

7ème colloque International du RAIFFET (Réseau Africain des Institutions de Formation de Formateurs de l'Enseignement Technique)
Éducation technologique, formation professionnelle et l'éducation pour une société durable
IPNETP d'ABIDJAN (CÔTE D'IVOIRE) 5 au 8 mai 2025
Axe 3 : Technologies de l'information et de la communication pour une société durable

Evaluation des apprentissages à l'ère de l'Intelligence Artificielle

Yadoho Noël COULIBALY,
Université Virtuelle de Côte d'Ivoire, GIRPE
yadoho.coulibaly@uvci.edu.ci

Abou FOFANA,
Ecole Normale Supérieure d'Abidjan
abou_fof@yahoo.fr

Résumé

L'évaluation a toujours occupé une place centrale dans les dispositifs d'apprentissage. Elle est désormais intégrée au processus d'enseignement-apprentissage-évaluation, démontrant qu'enseigner efficacement sans évaluation est presque impossible (Gérard, 2013). Selon Tardif (1993), citant Gardner (1992), l'évaluation consiste à collecter des informations sur les connaissances et compétences d'une personne, à lui fournir un retour significatif et à partager ces données avec la communauté.

L'essor de l'Intelligence Artificielle Générative (IAG) transforme les pratiques d'évaluation. Elle permet la génération massive de questions (QCM, QCU, textes à trous, etc.) et l'automatisation des feedbacks, répondant ainsi aux défis de la massification dans l'enseignement supérieur (Kamoun et al., 2025). Cependant, son utilisation dans l'évaluation académique suscite des préoccupations (Walbert & Polcheira, 2024). Une étude menée en Côte d'Ivoire (Coulibaly et Fofana, 2025) révèle que si les enseignants utilisent l'IAG pour concevoir des évaluations, les étudiants y ont également recours pour traiter ces évaluations. Cela soulève une question : les productions académiques des étudiants utilisant l'IAG suggèrent-elles réellement leurs acquis ?

Face à ces enjeux, certains enseignants voient l'IAG comme un assistant pédagogique, tandis que d'autres la perçoivent comme un substitut risqué (Walbert & Polcheira, 2024). Cette recherche explore les pratiques visant à encadrer l'usage de l'IAG dans l'évaluation afin de garantir que les résultats traduisent fidèlement le niveau de maîtrise des connaissances et compétences enseignées.

S'appuyant sur la théorie de l'évaluation formative de Black et Wiliam (2009), cette étude repose sur des entretiens semi-dirigés auprès d'enseignants d'universités ivoiriennes et une analyse qualitative. Les résultats montrent une priorisation des évaluations par projet ou exposé, ancrées dans un contexte local, au détriment des simples questions fermées ou ouvertes. Il apparaît également essentiel que les enseignants soient formés à l'usage de l'IAG pour concevoir des évaluations adaptées, rejoignant ainsi la position de Kamoun et al. (2025).

Mots clés : Intelligence Artificielle Générative ; Evaluation des apprentissages ; Technologies de l'Information et de la Communication ;

Abstract

Assessment has always occupied a central place in learning systems. It is now integrated into the teaching-learning-assessment process, demonstrating that teaching effectively without assessment is almost impossible (Gérard, 2013). According to Tardif (1993), citing Gardner (1992), assessment consists of collecting information about an individual's knowledge and skills, providing meaningful feedback, and sharing this data with the community.

The rise of Generative Artificial Intelligence (GAI) is transforming assessment practices. It enables the massive generation of questions (MCQs, MCQs, fill-in-the-blank texts, etc.) and the automation of feedback, thus addressing the challenges of massification in higher education (Kamoun et al., 2025). However, its use in academic assessment raises concerns (Walbert & Polcheira, 2024). A study conducted in Côte d'Ivoire (Coulibaly and Fofana, 2025) reveals that while teachers use GAI to design assessments, students also use it to process these assessments. This raises a question: do the academic productions of students using GAI truly suggest their prior learning?

Faced with these challenges, some teachers see AGI as a teaching assistant, while others perceive it as a risky substitute (Walbert & Polcheira, 2024). This research explores practices aimed at framing the use of AGI in assessment in order to ensure that the results faithfully reflect the level of mastery of the knowledge and skills taught.

Drawing on Black and Wiliam's (2009) theory of formative assessment, this study is based on semi-structured interviews with Ivorian university teachers and a qualitative analysis. The results show a prioritization of project-based or presentation-based assessments, anchored in a local context, to the detriment of simple closed or open questions. It also appears essential that teachers be trained in the use of AGI to design appropriate assessments, thus joining the position of Kamoun and al. (2025).

Keywords: Generative Artificial Intelligence; Learning Assessment; Information and Communication Technologies;

1. Introduction

Depuis la large diffusion de l'Intelligence Artificielle Générative (IAG), notamment ChatGPT, auprès du grand public en 2022 (Bordat-Teeuwen & Gonin, 2024 ; Livian, 2024), tous les secteurs d'activité, et en particulier celui de la formation, ont connu un bouleversement. La démocratisation massive de ChatGPT, un agent conversationnel basé sur l'intelligence artificielle, a transformé les usages des étudiants de l'enseignement supérieur ainsi que ceux de leurs enseignants à travers le monde.

Des universités de renom, telles que l'Université de Strasbourg et Sciences Po Paris, en ont fait les frais. En 2022, des étudiants de l'Université de Strasbourg ont été pris en flagrant délit d'utilisation de ChatGPT sans le mentionner et ont été contraints de repasser un examen. À la même période, Sciences Po Paris annonce son intention d'interdire l'utilisation de ChatGPT¹. En réalité, l'établissement interdisait non pas son utilisation en tant que telle, mais son usage dissimulé, c'est-à-dire sans transparence de la part des étudiants.

Parmi les nombreux usages de l'IAG dans le domaine de l'éducation, l'un d'eux semble poser problème : son utilisation dans l'évaluation, notamment de la part des étudiants. En effet, alors que les enseignants l'emploient pour concevoir des outils d'évaluation des travaux académiques, les étudiants, de leur côté, s'en servent également pour traiter ces évaluations (Coulibaly et Fofana, 2025). Ainsi, l'évaluation, qui est censée collecter des données sur les connaissances et compétences d'un apprenant (Tardif, 1993), mesure en réalité les capacités de l'outil d'intelligence artificielle à travers l'utilisateur qui l'exploite et non les capacités de son utilisateur. En d'autres termes, ce sont les réponses fournies par l'IA que l'apprenant copie et soumet à l'enseignant. Se pose alors la question de la crédibilité des évaluations, dans la mesure où elles ne permettent plus de vérifier le niveau réel de maîtrise des connaissances et des compétences de l'apprenant. Comment évaluer les acquis des apprenants à l'ère de l'intelligence artificielle générative ?

En nous appuyant sur la théorie de l'évaluation formative de Black et Wiliam (2009) – selon laquelle l'évaluation est un outil d'apprentissage plutôt qu'un simple moyen de mesure, mettant en avant la production de feedback et l'ajustement des contenus de formation – nous avons mené une étude exploratoire par le biais d'entretiens semi-dirigés auprès de plusieurs enseignants des universités de Côte d'Ivoire. L'objectif étant d'analyser leurs pratiques d'évaluation, en tenant compte du fait que les étudiants peuvent utiliser les outils d'IA pour traiter leurs examens.

Les résultats indiquent que les enseignants optent désormais pour des questions de type "texte à trous" ainsi que pour des projets de groupe présentés devant un jury, afin de limiter l'usage de l'IA dans les compositions et de mieux évaluer les compétences réelles des étudiants.

2. Contexte

2.1. L'IA et l'IAG

La première définition de l'intelligence artificielle a été énoncée en 1956. Elle stipule que l'intelligence artificielle est « la construction de programmes informatiques capables d'accomplir des tâches que, pour l'instant, les êtres humains réalisent de manière plus satisfaisante, car elles ont nécessité des processus mentaux de haut niveau tels que

¹ <https://www.letudiant.fr/educpros/enquetes/comment-adapter-les-evaluations-dans-le-superieur-pour-dejouer-chatgpt.html>

l'apprentissage perceptuel, l'organisation de la mémoire et le raisonnement critique » (Soulez, 2018).

L'IA vise ainsi à développer des systèmes informatiques capables de simuler certaines fonctions cognitives humaines, telles que la reconnaissance de la parole et des images, le raisonnement et la prise de décision (Haton, 2024). C'est dans ce contexte que Stefania Giannini, sous-directrice générale pour l'éducation de l'UNESCO, a affirmé que l'IA, s'appuyant sur une grande puissance de calcul, des réseaux neuronaux artificiels et de grands modèles de langage, parvient à imiter avec une habileté impressionnante le « pivot » de la civilisation humaine : le langage (Giannini, 2023).

L'Intelligence Artificielle Générative (IAG), quant à elle, se définit comme une technologie capable de générer automatiquement du contenu en réponse à des messages rédigés dans des interfaces conversationnelles en langage naturel (prompts). Le contenu produit peut prendre plusieurs formes : texte, images (photographies, peintures numériques, dessins animés), vidéos, musique ou encore code informatique (Organisation des Nations Unies pour l'éducation, la science et la culture [UNESCO], 2024, [Orientations pour l'intelligence artificielle générative dans l'éducation et la recherche](#)).

L'IAG fonctionne en analysant statistiquement la distribution des mots, des pixels ou d'autres éléments dans les données qu'elle a intégrées. Elle identifie et reproduit des modèles récurrents, par exemple en déterminant quels mots suivent, en règle générale, d'autres mots dans une séquence donnée.

2.2. Les capacités et les limites de l'IAG dans l'éducation

L'Intelligence Artificielle Générative (IAG) soulève de nombreuses préoccupations, bien qu'elle apporte également son lot d'avantages. Ces inquiétudes sont d'autant plus justifiées qu'elles concernent le domaine de l'éducation, un secteur essentiel pour l'avenir de nos sociétés.

En effet, cette technologie a une répercussion directe sur la société et influencera les pratiques éducatives pendant de nombreuses années à venir. Comme toute création humaine, l'IA présente à la fois des opportunités et des défis, tant pour les enseignants que pour les apprenants.

Selon l'Organisation des Nations Unies pour l'éducation, la science et la culture (UNESCO) (2024), ([Orientations pour l'intelligence artificielle générative dans l'éducation et la recherche](#)), l'IAG offre aux enseignants de nombreuses opportunités d'innovation en matière d'enseignement et permet d'améliorer les pratiques pédagogiques actuelles. Elle peut assister le personnel enseignant dans diverses tâches, telles que la génération d'idées (remue-méninge), la rédaction de courriels ou de textes, la reformulation de phrases ou encore la vulgarisation de concepts complexes.

L'IAG peut également accompagner les enseignants dans la préparation du matériel pédagogique, les aider à développer des compétences fondamentales à leur propre rythme et faciliter l'apprentissage par enquête ou par projet.

Par ailleurs, elle constitue un soutien précieux pour les apprenants ayant des besoins spécifiques. En effet, comme l'indiquent Bordat-Teeuwen & Gonin (2024), l'IAG est capable de répondre à des questions variées dans tous les domaines, de produire du contenu, de rédiger des écrits divers (*rapports, mémoires, lettres, QCM, etc.*), d'adapter et de traduire du contenu, de synthétiser et mettre en forme des informations, ainsi que de générer des illustrations et des plans, analyser des données, générer du code et résoudre des équations.

Toutefois, l'IAG présente certaines limites. Selon Bordat-Teeuwen & Gonin (2024). Les technologies d'IAG ne génèrent pas de savoir procédural, mais uniquement du savoir déclaratif, et ne sont pas capables de raisonner. De plus, elles disposent de données limitées pour les sujets très spécifiques, car leurs réponses reposent principalement sur les idées et les thématiques les plus courantes, voire dominantes, d'une discipline.

Par conséquent, les réponses fournies sont souvent génériques et ne tiennent pas toujours compte du contexte spécifique. De plus, l'IAG ne peut pas s'auto-corriger sans intervention humaine et ne garantit pas le respect du droit d'auteur.

2.3. IAG et évaluation des apprentissages

2.3.1. Evaluation académique

De manière générale, l'évaluation des apprentissages dans le domaine académique repose sur différents types de questions : questions à développement court ou long, questions vrai/faux, questions à choix multiples (QCM), questions d'association ou encore textes à trous (Lez et al, 2023). Qu'elle prend la forme d'un examen, d'un questionnaire, d'un test, d'un quiz ou d'un mini-quiz, l'évaluation est généralement soumise à des contraintes de temps et d'espace lors de sa passation. La littérature souligne que l'évaluation des apprentissages est un processus plus ou moins formel, structuré en trois étapes principales :

- a) Le recueil d'informations portant sur les apprentissages réalisés par l'étudiant, aussi bien pendant qu'à l'issue d'une séquence de formation.
- b) L'interprétation de ces informations, permettant d'établir un jugement sur les acquis de l'apprenant.
- c) La prise de décision, fondée sur ce jugement, pouvant aboutir à différentes actions selon le contexte (Tardif, 1993 ; Jouquan, 2002).

2.3.2. Pratiques d'évaluation dans les universités ivoiriennes

En Côte d'Ivoire, deux principales pratiques d'évaluation des apprentissages se dégagent. Ces pratiques sont directement liées aux modes d'enseignement adoptés par les établissements.

La première concerne les établissements dispensant des formations en présentiel, où les effectifs sont supposés être moins élevés que ceux des formations à distance. Dans ce contexte, les étudiants sont évalués en salle de classe, sur papier, sous la surveillance stricte d'examineurs veillant à empêcher toute communication entre candidats ou l'utilisation de ressources non autorisées. Ces établissements sont conçus pour faciliter ce type d'évaluation, avec des infrastructures adaptées.

La seconde pratique concerne les établissements fonctionnant à distance et en ligne. À ce jour, l'Université Virtuelle de Côte d'Ivoire (UVCI) est le seul établissement public en Côte d'Ivoire à proposer un enseignement à distance. Ces établissements accueillent généralement un grand nombre d'apprenants, une solution mise en place par les États pour répondre aux défis de la massification. Par exemple, l'UVCI est passé de 3 000 étudiants en 2017 à 17 920 en 2022 (Plumat, 2023). L'évaluation dans ces établissements se fait selon deux modalités :

- 1) La composition en ligne : les étudiants composent à distance, depuis le lieu de leur choix, à l'aide d'un terminal numérique (ordinateur ou smartphone) de leur choix.
- 2) La composition en présentiel : les étudiants se rendent dans des amphithéâtres et composent sur leurs propres terminaux connectés à Internet sous la surveillance d'examineurs.

Dans le cadre des évaluations en présentiel, l'accès aux ordinateurs et aux outils d'intelligence artificielle est strictement interdit. En revanche, dans les établissements fonctionnant en FOAD, l'utilisation de terminaux numériques connectés à Internet offre aux étudiants la possibilité d'accéder aux outils d'IA pendant leurs examens, soulevant ainsi de nouvelles questions quant à la fiabilité et à l'éthique des évaluations.

2.3.3. Evaluation avec l'IAG

L'intelligence artificielle générative (IAG) offre aujourd'hui aux enseignants de nombreuses possibilités pour concevoir tous les types d'évaluations, sans exception. Grâce à ces outils, la production massive de questions fermées est désormais possible, avec la possibilité d'indiquer le niveau de difficulté et les caractéristiques souhaitées par l'enseignant.

De manière générale, l'IAG permet de générer un grand nombre de questions à partir de diverses sources selon deux principales approches. La première possibilité est celle de l'utilisation de prompts via des outils tels que ChatGPT (Service universitaire de pédagogie – UBS, 22 nov. 2023). La seconde possibilité est celle des applications d'IA telles que Quiz Wizard et Wooclap (De Lièvre et al., 2023) qui disposent d'une interface utilisateur permettant à l'enseignant d'indiquer aux enseignants d'indiquer leurs sources et les types de questions à générer.

Dans ce contexte, il devient de plus en plus difficile d'ignorer l'influence des IAG dans la conception des modalités d'apprentissage et d'évaluation. Cependant, la prolifération de ces outils et leur accessibilité aux apprenants leur offrent désormais la possibilité de répondre aux évaluations sans nécessairement mobiliser leurs propres compétences intellectuelles (Coulibaly et Fofana, 2024). Cela soulève de nombreuses interrogations quant au rôle de l'enseignant et, surtout, sur l'avenir de l'évaluation des acquis des apprenants. L'essor des IAG pose une question cruciale : comment garantir la fiabilité des évaluations, alors que ces outils sont désormais capables de réussir avec aisance des examens autrefois comme infaillibles, tels que les tests de certification professionnelle destinés aux médecins, ingénieurs ou avocats (Giannini, 2023) ?

Face à ces défis, les enseignants doivent repenser et adapter leurs évaluations, en les orientant davantage vers l'évaluation des compétences². A cet effet, trois modalités d'intégration de l'IAG pour l'évaluation ressortent dans la littérature. Outre la prise en compte de l'IAG dans l'alignement pédagogique (entre les objectifs du cours, les activités et les évaluations), l'évaluation pourrait elle-même se faire sans l'IAG, ou avec l'IAG mais en limitant la pertinence de son utilisation ou encore en intégrant complètement avec l'IAG par le biais du sujet³.

L'évaluation sans les IAG implique une évaluation des connaissances et des compétences dans un environnement contrôlé et sans accès aux IAG. Il peut s'agir d'examens écrits sur table en présence, d'examens oraux en présence ou en ligne, d'évaluations au moyen d'ordinateur avec restriction d'accès des limitations de temps, d'exigences d'écriture manuscrite, de réalisation pratique, de questions à développement court ou long, questions vrai/faux, questions à choix multiples, question d'association ou textes à trou (Vangrunderbeeck et al., 2024).

L'influence de l'IAG peut être minimisée dans les évaluations en contextualisant les exercices et les questions (en utilisant des éléments spécifiques mentionnés en cours). Cela peut se faire également en demandant une réflexion personnelle et des exemples d'expériences vécues aux apprenants.

² <https://sup-ubs.fr/documentation/ia-generatives-repenser-ses-evaluations-fiche-recapitulative/>

³ <https://sup-ubs.fr/documentation/ia-generatives-repenser-ses-evaluations-fiche-recapitulative/>

Selon Vangrunderbeeck et al. (2024), pour minimiser l'influence de l'IAG dans les évaluations, l'enseignant doit contextualiser les exercices, demander des réflexions personnelles et des exemples d'expériences vécues aux étudiants; demander des réflexions critiques et des solutions nuancées, proposer des situations complexes où la véracité et la qualité doivent être examinées, concevoir des questions qui nécessitent une compréhension approfondie du contexte. Toujours dans la volonté de minimiser l'influence de l'IAG dans les évaluations, les enseignants pourraient aussi découper le processus d'évaluation en une suite d'évaluations formatives avant l'évaluation sommative. Ils pourraient également choisir des sujets d'évaluation en privilégiant des thématiques peu abordées sur le web; privilégier des sujets qui mobilisent des situations authentiques; adopter une démarche d'évaluation des compétences; diversifier les livrables attendus ou intégrer des modalités d'évaluations des travaux collaboratifs et une démarche argumentaire et réflexive.⁴

Enfin, dans le but d'évaluer les apprenants en tenant compte de l'IAG, l'enseignant peut encourager les étudiants à repérer des erreurs et/ou des biais dans des productions générées par ChatGPT et à retravailler les contenus générés avec références à l'appui. Ils pourraient aussi amener les apprenants à repérer ce qui peut être délégué à la machine ou pas ou encore inciter les apprenants à expliquer et illustrer leur démarche évolutive de conception de prompts pour affiner, ajuster les réponses obtenues.⁵

3. Problématiques et considérations théoriques

Au regard de tout ce qui précède, nous nous interrogeons sur les marges de manœuvre dont disposons l'enseignant dans un contexte où les étudiants utilisent également les outils d'IA pour traiter leurs évaluations. Comment évaluer les étudiants à l'ère de l'IA tout en mettant en évidence leurs compétences réelles? Faut-il combattre cette technologie ou l'appriivoiser? Cette réflexion soulevée par Lez et al. (2023) a mis en lumière la question cruciale de l'intégration de l'IA dans le processus d'évaluation des apprentissages.

La théorie de l'évaluation formative de Black et Wiliam (2009), repose sur les cinq principes que sont le rôle du feedback; l'implication de l'apprenant; l'ajustement de l'enseignement; le partage des objectifs d'apprentissage et des critères de réussite et l'interaction entre l'enseignant et ses élèves. Cette théorie s'inscrit parfaitement dans notre étude, dans la mesure où l'analyse des pratiques d'évaluation des apprentissages à l'ère de l'IA nous amène à nous interroger sur la manière dont les enseignants utilisent l'IA pour évaluer leurs apprenants et comment ces derniers exploitent les feedbacks issus de ces évaluations pour améliorer leurs contenus d'apprentissage et d'enseignement.

4. Méthodologie

Pour répondre à l'épineuse question de l'évaluation des apprentissages à l'ère de l'IA, nous avons mené des entretiens semi-dirigés auprès de huit auditeurs, dont sept enseignants issus d'universités ivoiriennes et un ingénieur techno-pédagogue. Parmi ces enseignants, cinq sont issus de l'Université Virtuelle de Côte d'Ivoire (UVCI) et deux de l'École Supérieure Africaine des TIC (ESATIC). L'ingénieur techno-pédagogue provient également de l'Université Virtuelle de Côte d'Ivoire. Le choix de ces établissements repose sur leur spécialisation dans la formation aux métiers du numérique. Les

⁴ <https://sup-ubs.fr/documentation/ia-generatives-repenser-ses-evaluations-fiche-recapitulative/>

⁵ <https://sup-ubs.fr/documentation/ia-generatives-repenser-ses-evaluations-fiche-recapitulative/>

enseignants qui y exercent sont confrontés à des apprenants utilisant quotidiennement des outils numériques, y compris des outils d'intelligence artificielle générative. Il est donc particulièrement pertinent d'étudier ces institutions, car leurs étudiants sont parmi les plus susceptibles d'utiliser de manière avancée l'intelligence artificielle générative pour traiter leurs devoirs et examens.

Le guide d'entretien invite les enseignants à expliquer leurs pratiques d'évaluation en tenant compte du fait que les étudiants utilisent l'intelligence artificielle générative pour composer ces évaluations. L'enseignant est amené à décrire l'ensemble de son processus de conception et d'administration des évaluations en précisant, à chaque étape, l'implication de l'IA ainsi que les outils spécifiques qu'il utilise.

Parmi les sept enseignants intégrés, trois enseignent l'informatique sans être spécialistes de l'IA, deux sont des spécialistes de l'intelligence artificielle et les deux autres sont issus du domaine des sciences sociales, plus précisément de l'anglais et de la communication. L'ensemble des entretiens s'est déroulé en ligne par visioconférence via l'application Zoom. Ces échanges ont été enregistrés afin de faciliter leur retranscription et leur analyse ultérieure.

Au total, huit verbatim ont été enregistrés et retranscrits sous format .doc. La retranscription a été réalisée à l'aide d'une intelligence artificielle, Riverside⁶, qui a fourni un texte finalisé à plus de 90 %. Une réécoute attentive des entretiens a ensuite été effectuée afin de vérifier l'exactitude de la transcription et d'apporter les corrections nécessaires. Ces verbatim ont ensuite fait l'objet d'analyse qualitative.

5. Résultats

5.1. Mode d'évaluation de l'institution

Les entretiens révèlent que l'organisation des compositions varie en fonction du mode d'enseignement de l'institution. À l'Université Virtuelle de Côte d'Ivoire, les examens se déroulent en ligne avec un nombre très élevé de candidats, souvent supérieur à un millier. Les étudiants composent sur des plateformes de formation à l'aide de leurs propres terminaux numériques, tels que des ordinateurs, tablettes ou smartphones, et accèdent à Internet depuis l'endroit de leur choix. Ces évaluations se déroulent ainsi sans aucune surveillance directe.

À l'École Supérieure Africaine des TIC, en revanche, les compositions ont lieu en présentiel avec un effectif restreint, généralement inférieur à trente candidats. Elles se déroulent dans des salles de classe, sur papier, et sous la surveillance des enseignants.

5.2. Outils d'évaluation et outils d'IA utilisés par les enseignants

La quasi-totalité des enseignants dispense principalement les questions de type texte à trous pour évaluer les apprentissages de leurs étudiants. Toutefois, en complément de cette approche, ils ont également recours, de manière variée, aux questions ouvertes ou aux projets de groupe. Le choix de ces types d'évaluation est principalement motivé par la volonté de limiter l'utilisation de l'intelligence artificielle par les étudiants.

Il ressort également que les questions à choix multiples, les questions à choix uniques et les questions d'ordonnance sont rarement utilisées. Pour concevoir leurs évaluations, les enseignants s'appuient sur divers outils d'intelligence artificielle. Dans l'ensemble, ils utilisent principalement ChatGPT et Quiz Wizard, mais font également appel à d'autres outils tels que Gemini et Copilot.

⁶ <https://riverside.fm/>

5.3. Action pour mitiger l'usage de l'IA par les étudiants lors des compositions

Face à la tendance des étudiants à utiliser des outils d'intelligence artificielle pour réaliser leurs devoirs, les enseignants mettent en place différentes stratégies afin de limiter ce recours lors des épreuves. Ce phénomène est pratiquement inexistant dans le cadre des compositions surveillées, qui se déroulent en présentiel et où les étudiants rédigent leurs réponses sur papier.

En revanche, il est particulièrement accentué lors des évaluations en ligne, effectuées sur des terminaux personnels avec accès à Internet. Pour contrer cette pratique, les enseignants privilégient un nombre plus important de questions, comportant davantage d'éléments à compléter, tout en réduisant le temps imparti pour la composition. En combinant ces deux approches, ils estiment que l'étudiant aura moins recours aux outils d'intelligence artificielle pour répondre aux questions. Selon eux, un étudiant qui tenterait d'utiliser ces outils pendant l'examen risquerait de ne pas traiter un nombre suffisant de questions, car les allers-retours entre l'intelligence artificielle et l'interface de composition pourraient perturber sa concentration et lui faire perdre un temps précieux.

D'autres enseignants recommandent l'utilisation de travaux de groupe ou de projets, qui doivent être présentés devant un jury. Selon eux, ce type d'évaluation limite fortement l'usage des outils d'intelligence artificielle, car il exige une implication active des étudiants et une démonstration concrète de leurs compétences.

Il ressort des entretiens que les enseignants sont conscients que leurs efforts pour limiter l'usage de l'intelligence artificielle par les étudiants lors des évaluations en ligne ne sont pas totalement efficaces, mais qu'ils contribuent néanmoins à en réduire l'utilisation. Selon eux, ces mesures seraient plus performantes si les évaluations en ligne se déroulaient dans des salles spécialisées, avec des outils dédiés à la composition et sous surveillance. Par ailleurs, les travaux de groupe, les projets et les exposés présentés devant un jury sont perçus comme une alternative plus efficace pour limiter le recours aux outils d'intelligence artificielle dans la réalisation des devoirs.

5.4. Crédit accordé aux productions de l'IA

S'il y a un point sur lequel les enseignants interrogés s'accordent à l'unanimité, c'est l'importance de la vérification des productions générées par les outils d'intelligence artificielle. La majorité d'entre eux affirme qu'ils prennent systématiquement le temps d'examiner les propositions de l'IA avant de les intégrer dans leurs outils d'évaluation.

Plusieurs témoignages illustrent cette position. L'un des enseignants souligne qu'« *il faut toujours évaluer ce que l'IA propose, prendre le temps de lire attentivement l'ensemble du contenu et procéder aux ajustements nécessaires avant l'évaluation. Il est essentiel de s'assurer que les suggestions de l'IA sont en adéquation avec ce qui a été enseigné aux apprenants.* » Un autre précise : « *Je n'utilise jamais les questions telles qu'elles sont produites par l'IA. J'y ajoute ma touche personnelle pour humaniser et personnaliser les évaluations, afin que les étudiants puissent percevoir la spécificité de l'enseignant dans les propositions générées. J'ai également tendance à complexifier les évaluations fournies par l'IA.* » Un troisième enseignant insiste sur le fait qu'« *il est impératif de vérifier chaque question générée par l'outil, car des erreurs peuvent s'y glisser avant même l'étape d'importation.* » Enfin, un dernier rappelle que « *l'enseignant ne doit pas être passif face à l'IA, il ne doit ni s'en contenter totalement ni lui accorder une confiance aveugle à 100 %.* »

5.5. Production de feedback et ajustement des enseignements

Il ressort des échanges avec les enseignants que, bien que la production de feedback poursuive les mêmes objectifs, elle se manifeste différemment selon qu'il s'agit d'un enseignement en présentiel ou à distance.

Dans le cadre de l'enseignement en présentiel, les copies des étudiants sont sous format papier. Les enseignants procèdent à une correction manuelle, attribuent une note à chaque production et y ajoutent un feedback personnalisé. En formation à distance, notamment dans le modèle de l'UVCI, le nombre élevé d'étudiants impose une autre approche. Les sujets d'évaluation, majoritairement composés de questions fermées, sont produits par l'enseignant et implémentés sur la plateforme de formation afin que les étudiants puissent composer. Les copies, entièrement numériques, sont accessibles via un terminal. La correction est automatisée par la plateforme de formation (Moodle). Ce qui permet de gérer efficacement le volume important de copies à traiter.

Dans ce contexte, le dispositif de l'UVCI recommande à chaque enseignant responsable d'une évaluation de consulter un échantillon d'au moins cinq copies sélectionnées aléatoirement. L'objectif est d'identifier les éléments les plus réussis et les difficultés récurrentes rencontrées par les étudiants. Sur la base de ces observations, l'enseignant est amené à produire une vidéo de remédiation destinée aux étudiants ayant obtenu de faibles performances et à mettre à jour son contenu de formation afin d'améliorer la compréhension des apprenants.

6. Discussion

La question de l'évaluation des apprentissages constitue une préoccupation majeure pour les acteurs de l'éducation. Les résultats issus de nos entretiens viennent confirmer nos analyses. En nous appuyant sur les travaux de Vangrunderbeeck et al. (2024), nous constatons que les pratiques d'évaluation des enseignants interrogés s'inscrivent dans les trois grandes approches d'évaluation intégrant l'intelligence artificielle.

Certains enseignants, en particulier ceux exerçant en présentiel, privilégient une évaluation sans implication de l'IA. Ceux évoluant en formation en ligne évaluent les acquis de leurs apprenants soit en minimisant l'impact de l'IA, soit en l'intégrant pleinement dans le processus d'évaluation (Lez et al., 2023). L'utilisation de textes à trous ou de questions ouvertes permet de limiter l'influence de l'IA, tandis que les évaluations basées sur des projets ou des présentations de travaux de groupe devant un jury prennent en compte l'intelligence artificielle.

Comme l'indiquent nos résultats, la formation des enseignants à l'usage des outils d'IA apparaît comme une nécessité. Cette recommandation est d'ailleurs soutenue par plusieurs auteurs, dont le Groupe de travail (2024).

La majorité des enseignants interrogés utilisent ChatGPT pour générer les questions intégrées à leurs évaluations. Cette pratique est également rapportée par De Lièvre et al. (2024) ainsi que par le Service universitaire de pédagogie – UBS (2023). Par ailleurs, les réserves émises par les enseignants quant aux réponses fournies par l'IA à la suite d'une requête sont partagées par Lez et al. (2023) et aussi De Lièvre et al. (2024). Bien qu'ils estiment que l'intelligence artificielle permet un gain de temps, ces auteurs soulignent que son utilisation nécessite une vérification et des ajustements afin d'assurer un alignement pédagogique optimal.

Pour limiter l'usage de l'IA par les étudiants lors des compositions, plusieurs stratégies sont mises en place par les enseignants. Plutôt que de s'opposer frontalement à son utilisation, comme le souligne Lez et al. (2023), ils préfèrent l'intégrer intelligemment dans leurs pratiques pédagogiques. Dans cette optique, ils cherchent à mieux comprendre et maîtriser ces outils afin d'en tirer le meilleur parti pour l'élaboration de

leurs évaluations. Les enseignants s'appuient ainsi sur plusieurs approches, notamment la contextualisation des évaluations générées par l'IA, la personnalisation des questions et l'instauration d'une logique pédagogique cohérente. Ces pratiques s'inscrivent pleinement dans les recommandations formulées par Lez et al. (2023).

En ce qui concerne les résultats précédemment mentionnés, l'évaluation ne semble plus répondre aux principes définis par Black et Wiliam (2009), dans la mesure où la base de référence servant à produire un feedback et à ajuster les enseignements est faussée. En effet, si le feedback fourni par l'enseignant repose sur une évaluation qui ne reflète pas fidèlement les compétences réelles de l'apprenant, il perd toute valeur constructive. Par ailleurs, l'enseignant, constatant de bons résultats aux évaluations, pourrait être conforté dans l'idée que la qualité de son enseignement est satisfaisante. Il ne remettrait donc pas en question sa pédagogie et n'envisagerait pas d'ajustements. Cette situation traduit également une implication réduite de l'étudiant dans son apprentissage, puisqu'il peut percevoir l'intelligence artificielle comme une solution lui permettant de contourner l'enseignant et d'éviter un véritable engagement dans le processus de formation. Ce qui contribuerait malheureusement à créer une dépendance cognitive vis-à-vis de l'IA⁷.

Toutefois, le potentiel offert par l'intelligence artificielle aux enseignants reste insuffisamment exploité pour redonner à l'évaluation son rôle central de mesure et d'ajustement. Comme le recommande Black et Wiliam (2009), l'évaluation doit permettre aux enseignants d'identifier les besoins individuels des apprenants afin d'adapter leur enseignement en conséquence. Elle doit également leur offrir la possibilité d'examiner et de suivre les progrès des étudiants tout en évaluant l'efficacité de leur propre pédagogie. Car en effet l'Intelligence Artificielle offre des possibilités d'adaptation des enseignements et même des évaluations à chaque apprenant (Allal & Mottier Lopez, 2005).

7. Recommandations

Deux recommandations majeures émergent des propositions des enseignants interrogés. La première concerne la formation des enseignants à l'utilisation des outils d'intelligence artificielle, tandis que la seconde porte sur la nécessité de vérifier les informations fournies par l'IA avant de les intégrer dans leurs contenus pédagogiques.

En effet, il est revenu de manière récurrente chez la quasi-totalité des enseignants interrogés que les enseignants eux-mêmes devraient être formés à l'usage des outils d'IA existant et sur leur fonctionnalité dans les pratiques d'évaluation. La formation devrait porter sur le fonctionnement des IA de manière générale et l'écriture des prompts en particulier.

Par ailleurs, il est recommandé aux enseignants utilisant les outils d'IA pour évaluer le niveau de maîtrise de leurs étudiants, de prendre soin de bien vérifier les questions que leur fournissent les outils d'IA avant de les intégrer à leur devoir. Cette vérification est essentielle pour adapter les questions au contexte de l'évaluation, identifier d'éventuels biais et les corriger afin d'assurer une évaluation plus pertinente et fiable.

8. Conclusion

« Combattre la bête IA ou l'apprivoiser » (Lez et al., 2023 ; De Lièvre et al., 2023) constitue la question centrale qui alimente cette étude. Cette problématique prend tout

⁷ <https://www.beaboss.fr/Thematique/digital-innovation-1074/intelligence-artificielle-2023/Breves/fascination-l-dependance-veritable-danger-466945.htm>

son sens lorsqu'elle est rapportée à la question de l'évaluation dans le domaine académique. L'utilisation des outils d'intelligence artificielle générative par les enseignants, d'une part, et par les étudiants, d'autre part, remet en question le processus d'évaluation des travaux académiques. En effet, il ressort de plusieurs enquêtes (Coulibaly et Fofana, 2024) que les étudiants utilisent les outils d'IA pour traiter les devoirs qui leur sont soumis pour l'évaluation de leurs acquis. De ce fait, les évaluations qui sont censées retourner des informations sur la qualité de maîtrise des étudiants concernant les connaissances et compétences enseignées, renvoient plutôt des informations sur la maîtrise qu'a l'IA de ces connaissances et de ces compétences.

« Comment évaluer les apprentissages à l'ère de l'IAG ? » est la question centrale à laquelle cette étude a tenté de répondre. Des entretiens semi-dirigés ont été menés auprès de sept enseignants et d'un ingénieur techno-pédagogique issu d'universités ivoiriennes. En nous appuyant sur la théorie de l'évaluation formative de Black et Wiliam (2009) et après une analyse qualitative de contenu, il ressort que les enseignants utilisent principalement ChatGPT et Quiz Wizard pour générer les questions intégrées à leurs évaluations.

Ils privilégient les textes à trous avec un nombre élevé d'éléments à compléter et un temps très réduit de composition afin de limiter l'usage des outils d'IA par les étudiants lors des compositions. Ils ont également recours à des questions de production, des projets de groupe et des exposés pour évaluer leurs étudiants et s'assurer que l'intelligence artificielle n'influence pas les résultats des évaluations.

En fin, bien que l'utilisation de l'IAG représente un gain de temps considérable pour les enseignants, il est recommandé qu'ils reçoivent une formation à l'usage de ces outils et qu'ils vérifient systématiquement les questions générées par l'IAG avant de les intégrer à leurs sujets d'évaluation.

9. Bibliographie

- Allal, L., & Mottier Lopez, L. (2005). L'évaluation formative de l'apprentissage: revue de publications en langue française. *OCDE, L'évaluation formative: pour un meilleur apprentissage dans les classes secondaires*, 265-290.
- Black, P., Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education)* 21(1), 5 (2009)
- Bordat-Teeuwen, D., & Gonin, V. (2024). Comment apprivoiser l'intelligence artificielle (IA) quand on est formateur en IFSI?
- Coulibaly, Y. N., et Fofana, A. (2024). Outils d'IA utilisés par les enseignants et les étudiants en Côte d'Ivoire. In Tome 4 : Education, Enseignement, Pédagogie, Didactique (pp.173-189). Della/Afrique. Vol N° Spécial / Fin Campagne Editoriale Décembre 2024.
- De Lièvre, B., Temperman, G., Decamps, S., Kumps, A., Dragone, L., & Housni, S. (2023, November). Évaluation avec l'Intelligence Artificielle: pistes pour l'action et la réflexion. In *Forum pédagogique*.
- De Travail, G. (2024). Utilisation Responsable de l'Intelligence Artificielle Générative.
- Gérard, F. M. (2013). L'évaluation au service de la régulation des apprentissages: enjeux, nécessités et difficultés. *Revue française de linguistique appliquée*, 18(1), 75-92.
- Giannini, S. (2023). *L'IA générative et le futur de l'éducation*. Paris, UNESCO.
- Soulez, M. (2018). Questions juridiques au sujet de l'intelligence artificielle. *Annales des Mines-Enjeux numériques*, 1(1), 81-85.
- Haton, J. P. (2024). L'intelligence artificielle générative.
- Jouquan, J. (2002). L'évaluation des apprentissages des étudiants en formation médicale initiale. *Pédagogie médicale*, 3(1), 38-52.
- Kamoun, N., Akhasbi, H., Lakrami, F., Gilles, J. L., Rigo, J. M., & Aliss, E. (2025). L'intelligence artificielle générative: une opportunité pour améliorer l'évaluation formative dans un contexte universitaire massif.
- Lez, A., Dubé, E. et Beaulieu, M. (2023). Évaluer à l'ère des IA : Combattre la bête ou l'apprivoiser. un guide de réflexion. Service de soutien à la formation, Université de Sherbrooke. Sous licence CC BY 4.0. repéré en ligne à chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.usherbrooke.ca/sf/fileadmin/sites/ssf/Enseignement/ia/Evaluer_ere_IA-V1-3.pdf
- Livian, Y. (2024, September). Utiliser l'IA de manière raisonnée et éthique: le cas de la rédaction des mémoires. In *Technologies et gouvernance des universités*.
- Plumat, J. (2023). Analyse comparative des universités Virtuelles francophones en Afrique. Rapport d'analyse pour l'AUF.
- Service universitaire de Pédagogie – UBS. (2023, 22 novembre). Utiliser ChatGPT pour concevoir un QCM et l'intégrer dans Moodle [vidéo]. YouTube. <https://youtu.be/4XVOEXaxB8I>
- Tardif, J. (1993). L'évaluation dans le paradigme constructiviste. *L'évaluation des apprentissages. Réflexions, nouvelles tendances et formation*, 27-56.
- Vangrunderbeeck, P., Deville, Y., Guisset, M., Baur, M., Biot, M., Raucet, B., & Wouters, P. (2024). Intégrer l'IA générative dans les stratégies pédagogiques.
- Walbert, L., & Polcheira, É. (2024). Risques de l'évaluation avec l'IA: craintes et solutions.