

LearnLens: An AI-Enhanced Dashboard to Support Teachers in Open-Ended Classrooms

Anonymous

No Institute Given

Abstract. Exploratory learning environments (ELEs), such as simulation-based platforms and open-ended science curricula, promote hands-on exploration and problem-solving but make it difficult for teachers to gain timely insights into students’ conceptual understanding. This paper presents LearnLens, a generative AI (GenAI)-enhanced teacher-facing dashboard designed to support problem-based instruction in middle school science. LearnLens processes students’ open-ended responses from digital assessments to provide various insights, including sample responses, word clouds, bar charts, and AI-generated summaries. These features elucidate students’ thinking, enabling teachers to adjust their instruction based on emerging patterns of understanding. The dashboard was informed by teacher input during professional development sessions and implemented within a middle school Earth science curriculum. We report insights from teacher interviews that highlight the dashboard’s usability and potential to guide teachers’ instruction in the classroom.

Keywords: GenAI · Teacher-facing Dashboard · Open-Ended Learning Environments · LLM · AI in Education · Science Education

1 Introduction and Background

Exploratory Learning Environments (ELEs), such as simulation-based platforms and game-based settings, provide students with opportunities to explore diverse problem spaces and experiment with various learning strategies. These environments prioritize conceptual understanding and the development of critical thinking skills over the mere acquisition of procedural knowledge [11]. While ELEs offer valuable opportunities for students to gain a deeper understanding of a domain and develop problem-solving skills by encouraging them to make decisions, test ideas, and construct meaning through hands-on exploration, they also pose significant challenges for students and teachers.

In real classrooms, it can be challenging for teachers to gain a comprehensive understanding of each student’s progress, thought processes, and misconceptions, especially when students are working independently or in small groups on open-ended tasks [5,12]. Since teachers can only interact with a limited number of students at a time, much of the students’ thinking remains hidden from view. Prior research suggests that utilizing interactive technologies to make students’

ideas more visible can help instructors provide more effective guidance, thereby enhancing student learning [13,7,8]. However, there is still a lack of scalable and pedagogically effective tools that assist teachers in interpreting open-ended responses in online learning environments in a timely manner.

Recent advances in generative AI (GenAI), particularly large language models (LLMs), offer significant opportunities for analyzing open-ended student responses. This includes tasks such as grading short answers and essays, reviewing student self-explanations, and identifying misconceptions [6,17]. These models can provide valuable feedback for both formative and summative assessments, making them especially useful in open-ended and exploratory learning environments. GenAI’s capability to identify patterns across diverse student responses and offer concise summaries of collective class understanding makes it a valuable tool for instructional decision-making.

A teacher-facing dashboard powered by GenAI provides a promising design model for this kind of support [7]. Dashboards that aggregate student-generated data with interactive visualizations and AI-generated insights can more effectively support educators, particularly when co-designed with them [9].

In this paper, we present LearnLens, a GenAI-supported interactive teacher-facing dashboard that offers multiple layers of insights into students’ thinking in an exploratory learning classroom. The dashboard ingests students’ open-ended responses from digital classroom activities (such as check-ins, exit tickets, and formative assessments). It generates interactive visualizations and AI-generated summaries and insights to make students’ thinking more visible and interpretable in real time, allowing teachers to adapt their instructions accordingly.

LearnLens was implemented as a high-fidelity prototype within a middle school Earth science curriculum that integrates science, computing, and engineering concepts together in a 15-lesson curriculum [1]. The dashboard’s features were informed by researcher-teacher interactions during the professional development sessions before running the classroom studies and were designed to address the challenges teachers face when interpreting open-ended responses in real-time. To evaluate the usability and instructional relevance of the dashboard, we conducted semi-structured interviews with two science teachers, focusing on how the dashboard could support their decision-making while teaching.

We present findings from one of our teacher interviews. The teachers appreciated the dashboard’s clear visualizations and modular interface. They also found the GenAI component useful for identifying gaps and misconceptions in students’ understanding. Our teacher-facing dashboard aligns with several design principles identified by Poh et al. (2023) [14] for supporting in-class learning. These principles include providing real-time insights to inform instruction, supporting teachers’ pedagogical goals, and minimizing the effort needed to interpret and act on the data.

Overall, this work supports the workshop’s goals by exploring how AI, particularly Generative AI (GenAI), can enhance teacher decision-making in exploratory learning environments. It contributes to ongoing efforts to integrate GenAI to support STEM education [15].

2 Designing the LearnLens Dashboard

2.1 Study Context: Earth Science Curriculum

The dashboard was implemented as a high-fidelity prototype within a 15-lesson middle school NGSS-aligned Earth science curriculum that integrates science, computing, and engineering concepts [1]. The curriculum covered topics such as conservation of mass, rainfall, absorption, and runoff, helping students understand the challenges of urban water runoff. Students began the curriculum with science lessons on rainwater absorption by different surface materials, before transitioning to computational modeling, utilizing a block-based Domain-Specific Modeling Language (DSML) [10]. Finally, students engaged with a computer-based Engineering Design Challenge, where they applied their knowledge to design a schoolyard that minimizes runoff after heavy rainfall.

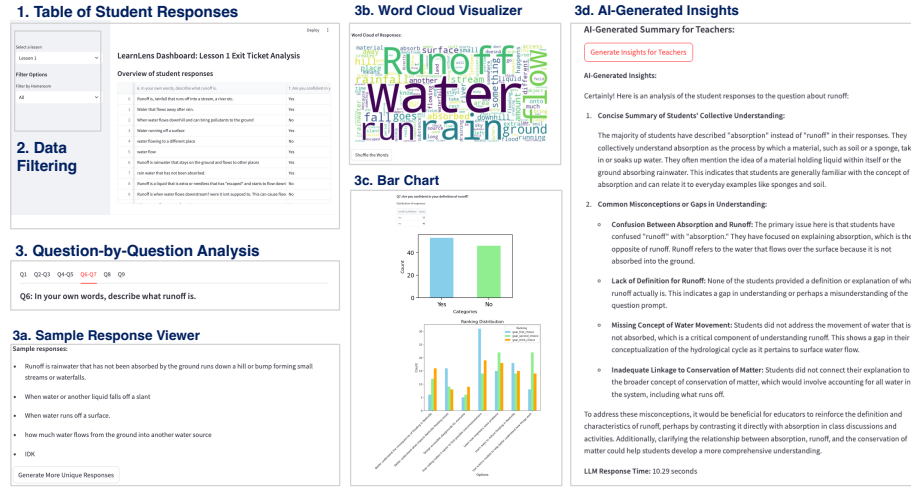
Daily assessments included exit tickets ($N = 8$), check-ins ($N = 3$), and formative assessments ($N = 4$), administered digitally as Google Forms. Exit tickets and check-ins were ungraded and intended solely to monitor students' learning; exit tickets captured student reflections, vocabulary development, and areas of confusion, while check-ins assessed understanding of key concepts discussed in class. In this paper, we demonstrate the potential of an AI-enhanced dashboard using data from a student check-in conducted after the end of Lesson 1.

2.2 Dashboard Implementation and Design

The dashboard components were guided by professional development sessions with expert teachers, who emphasized the need to assess students' understanding of core curriculum concepts at the end of each lesson to adapt their lesson plans. In previous implementations of the curriculum, the researchers were only able to demonstrate limited aspects of students' conceptual understanding (e.g., showing sample responses and word clouds on a PowerPoint slide).

The LearnLens dashboard was designed to build upon these earlier efforts. It was developed using Python to display data online (after the end of each class) and includes the following core features (see Figure 1 for an overview):

- **Table of Student Responses:** This table displays all student responses for an assessment, helping teachers quickly review the data.
- **Data Filtering:** The dashboard features a filter to allow teachers to view the performance of individual class sections by homeroom.
- **Question-by-Question Analysis:** Each question appears in its own tab to reduce information overload. For every question, the dashboard includes:
 - *Sample Response Viewer:* This function allows teachers to scroll through a range of student responses. Only five responses are shown at a time, with the option to view more by clicking the “Generate More Unique Responses” button.
 - *Word Cloud Visualizer:* A word cloud of students' responses was generated using the Python WordCloud package. The visualization highlights common vocabulary and themes across student responses. It also has the option to shuffle the responses.

**Fig. 1.** Features of LearnLens Dashboard

- *Bar Chart (for non-open-ended questions)*: A bar chart provided a quick summary of how students responded to multiple-choice or checkbox items.
- *AI-Generated Insights*: LLM model summarized students' understanding and identified common misconceptions in their understanding.

Among these, the AI-generated insights feature was the most novel and required careful refinement. We describe its development in the next section.

2.3 Development of the GenAI Component

The AI-generated insights in LearnLens were powered by the OpenAI GPT-4o model (version gpt-4o-2024-08-06)¹. The GenAI component was designed to (1) provide a concise summary of students’ collective understanding and (2) identify common misconceptions or gaps in that understanding. To develop effective prompts, we followed an iterative prompt engineering process with human validation at each step, focusing on relevance, specificity, and alignment with the curriculum’s context and goals.

In most instances, the GPT model generated accurate and reliable summaries. However, it struggled with identifying common misconceptions or gaps, occasionally producing hallucinated content. For example, in a simpler version of the prompt (zero-shot), GPT confused the curriculum’s focus on rainfall and runoff with the broader water cycle, incorrectly suggesting that teachers should emphasize the water cycle, which was not part of the curriculum.

To address these issues, we refined the prompt to include additional context, such as the agent’s role, the curriculum’s goals, relevant NGSS crosscutting concepts, and specific assessment questions using guidance from literature [16]. For

¹ <https://openai.com/index/hello-gpt-4o/>

questions targeting scientific knowledge (e.g., definitions of rainfall, runoff, and absorption), we included correct example responses as few-shot instances to help guide the model. The improved prompt was then tested using real classroom data from a previous implementation of the same curriculum. Overall, this human-in-the-loop design ensured that the GenAI component would serve as a useful instructional aid rather than a black-box solution [3].

3 Method

3.1 Participants

Two 6th-grade teachers (1 male, 1 female) from an urban public school with over 34 years of STEM teaching experience participated in the study. Both teachers had previously implemented the Earth science curriculum in their classrooms, making them well-suited for providing informed feedback. They provided informed consent to participate in the study. Teacher 1 implemented the Earth science curriculum across 3 class sections, while Teacher 2 taught one section. All study procedures and teacher interviews were approved by the [Anonymous Institution] IRB Committee.

3.2 Data sources and Analysis

We conducted semi-structured interviews with the two participating teachers three times during the implementation to evaluate the potential of the LearnLens dashboard. During each session, a researcher guided the teachers through the dashboard using a structured walkthrough; the teachers did not interact with the dashboard independently.

The interviews focused on two key areas: (1) teachers' current data practices, including the types of reports or dashboards they typically use, and (2) feature-specific feedback on the LearnLens dashboard, including sample responses, word clouds, bar graphs, and AI-generated summaries and insights. Teachers were encouraged to share suggestions for improvement and reflect on the potential instructional use of each feature.

All sessions were audio- and screen-recorded. This paper analyzes the interview when the teachers were introduced to the dashboard for the first time. The 1-hour-long audio recording was transcribed and analyzed thematically, using a deductive coding scheme aligned with the interview structure[4], sample codes include sense-making of presented analysis, clarifying analysis methods, connection with classroom experience, perception of features, and instructional modification. The major themes that emerged are discussed in the next section.

4 Results and Findings from Teacher Interviews

Current Data Practices: The teachers reported using a range of tools in their classroom instruction. Teacher 1 relied on more traditional methods, such as

paper-based exit tickets, whereas Teacher 2 incorporated several digital tools, including Google Forms, Schoology (a Learning Management System), and Ed-Puzzle. Despite the variety of tools, both teachers emphasized a standard limitation, i.e., these platforms do not automatically generate summaries or insights. For instance, Teacher 2 said, *“None of those things do anything except just organize it in a place. They don’t give you any feedback.”*

Dashboard Features Evaluation: Teachers appreciated the variety of visualizations we provided on the dashboard, particularly the word clouds for questions related to the definitions of scientific phenomena. Teacher 1 remarked, *“I love word clouds!”*. Additionally, they found the bar graphs and filtering feature helpful, as these tools allowed them to examine students’ understanding across different class sections. Displaying actual student responses verbatim was valuable because it enabled them to see exactly what students had written. The AI-generated misconceptions were also recognized as a beneficial feature:

Teacher 1: (Reading from screen) One student described absorption as involving electrons, which may indicate confusion with chemical processes like reactions or bonding. They don’t even know about reactions or bonding.

Teacher 2: So that, this misconception piece is super helpful here.

Connection with Classroom Instruction & Instructional Modifications:

While reviewing the dashboard, teachers could relate trends in student responses to the classroom discussions and shared how they would adjust their future instruction. For instance, Teacher 1 saw the word cloud as an accessible visualization and proposed printing and displaying it in the classroom to prompt student reflection (*Teacher 1: “seriously, we need to take some of these, if we can print them out and put the question with them, and just put them in various places around the classroom.”*)

The AI-generated insights encouraged teachers to focus more on reinforcing vocabulary, particularly with key terms like “absorption” and “runoff,” which are central components of the science curriculum. As Teacher 1 noted while reading from the screen: *“There are instances of incorrect spelling or uses of terms such as ‘absorps’ or ‘seps’, which could lead to misconceptions and misunderstandings or difficulty with terminology. ‘Absorption’ is a hard word for kids, and we probably need to put it up on our vocabulary.”*

Trust in AI-generated Insights: When reviewing the AI summaries with the teachers, the researchers asked them whether they trusted the AI-generated insights, noting that the LLMs can occasionally produce inaccurate or unsupported information. Teachers expressed confidence in AI when it summarized student responses or identified gaps based on those responses. However, they were skeptical of AI-generated suggestions that went beyond the data, such as recommendations for modifying lesson plans, which should draw from broader instructional knowledge. Thus, while the teachers trusted the descriptive capabilities of the AI, they were cautious about its prescriptive use:

Teacher 2: Well, in the place, I would come up with the trust pieces like when it gave suggestions on how to help misconceptions. I would have

to see what they would say there. [Researcher: Oh, okay!] ... it is telling us what they got from the data. But in that case, if you're asking for a prescription of what to do that's coming, that wouldn't be coming from this data, so that would be a different.

Teacher 1: But all this is from the data, so...

Teacher 2: So that's trustable for me. [Teacher 1: Right!] It's not the same as just getting stuff like you're asking it for a prescription of what to do when this and this.

5 Discussion and Future Work

Our findings demonstrate that GenAI-enhanced dashboards can help teachers make sense of students' open-ended responses in exploratory learning environments. Both teachers emphasized the need for such dashboards as the current instructional support tools lack summary or feedback insights. When shown the dashboards' functionalities, the teachers valued its modular design, visualizations (e.g., word cloud, bar chart), and especially the AI-generated insights that highlighted common gaps or misconceptions in students' understanding. These features are well-aligned with several design principles proposed by Poh et al. (2023) [14] for developing effective teacher-facing dashboards, such as simpler design, glanceable summaries, modular design, supporting sense-making and using real student data for evaluations.

Interviews, revealed that teachers could connect the analysis with their own instruction, using the insights to inform future instruction. They found visualizations, like word clouds, accessible for the students and decided to display them in the classroom. While teachers trusted summaries grounded in student data, they were cautious about prescriptive suggestions extending beyond the data, such as instructional recommendations. The findings resonate with the topology outlined by Campos et al. [2]. In the *Emotional* dimension, teachers' response was *Satisfactory* for conventional and *Positive* for the AI-driven insights. In the *Analytical* dimension, teachers could *Recall* events from classes and *Attribute* them to patterns in the data. In the *Intentional* dimension, we could observe *Action Intention* and *Planning* of instructional adaptations. These results align with previous work, which shows that interactive technologies can support instructional adaptation by surfacing patterns in student understanding [13,7].

The study reported in this paper is limited to a single interview with two teachers. Future work will incorporate additional data sources and longitudinal analysis to gain a deeper understanding of the sustained impact of AI-powered feedback on instruction. While there are growing efforts to integrate AI tools into educational practice [15], fewer studies evaluate the design with teacher feedback. This paper contributes to addressing this gap, providing evidence of teachers' positive experiences with AI-powered dashboards that offer timely, data-driven feedback. This work highlights the value of simple yet timely insights that can effectively support teachers in addressing emerging classroom needs.

References

1. Basu, S., McElhaney, K.W., Rachmatullah, A., Hutchins, N., Biswas, G., Chiu, J.: Promoting computational thinking through science-engineering integration using computational modeling. In: Proceedings of the 16th International Conference of the Learning Sciences (ICLS) (2022), <https://dx.doi.org/10.22318/icls2022.743>
2. Campos, F.C., Ahn, J., DiGiacomo, D.K., Nguyen, H., Hays, M.: Making Sense of Sensemaking: Understanding How K–12 Teachers and Coaches React to Visual Analytics. *Journal of Learning Analytics* **8**(3), 60–80 (Jul 2021). <https://doi.org/10.18608/jla.2021.7113>
3. Cohn, C., Hutchins, N., Le, T., Biswas, G.: A chain-of-thought prompting approach with llms for evaluating students’ formative assessment responses in science. Proceedings of the AAAI Conference on Artificial Intelligence **38**(21), 23182–23190 (Mar 2024). <https://doi.org/10.1609/aaai.v38i21.30364>, <https://ojs.aaai.org/index.php/AAAI/article/view/30364>
4. Deterding, N.M., Waters, M.C.: Flexible coding of in-depth interviews: A twenty-first-century approach. *Sociological methods & research* **50**(2), 708–739 (2021)
5. Dillenbourg, P.: Design for classroom orchestration. *Computers & Education* **69**, 485–492 (2013). <https://doi.org/https://doi.org/10.1016/j.compedu.2013.04.013>
6. Giannakos, M., Azevedo, R., Brusilovsky, P., Cukurova, M., Dimitriadis, Y., Hernandez-Leo, D., Järvelä, S., Mavrikis, M., Rienties, B.: The promise and challenges of generative AI in education. *Behaviour & Information Technology* pp. 1–27 (Sep 2024). <https://doi.org/10.1080/0144929X.2024.2394886>
7. He, P., Shin, N., Zhai, X., Krajcik, J.: A Design Framework for Integrating Artificial Intelligence to Support Teachers’ Timely Use of Knowledge-in-Use Assessments. In: Zhai, X., Krajcik, J. (eds.) *Uses of Artificial Intelligence in STEM Education*, pp. 348–370. Oxford University Press Oxford, 1 edn. (Oct 2024). <https://doi.org/10.1093/oso/9780198882077.003.0016>
8. Hutchins, N., Biswas, G.: Using Teacher Dashboards to Customize Lesson Plans for a Problem-Based, Middle School STEM Curriculum. In: LAK23: 13th International Learning Analytics and Knowledge Conference. pp. 324–332. ACM, Arlington TX USA (Mar 2023). <https://doi.org/10.1145/3576050.3576100>
9. Hutchins, N.M., Biswas, G.: Co-designing teacher support technology for problem-based learning in middle school science. *British Journal of Educational Technology* **55**(3), 802–822 (May 2024). <https://doi.org/10.1111/bjet.13363>
10. Hutchins, N.M., Biswas, G., Zhang, N., Snyder, C., Lédeczi, Á., Maróti, M.: Domain-Specific Modeling Languages in Computer-Based Learning Environments: a Systematic Approach to Support Science Learning through Computational Modeling. *International Journal of Artificial Intelligence in Education* **30**, 537–580 (2020). <https://doi.org/10.1007/s40593-020-00209-z>
11. Mavrikis, M., Geraniou, E., Gutierrez Santos, S., Poulouvassilis, A.: Intelligent analysis and data visualisation for teacher assistance tools: The case of exploratory learning. *British Journal of Educational Technology* **50**(6), 2920–2942 (Nov 2019). <https://doi.org/10.1111/bjet.12876>
12. Mavrikis, M., Gutierrez-Santos, S., Poulouvassilis, A.: Design and Evaluation of Teacher Assistance Tools for Exploratory Learning Environments. In: Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK ’16. pp. 168–172. ACM Press, Edinburgh, United Kingdom (2016). <https://doi.org/10.1145/2883851.2883909>

13. Moed-Abu Raya, K., Olsher, S.: Teachers' Formative Assessment Practices in Their Mathematics Classroom Using Learning Analytics Visualizations. *Digital Experiences in Mathematics Education* **10**(3), 395–417 (Dec 2024). <https://doi.org/10.1007/s40751-024-00148-7>
14. Poh, A., Castro, F.E.V., Arroyo, I.: Design Principles for Teacher Dashboards to Support In-Class Learning. pp. 736–743 (Oct 2023). <https://doi.org/10.22318/icls2023.197114>
15. Price, J.F., Grover, S.: Generative AI in STEM Teaching: Opportunities and Trade-offs. *Community for Advancing Discovery Research in Education (CADRE)* (2025), <https://eric.ed.gov/?id=ED672718>
16. White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., Schmidt, D.C.: A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT (2023), <https://arxiv.org/abs/2302.11382>
17. Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., Gašević, D.: Practical and ethical challenges of large language models in education: A systematic scoping review. *British Journal of Educational Technology* **55**(1), 90–112 (2024). <https://doi.org/10.1111/bjet.13370>