17th International Scientific Conference 2025

Sessions of the Faculty of Education

https://doi.org/10.36007/5390.2026.262

ENHANCING INQUIRY-BASED LEARNING WITH GENERATIVE AI: INSIGHTS FROM STUDENT INTERACTIONS

Mihály KOVÁCS [55] – Zoltán MURÁNYI [56]

ABSTRACT

This pilot study explores the potential of using generative artificial intelligence (AI) in chemistry education to support inquiry-based learning (IBL) and attempts to model AI-assisted chemistry research practice in the classroom. It fills a gap in the literature, which covers AI-supported IBL in other fields but not in chemistry. N=23 9th and 10th grade students used Mistral Le Chat to design an experiment to determine what causes cherry cake dough to turn green. The dialogues between the students and the AI and the students' exit cards were analysed, suggesting that AI can function as a supportive tool, e.g. in formulating hypotheses. Students found AI useful, e.g., for clear explanations and experiment design, although some reported that AI solved the problem for them. Based on these results, AI-supported IBL is a promising method in chemistry education, but further large-scale research is needed on how to use it effectively.

KEYWORDS

Inquiry-Based Learning (IBL), Generative AI, Chemistry Education, Scaffolding, Student-AI interaction

INTRODUCTION

Machine learning is used in chemical research and material development (e.g. Fehér et al., 2023). The starting point of our research is that large language models seem to be quite good at predicting the reactions of simple materials, so we can model this process in chemistry classes, i.e., students can use chatbot type generative artificial intelligence (AI) to plan their experiments.

From the perspective of education, this means that during Inquiry-Based Learning (IBL) lessons, AI partially takes over the role of teachers as mentors and the role of chemistry literature. There have been previous studies that examined how the IBL process can be supported with the help of computers, e.g., in biology with digital games (Bónus et al., 2024) or the impact on the planning process within the teacher's work was examined primarily in the context of physics (Moundridou et al., 2024), and in the case of physics lessons, the planning of experiments with the help of an AI assistant was also examined (Min et al., 2025). We have not yet found a similar

[55] Mihály Kovács, Eszterházy Károly Catholic University, Observatory and Science Experience Centre, kovacs2.mihaly@uni-eszterhazy.hu

[56] Dr. Zoltán Murányi, Eszterházy Károly Catholic University, Faculty of Natural Sciences, Chemistry, Institute of Chemistry and Physics, muranyi.zoltan@uni-eszterhazy.hu

publication in the case of chemistry. Hence, the aim of this pilot study is to lay the groundwork for a subsequent study conducted on a larger sample size to fill this research gap.

We analyzed the feedback collected on exit cards to examine the research questions. We asked participants about their experiences communicating with AI, the extent to which this tool helped them in their learning and in planning experiments, and how participating in the class affected their interest in chemistry.

We also analyzed the prompts written by students, which they used to interact with the generative AI in order to gather ideas and suggestions, primarily in the process of designing experiments, and the quality of the responses they received from the AI.

At the beginning of the lesson, students were warned not to ask questions about the entire experiment, because we had chosen an inquiry for which the AI had given us an incorrect answer. Their task was to find out, based on a cake recipe, what causes the dough around the sour cherries to turn green depending on the leavening agent used: baking soda, baking powder, or ammonium bicarbonate.

LITERATURE REVIEW

"IBL may be defined as a teaching approach which intends to promote learning by engaging students in any of the processes or activities typically involved in scientific research. These include: making observations, formulating hypothesis, defining problems and key guiding questions, designing and performing experiments and communicating results and evidence-based conclusions." (PRIMAS, 2017, 3.)

The following chapter summarizes the purpose of inquiry-based learning (IBL) and the challenges facing this approach, and possible solutions, with particular emphasis on conceptual understanding, motivation, soft -skill development, the importance of teacher support, and the potential role of AI in the process.

The role of IBL in conceptual understanding and motivation

Inquiry-based learning (IBL) can significantly increase understanding of concepts, as students acquire knowledge through active knowledge building. During learning, deeper understanding can be achieved by engaging students as deeply as possible in the research process. Important elements of the research process are asking questions, designing experiments, collecting data, deducing conclusions, and communicating results. It is especially effective when students can perform the experiments themselves, which helps to increase their understanding. (Minner et al., 2010)

Experiments also play a key role in terms of motivation, as even occasional teacher demonstrations can motivate students to learn. (Rogers & Fraser, 2023). Besides that, IBL also supports the development of soft skills like independent learning and teamwork through group work and autonomous activities. Even though these skills are hard to measure quantitatively and do not count much in school results, they prove to be extremely useful in the labor market. (Hmelo-Silver et al., 2007).

Challenges of IBL and possibilities for its scaffolding

One of the main criticisms of IBL is that the lack of teacher guidance can lead to learning difficulties, moreover, the complex nature of the method can increase cognitive load, making it difficult to acquire knowledge. (Kirschner et al., 2006).

However, guidance and scaffolding can also be implemented within the framework of IBL. To support students in the learning process, it is possible to share research methodology aids with

students, provide instructions on research and argumentation strategies, and incorporate expert information and guidance into task descriptions or the learning environment. (Hmelo-Silver et al., 2007)

Based on teachers' opinions, structured and guided research is the most suitable of the IBL methods for implementation in a secondary school environment (Tomperi & Aksela, 2014), which also highlights the importance of scaffolding mentioned above, as a possible way of carrying it out. To give a concrete example from the literature, Szalay and colleagues (2023) used three different levels of worksheets to implement the leveling. One group received a traditional, recipe-like experiment description, another group received an experiment design task in addition to the traditional experiment description, while the third group had to design the entire experiment. It is clear that these offer different levels of support to students.

Other methods mentioned for reducing cognitive load include whiteboards, on which students can record their achievements during a given activity, and action plans, which remind them of the subtasks they need to complete (Hmelo-Silver et al., 2007).

The potential role of generative artificial intelligence (AI) in IBL

In addition to being able to predict the results of certain experiments, generative AI also appears to be useful in other areas of IBL. Min and colleagues (Min et al., 2025) found that students relied on AI primarily in the early stages of their inquiry during their physics lessons, when they were facing difficulties in formulating hypotheses. They asked for help in defining variables, designing experimental procedures, and refining their plans.

Therefore, our research aims to examine whether AI can replace the above scaffolding forms, thereby enabling students to perform higher-level research tasks in the field of chemistry. With a well-formulated prompt, students can receive research methodology suggestions, expert guidance, task lists, or explanations of concepts from AI, which can help reduce cognitive load and make the learning process more efficient.

RESEARCH PROCESS AND METHODS

The pilot experiment lasted two weeks, during which students (N=23, 9th and 10th graders) who had never used generative AI in chemistry class before worked in groups of 3-4 on the previously mentioned research problem in two consecutive lessons: which leavening agent causes the dough around the cherries in the cherry cake to turn green, and why. During the first week, they communicated with Mistral Le Chat's chatbot, during which they had to develop an experiment to answer the question, and then in the second week, they were able to try out their experiments and test their hypotheses in practice. Before the first session, they were instructed not to ask for a complete experiment description, but to try to create one using the AI.

This software had various advantages from the perspective of our research. Firstly, it was developed in Europe, making it more suitable than other products from a data protection perspective, as all stored data is covered by the GDPR.

Secondly, it has a version that can be used without registration, so we did not have to oblige students to create user accounts. However, it is true that the free version only allows ten questions per day per device, which was also communicated to students before they started their work.

Finally, the software also provides links to individual dialogues, which facilitates data processing. Unfortunately, none of the groups managed to save and send the links, thus we only received screenshots, which were incomplete in some cases, which severely limits the results presented here.

We tokenized and lemmatized the student AI dialogues and removed the stopwords in accordance with natural language processing methods (Chary et al., 2018), then used Voyant Tools to analyze word frequency and to find the most characteristic collocations in both the students' prompts and the AI's responses. In addition, we also present some comments that are more interesting from the perspective of later large-scale research, such as those that can be classified into one of the categories mentioned in scaffolding.

At the end of the second lesson, students were also asked to fill out an exit card online, but the dropout rate was high, with only six students completing it. It contained the following questions:

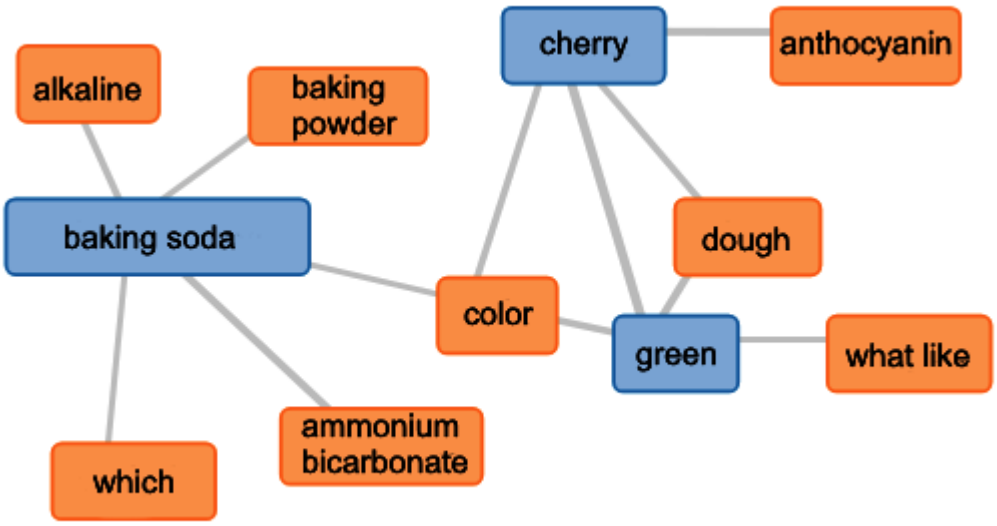
1. How did you feel during the lessons?
2. Write down a positive experience from the last classes.
3. Write down a positive experience regarding communication with AI.
4. Write down a difficulty regarding communication with AI.
5. What would you do differently in the previous class?
6. What would you suggest improving these lessons? What should we do differently if we were to teach these lessons in another class?
7. How much would you like to learn chemistry this way in the future? (1-5)

The 5th question referred to the class in which use of AI was permitted. Their responses were short sentences, so we could not use them for deeper analysis, but we will quote the most relevant ones for the large-scale research we are planning to carry out later.

RESULTS

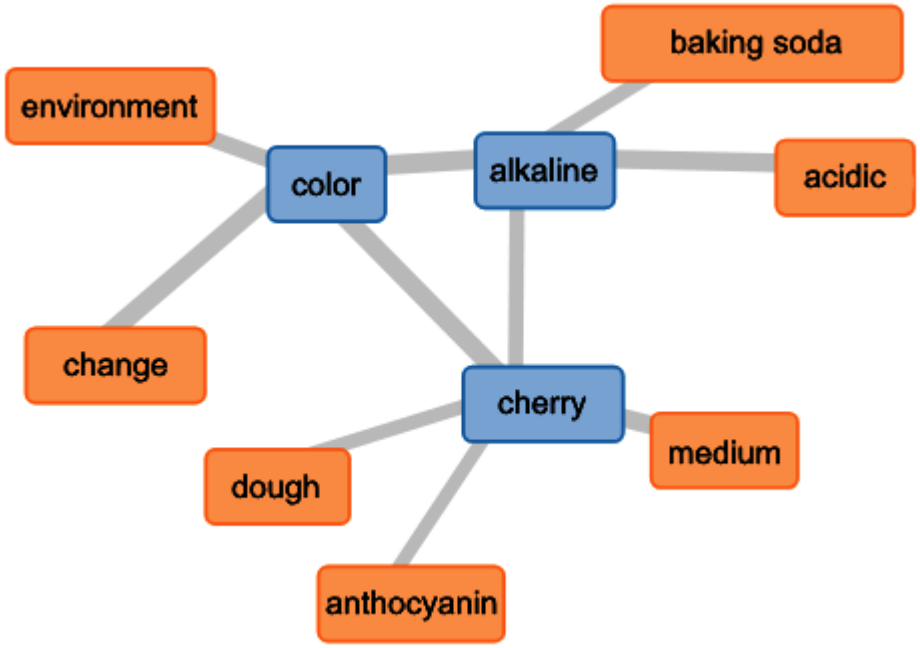
Quantitative analysis of the interactions

We found that most frequent words in the students' prompts were: "cherry" (10); "green" (7); "baking soda" (7); color (6); "what like" (6) from 94 words. The most common collocations are shown in *1*. Figure, the most frequent of them is the cherry-green. The occurrence of baking soda with the words sour cherry and green is not among the most common, but all three words form frequent collocations with the word color.



1. Figure: Most common collocations in prompts

In the answer of Le Chat, the most common words were: “color” (79); “cherry” (77); “alkaline” (57); “baking soda” (48); “acidic” (42) from 533 unique word forms. The most common collocations are shown in 2. Figure**Chyba! Nenašiel sa žiaden zdroj odkazov.**, and color-cherry is the most frequent. In this case, the most common words have very frequent collocations.



2. Figure: Most common collocations in chatbot's answer

Responses to exit cards

Most students had a positive experience with IBL lessons, based on the feedback shared on their exit cards. One student reported feeling puzzled. One student was happy about a successfully completed experiment, and several wrote that they enjoyed the opportunity because they “didn't have to study.” In contrast, one student wrote that it was a positive experience because “I learned how to plan an experiment.” Lastly, one student was unable to name any positive experiences.

Feedback on working with AI was varied. One student was dissatisfied with artificial intelligence, while several others said that the machine solved the problem instead of them. Two students found AI particularly useful because it explained the answers clearly.

In terms of challenges, students reported several problems with this chatbot. Many complained that they could only ask a limited number of questions, while others complained about the un-Hungarian-like responses. Reading through the dialogues, we found that the registration-free version of Le Chat indeed had problems with the Hungarian language occasionally, and in one case, a few words in English appeared in the response. However, the mention of few questions confirms our suspicion that we did not receive a significant portion of the dialogue, as there was only one group that came close to the ten questions based on the screenshots we received. One student also noted that the AI "could not really answer" their questions, but unfortunately, we were unable to link this comment to the submitted dialogues due to anonymization; it is possible that this team did not submit their dialogues either.

Several suggestions were made regarding what could be done differently. Some suggested using a different AI program to address the problems mentioned above, while others wanted more

interesting or more complex experiments because they found the current task too easy. Most of the students would change their own questions if they could do the task again, but one would not do anything differently.

However, it is worth noting that the students reported that the task was too easy and claimed that AI solved the task instead of them, before their teacher told them that their solutions were more or less incomplete.

We have not analyzed the quantitative question (7th) because of the extremely small sample size.

DISCUSSION

Among the leavening agents, baking soda was mentioned most frequently by students in their prompts, which is consistent with what we observed in the "typical" experiment plans. They primarily suspected that this substance caused the green discoloration of the dough since its solutions alkalinity is well known among the students. This is also the main motive of the AI's answer.

Nevertheless, the solution above is incomplete, as it does not take into consideration changes caused by heat. Although the chatbot suggested heating to several groups, either by recommending baking pieces of dough as an experiment or by explicitly mentioning that heat causes ammonia to be released from ammonium bicarbonate, which students know to produce an alkaline solution, none of the experiment plans included heating. Therefore, the intensity of the green color may have depended primarily on the amount of baking soda and ammonium bicarbonate used by each group. In addition, most groups did not take into account the effect of the lactic acid content of the yogurt mixed into the dough, even though the AI also mentioned it in certain interactions.

However, we found some experimental designs that examine the above questions but do not consider the whole picture. One team mixed yogurt into the test tubes, while another team identified ammonium bicarbonate as the main cause of the green color, specifically deducing the reason from a dialogue with AI, which was, however, incorrect. According to Le Chat, its solution is the most alkaline, which is not true at the same concentrations, but the ammonia gas released by heat can indeed turn the dough around the cherry green, as the cherry's red pigments – anthocyanins – are acid-base indicators.

Most groups realized this, although some tested cinnamon as well, since the AI stated that cinnamon polyphenols are also acid-base indicators, but added that at this concentration, it is probably invisible, as it is suppressed by the white color of the flour in the dough. This brings up the question of whether we were wrong to tell the children that the AI was giving the wrong answer, as they had no way of knowing what to believe and what not to believe.

In the prompts, we identified the same types of questions as those we cited from Min and colleagues' study (2025). Defining variables: "Could you tell me which of the following can act as an indicator: eggs, butter, powdered sugar, plain yogurt, flour, cinnamon, and almonds?", designing experimental procedures: "Design an experiment to make us understand why the dough turns green next to the cherry after baking.", and refining their plans: "How many cubic centimeters is a test tube?"

In addition, we found that students also used AI to test hypotheses, which means that in some groups we were able to model the process that inspired our research, namely that chemical properties and reactions can be predicted using AI in research practice. Namely: "Can you tell me if cinnamon is added to dissolved baking soda, will it change color?"

We found only one team that used prompts that asked questions to gain a deeper understanding of individual elements of the process, as if seeking expert assistance: "In what circumstances do alkaline and acidic substances come into contact, causing discoloration?"

Furthermore, we found another exciting question that cannot be classified into the above categories: “If we removed the anthocyanin that makes cherries red, what color would they be?” This question is completely irrelevant to the solution, as the task is not related to the color change of the cherries, but to that of the dough. However, it suggests that the students intended one of the ten questions to be answered in order to find out the answer to a question they personally found interesting. They also knew that this was not relevant to the topic, as their next question was about the color of anthocyanins in acidic and alkaline environments. It seems that they were simply curious about the answer, so this team was probably engaged by the activity. One of our research questions was whether AI-supported IBL helps to increase interest in chemistry, and from this perspective, this is a promising example.

CONCLUSION

During large-scale research, we need to place greater emphasis on properly instructing students, as several teams requested a complete experiment plan despite our warnings that it would be flawed. We believe that we will get better results if we do not simply prohibit students from asking this question, but give them ideas on how they can use AI, as none of the teams used it, for example, to get tips on research strategies, or a list of tasks to be completed. (Hmelo-Silver et al., 2007) For this purpose, we plan to give students some kind of handout, like an infographic, about possible ways of using the chatbot.

Furthermore, as mentioned earlier, students were unable to save the links to the dialogues, so they sent screenshots instead, even though this was a slower and more complicated process for them, too. A solution to this could be to create a slide show or a short instructional video to help teachers introduce the AI user interface to their students.

We do not plan to replace the chatbot, as data security is still a factor in its favor, and groups contain 3-4 people, which means that students can ask up to 30-40 questions during a 45-minute lesson, which should be sufficient. It is also worth considering from a research perspective when students get the complete, correct solution, as they wrote their exit cards knowing that their solutions were correct and the task was too easy for them, and we do not know how their opinion changed after they learned the evaluation of their solutions.

All in all, the implication of this research is that scaffolding of IBL is achievable with the help of AI, so IBL with AI support in chemistry education is a promising possibility. However, the results are severely limited by the fact that this was pilot research conducted on a small sample and we did not receive the full text of the student-AI dialogues.

REFERENCE

Bónus, L., Antal, E., & Korom, E. (2024). Digital Game-Based Inquiry Learning to Improve Eighth Graders’ Inquiry Skills in Biology. *Journal of Science Education and Technology*, *33*(4), 1–17. https://doi.org/10.1007/s10956-024-10096-x

Chary, M., Parikh, S., Manini, A., Boyer, E., & Radeous, M. (2018). A Review of Natural Language Processing in Medical Education. *Western Journal of Emergency Medicine*, *20*(1), 78–86. https://doi.org/10.5811/westjem.2018.11.39725

Fehér, P. P., Madarász, Á., & Stirling, A. (2023). Prediction of Redox Power for Photocatalysts: Synergistic Combination of DFT and Machine Learning. *Journal of Chemical Theory and Computation*, *19*(13), 4125–4135. https://doi.org/10.1021/acs.jctc.3c00286

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and Achievement in Problem-Based and Inquiry Learning: A Response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, *42*(2), 99–107. https://doi.org/10.1080/00461520701263368

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. *Educational Psychologist*, *41*(2), 75–86. https://doi.org/10.1207/s15326985ep4102_1

Min, T., Lee, B., & Jho, H. (2025). Integrating generative artificial intelligence in the design of scientific inquiry for middle school students. *Education and Information Technologies*, *30*. https://doi.org/10.1007/s10639-025-13410-1

Minner, D. D., Levy, A. J., & Century, J. (2010). Inquiry-based science instruction—what is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of Research in Science Teaching*, *47*(4), 474–496. https://doi.org/10.1002/tea.20347

Moundridou, M., Matzakos, N., & Doukakis, S. (2024). Generative AI tools as educators' assistants: Designing and implementing inquiry-based lesson plans. *Computers and Education: Artificial Intelligence*, *7*, 100277. https://doi.org/10.1016/j.caeai.2024.100277

Primas, P. (2017). *PRIMAS Guide for professional development providers*. https://primas-project.eu/wp-content/uploads/sites/323/2017/11/FINAL_WP4_Guide_PD_providers_licence_150708.pdf

Rogers, J. R., & Fraser, B. J. (2023). Sex and frequency of practical work as determinants of middle-school science students' learning environment perceptions and attitudes. *Learning Environments Research*, *26*(2), 315–336. https://doi.org/10.1007/s10984-022-09426-w

Szalay, L., Tóth, Z., Borbás, R., & Füzesi, I. (2023). Scaffolding of experimental design skills. *Chemistry Education Research and Practice*, *24*(2), 599–623. https://doi.org/10.1039/D2RP00260D

Tomperi, P., & Aksela, M. (2014). In-Service Teacher Training Project On Inquiry-Based Practical Chemistry. *Lumat: International Journal of Math, Science and Technology Education*, *2*(2), 215–226. https://doi.org/10.31129/lumat.v2i2.1075