Bloom's Taxonomy in the Age of AI

Hafiz Hanif

Faculty of Human Development, Sultan Idris Education University · ORCID 0000-0001-9943-6557

CITE AS	Hanif, H. (2026). Bloom's Taxonomy in the Age of AI (Working Paper No. 2026-01). Hack(edu).tech Chronicle. https://hackedu.tech/articles/blooms-taxonomy-in-the-age-of-ai
PUBLISHED	23 November 2026
DOI	10.5281/zenodo.24458690
LICENCE	CC-BY-4.0
AVAILABLE AT	https://hackedu.tech/articles/blooms-taxonomy-in-the-age-of-ai

ABSTRACT

Opening a series that rereads the frameworks teachers were trained on, this article argues that generative AI does not make Bloom's taxonomy wrong but cuts the link between a product and the thinking behind it, so a competent essay no longer evidences analysis or evaluation. Drawing on the 2001 revision's knowledge dimension and recent evidence on AI use and critical thinking, it argues that factual and conceptual knowledge become the precondition for supervising AI, that metacognitive knowledge becomes central, and that teachers should assess process rather than product.

KEYWORDS

Frameworks Revisited, Bloom's Taxonomy, Assessment, Artificial Intelligence, Pedagogy

This morning, the written papers of SPM 2026 begin. Over the next four weeks, 420,879 candidates will sit in exam halls across the country, pens in hand, phones surrendered at the door (SPM 2026 written examination to run from Nov 23 to Dec 17 - MOE 2026). Many of the questions they answer were written to target a particular level of thinking. Remember. Understand. Apply. Analyse. Evaluate. Create. Our teachers know the upper four by another name, KBAT, and many of us spent a good part of the last decade being trained to write questions that climb that ladder.

So it seems a fitting week to start a series I have been wanting to write for a while: a rereading of the frameworks we were all trained on, held up against a technology none of their authors imagined. I begin with the oldest and the most familiar. Bloom's taxonomy.

Most teachers carry Bloom's around as a pyramid, wide at the bottom with remembering and narrow at the top with creating, and with an unspoken moral attached: lower is basic, higher is better, good teaching drags students up the staircase. It is worth knowing that the pyramid is not in the original. The 1956 handbook was a classification of educational objectives, produced by a committee of university examiners who wanted a common language for describing what their test items were actually testing (Bloom et al. 1956). It was, at heart, an assessment tool. When Anderson, Krathwohl and colleagues revised it in 2001, they did two things that most of our training glossed over (Anderson et al. 2001). They renamed the categories as verbs and

moved create to the top, which we all noticed. And they added a second dimension, the type of knowledge involved: factual, conceptual, procedural and metacognitive. Krathwohl was explicit that the revised version relaxed the strict cumulative hierarchy of the original, so that the categories were allowed to overlap (Krathwohl 2002). The taxonomy became a grid. We kept drawing a pyramid.

Now bring in the machine.

Ask a generative AI to analyse the causes of the Malacca Sultanate's decline, evaluate two arguments about it, and create a new interpretation, and it will do all three in about twenty seconds, in fluent Malay or English, at a length of your choosing. Every verb at the top of our pyramid, produced on demand. The products that we spent years teaching students to make, and that we used as evidence that they could think at those levels, can now be made by something that does not think at all, in the sense that matters.

This is the first thing the age of AI does to Bloom's taxonomy, and it is not that the taxonomy is wrong. It is that the link between a product and the thinking behind it has been cut. For seventy years, if a student handed in a competent evaluation, we could reasonably infer that the student had evaluated. That inference was the whole basis of using the taxonomy to grade work. A well-built essay was evidence of the processes that built it. The essay is no longer evidence of anything except that an essay exists.

Which leads to a slightly uncomfortable observation about this week. The SPM hall, with its invigilators and its sealed papers and its rule against phones, is now one of the few places left in a student's year where a create-level answer is still good evidence of create-level thinking. The exam we often criticise for rewarding memorisation is, by accident of its old-fashioned setting, more trustworthy about higher-order thinking than most of the homework and coursework we set during the year. The problem is not SPM. The problem is everything else we assess, where the conditions that made the product meaningful have quietly disappeared.

The second thing AI does is more interesting, and it turns the pyramid's moral upside down.

We learned to treat the bottom of the taxonomy as the boring part. Remembering facts, understanding concepts, the stuff we were told not to dwell on in an age of search engines. But consider what it actually takes to evaluate an AI's answer. To judge whether the machine's analysis of the Malacca Sultanate is sound, you need to know enough about the Malacca Sultanate to notice when it is wrong. Evaluation without knowledge is not evaluation. It is guessing with confidence. A student with nothing in the factual and conceptual cells of the grid can only accept or reject the machine's output on vibes, and machines are very good at producing output that feels right.

The research is beginning to say this out loud. In a large field experiment with nearly a thousand high school mathematics students, students given unrestricted access to a GPT-4 tutor improved their practice scores markedly, then did worse than students who never had access once the tool was taken away; a version designed with guardrails, which pushed students to work rather than handing over answers, largely avoided the harm (Bastani et al. 2025). A survey of knowledge workers found that the more people trusted generative AI to do a task, the less critical thinking they reported applying to it, while those more confident in their own abilities reported more (Lee et al. 2025). And a study of 666 participants found that heavier use of AI tools was associated with lower critical thinking scores, mediated by cognitive offloading, with younger users the most dependent (Gerlich 2025). Read together, these are not arguments against AI in the classroom. They are arguments that the foundation matters more than ever, because the foundation is what lets a learner stay in charge of the tool instead of being carried by it.

So the bottom of the pyramid is not the cheap part any more. It is the part that makes the top of the pyramid possible to supervise.

The third change concerns the cell of the grid that most of us skipped entirely: metacognitive knowledge. Anderson and Krathwohl put it there in 2001, knowledge about one's own thinking, about which strategies to use, about what one does and does not know (Krathwohl 2002). It always looked a little abstract on the chart. It is not abstract now. Knowing when to use the machine, when not to, what it is likely to get wrong, and when your own understanding is too thin to check its work, is metacognitive knowledge in its purest form. It might be the single most important cell in the whole grid for a student growing up with AI, and it is the one we are least practised at teaching or assessing.

What does this mean for a teacher on an ordinary Tuesday? A few things, none of them requiring new software.

First, assess the process, not only the product. If the essay is no longer evidence, the drafting is. Ask students to show the path: the first attempt, the notes, the questions they asked, the part they changed after feedback, the part where they disagreed with the AI and why. A short oral defence of a written piece, five minutes at the teacher's desk, will tell you more about a student's analysis than the piece itself.

Second, stop apologising for knowledge. Teaching facts and concepts well is not old-fashioned in the age of AI. It is the precondition for every higher-order use of the tool. A student who knows the material can use AI to go further. A student who does not can only use it to go around.

Third, teach the metacognitive layer explicitly. When students use AI in a task, build in a moment where they must say what they asked, what came back, what they checked and how, and what they decided not to trust. This is the evaluate level and the metacognitive dimension working together, and it is assessable.

Fourth, when you write higher-order questions, ask whether the question can be answered by a machine in a way that looks like thinking. If it can, the question still has value as a learning task, but it no longer works as evidence on its own outside a supervised setting.

None of this retires Bloom's taxonomy. If anything, it rescues it from the pyramid. The taxonomy was always better understood as a map of different kinds of thinking than as a staircase of increasingly impressive products, and that is exactly the version we need now. The machine can produce the products at every level of the map. It cannot do the thinking for the student, and it cannot tell us whether the student has done it. That part remains ours.

To the 420,879 candidates walking into their halls this week: for these few hours, the work on your paper is yours alone, and everyone reading it will know that. It is a strange gift, in the age you are growing up in. Use it well.

Next Monday: SAMR, and what "redefinition" means when the tool redefines itself every few months.

References

[1] Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (2001). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. Longman.

[2] Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. Proceedings of the National Academy of Sciences, 122(26). https://doi.org/10.1073/pnas.2422633122

[3] Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain. David McKay.

[4] Gerlich, M. (2025). AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking. Societies, 15(1), 6. https://doi.org/10.3390/soc15010006

[5] Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. Theory Into Practice, 41(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2

[6] Lee, H.-P. (Hank), 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. In Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (pp. 1–22). ACM. https://doi.org/10.1145/3706598.3713778

[7] SPM 2026 written examination to run from Nov 23 to Dec 17 - MOE. (2026). Bernama. https://www.bernama.com/en/general/news.php?id=2596007