

L'intelligence artificielle générative : une opportunité pour améliorer l'évaluation formative dans un contexte universitaire massif

N. El Kamoun ¹, H. Akhasbi ¹, F. Lakrami ¹, J.L. Gilles ², J.M. Rigo ³, E. Aliss ⁴

¹ Université Chouaib Doukkali (MOROCCO)

² Haute Ecole Pédagogique de Lausanne (SWITZERLAND)

³ Université Hasselt (BELGIUM)

⁴ Universidad Simón I (BOLIVIA)

Résumé long de la communication

Dans le contexte actuel de l'enseignement supérieur, marqué par une évolution constante de l'environnement socio-technologique, l'évaluation formative avec feedback présente un potentiel considérable. Elle vise à offrir un retour d'information continu aux étudiants, favorisant ainsi l'amélioration de leurs apprentissages. Le concept d'« Assessment for Learning » (Assessment Reform Group, 1999 ; Black & Wiliam, 2006 ; Sambell et al., 2012 ; Laveault & Allal, 2016) constitue un tournant significatif dans ce domaine de recherche. Ce cadre pédagogique met l'accent sur l'engagement actif des étudiants. Il favorise l'usage de dispositifs permettant aux étudiants de mieux comprendre leurs forces et faiblesses, de s'autoréguler en développant leurs compétences métacognitives, et ainsi de progresser de manière autonome (CCSSO, 2009 ; McMillan, 2007 ; Tillema, Leenknecht & Segers, 2011 ; Reinholz, 2016 ; Gilles & Charlier, 2020). Pour maximiser l'efficacité de cette démarche, les questions à choix multiple, alignées avec les objectifs pédagogiques, jouent un rôle crucial. L'intégration des degrés de certitude dans ces dispositifs enrichit encore le processus, permettant une analyse plus nuancée des performances des étudiants (Leclercq, 1982 ; Leclercq & Gilles, 2003 ; Gilles, 2012).

Pour garantir l'efficacité de ces outils d'évaluation, il est primordial que le feedback soit à la fois précis, constructif et spécifiquement orienté (Yorke, 2003). Ce retour doit permettre aux étudiants de comprendre leurs réussites tout en identifiant les domaines à améliorer. L'encouragement joue également un rôle central dans ce processus, en renforçant la motivation des apprenants à travers la valorisation de leurs efforts et la reconnaissance de leurs progrès (Shute, 2008). Cela les incite à adopter une réflexion plus approfondie sur leurs apprentissages, consolidant ainsi leur motivation et leur compréhension (Nicol, 2014). Dans ce contexte, le modèle de Hattie et Timperley (2007) propose un cadre théorique solide pour définir les conditions optimales favorisant un impact positif sur l'apprentissage.

Cependant, la mise en œuvre efficace de cette approche dans des environnements d'enseignement de masse pose des défis logistiques et pédagogiques importants (Nicol, 2014). La conception de

QCM de qualité, l'analyse des résultats en tenant compte des degrés de certitude, l'exploitation de divers indicateurs statistiques, ainsi que la fourniture de feedbacks personnalisés à un grand nombre d'étudiants, augmentent considérablement la charge de travail des enseignants (Akhasbi et al., 2023). Ces exigences en termes de temps et de ressources peuvent limiter l'efficacité de cette approche à grande échelle.

Face à ces contraintes, l'émergence de l'intelligence artificielle générative (IAG) offre des perspectives prometteuses pour assister les enseignants et alléger leur charge de travail (Popenici & Kerr, 2017). L'IAG propose des solutions innovantes qui permettent une meilleure gestion des évaluations à grande échelle. Par exemple, elle peut faciliter l'élaboration de tables de spécifications pertinentes pour les évaluations (DiDonato-Barnes, Fives, & Krause, 2014), tout en aidant à concevoir des QCM de qualité, en veillant au respect des normes de rédaction et en générant des réponses alternatives plausibles (Przymuszała et al., 2020). De plus, l'IAG est capable de traiter et d'analyser rapidement de grandes quantités de données à haute dimensionnalité, permettant ainsi d'identifier les facteurs de réussite ou les obstacles rencontrés par les étudiants dans leurs parcours d'apprentissage. Cette analyse rapide et approfondie facilite la génération de feedbacks personnalisés, favorisant une individualisation plus efficace du suivi pédagogique (González-Calatayud et al., 2021). En automatisant les tâches répétitives, l'IAG permet aux enseignants de se concentrer davantage sur l'accompagnement pédagogique et le développement des compétences des étudiants, tout en réduisant leur charge de travail. Ainsi, en assistant les enseignants tout au long du processus d'évaluation, l'IAG contribue à optimiser l'expérience d'apprentissage en fournissant des retours plus rapides, précis et constructifs, répondant ainsi aux défis posés par l'enseignement de masse.

Cependant, malgré ces avantages potentiels l'intégration de l'IAG dans le processus d'évaluation formative doit être abordée avec prudence, en tenant compte des implications éthiques ainsi que les défis techniques et méthodologiques (Zawacki-Richter et al., 2019). Il est crucial que les enseignants soient bien formés à l'utilisation de cette technologie, afin de comprendre pleinement ses capacités et ses limites. Cette adoption doit se faire de manière réfléchie, en prenant en compte les risques potentiels liés à la protection des données personnelles, aux biais algorithmiques, à la fiabilité des rétroactions automatisées et à la possible réduction des interactions humaines (Williamson & Eynon, 2020). La gestion de la qualité des feedbacks générés par l'IAG nécessite une approche hybride, combinant des systèmes de surveillance automatisés capables de détecter les anomalies avec des mécanismes de signalement par les étudiants, permettant ainsi des interventions humaines au besoin. Cette stratégie vise à équilibrer l'efficacité de l'automatisation avec un contrôle humain pertinent, garantissant que les feedbacks restent fiables et précis tout en assurant la scalabilité du processus (García-Peñalvo & Corell, 2024). Cette stratégie vise à équilibrer l'efficacité de l'automatisation avec un contrôle humain pertinent, garantissant que les feedbacks restent fiables, précis et pertinents tout en assurant la scalabilité du processus (García-Peñalvo & Corell, 2024).

Afin de mieux comprendre ces enjeux, notre recherche explore l'application concrète de l'IAG dans l'évaluation formative, en étudiant ses avantages pratiques et les défis associés à travers une série d'expérimentations exploratoires. Dans cette communication, nous présenterons nos premières conclusions, qui visent à intégrer l'IAG à chaque étape du processus, depuis l'élaboration des tables de spécifications jusqu'à la génération des feedbacks. Ce travail s'inscrit dans la continuité du projet Erasmus+ CORETEV, axé sur l'intégration de l'approche qualité dans l'évaluation des apprentissages (Gilles et al., 2024). Bien que notre étude en soit encore à ses débuts, les résultats obtenus sont prometteurs. L'IAG s'est révélée particulièrement efficace pour accélérer et affiner des tâches telles que la création d'évaluations alignées et la production de feedbacks personnalisés. Cependant, des défis persistent, notamment en matière de contrôle de la qualité et d'adaptation. Des recherches supplémentaires seront nécessaires pour confirmer ces premiers constats, surmonter les obstacles identifiés et exploiter pleinement le potentiel de cette technologie dans l'évaluation des apprentissages. Nous restons convaincus que, tout comme dans d'autres domaines, l'IA a le potentiel de transformer l'évaluation en une expérience d'apprentissage à forte valeur pédagogique.

Bibliographie

- Akhasbi, H., El Kamoun, N., Lakrami, F., Gilles, J.-L., Rigo, J.-M., & Aliss, E. (2023). Remote Formative Assessment With Multiple-Choice Questions and Levels of Certainty: The Impact of Feedback on Learning. In L. Gómez Chova, C. González Martínez & J. Lees (Eds.), *Conference Proceedings of 16th International Conference of Education, Research and Innovation (ICERI 2023)* (pp. 3709-3719). Retrieved from <http://hdl.handle.net/20.500.12162/7429>
- Assessment Reform Group (ARG). (1999). *Assessment for learning: Beyond the black box*. Cambridge: University of Cambridge.
- Black, P., & William, D. (2006). Developing a Theory of Formative Assessment. In J. Gardner (Ed.), *Assessment and Learning* (pp. 81-100). London: Sage.
- CCSSO. (2009). *Report of the Council of Chief State School Officers. Vision for developing assessment systems that support high quality learning*. Washington, DC.
- DiDonato-Barnes, N., Fives, H., & Krause, E. S. (2014). Using a Table of Specifications to Improve Teacher-Constructed Traditional Tests: An Experimental Design. *Assessment in Education: Principles, Policy & Practice*, 21(1), 90-108.
- García-Peñalvo, F. J., & Corell, A. (2024). Artificial Intelligence in Education: A Review of the State of the Art and Future Directions. *Education and Information Technologies*, 29(1), 1-25.
DOI: 10.1007/s10639-023-11104-5
- Gilles, J.-L. (2012). *Conditions méthodologiques d'utilisation des degrés de certitude pour l'auto estimation des compétences dans le cadre du test de maîtrise de français EFES à la HEP Vaud*. Lausanne, Suisse : UER EN. <http://hdl.handle.net/20.500.12162/3777>
- Gilles, J.-L., et Charlier, B. (2020). Dispositifs d'évaluation à distance à correction automatisée versus non automatisée : analyse comparative de deux formes emblématiques. *Evaluer : journal international de recherche en éducation et formation - e-JIREF*, Numéro Hors-série(1), 143-154. <http://hdl.handle.net/20.500.12162/3865>
- Gilles, J.-L., Aliss, E., et Rigo, J.-M. (2024, janvier). Symposium « Retombées en ingénierie des évaluations du projet CORETEV ». Communication présentée à 35e Colloque de l'ADMEE-Europe "L'évaluation face aux défis de la diversité et de l'inclusion : entre normes et différenciations", Braga, Portugal. <http://hdl.handle.net/20.500.12162/7350>
- González-Calatayud, V., Prendes-Espinosa, P., et Roig-Vila, R. (2021). Artificial intelligence for student assessment: A systematic review. *Applied sciences*, vol. 11, no 12, p. 5467.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112. DOI: 10.3102/003465430298487
- Laveault, D., & Allal L. (2016). Implementing Assessment for Learning: Theoretical and Practical Issues. In D. Laveault & L. Allal (Eds.), *Assessment for Learning: Meeting the Challenge of Implementation*. Bern: Springer International Publishing Switzerland.

- Leclercq, D. (1982). *Confidence marking: Its use in testing*. *Evaluation in Education*, 6(2), 161-287.
- Leclercq, D., et Gilles, J.-L. (2003). Analyses psychométriques des questions des 10 check-up MOHICAN. Vue d'ensemble. Dans D. Leclercq (Ed.), *Diagnostic cognitif et métacognitif au seuil de l'université - Le projet MOHICAN mené par les 9 universités de la Communauté française Wallonie Bruxelles* (pp. 173-180). Liège, Belgique: Presses universitaires de l'Université de Liège. <http://hdl.handle.net/20.500.12162/757>
- McMillan, J. H. (2007). *Formative classroom assessment: Research, theory and practice*. New York: Teachers College Press.
- Nicol, D. (2014). From monologue to dialogue: improving written feedback processes in mass higher education. In: *Approaches to assessment that enhance learning in higher education*. Routledge, p. 11-27.
- Popenici, S. A. D., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning*, 12(1), 1-12.
- Przymuszala, P., Piotrowska, K., Lipski, D., et al. (2020). Guidelines on writing multiple choice questions: A well-received and effective faculty development intervention. *SAGE Open*, vol. 10, no 3.
- Reinholz, D. (2016). The assessment cycle: A model for learning through peer assessment. *Assessment & Evaluation in Higher Education*, vol. 41, no 2, p. 301-315.
- Sambell, K., McDowell, L. and Montgomery, C. (2012). *Assessment for Learning in Higher Education*. London, Routledge.
- Shute, V. J. (2008). Focus on Formative Feedback. *Review of Educational Research*, 78(1), 153-189
- Tillema, H., Leenknecht, M., & Segers, M. (2011). Assessing assessment quality: Criteria for quality assurance in design of (peer) assessment for learning – A review of research studies. *Studies in Educational Evaluation*, Vol 37, 25-34. <https://doi.org/10.1016/j.stueduc.2011.03.004>
- Williamson, B. & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology*, vol. 45, no 3, p. 223-235.
- Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher education*, vol. 45, p. 477-501.
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), 39. DOI: 10.1186/s41239-019-0171-0