenir l'apprentissage

Bien qu'il n'existe aucune définition qui fasse consensus dans la littérature quant au concept de rétroaction fournie par l'enseignant ou l'enseignante (Evans, 2013), celle-ci sera définie dans ce texte par une forme de communication, écrite ou orale, qui porte sur l'appréciation d'une production étudiante (Boucher, 2015). En enseignement supérieur, la rétroaction peut être un levier pour soutenir les apprentissages, notamment lorsqu'elle est constructive et qu'elle permet effectivement à l'étudiant ou l'étudiante de s'améliorer (Boucher, 2015; Henderson *et al.*, 2019). L'importance de la rétroaction comme geste pédagogique en enseignement supérieur ressort de la synthèse de l'ensemble des méta-analyses en éducation réalisée par Hattie (2024).

La théorie de l'étayage de Bruner (1983) conceptualise cette rétroaction constructive en postulant que le processus d'apprentissage repose principalement sur le dialogue entre personnes étudiantes et enseignantes (ou tutrices). Dans ce processus qui s'amorce par l'établissement d'une cible d'apprentissage commune entre ces deux parties, la personne enseignante joue un rôle essentiel en évaluant le niveau de maîtrise de la personne étudiante à intervalles réguliers, pour lui fournir une rétroaction personnalisée et adaptée qui lui permettra de se situer, de déterminer ses forces et ses faiblesses et, ainsi, de s'améliorer. Alors que la personne étudiante progresse, la personne enseignante réduira sa rétroaction jusqu'à ce que la première devienne autonome et atteigne la cible d'apprentissage. Ce moment représente ce que Bruner (1983) appelle l'internalisation : l'étudiant ou l'étudiante est désormais capable de réaliser la tâche de façon autonome, ayant mémorisé le dialogue avec son enseignant ou enseignante, lui permettant de s'autoévaluer et de se donner de la rétroaction. La théorie de l'étayage mise donc essentiellement sur la qualité du dialogue personne étudiante-personne enseignante, qui représente une rétroaction personnalisée permettant de soutenir l'apprentissage.

2. Le potentiel de l'IAg pour fournir de la rétroaction et de la rétroaction personnalisée

Depuis de nombreuses années, des recherches en technologies éducatives ont visé à substituer la rétroaction fournie par l'enseignant ou l'enseignante par une rétroaction automatisée fournie par des outils numériques, notamment intégrés à des systèmes de tuteurs intelligents (Belland *et al.*, 2017; Kim et Lim, 2019; Zheng, 2016). Néanmoins, une limite importante à ces outils est qu'ils ne peuvent rétroagir sur le discours de l'étudiant ou l'étudiante. De fait, la rétroaction fournie est

parfois générique, parfois adaptée selon des réponses formulées par l'étudiant ou l'étudiante à des questions fermées ou à choix multiples (Belland *et al.*, 2017; Kim *et al.*, 2018). De surcroît, la programmation d'un système de tuteur intelligent pour offrir une rétroaction plus adaptée nécessite une compréhension fine de ces systèmes et de modèles d'étudiants ou étudiantes pour être en mesure de programmer tous les cas de figure possibles (Boissière et Bruillard, 2021, chap. 13).

À l'opposé, l'IAg pourrait fournir une telle rétroaction à l'étudiant ou l'étudiante par sa capacité à réagir sur une production écrite, un texte librement écrit dans ses propres mots. Elle peut donc fournir une rétroaction non seulement sur une réponse à une question ouverte, mais également sur la démarche, les justifications, les arguments. Cette rétroaction sera alors personnalisée pour l'étudiant ou l'étudiante et le texte généré par l'IAg est unique. Qui plus est, l'IAg pourrait fournir une rétroaction adaptative à l'étudiant ou l'étudiante lorsqu'elle est utilisée à plusieurs reprises pour évaluer sa production grâce à sa capacité d'apprendre des productions et des rétroactions précédentes. Bref, avec des travaux de configuration adéquats, la rétroaction fournie par les systèmes fondés sur l'IAg pourrait se rapprocher davantage de celle préconisée dans la théorie de l'échafaudage de Bruner, car elle permet une meilleure impression de dialogue que les autres outils d'échafaudage numériques non fondés sur l'IAg. Cette rétroaction devrait donc être plus efficace pour l'apprentissage, car elle aidera davantage l'étudiant ou l'étudiante à s'autoévaluer et à déterminer ce qui doit être amélioré.

Dans l'intention d'éventuellement développer des systèmes d'IAg visant à offrir de telles rétroactions, la présente recherche recense les études qui permettent d'avoir une compréhension de l'état actuel des connaissances dans ce champ. Elle vise donc à : 1) décrire les caractéristiques des études qui avaient pour objet l'étude de la rétroaction fournie par une IAg, incluant la démarche méthodologique de conception de l'IAg; et 2) présenter les caractéristiques des rétroactions fournies par ces IAg.

S'inspirant de la démarche méthodologique d'un examen de la portée, elle a permis d'analyser 12 articles qui traitent de l'usage d'une IAg pour fournir une rétroaction sur la production d'un étudiant ou une étudiante universitaire, afin de l'aider à s'autoévaluer et à soutenir son apprentissage.

3. La démarche de recension inspirée d'un examen de la portée

L'examen de la portée est un type de recension des écrits dont l'intention est de comprendre l'étendue de la littérature scientifique dans un champ spécifique, généralement nouveau et peu documenté, pour en présenter une synthèse explicative (Arksey et O'Malley, 2005; Grimshaw, 2010). Ce type de recension est donc pertinent lorsque l'intention est de cibler les manques ou les limites des études actuelles pour proposer des pistes de recherche futures. Il se distingue d'autres formes de recensions, car il ne vise pas à évaluer la qualité des études ni à agréger les résultats d'études antérieures pour vérifier des hypothèses, par exemple concernant l'efficacité d'un outil ou d'une méthode pédagogique.

La démarche méthodologique proposée par Arksey et O'Malley (2005) inclut cinq étapes principales, dont la première consiste à formuler la question de recherche. Dans cette recension, la question est : « Quelles sont les caractéristiques des études qui portent sur l'usage d'une IAg pour fournir de la rétroaction sur une production d'un étudiant ou d'une étudiante, afin de soutenir son apprentissage? » Cette comparaison porte sur les éléments contextuels des études (lieu, population étudiante, discipline), mais surtout sur les caractéristiques de la rétroaction fournie (moments, format, relation avec la rétroaction enseignante, personnalisation, adaptabilité). Pour ces dernières,

les justifications apportées concernant les choix effectués durant la conception de l'IAg sont également rapportées lorsque connues, pour mieux comprendre les facteurs à considérer pour fournir une rétroaction de qualité à l'aide d'IAg.

La deuxième étape porte sur la recension du corpus pertinent pour répondre à la question de recherche. Pour cela, les bases de données ERIC, IEEE Xplore, UNESCO, Web of Science et Érudit ont été mobilisées. Considérant l'évolution rapide de l'IA, la recherche d'articles s'est limitée à ceux publiés depuis 2019, soit les cinq années précédant le moment où la recherche a eu lieu (hiver 2024). Les mots-clés utilisés (en français et en anglais) sont : intelligence artificielle, étayage, rétroaction, cognition et enseignement supérieur. La stratégie de sélection des études pertinentes (3^e étape de l'examen de la portée) est illustrée à la figure 1. D'abord, la recension initiale a mené à repérer 1 091 articles en excluant les doublons. Un premier tri effectué par un codeur a permis de retirer 972 articles dont les titres et résumés montraient que leur sujet principal ne correspondait pas à l'étude de la rétroaction fournie par une IA. Cette personne a relu les résumés des 119 études restantes et retiré 55 autres articles du corpus pour des motifs de mauvaise population à l'étude ou de sujet principal non pertinent.

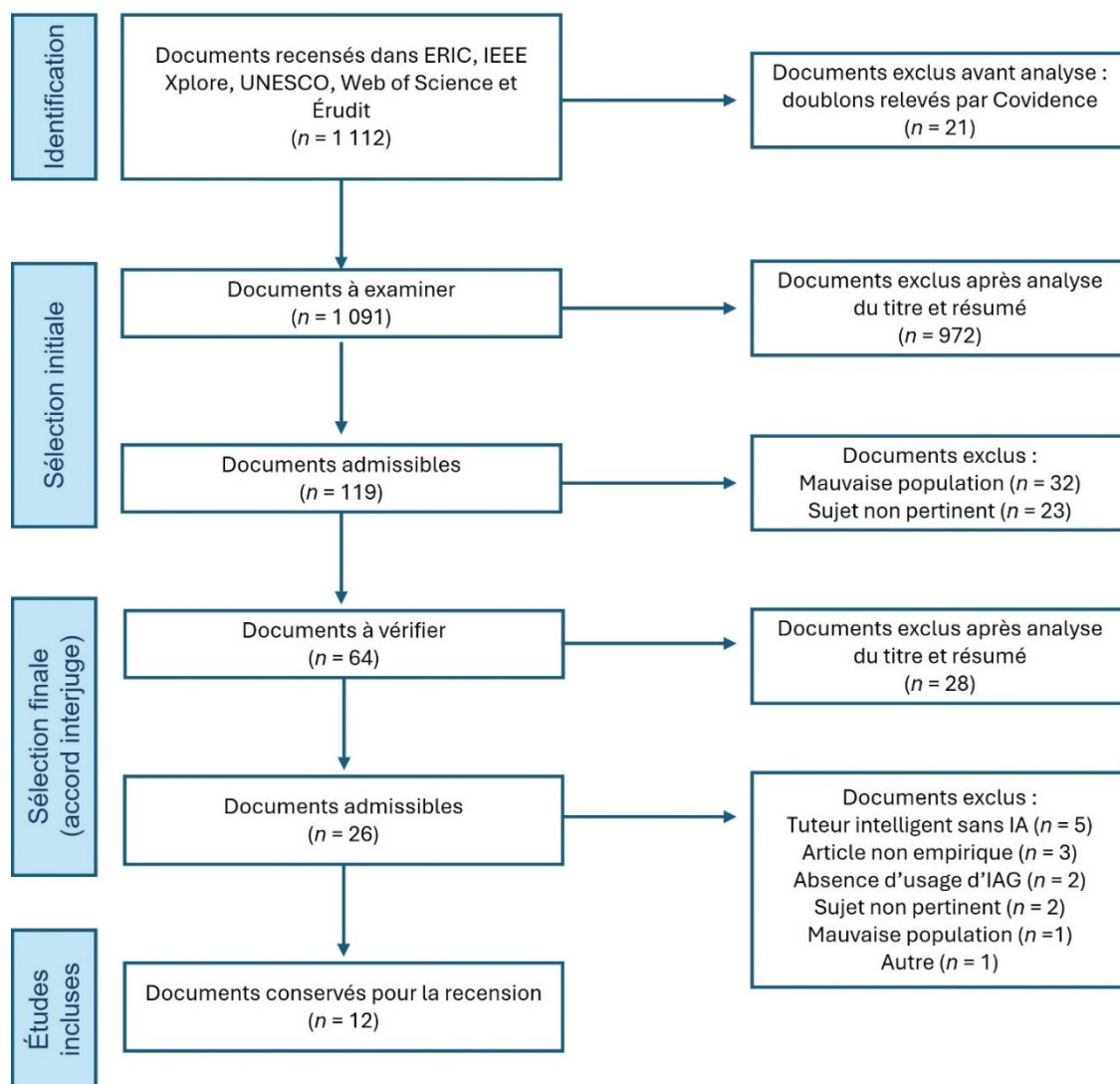


Figure 1
Processus de sélection des études de la recension

Dans un deuxième temps, l'application Covidence a été exploitée pour effectuer le tri final des articles. Cette application facilite la sélection des études pertinentes en fonctionnant selon le modèle PRISMA, qui utilise aussi les principes d'accord interjuges. Ainsi, il est possible d'y importer les références complètes des articles à trier, incluant le résumé et le DOI lorsque fournis. Ensuite, les personnes codeuses doivent élaborer des critères d'inclusion et d'exclusion qu'elles utiliseront pour évaluer chaque article à trier. Individuellement, chacune d'elles lit le titre et le résumé de chaque article pour déterminer s'il doit être inclus ou exclu de la recension, en justifiant son choix à partir des critères préalablement établis. Il est également possible pour la personne codeuse de signifier son incertitude si elle ne parvient pas à se prononcer. Au terme de cette activité de codage, Covidence agrège les résultats, ce qui permet de repérer rapidement les articles qui font consensus et ceux pour lesquels il n'y a pas d'accord interjuges.

Pour cette recension, deux personnes codeuses ont échangé entre elles pour formuler des critères d'inclusion qui reposaient sur la population à l'étude (étudiants et étudiantes universitaires), l'usage d'une IAg pour offrir de la rétroaction et la présence de résultats empiriques. La lecture des titres et des résumés des 64 articles importés dans Covidence a permis d'en exclure 28 autres sur la base de ces critères. Enfin, une seconde session de codage a été menée sur les 36 articles restants, où une lecture approfondie de leur méthodologie a permis d'en exclure 24. Il s'agissait principalement d'articles dont les systèmes de tuteurs intelligents ou applications numériques étudiés ne permettaient pas que la rétroaction fournie à l'étudiante ou l'étudiant soit créée à partir d'une IAg. Au total, 12 articles ont donc été analysés pour cet examen de la portée. La synthèse du corpus (4^e étape de l'examen de la portée) et la présentation des résultats (5^e étape) sont exposées à la section suivante.

4. Les résultats : caractéristiques des rétroactions fournies par une IAg

Les résultats sont présentés dans l'ordre des deux objectifs de cette recension, dont le premier décrit les caractéristiques des études recensées. Le deuxième présente les caractéristiques des rétroactions fournies par les IAg.

4.1 Les caractéristiques des études recensées

Le tableau A.1 (annexe A) présente une synthèse des caractéristiques des études recensées. Bien qu'une seule d'entre elles ait été publiée en 2019, on remarque que la majorité l'ont été entre 2022 et 2024 ($n = 8$), ce qui est cohérent avec l'arrivée de ChatGPT en 2022 et de la hausse généralisée des usages de l'IAg depuis. Les disciplines pour lesquelles ces IAg ont été conçues sont variées, incluant certaines dans ces domaines : arts ($n = 2$), science, technologie et mathématique ($n = 2$), sciences de l'éducation ($n = 2$) et langues et communication ($n = 3$). La plupart des personnes participantes poursuivaient des études de 1^{er} cycle universitaire ($n = 7$), bien que deux études portent sur des étudiants et étudiantes de cycles supérieurs (Kim *et al.*, 2022; Nazari *et al.*, 2021) et que deux études menées dans des cours du domaine des langues et communication (Escalante *et al.*, 2023; Yu, 2023) étaient ouvertes à tous. Toutes ces études sauf deux (Taub et Azevedo, 2019; Zheng *et al.*, 2024) se sont déroulées dans un contexte authentique de cours universitaire. Zheng *et al.* (2024) ont plutôt utilisé un cadre expérimental où les personnes participantes volontaires ont réalisé des tâches d'apprentissage dans un environnement numérique collaboratif dans lequel l'IAg fournissait de la rétroaction. L'étude de Taub et Azevedo (2019) s'est déroulée sur une période de deux jours durant lesquels les personnes participantes effectuaient des activités

d'apprentissage liées à des contenus de biologie avec l'assistance d'un tuteur intelligent pour une durée allant d'une à quatre heures.

Un résumé du principal objectif de ces études présenté au tableau A.1 permet de dégager une tendance. La majorité d'entre elles correspondent à des études comparatives ($n = 8$), où les effets de la rétroaction fournie par l'IAg (groupe expérimental) sont comparés à une situation où la rétroaction est fournie par une personne enseignante ou tutrice ($n = 3$), où la rétroaction fournie est générique et la même pour tous ($n = 2$) ou à une situation sans rétroaction ($n = 3$). Dans le premier cas, les groupes témoins reçoivent une rétroaction personnalisée par une personne enseignante ou tutrice durant la réalisation de la tâche (Chiu *et al.*, 2024; Essel *et al.*, 2022) ou après avoir soumis une production (Escalante *et al.*, 2023). Dans le deuxième cas (Mousavi *et al.*, 2021; Sailer *et al.*, 2023), la rétroaction de l'IAg est plutôt comparée à une rétroaction générique prédéterminée qui est la même, quelle que soit la production réalisée. Dans le troisième cas (Darvishi *et al.*, 2022; Nazari *et al.*, 2021; Taub et Azevedo, 2019), les deux groupes recevaient une rétroaction par les pairs, tandis que le groupe expérimental recevait en plus une rétroaction de l'IAg. La comparaison porte sur les apprentissages réalisés (p. ex., Chiu *et al.*, 2024), la performance scolaire (p. ex., Escalante *et al.*, 2023) ou l'engagement de l'étudiant ou l'étudiante (p. ex., Nazari *et al.*, 2021). Les méthodologies de ces huit études s'appuient donc sur des données quantitatives, où l'utilisation de pré- et post-tests est prédominante. Toutes ces études, sauf celle d'Escalante *et al.* (2023), présentent des résultats statistiquement significatifs qui montrent que les apprentissages, la performance ou l'engagement des groupes expérimentaux ayant utilisé les IAg étaient plus élevés que ceux des groupes témoins.

Yu (2023) et Zheng *et al.* (2024) exploitent plutôt des données mixtes pour comprendre l'impact de la rétroaction de l'IAg sur les apprentissages. Dans ces études, les perceptions étudiantes à l'égard de l'IAg (données qualitatives) sont recueillies par questionnaire (Yu) ou entretiens (Zheng *et al.*). Ces études montrent des résultats globalement positifs de l'usage des IAg, bien que perçus uniquement par les étudiants et étudiantes à performance élevée dans l'étude de Yu. La deuxième étude d'Escalante *et al.* (2023) utilise également des questionnaires pour recueillir de telles perceptions, qui les amènent à conclure que les étudiants et étudiantes n'ont pas de préférence quant au type de rétroaction (par l'IAg ou la personne enseignante). Enfin, Kim *et al.* (2022) sont les seuls à utiliser une méthodologie qualitative en analysant le contenu de productions écrites des personnes participantes pour comprendre l'influence de la rétroaction de l'IAg. En utilisant une grille d'évaluation, ils ont constaté que la qualité des arguments était améliorée après l'usage de l'IAg.

4.2 Les caractéristiques des rétroactions fournies par les IAg

Les principales caractéristiques des rétroactions fournies par les IAg sont présentées au tableau A.2 (annexe A). Un premier élément qui s'en dégage est la variabilité dans les études concernant leur description du processus de conception du système d'IAg fournissant de la rétroaction, alors que cette donnée est très importante. Sept études indiquent avoir conçu une IAg pour la recherche en présentant les modèles, algorithmes ou logiciels utilisés pour ce faire, mais seulement trois d'entre elles précisent comment l'IAg a été configurée pour générer de la rétroaction aux étudiants et étudiantes. Cinq études ont plutôt utilisé une IAg déjà existante : GPT-4 (Escalante *et al.*, 2023), SARA (Mousavi *et al.*, 2021), MetaTutor (Taub et Azevedo, 2019), Grammarly (Nazari *et al.*, 2021; Yu, 2023), Packback (Yu, 2023) et SafeAssign (Yu, 2023).

Considérant la diversité des IAg mobilisées pour ces études, il a semblé pertinent de s'interroger sur les justifications apportées par les auteurs et autrices pour expliquer le choix de leur IAg. Ces

justifications sont variées et portent sur des éléments distincts. Taub et Azevedo (2019) se distinguent en justifiant leur choix de MetaTutor par sa conception précisément axée sur le soutien à l'apprentissage autorégulé où la rétroaction joue un rôle essentiel. Quatre études justifient leur choix en expliquant que le modèle employé est supérieur à d'autres, soit parce qu'il est plus juste (Chiu *et al.*, 2024), plus précis (Kim *et al.*, 2022; Zheng *et al.*, 2024) ou plus performant (Sailer *et al.*, 2023). Deux études appuient leur choix par leur intention de comparer la performance d'une IAg particulière à d'autres qui offrent des fonctionnalités équivalentes (Escalante *et al.*, 2023; Mousavi *et al.*, 2021). Yildiz Durak (2023) soutient la pertinence de Flow XO en raison de son interface simple et de la possibilité d'intégrer l'agent à Slack. Cependant, quatre études (Darvishi *et al.*, 2022; Essel *et al.*, 2022; Nazari *et al.*, 2021; Yu, 2023) présentent une justification plutôt implicite en montrant que les fonctionnalités de l'algorithme, du modèle, du logiciel ou de l'IAg utilisée sont adéquates pour répondre aux besoins de la recherche.

Par ailleurs, une seule étude (Chiu *et al.*, 2024) présente explicitement la démarche employée pour tester la fiabilité de la rétroaction fournie par l'IAg. Seulement trois études exposent des résultats où les étudiants et étudiantes indiquent avoir reçu de la rétroaction non pertinente de l'IAg (Essel *et al.*, 2022; Kim *et al.*, 2022) ou pas assez précise (Zheng *et al.*, 2024). Néanmoins, aucune étude ne fait mention d'une procédure à suivre par l'étudiant ou l'étudiante recevant une rétroaction qui ne lui semble pas adéquate, pertinente ou juste.

Par opposition, soulignons la présentation hautement détaillée de l'article de Sailer *et al.* (2023), dont la section de la méthodologie justifie le choix du modèle et détaille le processus de préentraînement en précisant que 118 productions étudiantes ont été analysées par l'équipe enseignante et que cela a permis de distinguer 40 cas sur lesquels reposait l'analyse de l'IAg. Cette étude précise également comment l'IAg a été intégrée à un logiciel de simulation déjà existant. La description de la séquence de l'activité pédagogique clarifie comment les étudiants et étudiantes recevaient la rétroaction, tant dans le groupe expérimental que dans le groupe témoin.

Dans la majorité des études ($n = 9$), la rétroaction porte sur l'adéquation entre la production étudiante et les contenus disciplinaires visés. Le format de ces rétroactions est personnalisé en fonction de la production étudiante : l'IAg analyse le contenu de la production et génère une rétroaction unique. De plus, la rétroaction fournie dans l'étude de Sailer *et al.* (2023) est également adaptative, c'est-à-dire que l'IAg évalue la progression de l'étudiant ou l'étudiante en comparant sa performance actuelle avec les précédentes pour produire la rétroaction.

La rétroaction est le plus souvent promulguée pendant l'activité d'apprentissage ($n = 6$) : l'étudiant ou l'étudiante réalise une première production, puis reçoit la rétroaction de l'IAg pour ensuite l'améliorer. Dans trois études, la rétroaction est plutôt fournie immédiatement après la fin de l'activité d'apprentissage, afin que cela puisse aider l'étudiant ou l'étudiante à s'améliorer lors d'activités ultérieures. Enfin, la rétroaction dans la recherche d'Escalante *et al.* (2023) était fournie à un moment prédéterminé chaque semaine. Il est à noter que dans leur étude 2, les étudiantes et étudiants recevaient deux rétroactions hebdomadaires (tuteur ou tutrice et IAg) indépendantes, afin de pouvoir les comparer et exprimer leurs préférences de rétroaction.

Deux autres objets de rétroaction fournie par l'IAg sont observés dans ces études, soit la qualité de la rétroaction fournie par l'étudiant ou l'étudiante en considérant des indicateurs d'une rétroaction constructive (Darvishi *et al.*, 2022) et des données d'apprentissage variées en tant que prédicteurs de la performance (Mousavi *et al.*, 2021). Enfin, Essel *et al.* (2022) ont conçu un agent qui répondait (rétroaction) aux questions de l'étudiant ou l'étudiante. Pour cette étude, la rétroaction était standardisée, car le texte produit par l'agent avait majoritairement été rédigé lors de sa conception.

5. Discussion

La comparaison des processus de conception et des caractéristiques des rétroactions de ces études amène à formuler deux éléments de discussion pour orienter la recherche future.

5.1 Des processus de conception des IAg à expliciter et à justifier davantage

La diversité des IAg utilisées souligne les besoins de mieux justifier les choix effectués et d'expliquer le processus de conception, incluant les phases de conception et de configuration et la séquence pédagogique. En effet, plusieurs études nomment le modèle utilisé (p. ex., ChatGPT, SBERT, etc.), sans nécessairement présenter les avantages ou les limites de leur choix. Or, cet accent unique sur la performance soulève des questionnements sur les plans éthique et juridique, notamment, qui ne sont pas mentionnés dans ces études. S'il est généralement reconnu que les modèles de ChatGPT sont les plus performants, ce sont aussi les modèles qui ont été entraînés sur les plus larges ensembles de données et qui contiennent le plus de paramètres. L'empreinte carbone de leur entraînement et de leur utilisation est donc la plus élevée. Par ailleurs, ce modèle réside aux États-Unis, et les législations en place au Québec et en Europe exigent des établissements d'enseignement supérieur que les données étudiantes restent dans le territoire national. Enfin, selon le classement établi par Hugging Face (2024), d'autres modèles sont presque aussi performants, mais obtiennent une cote éthique moyenne plus élevée.

Par ailleurs, les préoccupations liées aux confabulations de ces systèmes, par exemple les répercussions sur l'apprentissage liées à une rétroaction erronée, sont plutôt rares dans les études recensées. Aussi, il n'y a pas systématiquement de mention qui vise à justifier leur modèle plutôt qu'un autre équivalent, et la réflexion sur le choix du modèle semble, au mieux, limitée à la performance. Plusieurs études présentent leur méthodologie très brièvement, notamment la phase de conception de l'application dotée d'IAg. On pourrait dire qu'elles manquent de transparence. Conséquemment, il est parfois ardu de comprendre le processus par lequel la production étudiante est analysée par l'IAg et les paramètres employés par celle-ci pour fournir une rétroaction personnalisée. Cette difficulté est aussi parfois attribuable au manque d'information quant à l'interface utilisée par les étudiants et étudiantes pour réaliser l'activité d'apprentissage. Le dévoilement du mécanisme utilisé pour donner de la rétroaction aux personnes utilisatrices elles-mêmes pourrait représenter un avantage considérable.

Ces éléments amènent alors à recommander que la place accordée à la présentation de la démarche de conception dans ce type d'étude soit augmentée pour y faire ressortir des éléments essentiels à une compréhension claire des actions et des choix qui ont été effectués. Dans la foulée des nombreuses préoccupations éthiques et légales relatives à l'utilisation de l'IAg, ces aspects devraient aussi idéalement être traités dans la section de l'article portant sur la méthodologie. Cela permettrait ainsi à la personne lectrice de se représenter clairement et en détail la rétroaction qui a été fournie, incluant les caractéristiques de l'IAg et son processus de conception, d'affinage et parfois même d'entraînement pour formuler une rétroaction. De surcroît, cela amènerait à mieux justifier ou argumenter les choix effectués pour la recherche : la pertinence de l'IAg choisie pour fournir de la rétroaction dans ce contexte précis d'apprentissage serait mieux détaillée. Enfin, cela permettrait de mieux apprécier la rigueur scientifique de la recherche, notamment sur le plan de l'applicabilité, soit la validité externe (méthodologies quantitatives) ou la transférabilité (méthodologies qualitatives) (Fortin et Gagnon, 2022). Autrement dit, une description détaillée et argumentée de la méthodologie de conception et d'utilisation de l'IAg faciliterait son appropriation dans les milieux de pratique, car il serait possible de comprendre dans quels contextes

pédagogiques (domaines d'études, profils étudiants, tâches et activités d'apprentissage, etc.) la rétroaction fournie par cette IAg est bénéfique pour soutenir les apprentissages. Les chercheurs et chercheuses du domaine devraient aussi se pencher sur les éléments éthiques à considérer dans le processus de conception. En somme, considérant le contexte émergent et évolutif de l'utilisation de l'IAg pour fournir de la rétroaction personnalisée aux étudiants et étudiantes, il semble essentiel de fournir une description complète du processus de conception de l'IAg pour faciliter la compréhension de l'étude, justifier sa pertinence et établir la portée des résultats obtenus.

5.2 Des bénéfices pédagogiques à utiliser l'IAg pour fournir de la rétroaction à peaufiner

Une limite observée dans plusieurs de ces études concerne la justification pédagogique permettant le choix d'une IAg plutôt que d'une autre. Bien que la plupart d'entre elles aient exposé la pertinence pédagogique de la rétroaction personnalisée parce qu'elle permet à l'étudiant ou l'étudiante de se situer, de s'améliorer ou de s'autoréguler, peu ont mobilisé une théorie de l'apprentissage pour justifier d'une part le recours à l'IAg pour offrir de la rétroaction et d'autre part le choix de l'IAg employée. Toutefois, certaines études dont le principal objectif est de comparer la performance d'étudiants et étudiantes à une évaluation selon le type de rétroaction reçue soulignent la pertinence d'une rétroaction personnalisée comparativement à une rétroaction standardisée.

Néanmoins, aucune de ces études ne mobilise un cadre théorique clair pour appuyer les choix à effectuer et la manière de dialoguer avec une IAg. Pour la recherche future, il serait intéressant d'explorer le potentiel de ces outils pour simuler un dialogue entre une personne tutrice et une personne étudiante. Bien qu'il puisse y en avoir d'autres, la théorie de l'échafaudage de Bruner (1983) nous semble particulièrement appropriée pour appuyer la pertinence de dialoguer avec une IAg tout en donnant des orientations assez précises quant à la configuration de ces interactions. Autrement dit, ces études explorent peu ou pas les transformations dans le processus d'apprentissage qui peuvent être induites par l'usage des IAg pour fournir de la rétroaction personnalisée, voire adaptative. Cela est potentiellement dû aux méthodologies employées : la plupart constituent des études quantitatives visant à comparer la performance à une tâche évaluative d'un groupe expérimental à celle d'un groupe témoin. La recherche future pourrait alors s'intéresser à mieux comprendre le rôle ou l'influence des IAg durant le processus d'apprentissage à partir d'études qualitatives, ce qui amènerait à montrer avec des données empiriques à l'appui comment elles peuvent générer des retombées positives sur l'apprentissage. De fait, il serait pertinent d'approfondir les facteurs, conditions et caractéristiques de l'IAg ou des personnes étudiantes qui favorisent des retombées positives de la rétroaction par l'IAg sur l'apprentissage comparativement à d'autres formes de rétroaction, notamment en s'appuyant sur des cadres théoriques qui réfèrent à la rétroaction dans le processus d'apprentissage. Aucune de ces études n'explicite un cadre conceptuel ou des hypothèses de conception traditionnellement utilisés pour créer des tuteurs adaptatifs afin d'expliquer comment les rétroactions peuvent être adaptées à l'aide de l'IAg, alors qu'il s'agit pour nous d'une des pistes les plus prometteuses.

Conclusion

Bien que la recherche sur les systèmes de tutoriels intelligents date déjà d'une trentaine d'années, l'IAg pourrait amener une transformation des manières de concevoir ces systèmes et promet des pistes de développement rapide. Les méthodes traditionnelles de développement des tuteurs intelligents reposent sur une démarche précise et bien documentée, mais qui pourrait être revue en

fonction des contributions potentielles de l'IAg : développement d'une ontologie des connaissances du domaine, d'un modèle de la personne apprenante, des interactions entre la personne apprenante et le domaine de connaissances, etc.

Par exemple, l'utilisation de la génération augmentée par récupération (*retrieval-augmented generation* [RAG]), qui consiste essentiellement à vectoriser une base documentaire et à faire passer les requêtes et les réponses par cette base vectorisée, pourrait possiblement remplacer la longue et complexe étape de préparation d'une ontologie du domaine. Des techniques plus affinées pourraient aussi être mises à contribution pour améliorer la fiabilité de la rétroaction fournie ou s'assurer d'un style d'écriture adéquat et compréhensible par l'étudiant ou l'étudiante (p. ex., OpenAI, s.d.). La réflexion sur les méthodes permettant aux systèmes fondés sur les IAg de remplacer les modèles étudiants demeure aussi à faire.

Ainsi, cette étude visait à effectuer un examen de la portée, afin de mieux comprendre l'état actuel de la recherche sur l'usage d'IAg pour offrir de la rétroaction aux étudiants et étudiantes universitaires. Plus précisément, elle visait à décrire les caractéristiques de ces études en mettant l'accent sur la démarche méthodologique (conception de l'IAg, objectifs poursuivis, méthodologie) et à distinguer les rétroactions fournies par ces IAg selon différents critères (objet, moment et format de la rétroaction). Parce qu'elle repose uniquement sur douze études et considérant le rythme rapide de publication d'écrits scientifiques dans le domaine, cette recherche n'a pas pu analyser d'autres études sans doute menées sur le sujet.

Néanmoins, les résultats présentés soutiennent deux éléments de discussion. D'abord, il a été argumenté qu'il serait pertinent que les études dans ce champ documentent mieux leur démarche méthodologique pour faciliter la compréhension de l'étude et pour en apprécier la rigueur scientifique, en invitant les autrices et auteurs à rendre plus explicites les raisons de leurs choix d'IAg et de grands modèles de langage, en appuyant aussi leur réflexion sur des critères éthiques et en étant plus transparents sur les processus de configuration des outils développés. Il serait nécessaire d'explicitier les critères de conception et de les lier à un cadre conceptuel montrant comment la rétroaction fournie par l'IAg peut soutenir l'apprentissage. Ensuite, il a été suggéré de mieux situer la pertinence de la rétroaction fournie par une IAg en menant des recherches dont l'intention serait de mieux comprendre comment elle influence le processus d'apprentissage, en appuyant ces recherches sur des théories de l'apprentissage ou de la rétroaction bien situées. En prenant la théorie de l'étayage à titre d'exemple, la perception du dialogue de l'IAg pourrait être étudiée : Est-il pertinent, adéquat, suffisant, etc.? D'autres caractéristiques de ces rétroactions pourraient également faire l'objet de recherches futures. Enfin, il pourrait être pertinent d'étudier les liens entre la rétroaction enseignante et celle de l'IAg.

Par ailleurs, une des limites observées par cette recension est l'absence dans ces recherches d'étude des limites de ces IAg. Par exemple, considérant les limites des modèles de traitement de langage, comment cela peut-il influencer la qualité de la rétroaction de l'IAg? De quelle manière les hallucinations ou les erreurs d'interprétation de l'IAg sont-elles considérées pour éviter que cela nuise à l'apprentissage? Qu'est-ce que ces IAg favorisent ou empêchent lorsqu'elles sont mobilisées pour fournir de la rétroaction? Quelle est l'attention donnée à des critères légaux ou éthiques comparativement aux critères de performance? Quelle complémentarité cette rétroaction peut-elle offrir avec les autres formes de rétroaction et de tutorat? En somme, si la recherche actuelle vise à montrer le potentiel des IAg pour offrir de la rétroaction, le besoin d'approfondir les conditions, contextes et caractéristiques de ces IAg qui justifient leur pertinence comparativement à d'autres formes de rétroaction demeure.

Notes

Usage d'outils d'intelligence artificielle

Le résumé a été traduit en anglais à l'aide de ChatGPT. La traduction a été ensuite révisée par l'autrice principale, puis par un traducteur professionnel.

Disponibilité des données

La liste de tous les articles recensés est disponible sur demande par courriel à l'autrice principale, **Chantal Tremblay**.

Références

Note. Les astérisques devant les notices désignent les douze articles de la recension.

Arksey, H. et O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, 8(1), 19-32.

<https://doi.org/bqnqnb>

Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakci, Ö. et Mariman, R. (2024). *Generative AI can harm learning* [rapport de recherche]. Wharton School, Université de Pennsylvanie.

<https://doi.org/10.2139/SSRN.4895486>

Belland, B. R., Walker, A. E., Kim, N. J. et Lefler, M. (2017). Synthesizing results from empirical research on computer-based scaffolding in STEM education: A meta-analysis. *Review of Educational Research*, 87(2), 309-344. <https://doi.org/ggdpcet>

Boissière, J. et Bruillard, É. (2021). *L'école digitale : une éducation à construire et à vivre. Numérique et transformations de l'école*. Armand Colin.

Boucher, C. (2015). La rétroaction : élément phare pour l'apprentissage de nos étudiants. *Le tableau*, 24(3). [https://pedagogie.quebec.ca/...](https://pedagogie.quebec.ca/)

Bruner, J. S. (1983). *Le développement de l'enfant. Savoir faire, savoir dire* (J. Michel et M. Deleau, trad.). Presses universitaires de France.

*Chiu, M., Hwang, G., Hsia, L. et Shyu, F. (2024). Artificial intelligence-supported art education: A deep learning-based system for promoting university students' artwork appreciation and painting outcomes. *Interactive Learning Environments*, 32(3), 824-842.

<https://doi.org/pk6h>

Conseil de l'innovation du Québec. (2024). *Prêts pour l'IA. Répondre au défi du développement et du déploiement responsables de l'IA au Québec*. [https://conseilinnovation.quebec/...](https://conseilinnovation.quebec.ca/)

*Darvishi, A., Khosravi, H., Sadiq, S. et Gašević, D. (2022). Incorporating AI and learning analytics to build trustworthy peer assessment systems. *British Journal of Educational Technology*, 53(4), 844-875. <https://doi.org/10.1111/bjet.13233>

*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, article 57. <https://doi.org/g82g6b>

- *Essel, H. B., Vlachopoulos, D., Tachie-Menson, A., Johnson, E. E. et Baah, P. K. (2022). The impact of a virtual teaching assistant (Chatbot) on students' learning in Ghanaian higher education. *International Journal of Educational Technology in Higher Education*, 19, article 57. <https://doi.org/gtm4gx>
- Evans, C. (2013). Making sense of assessment feedback in higher education. *Review of Educational Research*, 83(1), 70-120. <https://doi.org/gf82tm>
- Fortin, M.-F. et Gagnon, J. (2022). *Fondements et étapes du processus de recherche : méthodes quantitatives et qualitatives* (4^e éd.). Chenelière Éducation.
- Grimshaw, J. (2010). *Guide sur la synthèse des connaissances*. Instituts de recherche en santé du Canada. <https://cihr-irsc.gc.ca/f/41382.html>
- Hattie, J. (2024). Foreword: Advancing assessment capabilities in higher education. Dans C. Evans et M. Waring (dir.), *Research handbook on innovations in assessment and feedback in higher education – Implications for teaching and learning* (p. xvii-xxvii). Edward Elgar Publishing.
- Henderson, M., Ryan, T. et Phillips, M. (2019). The challenges of feedback in higher education. *Assessment & Evaluation in Higher Education*, 44(8), 1237-1252. <https://doi.org/fj6z>
- Hugging Face. (2024). *Open LLM Leaderboard archived*. Récupéré le 28 octobre 2024 de <https://huggingface.co/...>
- Kim, J. Y. et Lim, K. Y. (2019). Promoting learning in online, ill-structured problem solving: The effects of scaffolding type and metacognition level. *Computers & Education*, 138, 116-129. <https://doi.org/10.1016/j.compedu.2019.05.001>
- *Kim, M. K., Kim, N. J. et Heidari, A. (2022). Learner experience in artificial intelligence-scaffolded argumentation. *Assessment & Evaluation in Higher Education*, 47(8), 1301-1316. <https://doi.org/pk6m>
- Kim, N. J., Belland, B. R. et Walker, A. E. (2018). Effectiveness of computer-based scaffolding in the context of problem-based learning for STEM education: Bayesian meta-analysis. *Educational Psychology Review*, 30(2), 397-429. <https://doi.org/ghbvm>
- Lepage, A. (2024). *Étude de l'adoption des principaux types d'usages de l'intelligence artificielle par les enseignants et enseignantes du postsecondaire* [thèse de doctorat, Université de Montréal, Canada]. Papyrus. <http://hdl.handle.net/1866/40548>
- *Mousavi, A., Schmidt, M., Squires, V. et Wilson, K. (2021). Assessing the effectiveness of Student Advice Recommender Agent (SARA): The case of automated personalized feedback. *International Journal of Artificial Intelligence in Education*, 31(3), 603-621. <https://doi.org/gqsc3t>
- *Nazari, N., Shabbir, M. et Setiawan, R. (2021). Application of artificial intelligence powered digital writing assistant in higher education: Randomized controlled trial. *Heliyon*, 7(5), article e07014. <https://doi.org/10.1016/j.heliyon.2021.e07014>
- OpenAI. (s. d.). *Fine-tuning*. OpenAI platform. Récupéré le 11 mars 2025 de <https://platform.openai.com/...>

- Pardos, Z. A. et Bhandari, S. (2024). ChatGPT-generated help produces learning gains equivalent to human tutor-authored help on mathematics skills. *PLOS One*, 19(5), article e0304013. <https://doi.org/10.1371/journal.pone.0304013>
- Reporters sans frontières. (2023, 5 décembre). *Pierres angulaires de l'industrie de l'IA, les modèles de fondation doivent être régulés pour protéger le droit à l'information*. <https://rsf.org/...>
- Sabiston, C. M., Vani, M., de Jonge, M. et Nesbitt, A. (2022). Scoping reviews and rapid reviews. *International Review of Sport and Exercise Psychology*, 15(1), 91-119. <https://doi.org/kbsw>
- *Sailer, M., Bauer, E., Hofmann, R., Kiesewetter, J., Glas, J., Gurevych, I. et Fischer, F. (2023). Adaptive feedback from artificial neural networks facilitates pre-service teachers' diagnostic reasoning in simulation-based learning. *Learning and Instruction*, 83, article 101620. <https://doi.org/10.1016/j.learninstruc.2022.101620>
- *Taub, M. et Azevedo, R. (2019). How does prior knowledge influence eye fixations and sequences of cognitive and metacognitive SRL processes during learning with an intelligent tutoring system? *International Journal of Artificial Intelligence in Education*, 29(1), 1-28. <https://doi.org/gpd5hf>
- Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y. et Gašević, D. (2024). Practical and ethical challenges of large language models in education: A systematic scoping review. *British Journal of Educational Technology*, 55(1), 90-112. <https://doi.org/10.1111/bjet.13370>
- *Yildiz Durak, H. (2023). Conversational agent-based guidance: Examining the effect of chatbot usage frequency and satisfaction on visual design self-efficacy, engagement, satisfaction, and learner autonomy. *Education and Information Technologies*, 28, 471-488. <https://doi.org/gtz238>
- *Yu, E. (2023). Intelligent enough? Artificial intelligence for online learners. *Journal of Educators Online*, 20(1). https://thejeo.com/archive/2023_20_1/you
- Zheng, L. (2016). The effectiveness of self-regulated learning scaffolds on academic performance in computer-based learning environments: A meta-analysis. *Asia Pacific Education Review*, 17(2), 187-202. <https://doi.org/f8pv53>
- *Zheng, L., Fan, Y., Chen, B., Huang, Z. et Long, M. (2024). An AI-enabled feedback-feedforward approach to promoting online collaborative learning. *Education and Information Technologies*, 29, 11385-11406. <https://doi.org/pk6p>

Annexe A – Caractéristiques des douze articles de la recension

Le tableau A.1 présente les caractéristiques de chaque étude du corpus. Il expose la discipline et le cycle d'études du cours dans lequel l'étude a eu lieu. Ensuite, le principal objectif, les éléments méthodologiques et le principal résultat de l'étude sont décrits.

Tableau A.1

Caractéristiques des études du corpus

Article	Discipline	Cycle	Principal objectif	Méthodologie Nombre de participant(e)s, type de données, instruments	Principal résultat
Chiu <i>et al.</i> (2024)	Arts	1 ^{er} cycle	Comparer l'effet sur l'apprentissage de la rétroaction de l'IAg à celui d'une rétroaction de l'enseignant(e)	$n = 46$ (gr. expér. : $n = 24$) Données quantitatives Pré- et post-test	La performance (apprentissages) du groupe expérimental est significativement plus élevée que celle du groupe témoin.
Darvishi <i>et al.</i> (2022)	Sciences cognitives et technologies de l'information	1 ^{er} cycle	Comparer l'effet sur la qualité de la rétroaction étudiante de la rétroaction de l'IAg à celui d'une situation sans rétroaction fournie par l'IAg	$n = 374$ (gr. expér. : $n = 187$) Données quantitatives Traces informatiques (environnement numérique d'apprentissage)	La rétroaction fournie par le groupe expérimental est plus longue et plus appréciée par les pairs que celle du groupe témoin.
Escalante <i>et al.</i> (2023)	Anglais langue seconde	Non applicable	Étude 1. Comparer l'effet sur la performance de la rétroaction de l'IAg à celui d'une rétroaction fournie par un(e) tuteur(-trice) Étude 2. Comparer la préférence des étudiant(e)s entre la rétroaction de l'IAg et celle de l'enseignant(e)	$n = 48$ (gr. expér. : $n = 23$) Données quantitatives Pré- et post-tests (tests hebdomadaires et test final) $n = 43$ Données mixtes Questionnaire (items portant sur préférences, satisfactions, utilisabilité)	Aucun effet sur la performance associé au type de rétroaction reçue. Aucune préférence statistiquement significative (appréciation semblable des deux types).
Essel <i>et al.</i> (2022)	Multimédia	1 ^{er} cycle	Comparer l'effet sur la performance de la rétroaction de l'IAg à celui d'une rétroaction de l'enseignant(e) synchrone	$n = 68$ (gr. expér. : $n = 34$) Données quantitatives Pré- et post-tests	La performance du groupe expérimental est significativement plus élevée que celle du groupe témoin.
Kim <i>et al.</i> (2022)	Design pédagogique et technologies éducatives	2 ^e cycle	Expliquer dans quelle mesure la rétroaction de l'IAg influence la qualité des arguments rédigés par les étudiant(e)s	$n = 14$ Données qualitatives Productions écrites (pré- et post-intervention) et autoréflexions	La qualité des arguments a été améliorée après l'usage de l'IAg.
Mousavi <i>et al.</i> (2021)	Biologie	1 ^{er} cycle	Comparer l'effet sur la note finale du cours de la rétroaction de l'IAg à celui d'une rétroaction générique prédéterminée	$n = 1\,026$ (gr. expér. : $n = 499$) Données quantitatives Note finale et variables modératrices	La note finale du groupe expérimental est plus élevée que celle du groupe témoin (taille d'effet faible).

Article	Discipline	Cycle	Principal objectif	Méthodologie Nombre de participant(e)s, type de données, instruments	Principal résultat
Nazari <i>et al.</i> (2021)	Anglais langue seconde	3 ^e cycle	Comparer l'effet sur l'engagement de la rétroaction de l'IAg à celui d'une situation sans rétroaction de l'IAg	$n = 93$ (gr. expér. : $n = 51$) Données quantitatives Questionnaires (échelles d'engagement, d'autoefficacité et d'émotions)	L'engagement cognitif et émotionnel du groupe expérimental est significativement plus élevé que celui du groupe témoin.
Sailer <i>et al.</i> (2023)	Éducation	1 ^{er} cycle	Comparer l'effet sur la performance de la rétroaction de l'IAg à celui d'une rétroaction générique prédéterminée	$n = 178$ (gr. expér. 2 x 2 : $n = 57$ et $n = 30$) Données quantitatives Pré-test et résultats (notes) des productions écrites post-intervention	La performance (précision du diagnostic et justifications) du groupe expérimental est significativement plus élevée que celle du groupe témoin.
Taub et Azevedo (2019)	Biologie	1 ^{er} cycle	Comparer l'effet sur l'apprentissage et l'engagement de l'utilisation de MetaTutor à celui d'une situation sans tuteur(-trice)	$n = 194$ (gr. expér. : $n = 95$) Données quantitatives Pré- et post-test	Le gain d'apprentissage du groupe expérimental est significativement plus élevé que celui du groupe témoin. Pas d'effet sur l'engagement.
Yildiz Durak (2023)	Design graphique	Non précisé	Examiner les effets de la rétroaction fournie par l'IAg sur l'apprentissage, l'autoefficacité, l'engagement, la satisfaction et l'autonomie	$n = 86$ Données quantitatives Questionnaire (échelles de satisfaction, autonomie, engagement, sentiment d'autoefficacité)	Relation positive entre l'usage de l'IAg et l'autoefficacité. Aucune relation significative entre l'usage de l'IAg et les autres variables.
Yu (2023)	Communication	1 ^{er} cycle	Comprendre l'impact de la rétroaction de l'IAg sur les pratiques d'écriture	$n = 68$ Données mixtes Résultats (notes) de productions écrites Questionnaire (échelles d'attitudes et de style d'apprentissage, questions ouvertes sur la rétroaction de l'IAg)	Les étudiant(e)s hautement performant(e)s perçoivent que l'IAg améliore leurs pratiques d'écriture, ce qui est moins perceptible chez les moins performant(e)s.
Zheng <i>et al.</i> (2024)	Non applicable	1 ^{er} cycle	Comprendre l'impact de la rétroaction de l'IAg sur la compréhension des contenus et l'expérience de collaboration	$n = 153$ (2 gr. expér. : $n = 17$ et $n = 17$) Données mixtes Pré-test, traces informatiques (environnement numérique d'apprentissage) et entretiens	La performance (compréhension des contenus) et l'expérience de collaboration du groupe expérimental sont significativement plus élevées que celles des groupes témoins.

Le tableau A.2 montre les caractéristiques des rétroactions fournies par les IAg pour chacune de ces études. Il présente le système d'IAg utilisé, les arguments formulés par les chercheurs pour justifier le système d'IAg employé, le ou les objets de la rétroaction, ainsi que le moment où elle était fournie aux étudiants et son format.

Tableau A.2*Caractéristiques des rétroactions fournies par les IAg*

Article	IAg : utilisation, conception (entraînement, ingénierie, etc.)	Justification du choix de l'IAg	Objet(s) de la rétroaction fournie par l'IAg	Moment	Format
Chiu <i>et al.</i> (2024)	Conception de l'IAg avec ResNet50, avec phase d'entraînement pour 6 groupes d'arts sur 323 images et des images simulées à partir des premières, puis intégration à une application numérique de création artistique	Oui ResNet50 choisi pour la pertinence d'une architecture dotée d'un réseau neuronal convolutif et la justesse de son analyse des images	Caractéristiques de l'œuvre multimédia produite (couleur, espace, thème, taille)	Pendant l'activité	Personnalisée
Darvishi <i>et al.</i> (2022)	Conception de l'IAg avec <i>Google BiLingual Evaluation Understudy</i> (GLEU) et <i>Sentence Bidirectional Encoder Representations from Transformers</i> (SBERT), puis intégration à l'outil numérique RiPPLE	Implicite Explication des fonctions de GLEU et SBERT	Qualité de la rétroaction rédigée par l'étudiant(e)	Pendant l'activité	Personnalisée
Escalante <i>et al.</i> (2023)	Utilisation de GPT-4 (OpenAI) avec rédactique pour préciser le contexte, le rôle de l'IAg et les indicateurs à évaluer	Oui GPT-4 choisi pour comparer la performance d'une IAg générale à celle d'une IAg spécifique (p. ex., Grammarly)	Six indicateurs de la qualité de la langue d'une production écrite	Prédéterminée (hebdomadaire)	Personnalisée
Essel <i>et al.</i> (2022)	Conception de KNUSTbot (l'IAg) avec Flow XO, puis intégration à l'application WhatsApp	Implicite Explication des fonctions de Flow XO	Toute question formulée par l'étudiant(e) concernant les contenus de programmation informatique	Pendant l'activité	Standardisée
Kim <i>et al.</i> (2022)	Conception du système d'échafaudage appuyé par l'IA (AISS) à l'aide d'un algorithme conçu avec un modèle de traitement automatique des langues préentraîné avec 1,5 milliard de paramètres	Oui Apprentissage automatique permet une meilleure précision par sa capacité d'analyser plus de paramètres	Qualité des arguments (cohérence, pertinence, contenu, argumentaire) d'un court texte argumentatif	Pendant l'activité	Personnalisée
Mousavi <i>et al.</i> (2021)	Utilisation de l'assistant virtuel SARA, basé sur l'analytique des apprentissages : analyse de données d'apprentissage pour fournir la rétroaction dans un environnement numérique d'apprentissage	Oui SARA choisi pour évaluer sa performance, comparativement à d'autres IAg semblables	Données d'apprentissage variées (sociodémographiques, caractéristiques individuelles, activités réalisées)	Prédéterminée (hebdomadaire)	Personnalisée

Article	IAg : utilisation, conception (entraînement, ingénierie, etc.)	Justification du choix de l'IAg	Objet(s) de la rétroaction fournie par l'IAg	Moment	Format
Nazari <i>et al.</i> (2021)	Utilisation de la version premium de Grammarly	Implicite Explication des fonctions de Grammarly	Indicateurs de la qualité de l'écriture d'une production écrite	Pendant l'activité	Personnalisée
Sailer <i>et al.</i> (2023)	Conception de l'IAg avec Neural Web et INCEpTION Phase d'entraînement : 118 solutions codées 40 cas de rétroaction prédéfinis	Oui Neural Web choisi pour sa performance dans l'évaluation du raisonnement diagnostique	Diagnostic adéquat et pertinence des justifications rédigés dans une courte production écrite	Immédiatement après l'activité	Adaptative
Taub et Azevedo (2019)	Utilisation de Metatutor, qui a une fonctionnalité de rétroaction, entre autres	Oui Metatutor choisi car conçu spécifiquement pour soutenir l'apprentissage autorégulé	Cohérence avec la théorie d'une courte production écrite	Immédiatement après l'activité	Personnalisée
Yildiz Durak (2023)	Conception de l'IAg avec Flow XO, puis intégration à la plateforme Slack	Oui Flow XO choisi pour son interface simple et son intégration possible à Slack (facilité d'utilisation)	Toute question formulée par l'étudiant(e) concernant les contenus de design graphique	Pendant l'activité	Standardisée
Yu (2023)	Utilisation de la version premium de Grammarly, de Packback ou de SafeAssign, au choix de l'étudiant(e)	Implicite Explication des fonctions des trois IAg	Indicateurs de la qualité de l'écriture d'une production écrite	Pendant l'activité	Personnalisée
Zheng <i>et al.</i> (2024)	Conception de l'IAg avec BERT-BiLSTM-CFR	Oui BERT-BiLSTM-CFR choisi pour sa meilleure précision comparativement à d'autres algorithmes semblables	Contenu de la production, indicateurs d'efficacité de la collaboration	Pendant l'activité	Personnalisée