Emanuel Nestler

A matter of responsibility and disciplinary expertise: Using large language models in higher education

'Go outside. Identify animals, plants, and fungi. Investigate them with your students. Discuss your findings and enjoy biology.' If I were a large language model, this is how I might respond when prompted by a young pre-service teacher: 'I am a five-year-old child. Describe how to teach biology in no more than four sentences.'

When we discuss AI, particularly LLMs, in higher education, we often focus on their effects and outputs: faster completion of written assignments, better translations of academic papers, and further promises – or fears about the consequences of the new technology. Although the effects of LLMs are already apparent, neither the future shape of higher education nor their consequences for individual disciplines can yet be predicted reliably.

Drawing on research in biology teacher education [1, 2], this article examines the use of LLMs in the preparation of pre-service biology teachers and the school-based mentors who support them. The question of how to use LLMs in teaching and learning is also a recurring theme [3]. Before considering possible responses, however, we must clarify what an LLM is [4]. At a basic level, an LLM generates text by predicting probable linguistic sequences from patterns identified in large datasets. It simulates an answer based on a massive amount of text data. This is why users need to assign the LLM a role in the prompt, for example: "Answer as a scientist …". An LLM's answer is not based on understanding. Its outputs may simulate features of understanding without constituting understanding in the human or epistemic sense. LLM outputs may therefore display features associated with competent performance, although the model itself cannot be assumed to possess the corresponding competence.

An LLM-generated text may be of higher apparent quality than a text written by a student. For educational purposes, however, the relationship between the submitted product and the student's understanding is crucial. Learning involves a transformation of the learner. This transformation always requires both a reason and effort. There is no such thing as effortless learning. It can *feel* effortless, but it must involve personal transformation. There are effective ways to reduce the effort required for learning, but this effort cannot be eliminated entirely. For example, educators can break down complex learning goals into manageable steps, teach effective learning strategies, and make the purpose of learning explicit.



Fig. 1: Technologies such as games (pictured) or AI can enhance science education, but the core of the subject must remain the focus (Picture: Torsten Kreher)

The focus of higher education teachers and lecturers can be the learning of the student or the learning products like homework or papers. Examinations and written assignments are commonly treated as evidence of learning. The availability of LLMs weakens the inference from the quality of a submitted product to the quality of the student's learning. The validity of this inference must therefore be examined rather than assumed. The seemingly obvious solution is to ban LLMs or discontinue assessments based on work that LLMs can generate. But we live and research in a transforming world with LLMs. We should therefore ask which skills future researchers and university educators will need, and transform our teaching and assessment system accordingly. I can't answer that question for you (maybe ChatGPT can simulate an answer). Three principles nevertheless remain central to teaching and learning.

First, we need to take responsibility [5, 6] for our teaching and the transformation of teaching and learning in our subject. As noted above, transformation and learning require effort. We as educators need to accept that we are change agents that transform teaching in times of AI. The responsibility for teaching, feedback and assessment cannot be delegated to an LLM.

Second, in the educational discourse opportunity-use-models of teaching and learning are used to conceptualise both [7].

Dr. Emanuel Nestler
Institute of Bioscience
University of Rostock
Universitätsplatz 4, 18055 Rostock
Emanuel.nestler@uni-rostock.de

DOI-Nr.: 10.26125/d8ma-qt19

LLMs can influence both the design of learning opportunities and the ways in which students use them. While you take responsibility for the learning opportunity, your students also need to take responsibility for their own learning. The reflection on that should be part of the created learning opportunity. Students may minimise effort when they perceive little value in a task. Learning environments should therefore make the purpose and disciplinary relevance of demanding activities explicit.

Third, you need to think about the core ideas of your subject or domain as well. Subjects developed historically, and content was assigned to them. In German there is the term 'Fachlichkeit' [8, 9]. No single English term fully captures the German concept. In this article, disciplinary expertise refers to the discipline-specific ways of knowing, reasoning and acting, together with the standards by which disciplinary work is evaluated. What is important in this subject, and what are the best ways to ensure high quality? In science education, scientific inquiry and the nature of science provide examples of discipline-specific practices and metacognitive knowledge about how scientific knowledge is produced and evaluated. The opening of this article illustrates these ideas: Engage directly with what we study. Use scientific methods. Think and discuss.

If we think about our teaching, we can use these disciplinary principles as criteria for evaluating LLM use in teaching and research. LLMs may support disciplinary practices when they are used as critical readers, for example to suggest alternative perspectives, identify potentially relevant literature or improve linguistic clarity. Any sources suggested by an LLM must be verified against bibliographic databases and the original publications; otherwise, the evidential quality of the text is compromised.

These questions are not entirely new. Similar debates accompanied the introduction of machine-learning applications for image-based species identification. The discussions were very similar: Do we need to know species? Should we use books instead of these apps? Should we use apps even if they aren't as accurate as scientists like entomologists? Today, parents, teachers and pupils can use these applications to support the exploration of local biodiversity. These apps can also help scientists learn to identify unfamiliar species. When scientific accuracy is required, researchers consult an expert on the relevant taxonomic group. Where scientific validity or personal safety is at stake – for example, when identifying an edible wild mushroom – expert verification remains essential.

Responsibility, disciplinary core ideas and discipline-specific quality standards should guide decisions about LLM use. Because higher education prepares future researchers, students should learn to articulate and apply these criteria themselves.

References

[1] Nestler, E., Retzlaff-Fürst, C. & Groß, J. (2023). Assessing the Quality of Mentoring: Evidence-Based Development of a Training Programme for Biology Mentors and Pre-service Biology Teachers for Scientific Issues. In K. Korfiatis, M. Grace & M. Hammann (Eds.), *Contributions from Biology Education Research: Shaping the Future of Biological Education Research. Selected Papers from the ERIDOB 2022 Conference* (pp. 293–309). Springer International Publishing.

[2] Nestler, E. (2024). Die biologiespezifische Qualifizierung von Mentorinnen zur Förderung der naturwissenschaftlich-biologischen Erkenntnisgewinnung in den Schulpraktischen Übungen Biologie: Eine Mixed-Methodology-Studie zur Wirkung einer biologiespezifischen Mentorinnenqualifizierung [*Kumulative Dissertation*, Universität Rostock].

[3] Nestler, E. (in press). Eine Frage der Verantwortung – Eine Analyse zu KI und Mentoring für angehende Lehrkräfte.

[4] Li, S., Wong, K. W., Wang, G. & Duong, T.-T. (2025). A systematic review of multi-modal large language models on domain-specific applications. *Artificial Intelligence Review*, **58(12)**. https://doi.org/10.1007/s10462-025-11398-1

[5] Eckert, J. (2016). Tugend, Recht und Moral: Tendenzen der Verantwortungszuschreibung. Deutsche *Zeitschrift für Philosophie*, **64(2)**, 246–257. https://doi.org/10.1515/dzph-2016-0018

[6] Kuhlmann, N. (2023). „Verantwortung" als pädagogischer Topos: Anerkennungstheoretische Perspektiven (1st ed.). Beltz Juventa.

[7] Keller-Schneider, M., & Albisser, S. (2012). Grenzen des Lernbaren? Ergebnisse einer explorativen Studie zum Erwerb adaptiver Unterrichtskompetenz im Studium. In T. Hascher & G. H. Neuweg (Eds.), *Forschung zur (Wirksamkeit der) Lehrer/innen/bildung* (pp. 85–103). Wien: LIT Verlag. https://doi.org/10.5281/zenodo.7361334

[8] Reh, S. & Pieper, I. (2018). Die Fachlichkeit des Schulfaches: Überlegungen zum Deutschunterricht und seiner Geschichte zwischen Disziplinen und allgemeinen Bildungsansprüchen. In M. Martens, K. Rabenstein, K. Bräu, M. Fetzer, H. Gresch, I. Hardy & C. Schelle (Eds.), *Konstruktionen von Fachlichkeit: Ansätze, Erträge und Diskussionen in der empirischen Unterrichtsforschung* (pp. 21–41). Klinkhardt.

[9] Nestler, E., Malmberg, I., Heinrich, G. & Retzlaff-Fürst, C. (2021). Praxisphasen als Räume der Konstruktion von Fachlichkeit. In T. Leonhard, P. Herzmann & J. Košinár (Eds.), *„Grau, theurer Freund, ist alle Theorie"? Theorien und Erkenntniswege schul- und berufspraktischer Studien* (pp. 163–178). Waxmann.



Dr. Emanuel Nestler

Emanuel Nestler qualified as a secondary-school teacher of biology and social sciences at the University of Rostock and subsequently taught both subjects. In 2024, he earned his doctorate in biology education with a dissertation on biology-specific mentor training for scientific inquiry. As a research associate at the University of Rostock, he works on teacher professionalisation, digital and game-based learning, and STEM networks. He currently coordinates mentor training and the NORDMINKT cluster for STEM and STEAM education.