

Fichier 1

Les progrès de l'IA mettent en tension les systèmes éducatifs (par Maxime Cruzel)

par

Maxime Cruzel

Mezetulle, 27 juillet 2026

URL : <https://www.mezetulle.fr/les-progres-de-lia-mettent-en-tension-les-systemes-educatifs-par-maxime-cruzel/>

On pouvait croire que l'apparition de l'IA générative allait remettre les pendules à l'heure dans la doctrine scolaire, en soulignant le besoin de principes, de hiérarchisation des connaissances et donc d'instruction, la nécessité de faire travailler les élèves sur un contenu substantiel organisé. Mais non, on continue avec une doctrine éducative qui promeut les compétences vides et le comportementalisme, et qui maintenant s'autorise d'un apparent bon sens : à quoi bon instruire puisque l'IA sert les connaissances à la demande ? S'appuyant sur de solides et récentes références, analysant les progrès de l'IA et des exemples comparatifs, Maxime Cruzel¹ montre que cette position « inverse le rapport entre savoirs et compétences ». Les compétences ne précèdent pas les

connaissances, elles en résultent. À mesure que l'IA se développe et qu'elle change de nature, il est urgent d'instruire.

[Article initialement publié dans la revue en ligne *Telos* le 22 juillet 2026², repris ici avec l'aimable autorisation de l'auteur et des éditeurs de la revue. Je les remercie vivement. Les intertitres ont été ajoutés par Mezetulle.]

À quoi bon transmettre des connaissances que la machine restitue mieux et plus vite ? La question traverse le débat éducatif depuis plusieurs décennies et l'arrivée d'internet, avec souvent une préconisation qui découle du « bon sens » : puisque toute la connaissance est accessible avec internet, l'école ne doit plus dispenser des savoirs mais se concentrer sur les compétences, l'esprit critique et la capacité à chercher l'information. L'argument a trouvé dans l'IA générative une seconde jeunesse. Cette philosophie inverse pourtant le rapport réel entre savoirs et compétences, et la dernière mutation de l'IA, avec sa bascule vers l'autonomie, rend cette inversion plus malencontreuse que jamais : superviser une machine qui travaille seule exige précisément ce que l'on propose de moins enseigner.

C'est ce paradoxe qui place de nombreux systèmes éducatifs, notamment occidentaux, en tension au moment d'appréhender les spectaculaires progrès de l'IA générative depuis le milieu de l'année 2025.

Sommaire

1. Les progrès de l'IA générative

2. *« On ne vérifie bien que ce que l'on connaît »*
3. *Enseigner des connaissances organisées et non des « compétences générales » vides*
4. *Les savoirs disciplinaires, générateurs d'autonomie et d'esprit critique*
5. *Références*
6. *Notes*

Les progrès de l'IA générative

En quelques mois, en effet, l'IA générative a changé de nature. Le paradigme initial, une question, une réponse, une retouche, a cédé la place à des agents, ou boucles agentiques, qui enchaînent seuls les étapes d'une tâche : lire des documents, exécuter des actions, observer le résultat, corriger le tir. L'institut METR, qui mesure la durée des tâches humaines qu'un modèle mène à bien de façon autonome (le LLM Time Horizon), a récemment vu cette durée doubler tous les quatre à cinq mois ; les meilleurs modèles accomplissent désormais, avec une réussite d'une fois sur deux, des tâches qui occuperaient un professionnel plusieurs heures, quand ils plafonnaient à quelques minutes il y a deux ans. David Monniaux, chercheur en informatique au CNRS, note que les IA génératives les plus avancées, les modèles dits « frontier », s'attaquent désormais au concours d'entrée en mathématiques de l'ENS Paris-Lyon.

Yann Dubois, qui codirige le postentraînement des modèles chez OpenAI, expliquait en mai au MAD Podcast l'origine de ce sursaut : l'apprentissage par renforcement, cette méthode qui récompense la machine lorsqu'elle atteint le résultat visé au lieu de lui faire imiter des textes, a pris le

relais d'un entraînement sur des documents rédigés par des humains, arrivé à ses limites. Cantonnée à l'origine aux domaines où la réponse se vérifie mécaniquement (compétitions de mathématiques ou de code), la méthode par renforcement s'attaque - depuis que les modèles ont atteint en 2025 la taille critique nécessaire - aux tâches réelles, mal définies, où il faut d'abord rassembler contexte et ressources avant de raisonner.

« On ne vérifie bien que ce que l'on connaît »

Cette forte accélération des performances de l'IA générative a un revers qui concerne directement l'école. Les erreurs se raréfient, de près de 8% à environ 1% en trois ans sur des tâches de résumé, mais elles se déplacent vers les questions pointues, rares dans les données d'entraînement : celles, précisément, qu'un professionnel pose dans l'exercice de son métier. Elles changent aussi de forme : *au fil de la boucle agentique*, les modèles de raisonnement insèrent des affirmations fausses dans des chaînes de déduction cohérentes, où rien ne les distingue en surface des étapes valides qui les entourent.

La charge de la vérification repose alors sur l'utilisateur, et les premières études convergent sur ce qui conditionne son succès. En Chine, 142 internes de médecine générale confrontés aux réponses de ChatGPT sur des cas cliniques n'en ont repéré les erreurs qu'un peu plus d'une fois sur deux, moins encore sur les cas complexes. Au King's College de Londres, un cinquième seulement des étudiants d'économétrie a identifié l'erreur conceptuelle glissée par ChatGPT dans un devoir noté, détection que prédisaient le

bagage disciplinaire et les capacités d'interprétation, non l'application mécanique de procédures. À l'université Khalifa d'Abou Dabi, des étudiants en informatique n'ont repéré les hallucinations que dans leur spécialité, le code, et se sont montrés incapables, sans les bases d'un cours, d'évaluer les réponses de ChatGPT.

On ne vérifie bien que ce que l'on connaît. Daniel Willingham a montré que les compétences cognitives générales, résolution de problème ou pensée critique, ne s'exercent pas dans le vide : elles opèrent sur un contenu, et la qualité du raisonnement dépend des connaissances disponibles en mémoire pour interpréter la situation. *On ne pense pas de façon critique sur un sujet qu'on ignore, faute de concepts pour l'analyser et de repères pour juger qu'une affirmation est plausible.*

Enseigner des connaissances organisées et non des « compétences générales » vides

Les travaux d'Alain Lieury sur la mémoire des collégiens précisent le mécanisme : c'est la richesse du réseau sémantique, ces concepts reliés entre eux qu'enrichit l'école à chaque cours, exercice ou lecture, qui prédit la réussite scolaire. Un terme comme « hypoténuse » ne vaut pas comme mot retenu, il ouvre l'accès à un concept organisé, rattaché à des catégories, et Lieury a mesuré entre ce vocabulaire encyclopédique et la réussite au collège des corrélations élevées. Or, dans des travaux en psychologie de la lecture, Tobias Richter et ses collègues ont montré, en chronométrant des lecteurs face à des affirmations vraies ou fausses, que la validation s'opère au fil même de la

compréhension, sans décision délibérée, pourvu que la mémoire contienne les connaissances correspondantes ; quand elles manquent, rien n'achoppe, et le jugement se rabat sur les indices de surface, ceux que la machine optimise.

André Tricot et John Sweller en ont tiré une conséquence pour *la conception des programmes scolaires* : les compétences générales ne s'enseignent pas directement, ce que l'école peut construire, ce sont des connaissances organisées, dont la maîtrise produit ce qui apparaît de l'extérieur comme des compétences. Un courant international de psychologie cognitive prolonge aujourd'hui cette conclusion, sous la bannière du renouveau des connaissances, en plaidant pour *des programmes scolaires de nouveau riches en savoirs*.

Cette richesse de connaissances est précisément la ressource qu'exige un agent. Dans le même entretien, Yann Dubois (OpenAI) relevait que cadrer une tâche floue constitue une compétence à part entière. La remarque vaut pour l'utilisateur : cadrer une demande pour un agent, c'est mobiliser les concepts et les mots justes du domaine, riches en implicites ; vérifier ce qu'il produit, c'est disposer des repères pour juger de l'exactitude. Les deux opérations reposent sur le mécanisme cognitif précédemment décrit : *on ne voit pas passer l'erreur qu'on n'a pas les moyens de détecter*.

Une enquête de Microsoft Research auprès de 319 salariés montre que la confiance dans ses propres capacités, non dans l'outil, prédit le regard critique porté sur ses productions. Du reste, la faible capacité à spécifier, faute

de savoirs associés, une demande floue ou complexe limite le recours à l'IA sur des problèmes où elle pourrait pourtant apporter une forte valeur ajoutée. La Harvard Business Review soulignait début 2026 que les salariés maîtrisant mal leur métier sous-utilisent l'IA, en plus de subir ses erreurs. Concrètement, que ce soit pour une photographie prise à la volée ou générée par l'IA, un photographe professionnel générera toujours des contenus plus qualitatifs que l'utilisateur lambda, par sa capacité à la fois à bien cerner la tâche à accomplir, dans toute sa complexité, et à déceler les incohérences du rendu proposé pour itérer. Le chaînon complet, du déficit de connaissances scolaires à la moindre capacité de supervision, reste à établir tant le phénomène est récent ; ses deux extrémités le sont déjà.

Les savoirs disciplinaires, générateurs d'autonomie et d'esprit critique

L'école française, comme celle de nombreux pays occidentaux, aborde ce virage en mauvaise posture. En mathématiques, un élève français de quinze ans accuse, selon PISA, l'équivalent de près de trois années d'apprentissage de retard sur son homologue coréen et de cinq années sur son homologue singapourien, et l'écart se retrouve en sciences. Ces systèmes ont maintenu les savoirs disciplinaires au centre de l'enseignement et, à contre-courant des choix occidentaux, continuent d'évaluer les connaissances par des examens terminaux qui décident de l'orientation. Ce choix n'a rien d'archaïque : une évaluation de connaissances renseigne bien au-delà d'elles, sur l'autonomie qu'il a fallu pour les acquérir, sur les capacités cognitives

exercées et, dans une certaine mesure, sur les compétences, puisque le réseau conceptuel qu'elle mesure est ce qui les porte. Le même paradigme permet au passage d'évaluer, par inférence, les pratiques pédagogiques et les progrès des élèves.

La recherche en évaluation aboutit au même constat pour les épreuves orales, qu'elle recommande depuis l'arrivée de l'IA générative : elles se dérobent à toute assistance et exposent en direct le réseau conceptuel chez l'élève, terreau de l'utilisation qu'il fera plus tard de l'IA. De cette dépendance découle la priorité scolaire : transmettre davantage de connaissances organisées, non pas moins, parce que l'esprit critique, l'autonomie et la capacité à superviser des machines en procèdent. Les faits rassemblés ici convergent : des machines qui gagnent en autonomie à un rythme inédit, des erreurs qui migrent vers les questions expertes et une vérification qui dépend du bagage disciplinaire, une psychologie cognitive qui place les connaissances à la racine des compétences, des systèmes scolaires qui réussissent en les maintenant au centre.

Le calendrier n'attend pas : l'IA a changé de nature en dix-huit mois, les cycles de réforme scolaire se comptent en décennies. La préconisation de « bon sens » qui ouvre cet article a le mérite du confort, elle dispense de transmettre ce qui coûte à transmettre ; tout ce qui précède indique qu'elle prépare des élèves incapables de superviser les machines auxquelles on les aura confiés, relégués au rang de consommateurs « B2C » du produit fini, qui n'en tirent guère de valeur.

Références

- Dang, C. T., & Nguyen, A. (2025). Distinguishing fact from fiction: Student traits, attitudes, and AI hallucination detection in business school assessment [Préprint]. arXiv. <https://arxiv.org/abs/2506.00050>
- Kalai, A. T., Nachum, O., Vempala, S. S., & Zhang, E. (2025). Why language models hallucinate [Préprint]. arXiv. <https://arxiv.org/abs/2509.04664>
- Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems. <https://doi.org/10.1145/3706598.3713778>
- Lieury, A. (2020). Mémoire et réussite scolaire (4e éd.). Dunod.
- METR. (2025, 19 mars). Measuring AI ability to complete long tasks. <https://metr.org/blog/2025-03-19-measuring-ai-ability-to-complete-long-tasks/>
- Monniaux, D. (2026, 15 mai). L'IA générative face au concours d'entrée à l'École normale supérieure [Billet de blog]. Mediapart. <https://blogs.mediapart.fr/david-monnaux/blog/150526/liagenerative-face-au-concours-dentree-lecole-normale-superieure>
- OCDE. (2023). PISA 2022 results (Volume I): The state of learning and equity in education. OECD Publishing. <https://doi.org/10.1787/53f23881-en>
- Ranganathan, A., & Ye, X. M. (2026, 9 février). AI doesn't reduce work—it intensifies it. Harvard Business Review. <https://hbr.org/2026/02/ai-doesnt-reduce-work-it-intensifies-it>
- Redefining student assessment in AI-infused learning environments: A systematic review of challenges and strategies for academic integrity. (2025). AI and Ethics. <https://doi.org/10.1007/s43681-025-00871-w>
- Richter, Schroeder & Wöhrmann (2009), « You don't have to believe everything you read: Background knowledge permits fast and efficient validation of information », Journal of Personality and Social Psychology

- Shoufan, A. (2023). Can students without prior knowledge use ChatGPT to answer test questions? An empirical study. *ACM Transactions on Computing Education*. <https://doi.org/10.1145/3628162>
- Shoufan, A., & Esmaeil, A.-A.-A. (2026). AI hallucination from students' perspective: A thematic analysis [Préprint]. *arXiv*. <https://arxiv.org/abs/2602.17671>
- Sun, Z., Wang, Q., Wang, H., Zhang, X., & Xu, J. (2025). Detection and mitigation of hallucination in large reasoning models: A mechanistic perspective [Préprint]. *arXiv*. <https://arxiv.org/abs/2505.12886>
- Surma, T., et al. (2025). Developing curriculum for deep thinking: The knowledge revival. Springer (open access). <https://doi.org/10.1007/978-3-031-74661-1> A systematic critical review of generative AI's impact on authorship, pedagogy, and integrity (2023–2025). (2026). *Frontiers in Education*. <https://doi.org/10.3389/feduc.2026.1769680>
- Tricot, A., & Sweller, J. (2014). Domain-specific knowledge and why teaching generic skills does not work. *Educational Psychology Review*, 26(2), 265–283. <https://doi.org/10.1007/s10648-013-9243-1>
- Turck, M.. (2026, 21 mai). OpenAI's Yann Dubois: Why AI progress suddenly feels real [Épisode de podcast vidéo]. *The MAD Podcast*. <https://www.youtube.com/watch?v=DhD1zZ8w8Mw>
- Vectara. (2026). Hallucination leaderboard [Dépôt GitHub]. <https://github.com/vectara/hallucination-leaderboard>
- Willingham, D. T. (2009). Why don't students like school? A cognitive scientist answers questions about how the mind works and what it means for the classroom. Jossey-Bass.
- Zhou, J., Zhang, J., Wan, R., Cui, X., Liu, Q., Guo, H., Shi, X., Fu, B., Meng, J., Yue, B., Zhang, Y., & Zhang, Z. (2025). Integrating AI into clinical education: Evaluating general practice trainees' proficiency in distinguishing AI-generated hallucinations and impacting factors. *BMC Medical Education*, 25(1), 406. <https://doi.org/10.1186/s12909-025-06916-2>

Notes

1 – Chef de projet modern workplace et digital learning, Ionis Education Group.

2 – <https://www.telos-eu.com/fr/les-progres-de-lia-mettent-en-tension-les-systemes-educatifs.html>
