GenAI as a Partner for Lesson Planning: Insights from an Exploratory Study with Pre-Service Science Teachers

Daniela Fiedler

Department of Science Education
University of Copenhagen

Introduction

Generative Artificial Intelligence (GenAI) is rapidly driving a paradigm shift in the creation and dissemination of human intellectual work (Kim, 2026). While it offers transformative potential in education, it also poses significant ethical and pedagogical challenges (e.g., Cooper, 2023; Kim, 2026). Although GenAI is often described as a new "frontier" in science education, systematic reviews and survey studies indicate that its integration into teacher education curricula remains in its infancy (e.g., Prilop et al., 2025; Spasopoulos et al., 2025). Indeed, pre-service teachers encounter GenAI tools in dual roles: as higher education learners using GenAI (for instance, to clarify complex concepts), and as future educators who must determine how to integrate such technologies into their teaching practice (Gamlem et al., 2026). While research suggests that experienced educators increasingly use GenAI to enhance instruction (Prilop et al., 2025), many also remain cautious, expressing concerns about the erosion of core teaching competencies and ethical ambiguity (Gamlem et al., 2026).

Navigating these challenges requires sufficient GenAI literacy, understood as a multidimensional competence encompassing (1) technical proficiency, (2) critical evaluation, (3) communication skills (such as prompt engineering), (4) creative application, and (5) ethical-epistemic awareness (Huang et al., 2025; Valeri et al., 2025). Recent research highlights the triadic nature of this literacy, conceptualizing GenAI simultaneously as a teaching tool (for preparation), teaching content (AI as a topic), and a learning tool for students to navigate (Prilop

et al., 2025). Developing GenAI literacy is particularly vital because GenAI systems often present themselves as 'epistemic authorities', even though they produce hallucinated information and reflect historical societal biases (e.g., Arantes, 2025; Cooper, 2023; Cooper & Tang, 2024).

In science education, these issues are further amplified by a broader post-truth climate in which evidence and expertise are increasingly contested (Arantes, 2025; Cooper, 2023). Educators are therefore encouraged to move beyond viewing GenAI either as a threat to their teaching or as a mere shortcut for administrative tasks and instead engage with it collaboratively – as a partner in dialogue and 'co-design' of instructional settings (Spasopoulos et al., 2025; Tang & Putra, 2026). While GenAI can support the structuring of lesson components, for instance, within the 5E model (*Engage, Explore, Explain, Elaborate, Evaluate*; Tanner, 2010) or its Danish equivalent, the 6F model (*Forudsættning, Fang, Forsk, Forklar, Forlæng, Feedback*; Madsen et al., 2020), its effective use by (pre-service) teachers requires robust pedagogical judgment and subject-matter expertise to identify mistakes and align GenAI outputs with instructional intentions (Cooper, 2023; Kim, 2026).

In my pedagogical development project, I investigated how pre-service science teachers experience GenAI as a lesson-planning partner during a master's course for future upper secondary science teachers.

Course Context and Integration of GenAI

The *Advanced Science Didactics* course is an eight-week, Master's-level science education course designed for future upper secondary science teachers (Danish: *Gymnasielærer*) to develop pedagogical competence in planning, executing, and reflecting on high school science lessons. Classes meet weekly for approximately 1.5 days.

In this iteration, a colleague and I taught the course and jointly decided to integrate GenAI as a central pedagogical component by giving students structured opportunities to engage with GenAI-supported lesson planning. We both see GenAI as another digital tool, like interactive simulations, that teachers need to know and can integrate pedagogically

and reflectively into science lessons. We believe that future teachers need to move from being passive recipients of GenAI answers to explorers and active authors of their lesson designs by treating GenAI as a supportive tool rather than a replacement for pedagogical judgment (Kim, 2026). The course structure described below, including class exercises and assignments, shows how we intended to support the five GenAI literacy dimensions introduced above (i.e., technical proficiency, critical evaluation, communication skills, creative application, and ethical-epistemic awareness).

We formally introduced students to GenAI tools through a 6F lesson exercise exploring the epistemic-authority framing of GenAI tools in the second week of their teaching (Cooper, 2023; Cooper & Tang, 2024). Throughout the following weeks, students were encouraged to use GenAI as a sparring partner during class activities and for their own lesson planning. Therefore, students should actively engage in conversations with GenAI systems to generate or refine ideas and lesson components, such as drafting student activities or predicting misconceptions (Kim, 2026; Spasopoulos et al., 2025). In addition, course assignments included a task in which students actively used GenAI tools for lesson planning and reflected on GenAI-generated ideas or conversations while evaluating accuracy, epistemic quality, and alignment with pedagogical intentions. Students were also asked to design at least one science lesson plan that includes an activity in which their future pupils would work directly with GenAI, requiring them to adapt the tool creatively for student-facing contexts. We finalized the experiment in the last week of teaching through a 6F lesson exercise to reflect on students' use of GenAI throughout the course. This exercise was particularly meant to help students understand their strategies and habits, and support deeper thinking about their future teaching practice.

Pre- and Posttest Questionnaire

Since I was particularly interested in students’ development of GenAI literacy and school-related perspectives, as well as their perceived challenges and benefits of using GenAI to design innovative learning environments, I administered a short pretest questionnaire in the first

week of teaching (prior to any mention of GenAI in the course) and a posttest questionnaire in the final week (after the reflection exercise described above). Both questionnaires included Likert-scale items on GenAI literacy adapted from Valeri et al. (2025), along with multiple-choice and open-response questions about general GenAI use. The posttest also included items assessing perceived challenges and benefits of using GenAI in lesson planning, open-ended responses describing successful and less effective uses of GenAI, and their intention to use GenAI for future lesson planning (see Appendix for the full list of items used in the questionnaires).

While the original items from Valeri et al. (2025) were mainly meant for presenting situational distributions, I decided to construct two sum scales based on theoretical considerations that can be used for simple inferential statistics: First, the scale *GenAI Literacy* (7 items) assesses students' perceived ability to understand, evaluate, and effectively use GenAI in STEM teaching contexts, while the scale *GenAI School* (3 items) captures broader perspectives on the role of GenAI in school practices and institutional adoption (Fig.1).

GenAI Literacy	**GenAI School**
1. I can detect potential inaccuracies in responses generated by GenAI tools such as ChatGPT.	1. GenAI tools such as ChatGPT can be a substitute to traditional learning materials such as books.
2. I am familiar with the concept of machine learning, which is the technology behind GenAI tools such as ChatGPT.	2. The use of GenAI tools like ChatGPT in schools will increase in the coming years.
3. I adjust my questions to GenAI tools such as ChatGPT depending on the STEM subject I am working with.	3. I believe it is important to discuss how GenAI tools such as ChatGPT can be used in schools
4. GenAI tools such as ChatGPT help me understand difficult concepts in STEM subjects.	
5. I have a good understanding of how GenAI tools such as ChatGPT generate texts.	
6. I double-check the information generated by GenAI tools like ChatGPT against other sources	
7. I feel that I have sufficient knowledge to use GenAI tools such as ChatGPT to support my studies in the STEM subjects.	

Fig. 1. Statements used for the scales GenAI Literacy and GenAI School.

Each scale included only items with a unidirectional relationship, ensuring that higher agreement consistently reflects higher levels of literacy or awareness. Hence, items assessing trust in GenAI outcomes, sustainability, or data privacy

concerns were excluded, as both higher and lower levels may reflect either naïve acceptance or critically informed judgment. Including such items in a composite pre-post comparison could lead to ambiguous or misleading interpretations of change.

For each scale, I summed responses to the individual statements (e.g., from 1 = strongly disagree to 5 = strongly agree) into a single total score per participant and time point. Given the small sample size ($N = 10$) and the ordinal nature of the underlying data, I assessed pre-post changes using exact Wilcoxon signed-rank tests and reported effect sizes (r) to indicate the magnitude of change.

Results and Discussion

Before the course, students primarily used GenAI for operational purposes, such as *understanding concepts* (80%) and *correcting, summarizing, or translating materials* (60% each). The open-response answers also highlighted concerns about *inaccurate or unreliable output* (7 students) and difficulty *formulating effective prompts* (2 students).

Development of GenAI Literacy

GenAI Literacy sum scores increased from pretest ($M = 24.7$, $SD = 3.8$) to posttest ($M = 28.7$, $SD = 3.7$), with most participants showing gains over the course (Fig. 2). This change proved to be statistically significant with a large effect size ($p = .014$, $r = .82$).

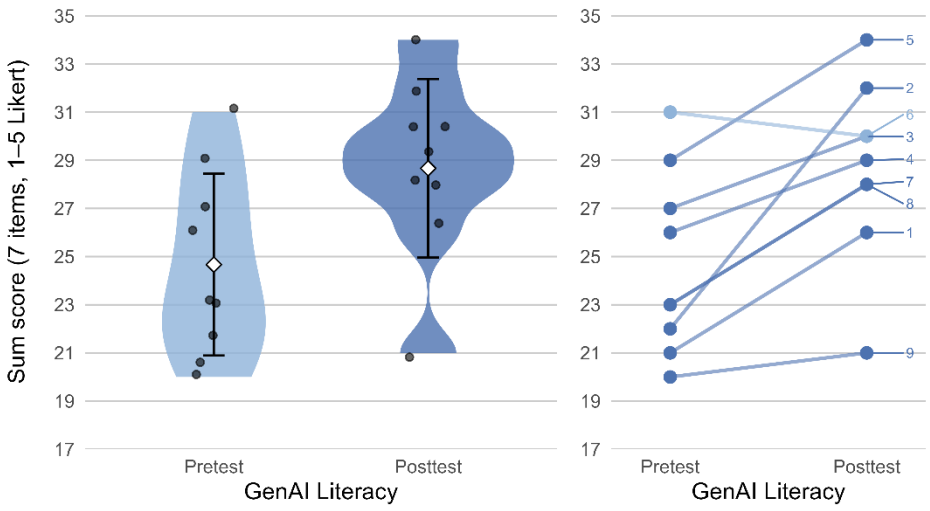


Fig. 2. Sum scores of GenAI Literacy in the first (pretest) and last (posttest) week of the course (n = 9). The right panel depicts individual changes.

Among the items showing the greatest changes, a pattern emerged linking increased knowledge to more reflective GenAI use. Students reported greater familiarity with the underlying technology (machine learning), along with a stronger tendency to cross-check GenAI-generated information against other sources and adapt their prompts to the STEM subject at hand. Hence, integrating GenAI throughout the course seemed to help build a stronger technical understanding of GenAI tools, while repeated prompt formulation, output evaluation, and iterative lesson planning may have strengthened their ability to use GenAI critically and reflectively in relation to subject-specific content and pedagogical goals.

Changes in GenAI School Perspectives

GenAI School sum scores showed no meaningful change from pretest ($M = 11.3$, $SD = 2.4$) to posttest ($M = 11.8$, $SD = 2.9$). However, individual trajectories showed slightly more increases (6 students) compared to decreases (2 students) or no change (1 student; Fig. 3).

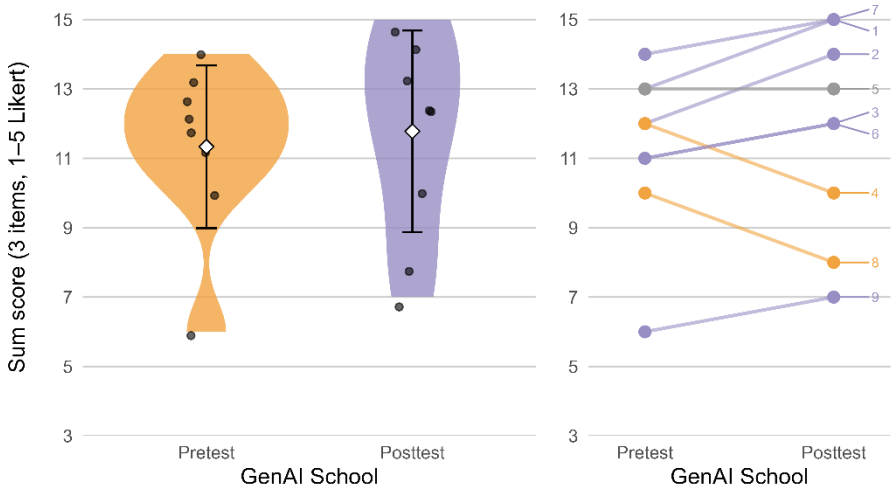


Fig. 3. Sum scores of GenAI School in the first (pretest) and last (posttest) week of the course (n = 9). The right side depicts individual changes for the students, with lilac indicating an increase, orange a decrease, and grey no change in the individual sum scores from pretest to posttest.

The item-level analyses also indicated that students in the pretest already mostly agreed that GenAI use will grow in schools but also needs sufficient discussion. In contrast, students were more divided on whether GenAI tools can substitute for traditional learning materials (evenly divided among agree, neutral, and disagree). The findings suggest that while students anticipated broader institutional adoption of GenAI, they remain cautious about replacing established learning materials and already recognize the need for ongoing discussion and critical reflection regarding GenAI use in schools.

Benefits and Challenges of GenAI Use for Lesson Planning

Participants indicated both clear benefits and notable challenges when using GenAI for lesson planning (Fig. 4).

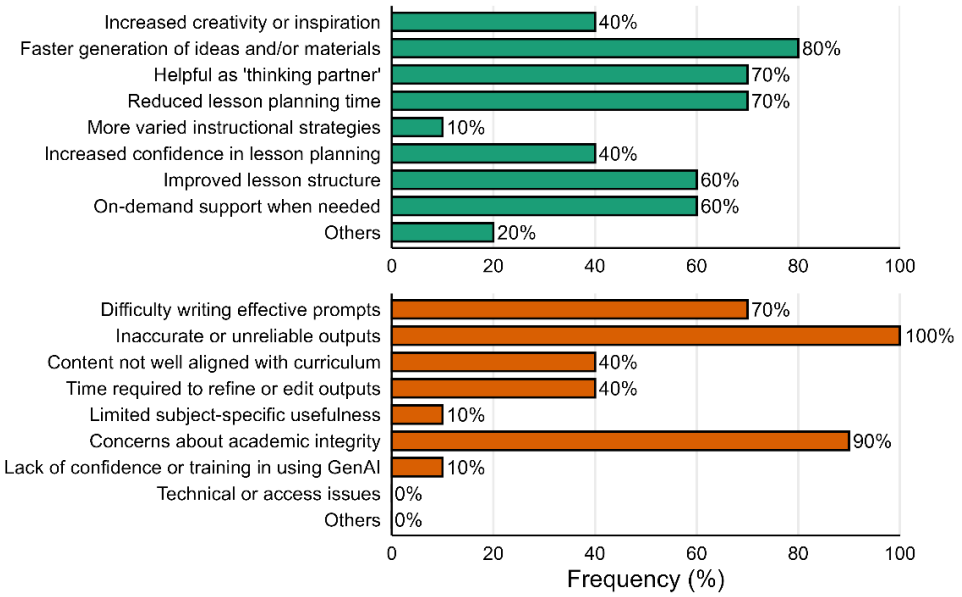


Fig. 4. Students' reported benefits (top panel) and challenges (bottom panel) experienced during GenAI-supported lesson planning (N = 10). Percentages reflect the proportion of students selecting each option in the posttest multiple-choice items.

Most found GenAI helpful for generating ideas and overcoming initial planning barriers. Several also described it as a useful ‘thinking partner’. One student wrote:

> “Creating a full 6F within my selected topic with great accuracy. Gave me a great start for creating my own lesson plan.” (dnv07)

At the same time, significant challenges persisted. All students reported encountering inaccurate or unreliable output, and many described the difficulty of adapting GenAI suggestions to fit their personal views or ideas. One student commented:

> “If you have a very specific idea, it takes a lot of time, if even possible, to get the result that you want.” (dnv02)

Overall, the perceived benefits outweighed the challenges for most students. Seven out of the ten students reported more benefits than challenges, whereas three students experienced the opposite pattern. However, all students encountered both advantages and limitations when working with GenAI. Thus, participants generally viewed GenAI as a valuable aid for lesson planning, while recognizing that effective use of AI-generated suggestions still requires critical evaluation, professional judgment, and subject-matter expertise.

Intention to Use GenAI in the Future

Eight of the ten participants indicated that they intend to use GenAI for future lesson planning. Students again emphasized its usefulness for idea generation and as a ‘dialogic partner’:

> “It saves time and helps you see things from a different angle or perspective when you don’t have anyone to talk to.” (dnv05)

Those expressing hesitation cited concerns about the time required to adapt GenAI outputs:

> “Because good to have as backup but faster to do self.” (dnv04)

Together, the responses indicate a pattern of selective, reflective adoption: students value GenAI as a complementary tool while maintaining a critical stance towards its limitations.

Conclusion and Implications

The results of this pedagogical development project show that structured, iterative engagement with GenAI can meaningfully support GenAI literacy development among pre-service science teachers. Students appeared to develop more reflective, pedagogically grounded interactions with GenAI tools, especially for lesson planning. This mirrors recent research on GenAI literacy more broadly (Huang et al., 2025; Prilop et al., 2025). The significant increase in students' literacy scores suggests that structured GenAI use can strengthen epistemic vigilance and help students position it as a tool that requires active professional judgment rather than passive consumption (Arantes, 2025; Cooper, 2023).

At the same time, students' perspectives on GenAI's broader role in schools remained largely stable across the course. Item-level patterns suggest that students anticipated broader institutional adoption of GenAI while remaining cautious about replacing established teaching resources, reflecting results that position GenAI as a supplementary rather than a replacement resource for pedagogical expertise (Valeri et al., 2025). Our students' strong endorsement of the need for ongoing discussions mirrors calls in the research community for institutional guidelines, transparency norms, and ethical frameworks to help educators navigate issues such as bias, hallucinations, and academic integrity (Cooper, 2023).

The coexistence of benefits and challenges reported by students further highlights the need for (teacher) education programs to prepare (future) teachers for GenAI-supported planning. Our students recognized GenAI as a productive 'thinking partner' for idea generation and early structuring, echoing findings that GenAI can reduce cognitive load during the initial phases of lesson planning (Huang et al., 2025; Kim, 2026; Tang & Putra, 2026). Yet they also faced substantial epistemic and practical difficulties, including inaccurate outputs, alignment issues, and the extensive time required for refinement. These challenges reaffirm that GenAI does not simplify lesson planning; instead, it shifts the teacher's

role toward evaluator, editor, and curator of AI-generated suggestions (Cooper, 2023; Kim, 2026). Nevertheless, the strong intention to continue using GenAI for lesson planning further suggests that students regarded GenAI as a valuable supplement to their professional practice despite recognized limitations.

It is important to note that the findings are based on questionnaire data from a small cohort ($N = 10$), which limits how far the results can be generalized and how deeply individual experiences can be understood. Replicating this study with a larger sample would help clarify whether the observed patterns hold more broadly. Future iterations could also combine the Likert-scale items with brief follow-up open-response questions to better understand the reasoning behind students' agreement or disagreement. Moreover, post-course interviews could offer richer insights into how individual students' perspectives developed over time. Finally, assessing students' prior experience with GenAI at the outset through more questions could help clarify how existing familiarity shapes literacy development over the course.

Taken together, the findings suggest several implications for (teacher) education. First, GenAI literacy should be treated as a core pedagogical competency, requiring explicit training in evaluating outputs, formulating effective prompts, and aligning AI-generated material with curricular and pedagogical goals. Second, educators should model responsible GenAI use by making epistemic risks transparent and by demonstrating quality-assurance strategies during instruction. This could include sharing prompts educators used to design their own exercises and openly discussing how they evaluated, revised, or discarded GenAI outputs. Finally, educators should support students through clear guidelines and opportunities for collaborative experimentation, helping them cultivate a professional stance that embraces GenAI as a supportive complement to, rather than a replacement for, subject-matter expertise and pedagogical decisions.

References

Arantes, J. (2025). Understanding intersections between GenAI and pre-service teacher education: What do we need to understand about

the changing face of truth in science education? *Journal of Science Education and Technology*, 34(6), 1311–1322.

Cooper, G. (2023). Examining science education in ChatGPT: An exploratory study of generative artificial intelligence. *Journal of Science Education and Technology*, 32(3), 444–452.

Cooper, G., & Tang, K.-S. (2024). Pixels and pedagogy: Examining science education imagery by generative artificial intelligence. *Journal of Science Education and Technology*, 33(4), 556–568.

Gamlem, S. M., Johnston, S.-K., Moltudal, S., McGrane, J., & Hopfenbeck, T. N. (2026). Generative AI in initial teacher education: Exploring the alignment of perceptions and experiences of pre-service teachers and their teacher educators. *Teaching and Teacher Education*, 172, Article 105382.

Huang, T., Wu, C., & Wu, M. (2025). Developing pre-service language teachers' GenAI literacy: An interventional study in an English language teacher education course. *Discover Artificial Intelligence*, 5(1), Article 163.

Kim, W. J. (2026). Teachers' use of generative artificial intelligence for designing science lessons in support of environmental science agency. *Research in Science Education*, 56(1), 203–222.

Madsen, L. M., Evans, R., & Bruun, J. (2020). Undersøgelsesbaseret undervisning: 6F-modellen – dens tilblivelse og udvikling i Danmark [Inquiry-based learning: The 6F model - its origins and development in Denmark]. *MONA - Matematik- Og Naturfagsdidaktik*, 2020(1), 24-44.

Prilop, C. N., Mah, D.-K., Jacobsen, L. J., Hansen, R. R., Weber, K. E., & Hoya, F. (2025). Generative AI in teacher education: Educators' perceptions of transformative potentials and the triadic nature of AI literacy explored through AI-enhanced methods. *Computers and Education: Artificial Intelligence*, 9, Article 100471.

Spasopoulos, T., Sotiropoulos, D., & Kalogiannakis, M. (2025). Generative AI in pre-service science teacher education: A systematic review. *Advances in Mobile Learning Educational Research*, 5(2), 1501–1523.

Tang, K.-S., & Putra, G. B. S. (2026). Generative AI as a dialogic partner: Enhancing multiple perspectives, reasoning, and argumentation in

science education with customized chatbots. *Journal of Science Education and Technology*, 35(1), 128–140.

Tanner, K. D. (2010). Order Matters: Using the 5E Model to Align Teaching with How People Learn. *CBE—Life Sciences Education*, 9(3), 159–164.

Valeri, F., Nilsson, P., & Cederqvist, A.-M. (2025). Exploring students' experience of ChatGPT in STEM education. *Computers and Education: Artificial Intelligence*, 8, Article 100360.

Appendix

Items Used in Pretest and/or Posttest Questionnaire

(1) **Please indicate your most often used GenAI tools, such as ChatGPT, Copilot, or any other:**
Tool 1: _____, Tool 2: _____, Tool 3: _____

(2) **For which of the following activities do you use GenAI tools such as ChatGPT?** *[Pretest Only]*

- o Correcting texts
- o Summarizing texts (e.g., journal articles)
- o Translation of words, texts or materials
- o Repetition before an exam
- o Understanding concepts
- o Assignments or project work
- o Other, please specify: _____

(3) **Please indicate how much you agree or disagree with the following statements:** *[based on Valeri et al., 2025]*
Rating scale: (1) strongly disagree – (2) somewhat disagree – (3) neither agree nor disagree – (4) somewhat agree – (5) strongly agree

1. I can detect potential inaccuracies in responses generated by GenAI tools such as ChatGPT. *[Literacy]*
2. I am familiar with the concept of machine learning, which is the technology behind GenAI tools such as ChatGPT. *[Literacy]*
3. GenAI tools such as ChatGPT can be a substitute to traditional learning materials such as books. *[School]*

4. I adjust my questions to GenAI tools such as ChatGPT depending on the STEM subject I am working with. *[Literacy]*
5. GenAI tools such as ChatGPT help me understand difficult concepts in STEM subjects. *[Literacy]*
6. I trust the answers that GenAI tools like ChatGPT generate.
7. I have a good understanding of how GenAI tools such as ChatGPT generate texts. *[Literacy]*
8. I double-check the information generated by GenAI tools like ChatGPT against other sources. *[Literacy]*
9. I feel that I have sufficient knowledge to use GenAI tools such as ChatGPT to support my studies in the STEM subjects. *[Literacy]*
10. The use of GenAI tools like ChatGPT in schools will increase in the coming years. *[School]*
11. I am concerned about the sustainability aspects when using GenAI tools such as ChatGPT for my studies.
12. I am interested in learning more about how GenAI tools such as ChatGPT can be used in a way that helps me for my studies.
13. I am concerned about privacy aspects when using GenAI tools such as ChatGPT for my studies.
14. I believe it is important to discuss how GenAI tools such as ChatGPT can be used in schools. *[School]*

(4) **If you have used GenAI such as ChatGPT before, what challenges did you encounter?** *[Pretest Only]*

[the following items were used in posttest only]

(5) Throughout the course, you were repeatedly encouraged to use GenAI tools, such as ChatGPT, as aids. **For which of the following activities did you actually use it?**

- o Correcting texts
- o Summarizing texts (e.g., journal articles)
- o Translation of words, texts or materials
- o Repetition before an exam
- o Understanding concepts
- o Lesson planning

- o Getting initial ideas for a science lesson
- o Other, please specify: _____

(6) We encouraged the use of GenAI tools, such as ChatGPT, for lesson planning. **What benefits have you experienced from using GenAI?**

- o Increased creativity or inspiration
- o Faster generation of ideas and/or materials
- o Helpful as 'thinking partner'
- o Reduced lesson planning time
- o More varied instructional strategies
- o Increased confidence in lesson planning
- o Improved lesson structure
- o On-demand support when needed
- o Other, please specify: _____

(7) **And what challenges have you experienced when using GenAI?**

- o Difficulty writing effective prompts
- o Inaccurate or unreliable outputs
- o Content not well aligned with curriculum standards
- o Time required to refine or edit AI-generated materials
- o Limited subject-specific usefulness
- o Concerns about academic integrity
- o Lack of confidence or training in using GenAI tools
- o Technical or access issues
- o Other, please specify: _____

(8) **Can you describe a time when GenAI supported your lesson planning and a time when it did not work as expected?**
Example for support:

_____

Example for not working as expected:

_____

(9) **How strongly do you agree or disagree with the following statement: "I plan to use GenAI tools, such as ChatGPT, for lesson planning in my future teaching."**
Rating scale: (1) Strongly disagree – (2) Somewhat disagree – (3) Neither agree nor disagree – (4) Somewhat agree – (5) Strongly agree

(10) **Why (or why not) do you plan to use GenAI for lesson planning or lesson preparation in the future? Please explain:**

_____