AI-ENABLED LEARNING PRODUCT STRATEGY

Navigating the Future of Learning Technology

Curtiss J. Shupe

AI-Enabled Learning Product Strategy

AI learning products won't win by sounding smarter—they'll win by proving people actually learn better.

The most effective AI-based products for learning won't succeed because they are able to produce more content more efficiently. They will succeed because they make the process of learning more effective. That means they will make more people start to learn, improve the identification of skill gaps, make the content more relevant, make the practice and assessment more effective, and improve retention of what has been learned over time. The Organisation for Economic Co-operation and Development's (OECD) 2026 guidelines are helpful here too. Their view is that general-purpose generative AI products are helpful when used in conjunction with a set of clear principles for teaching; however, more promising are those AI products that are specially designed for use in learning. The United States (U.S.) Department of Education and the United Nations Educational, Scientific and Cultural Organization (UNESCO) are also promoting the use of AI for learning from a governance perspective; they say it needs to be evidence-based, related to educational objectives, transparent so it can be inspected, and used to assist human judgment rather than replace it.

The distinction is important because so many product strategies for AI simply amount to adding a feature and declaring the platform intelligent. Add a chatbot, auto-gen some quiz questions, and summarize some content. Voilà. This is not much of a strategy. The actual warning from the OECD is more specific and less convenient. GenAI is not magic. It can enhance good pedagogy and bad pedagogy equally. Education systems should support the former and not the latter. In product terms, the question is not where to insert the AI. The question is where the AI can improve the learner journey without diminishing effort, trust, or signal quality.

The first place that matters is learner activation. In many products, the large drop-off occurs before the sustained learning even begins. Learners fail to complete the onboarding process, skip administrative tasks, or never build up enough steam to get into practice. AI helps here by reducing friction and improving timing. The chatbot results at Georgia State matter because they demonstrated positive effects on completion of required pre-matriculation tasks and timely enrollment. The results also demonstrated that messages related to specific registration holds had a strong effect on advising related to those holds. Messages had a weaker effect overall. A separate field experiment by the National Bureau of Economic Research found that using streaks and reminders improved use of an online math program. Each had a separate effect on a different margin of the outcome. The first point to note here is that activation is a design problem with different types of interventions. Activation is not a communication problem with a mascot and a 'send' button.

The second strategic domain is skills discovery and mapping. Clearly, a learning product has more potential value if it can develop a credible picture of what the learner knows and what the learner still needs to know, and how the learner might best get from the first to the second. OECD's skills-first work is interesting because it provides a picture of a labor market that increasingly signals skills directly. It also argues that a labor market with rapidly shifting demand requires upskilling opportunities that are rapid, targeted, and accessible, such as microcredentials, online learning, and workplace training. It also points to the increasingly prominent role of taxonomies, skills-based assessments, and digital records of capabilities. That's interesting for learning product strategy. Skills mapping is not trivial metadata. It should relate learning activity, practice, assessment, and recognition in ways that make capability claims more fine-grained and more believable.

That brings us to the third domain: Relevance. Relevance in AI-enabled learning products is not simply a measure of how good the recommendations are; it is whether or not the product provides the learner with the right content, practice, and guidance at the right time in their learning progression, at the appropriate level of difficulty, and in a way that makes sense within the learner's and the educator's context. The OECD's work on purpose-built education tools and the Department of Education's work on alignment with education goals are both headed in the same direction: good products will lower search costs for the learner, create targeted pathways for the learner, and adjust their support in ways that are easy to follow for both the learner and the educator. We are not trying to maximize personalization; personalization is futuristic for a demo. We are trying to maximize relevance in a way that makes sense to both the learner and the educator without obscuring how a particular piece of information was recommended.

The fourth domain is the design of practice across courses, labs, and simulations. This is where AI-powered learning strategy truly crosses the line from recommendation engines into efficacy. The research looks encouraging, but it's not a free pass. In a 2025 systematic review published in npj Science of Learning, for instance, it was found that intelligent tutoring systems had overall positive effects on learning and performance in kindergarten through 12th grade (K-12) settings, though the effects were qualified if the comparison was to a strong non-intelligent tutoring system instead of weaker baseline instruction. In another 2025 study, also designed as a randomized controlled trial, it was found that the research-based AI tutor produced significantly higher post-test performance in less time than in-class active learning in the setting, while students also reported higher engagement and motivation. The moral isn't that AI-powered tutoring will always beat instruction. It's that design of tutoring, labs, and practice, including sound pedagogy, not just generic dialogue with a confidence problem, can drive product advantage.

This is also where mobile matters, but again with more nuance than was suggested in the original draft. Mobile devices are not simply smaller screens or conduits for delivering content. In a good product for learning, they could be helpful for delivering prompts for spaced repetition, retrieval opportunities, reflection, and workflow reinforcement. But are reminders really helpful? According to a 2024 npj Science of Learning study, smartphone-based reminders were effective for increasing the probability of studying on those reminder days but also showed signs of over-reliance and did not improve test performance compared to control conditions. The more defensible product conclusion is that mobile devices could be helpful for orchestrating these types of spacing and retrieval activities, but only if reminder design is more supportive of self-regulated learning than it is substituting for it. The Institute of Education Sciences continues to support spacing of learning over time and the use of quizzes for retrieval practice and spaced presentation of information as evidence-based practices for reducing forgetting.

The fifth domain is assessment quality. This is where many AI-enabled learning products are strategically underbuilt. Assessment is more than simply speeding up scoring or question generation. The work of Educational Testing Service (ETS) on the future of assessment is important because it suggests more innovative forms of assessment, such as gamified and interactive experiences and assessments of important skills in immersive environments that are relevant to those skills. Their work also recognizes that Gen AI makes aspects of the test development cycle, such as personalized feedback, more efficient. For product leaders, it means that assessment is more than simply a measure of course completion. It is a measure of starting capabilities, next-step learning needs, performance in realistic environments, and interpretable results that have implications beyond the course or product itself. If a product cannot generate a reasonable level of evidence of skill development, then the recommendation and curation layers are also strategically half-built.

This necessity immediately raises the toughest problem with AI-enabled learning strategy: trust. ETS' model of a responsible AI approach to measurement and learning posits that AI "should improve validity, reliability, fairness, and utility," which "requires rigorous R&D, quality assurance, stakeholder engagement, continuous improvement, communication, and governance." The Department of Education's requirements also include inspectability, explainability, human alternative solutions, and the capacity of teachers to override AI suggestions if needed. The overall normative statement by UNESCO speaks to human-centered use that is "ethical, safe, and equitable." This means that recommendations, feedback, and skill inferences must be contestable and subject to human review if the stakes are high. Trust in learning products isn't created by a clean interface and soothing voice. Trust is created by evidence quality, transparency, and recourse.

The implication of this analysis for product managers of products used for learning is that AI-enabled learning must be designed as a workflow architecture, rather than a collection of features. That means that discovery, diagnosis, guidance, practice, feedback, assessment, and skill signaling must function as one system. Recent evidence from an NBER paper analyzing a field experiment with Khan Academy in India offers useful insights. The paper found that large gains are not driven by access to the platform but by the structure of implementation that ensures connectivity, practice time protection, supervision of use, coordination of content with teachers, and monitoring of progress. In other words, technology performance was inextricably linked with workflow and operations. This is the same implication for enterprise learning products. AI features will only have value if designed as part of a learning operating model, rather than draped over a fragmented experience and expected to look strategic.

That architecture also affects what we need to measure. Content views and completion rates are important operationally but are not sufficient for product strategy. A more robust measure needs to include metrics such as time to first meaningful action, activation after onboarding, recommendation relevance, frequency and quality of practice, adherence to spaced retrieval, assessment reliability, fairness drift, and verified skill progression. These are not all directly validated as a single framework, but they follow from the evidence on activation, practice quality, assessment trust, and implementation. Georgia State illustrates the need for targeted friction reduction for activation. The mobile reminder literature illustrates the potential for engagement without durable learning to mislead. ETS makes it clear that validity, reliability, fairness, and utility need to remain front and center. And the Khan Academy implementation study makes it clear that it is use, not access per se, that drives much of the learning gain.

So the strategic opportunity isn't to build a learning product that sounds intelligent. It's to build one that learns about the learner responsibly, guides them precisely, gives them better practice, produces more credible evidence, and does all of that at scale. The strongest AI-enabled learning products will look less like content warehouses with chat interfaces and more like governed performance systems for activation, progression, and skill recognition. That's the version of AI-enabled learning product strategy most likely to matter to serious buyers, enterprise operators, and hiring managers.