

Faculty of Education and Humanities  
Department of Pedagogy of Natural Sciences

“Admitted to defense”

Head of Department

Assoc. Prof. PhD



Zhangyl Abilbek

«06» қаңтар 2025

MASTER'S DEGREE DISSERTATION

AI in chemistry education and ethical considerations

7M01502 – Chemistry

Student:

Kiikov

Kiikov Bakytbek

Scientific advisor:

Zhangyl

PhD, Assoc. Prof. Zhangyl Abilbek

Format controller:

Nazgul

MSc. Nazgul Otegenova

Kaskelen, 2025

# CONTENTS

<b>INTRODUCTION</b>	1
<b>1. LITERATURE REVIEW</b>	6
1.1 AI tools	6
1.1.1 ChatGPT	6
1.1.2 Yandex AI(YaLM)	7
1.1.3 Gemini	8
1.1.4 DeepSeek	9
1.2 AI in chemistry education	12
<b>2. METHODOLOGY</b>	26
2.1 Research design	26
2.2 Participants and sampling criteria	26
2.3 Data collection	26
2.3.1 Survey for Research Question 1	26
2.3.2 Survey for Research Question 2	27
2.3.3 Solving Problems Using ChatGPT (Research Question 3)	27
2.4 Data Analysis	28
2.5 Ethical considerations	28
<b>3. RESULTS</b>	29
3.1 RQ1 The effectiveness and usability of AI integrated chemistry education	29
3.2 RQ2 Ethical and privacy concerns	35
3.3 RQ3 Advantages and disadvantages of using AI tools	40
<b>4. DISCUSSION</b>	58
<b>5. CONCLUSION</b>	66
<b>6. SHORTOMINGS AND LIMITATIONS</b>	68
<b>REFERENCES</b>	70
<b>APPENDICES</b>	74

## ABSTRACT

This dissertation examines the integration of Artificial Intelligence (AI) tools, specifically ChatGPT, in chemistry education in Kazakhstan, aiming to identify challenges, opportunities, ethical considerations, and the practical effectiveness of these tools in enhancing learning outcomes. The study explored three key areas: the effectiveness of AI integration into chemistry education, the ethical and privacy concerns associated with student use of AI tools, and the primary advantages and disadvantages of employing AI in chemistry instruction.

Using a mixed-method approach, the research combined quantitative surveys among 108 chemistry education students from Kazakh universities and qualitative expert evaluations of AI-generated chemistry solutions. The theoretical framework drew from existing literature on AI integration in education, ethical implications, and pedagogical impacts.

Findings indicated that students strongly prefer ChatGPT due to its efficiency, clarity, and ability to facilitate independent learning, primarily utilizing it for problem-solving and exam preparation. However, significant limitations were observed, including accuracy issues, logical inconsistencies, and inadequate linguistic adaptation to the Kazakh language. Ethical concerns highlighted were academic integrity, dependency on technology, and unequal access to premium AI features.

The dissertation contributes theoretically by providing empirical evidence of AI's educational benefits and limitations, and practically by recommending structured AI integration strategies, specialized training, enhanced linguistic localization, and ethical guidelines. Ultimately, this research informs educators, policy-makers, and developers aiming to harness AI responsibly and effectively in chemistry education.

**Key words:** *Artificial intelligence, Chemistry education, AI in chemistry education.*

## АНДАТПА

Бұл диссертация Қазақстандағы химия білім беру саласында жасанды интеллект (ЖИ) құралдарын, атап айтқанда ChatGPT құралын пайдалануды зерттеуге арналған. Зерттеудің негізгі мақсаты – химия біліміндегі ЖИ құралдарын пайдаланудың мүмкіндіктері мен қиындықтарын, этикалық мәселелерін және оқу нәтижелерін арттырудағы нақты тиімділігін анықтау. Зерттеуде химия білімінде ЖИ құралдарын қолданудың тиімділігі, студенттердің ЖИ құралдарын қолдану кезінде туындайтын этикалық және жеке деректерді қорғау мәселелері, сондай-ақ химия пәнін оқытуда ЖИ құралдарын пайдаланудың негізгі артықшылықтары мен кемшіліктері қарастырылды.

Зерттеу аралас әдіс арқылы жүргізілді, оған Қазақстан университеттерінде химия пәнін оқып жүрген 108 студент арасында жүргізілген сандық сауалнамалар мен сарапшылардың ЖИ арқылы шығарылған химиялық есептерге сапалық бағалауы кірді. Теориялық негіз ретінде ЖИ-дің білім беру процесіндегі интеграциясы, этикалық салдары және педагогикалық әсері туралы қолданыстағы әдебиеттер пайдаланылды.

Зерттеу нәтижелері бойынша, студенттер ChatGPT құралын оның тиімділігі, түсінікті түсіндіруі және өз бетінше білім алуды жеңілдету мүмкіндігі үшін белсенді түрде қолданатыны анықталды. Негізінен, студенттер бұл құралды химиялық есептерді шешу мен емтиханға дайындық барысында пайдаланады. Алайда, зерттеу барысында дәлдік мәселелері, логикалық сәйкессіздіктер және қазақ тіліне жеткілікті дәрежеде бейімделмеуі сияқты маңызды шектеулер де анықталды. Сонымен қатар, академиялық адалдықты бұзу, технологияға шамадан тыс тәуелділік және ЖИ құралдарының ақылы нұсқаларына қолжетімділіктің әркелкілігі секілді этикалық мәселелер де назарға алынды.

Бұл диссертация ЖИ құралдарын білім беру саласына енгізудің теориялық және практикалық аспектілерін нақты эмпирикалық мәліметтер арқылы толықтырады. Сонымен қатар, зерттеу нәтижелері негізінде ЖИ-ді оқу процесіне интеграциялаудың стратегиялары ұсынылып, арнайы оқыту бағдарламаларын әзірлеу, тілдік бейімделуді жақсарту және этикалық нормаларды қалыптастыру бойынша нақты ұсыныстар жасалды. Зерттеу нәтижелері химия білім беру саласындағы ЖИ-ді жауапкершілікпен және тиімді пайдалануға ұмтылған оқытушыларға, саясаткерлерге және әзірлеушілерге бағытталған.

***Кілт сөздер:*** жасанды интеллект, химияны оқыту, химия біліміндегі ЖИ.

## АННОТАЦИЯ

В данной диссертации рассматривается интеграция инструментов искусственного интеллекта (ИИ), в частности ChatGPT, в процесс обучения химии в Казахстане с целью выявления проблем, возможностей, этических аспектов и практической эффективности этих инструментов в повышении результативности обучения. В исследовании рассматривались три ключевые области: эффективность интеграции ИИ в химическое образование, этические проблемы и проблемы конфиденциальности, связанные с использованием студентами инструментов ИИ, а также основные преимущества и недостатки использования ИИ в преподавании химии.

Используя смешанный метод, исследование объединило количественные опросы 108 студентов-химиков из казахстанских университетов и качественные экспертные оценки решений по химии, созданных ИИ. Теоретическая основа была основана на существующей литературе по интеграции ИИ в образование, этическим последствиям и педагогическому воздействию.

Результаты показали, что студенты отдают предпочтение ChatGPT из-за его эффективности, ясности и способности облегчить самостоятельное обучение, в основном используя его для решения задач и подготовки к экзаменам. Однако были отмечены существенные недостатки, включая проблемы с точностью, логические несоответствия и недостаточную лингвистическую адаптацию к казахскому языку. Среди этических проблем были отмечены академическая добросовестность, зависимость от технологии и неравный доступ к премиум-функциям ИИ. Диссертация вносит теоретический вклад, предоставляя эмпирические данные о преимуществах и ограничениях ИИ в сфере образования.

**Ключевые слова:** *искусственный интеллект, химическое образование, ИИ в преподавании химии.*

## LIST OF TABLES

- Table 1.** *Comparative Summary Table*
- Table 2.** *Advantages and disadvantages of using AI tool in chemistry education*
- Table 3.** *Presentation name of the searched articles*
- Table 4.** *Activities used in AI integrated Instruction*
- Table 5.** *Most commonly used AI tools*
- Table 6.** *The usability and Effectiveness of AI tools*
- Table 7.** *Main purposes of using AI tools in chemistry learning*
- Table 8.** *Survey findings on ethical and privacy concerns*
- Table 9.** *The value of Kendall's  $W$*
- Table 10.** *Points and means of first question*
- Table 11.** *Scores of each expert in within-expert ranks for first question*
- Table 12.** *Points and means of second question*
- Table 13.** *Scores of each expert in within-expert ranks for second question*
- Table 14.** *Points and means of third question*
- Table 15.** *Scores of each expert in within-expert ranks for third question*
- Table 16.** *Points and means of fourth question*
- Table 17.** *Scores of each expert in within-expert ranks for fourth question*
- Table 18.** *Points and means of fifth question*
- Table 19.** *Scores of each expert in within-expert ranks for fifth question*
- Table 20.** *Advantages of AI tools*
- Table 21.** *Disadvantages of AI tools*

## LIST OF FIGURES

*Figure 1. Process Flow for AI Tool Selection*

*Figure 2. Frequency of AI tool usage in chemistry studies*

## **LIST OF ABBREVIATIONS**

*AI – Artificial intelligence*

## **LIST OF APPENDICES**

*Appendix 1. Survey Questions*

*Appendix 2. Questions from UNT in Kazakh language*

*Appendix 3. Questions from UNT and their solutions by ChatGPT in English language*

# INTRODUCTION

As Stephen Haking say, “Success in creating AI would be the biggest event in human history. Unfortunately, it might also be the last, unless we learn how to avoid the risks.” Due to the infinity of the human thinking system, new innovative and technological changes are happening and developing in every century. One such technological innovation is artificial intelligence. Artificial intelligence(AI) is an intelligent machine or software technology that can think at a human level thanks to a large database. Artificial intelligence (AI) is the basis of many service sectors that have adopted new information technologies (S. J. Russell, P. Norvig, 2016). People can use AI for different sides, like finding the answers of unclear questions; help to write assignments; editing the photos by writing details. Although the roots of artificial intelligence date back several decades, there is now a clear consensus that intelligent machines with learning, reasoning and adaptive capabilities are particularly important (Arrieta, A. B., Díaz-Rodríguez,2020). It is thanks to these features that AI methods are achieving unprecedented levels of performance when learning to solve increasingly complex computational tasks, which makes them crucial for the future development of human society (D. M. West, 2018). Recently, the complexity of AI-powered systems has grown so much that they do not require human intervention to design and deploy. When decisions from such systems ultimately affect people's lives, such as medicine, law or defense, there is a need to understand how such decisions can be supported by AI techniques.

Artificial intelligence (AI) is a field of research that combines applications of machine learning, algorithmic generation, and natural language processing. Artificial intelligence applications are changing educational tools. AI has various educational applications such as personalized learning platforms that promote student learning, developed virtual labs, and bots that help with homework and projects. Although artificial intelligence has many benefits and advantages in improving student knowledge and teacher experience, it has not been properly applied at the ethical level (Akgun, S., & Greenhow, C., 2022).

**Significance of the work.** This study is significant for several reasons, both practical and theoretical. As Artificial Intelligence becomes more accessible and widely implemented in educational environments, it is vital to understand its specific contributions, limitations, and ethical implications in subject-specific contexts such as chemistry. By examining how AI tools, particularly ChatGPT, are used in chemistry education, the study provides timely and critical insights into how these technologies affect the learning experience, academic integrity, and pedagogical strategies. The study holds value for educators, as it helps them understand how to integrate AI tools

effectively to enhance instruction, personalize learning experiences, and improve student engagement. It also supports students by clarifying the educational benefits and potential risks of using AI tools, helping them navigate challenges such as overreliance on technology or inadvertent breaches of academic honesty. For educational technologists and developers, the study offers an evidence-based evaluation of AI tool performance in chemistry, particularly in relation to task accuracy, content relevance, and the logic of solution processes. This can inform future development of more reliable, subject-specific AI applications tailored to classroom needs.

Importantly, the study also addresses the ethical dimension of AI use in education. It emphasizes the need for clear guidelines around fairness, data privacy, authorship, and responsible use, which are essential for maintaining trust and integrity in academic settings. By highlighting both the opportunities and the ethical risks associated with AI, the study contributes to the creation of informed policies and teaching practices. Overall, the findings of this research are expected to help shape best practices for using AI in chemistry education, support the development of ethical AI literacy among students and teachers, and guide future research in educational innovation and digital pedagogy.

**Background theories.** In many education systems, chemistry teachers now confront a dual challenge: they must grasp the discipline itself and learn to orchestrate artificial-intelligence (AI) tools that are evolving faster than curricula. International reviews show that while AI promises individualized instruction and vivid visualisations of molecules and reactions, most chemistry teachers still feel unprepared to weave these tools into lesson plans, citing the absence of subject-specific training and clear integration models . A.O. Alli’s meta-analysis of virtual labs and adaptive platforms likewise notes that technical costs and weak AI literacy leave many educators “flying blind” when they try to deploy intelligent tutors in the classroom . Reliability is another sticking point. Empirical studies report that chatbots handle recall-level multiple-choice items reasonably well yet stumble on multi-step, representation-heavy problems; Clark (2023) puts ChatGPT’s overall accuracy at just 44 percent on general-chemistry exams, a figure echoed by Fergus, Botha and Ostovar (2023) for higher-order assessment items . Language matters too: cross-version work by Daher, Diab and Rayan (2024) shows that error rates climb when prompts move from English to other languages, raising equity concerns in multilingual settings . Ethical questions layer onto these technical limits.

The national context mirrors the global picture: Kazakh educators report the same scarcity of AI-ready teaching materials and professional-development pathways that Yildirim & Akcan (2024) diagnose elsewhere, and integration into the chemistry

curriculum remains sporadic . Mapping these intertwined issues-teacher readiness, tool reliability, and academic integrity-creates the theoretical backdrop for the present study, which seeks to quantify benefits and risks of large-language models in chemistry education and to suggest ethically grounded adoption strategies.

**Relevance.** The integration of Artificial Intelligence (AI) into educational settings has become increasingly prevalent, reshaping traditional pedagogical practices. In the context of chemistry education-a discipline that combines theoretical knowledge with answering questions and complex problem-solving-AI offers significant potential to enhance learning through adaptive platforms, virtual laboratories, and intelligent tutoring systems. As educators and institutions begin to adopt AI-driven tools, it becomes essential to evaluate the effectiveness of these technologies in supporting student learning outcomes and engagement.

This study is particularly relevant at a time when digital learning tools are expanding rapidly, yet there remains a lack of comprehensive understanding regarding their pedagogical impact, influence on education, and ethical implications. By addressing how effectively AI is integrated into chemistry education, the study provides insights into the practical utility and educational value of such tools. In parallel, it is also important to keep in mind the importance of ethical issues that arise when using artificial intelligence tools, such as fair access, academic integrity and authorship. Additionally, the study seeks to balance the discourse by examining both the advantages-such as personalized learning and improved accessibility-and the disadvantages, including dependency on technology and the risk of academic dishonesty. Through this holistic approach, the research contributes to the development of responsible, effective, and ethical AI use in science education. The findings aim to inform educators, policymakers, and developers about best practices and potential pitfalls, making the study timely and valuable in the evolving educational landscape.

**Problem Statement.** The growing integration of Artificial Intelligence (AI) tools into education presents both opportunities and challenges, particularly in subject areas that require conceptual understanding, problem-solving, and precise reasoning, such as chemistry. While AI tools like ChatGPT are increasingly used by students to solve chemistry problems and assist in learning, there remains a significant gap in understanding how effectively these tools contribute to educational outcomes in chemistry. Moreover, there is limited research evaluating the reliability, clarity, and subject accuracy of AI-generated content in the context of science education. At the same time, the rapid adoption of AI in academic settings raises important ethical and privacy concerns, including issues related to academic integrity, authorship, data

security, and unequal access to technology. Students may use AI tools without fully understanding the ethical boundaries of such use, which can undermine learning objectives and compromise fairness in assessment.

Despite the potential benefits of AI for personalized learning and increased accessibility, the lack of critical evaluation regarding its effectiveness, limitations, and ethical use in chemistry education represents a pressing issue. Therefore, there is a need for comprehensive research that explores the educational value of AI tools, examines their advantages and disadvantages, and assesses the ethical implications of their use from both student and expert perspectives.

### **Research Questions**

1. How effectively is AI integrated into chemistry education?
2. What ethical and privacy concerns arise when students use AI tools in chemistry education?
3. What are the main advantages and disadvantages of using AI tools in teaching chemistry?

**Aim.** The aim of this study is to investigate the effectiveness of using AI tools in the education of chemistry learners and to explore the advantages and limitations of these tools. The research focuses on how ChatGPT can be used to solve chemistry questions, and how experts evaluate the quality of these tasks. The importance of ethical considerations among users is also emphasised.

### **Objectives**

- To analyze the effectiveness and usefulness of artificial intelligence tools in the teaching and learning of chemistry
- To identify areas of application of artificial intelligence tools in chemistry education
- To analyze the ethical considerations of students when using artificial intelligence tools
- To analyze the accuracy of ChatGPT in translating and solving chemistry questions.
- To evaluate the clarity and logic of solution processes provided by ChatGPT.
- To check how well ChatGPT's answers match the chemistry subject context.
- To measure the reliability of ChatGPT's answers based on expert opinions.
- To identify the strengths and weaknesses of using ChatGPT in chemistry education through expert evaluation.

**Theoretical value.** This research adds to the growing academic interest in the use of AI in education, especially in STEM subjects like chemistry. By focusing on ChatGPT, the study provides new insights into how AI-generated content can support teachers and students. It also helps researchers understand the limitations and risks of using AI tools in the classroom. The findings can be useful for teachers, curriculum designers, and AI developers who want to improve educational technologies.

# **1. Theoretical part**

## **1.1. AI tools**

As time goes by, science and technology are developing very fast, and many new tools related to artificial intelligence are being created. Some of these tools provide information that meets almost all human needs, while others are made for specific areas. Today, there are large and popular artificial intelligence tools that are widely used by people. For example: ChatGPT, Gemini, Yandex AI, and DeepSeek. In the following section, I will discuss the definitions, advantages, disadvantages, and areas of use of these tools.

### **1.1.1. ChatGPT**

ChatGPT is an AI-based conversational agent developed by OpenAI. It is well known for writing text that sounds natural, helping people learn with personal tutoring, and supporting work in areas like education, healthcare, marketing, and software development. ChatGPT is a well-known and popular artificial intelligence (AI) tool that many people use in different areas. One of the main reasons people like it is because it gives answers that are clear, human-like, and usually correct. For example, in a medical test related to brain scans (called neuroradiology), ChatGPT 4.0 got an accuracy of 64.89%. That is, it answered correctly more than half the time, which was better than at least some other AI systems tested at the same time (Rishi Gupta et al., 2024). It also performed well when it had to determine decisions based on ethics or what is right and wrong, particularly in the medical domain (Hillary Chu et al., 2024). Beyond being smart, ChatGPT is also insanely creative and its effect on education has been nothing short of helpful. It can write stories, innovate completely new concepts, and offer students fun and interesting ways to learn. Teachers could use it to create worksheets, quizzes and other educational materials built to fit the students' needs (Sukhpal Singh Gill et al., 2023). It's part of the reason why so many people are using ChatGPT in schools and with online learning. Another nice thing about ChatGPT is that it's quite versatile. Search by keyword. It can be applied to a wide range of fields. In education, for example, the technology can assist students with writing essays, provide feedback and even take on the aspect of a tutor that answers questions. Doctors, in health care, could use ChatGPT to assist in decision-making or discussing ethical questions. In commerce and computer programming, it can generate marketing copy, respond to emails or chats with customers, and even write software code (Mohammad Fraiwan & Natheer Khasawneh, 2023; Hillary Chu et al., 2024).

But useful as ChatGPT is, it also has its share of issues. It is wrong or biased in its responses, but it is a product of the intelligence as trained from information available on the internet, which at times may be false or not fair (Aamo Iorliam & Joseph

Abunimye Ingio, 2023). This is a very serious problem when people use it in real situations, such as in medicine or law, where errors can be deadly. Privacy is another concern, as ChatGPT was trained on large datasets that could contain personal or sensitive information. Also, ChatGPT only knows things until a certain point in time (e.g., to 2021), so it cannot always provide the latest news or updates (Sukhpal Singh Gill et al., 2023). Another problem is that ChatGPT is not a human. It can give answers quickly and sound smart, but it does not have feelings, emotions, or real understanding like a person. In fields like education and healthcare, this lack of human connection can be a problem because empathy and personal care are very important (David Baidoo Anu & Leticia Owusu Ansah, 2023).

Still, many people find ChatGPT very helpful. In education, teachers use it to plan lessons, make student exercises, and give personal feedback. This has changed how teaching works in both classrooms and online. In healthcare, it is used as an extra tool to help doctors think about diagnoses and ethical problems, though final decisions still need to be made by humans (Hillary Chu et al., 2024). In business and software engineering, ChatGPT facilitates marketing, customer support, and software development through content generation and code assistance (Mohammad Fraiwan & Natheer Khasawneh, 2023).

### **1.1.2. Yandex AI (YaLM)**

Yandex AI is a new artificial intelligence tool created by Yandex, a large technology company in Russia. It is built using a special system called YaLM, which stands for "Yet another Language Model." Yandex AI is made mainly for working with Russian. This means it is very good at reading, understanding, and writing in Russian, which makes it very helpful for apps and tools used by Russian speakers. One of its best advantages is that it works fine even on less powerful computers or devices. This makes it more affordable and easier to use in settings without high-tech access. One of its strong suits of Yandex AI is that it plugs into other Yandex services well. For instance, it contributes to making a better Yandex search engine and powers the voice assistant called Alice. This is to ensure a seamless and connected user experience when using different Yandex services. But Yandex AI is far from categorically pious. It's not very useful for other languages. It has only limited support for English, so it's not going to be a great option for those in need of something that works in multiple languages. First, Yandex AI is not as good at doing really hard tasks compared with more sophisticated AI systems like ChatGPT or Google's Gemini. It may do less well, for instance, in medical decision-making or writing creative stories. The lack of support for developers is another obstacle. Yandex AI has fewer resources, unlike ChatGPT or other major platforms with dozens of tutorials, guides, and community support. For

those who want to build new apps or solve problems based on Yandex AI, this becomes more challenging. Even with these problems, Yandex AI still works very well in many different areas. It's the main product Yandex has developed in its quest to improve the search results in its own search engine - and a technology that can help people ask questions in a more human way. It also has "Alice," the voice assistant that answers questions from users. Brands use Yandex AI to create Russian-speaking chatbots in customer service. It also assists in generating digital content, news items, or online posts to be made to/from Russian media websites.

### **1.1.3. Gemini**

Gemini is an intelligent language model created by Google AI. It is designed to be in competition with other advanced tools such as ChatGPT. One of the best features of Gemini is that it can find and utilize real-time information directly from Google Search. This means that it provides answers analyzing the most current and exact data available at the moment (Manali Shukla & Ishika Goyal, 2024). This is especially helpful for people working in schools, universities, or businesses, where having the latest information is very important. One useful part of Gemini is that it works well with popular Google tools like Google Docs, Google Sheets, and Gmail. This helps users write documents, send emails, and organize data in one place without moving between different apps. It saves time and helps to do tasks more easily (Aamo Iorliam & Joseph Abunimye Ingio, 2023; Manali Shukla & Ishika Goyal, 2024). Also, Gemini has powerful safety tools, which use smart protection systems to keep user information safe. It's important for companies, schools, and other groups that need to protect private data from being lost or taken. According to this benefit, many companies collaborate and work with Gemini to keep their data secure (Aamo Iorliam & Joseph Abunimye Ingio, 2023).

Unfortunately, Gemini has some problems. In medical tests, it did not perform as well as ChatGPT. For example, in one test about reading brain scans, it only got 55.73% of the answers correct, which is lower than ChatGPT 4.0 (Rishi Gupta et al., 2024). This shows that it might not be the best choice for very serious or detailed medical work. Also, some users say that Gemini is not as creative as ChatGPT. It can give useful answers, but it might not be as good at writing stories or coming up with new and imaginative ideas (Aamo Iorliam & Joseph Abunimye Ingio, 2023). Its performance in tasks that require deep thinking about ethics such as what is right or wrong in healthcare also seems weaker than other tools (Hillary Chu et al., 2024). Because of this, it is not the best option for jobs where clear moral thinking and expert decisions are needed.

Even with these issues, Gemini is used in many helpful ways. In schools and universities, students and teachers use it to help with writing assignments and doing research. Since it connects directly to Google Search, it can give updated and relevant information for school projects. In business, Gemini is useful for writing emails, creating documents, and organizing information in spreadsheets. This makes everyday work faster and easier. In healthcare, Gemini can still be used, but only as an extra tool. Its lower performance in medical and ethical tasks means that it should support human experts, not replace them.

#### **1.1.4. DeepSeek**

DeepSeek is a special type of artificial intelligence (AI) tool that is known for being open-source and affordable. It was created by a company called DeepSeek Lab. Unlike some other AI models that cost a lot of money to use, DeepSeek is designed to be much cheaper and more accessible to people around the world. It focuses on helping users save money without losing performance. In fact, studies show that DeepSeek-R1 can be up to 95% less expensive than other popular AI models, but still performs very well in tasks like solving math problems (Samuel Ariyo Okaiyeto et al., 2025). The best thing about DeepSeek is that it is an offline tool. That means it doesn't have to be connected to the internet to function. This will be particularly useful in hospital and government office settings, where privacy will be of the utmost importance (Abdulrahman Temsah et al., 2025). One advantage to AI tools that work offline is that they are better able to protect sensitive pieces of information; this is a large concern within these contexts. DeepSeek leverages a novel AI design methodology, called the Mixture of Experts (MoE). This setup makes sense because it only brings in the parts of the system needed to do a certain job. This makes it work faster and use less computer power. It also helps the AI to think step by step, a way of thinking called chain-of-thought reasoning. This helps DeepSeek solve tasks better and check answers (Samuel Ariyo Okaiyeto et al., 2025). Another major point of DeepSeek is that it is free and open-source. This means everyone can use it for studying, learning, searching information and others. DeepSeek is beneficial for people in developing countries who want to research AI or create their own AI tools but don't have much money. It opens up opportunities for humans, who want to join and influence the world of AI.

DeepSeek has many positive sides, but it also has some limits. For example, until now we don't know enough about how well it works in major areas like medicine for decision-making. More experiments and real-life testing are needed to find out if it can truly be trusted in such fields (Abdulrahman Temsah et al., 2025). In addition, even if DeepSeek can work offline without an internet connection, the company says that some

user data will still be used for developing the system. It may be an anxiety for consumers who are taking a strong stance on privacy.

Even so, DeepSeek is becoming more popular in other places. It is already being used in healthcare and medical learning, as keeping data private and working offline are very important features. Deepseek is comfortable for schools, small businesses, and organizations without marketing budgets, particularly in third-world countries. People use it for teaching, research, and other everyday uses. Finally, DeepSeek is available for everyone in the world. It's useful tool for learning and an area for creativity and innovation. It enables researchers and students to investigate how AI functions, experiment with new ideas, and build better systems. It is thus a useful instrument for the next level of AI development worldwide (Samuel Ariyo Okaiyeto et al., 2025)

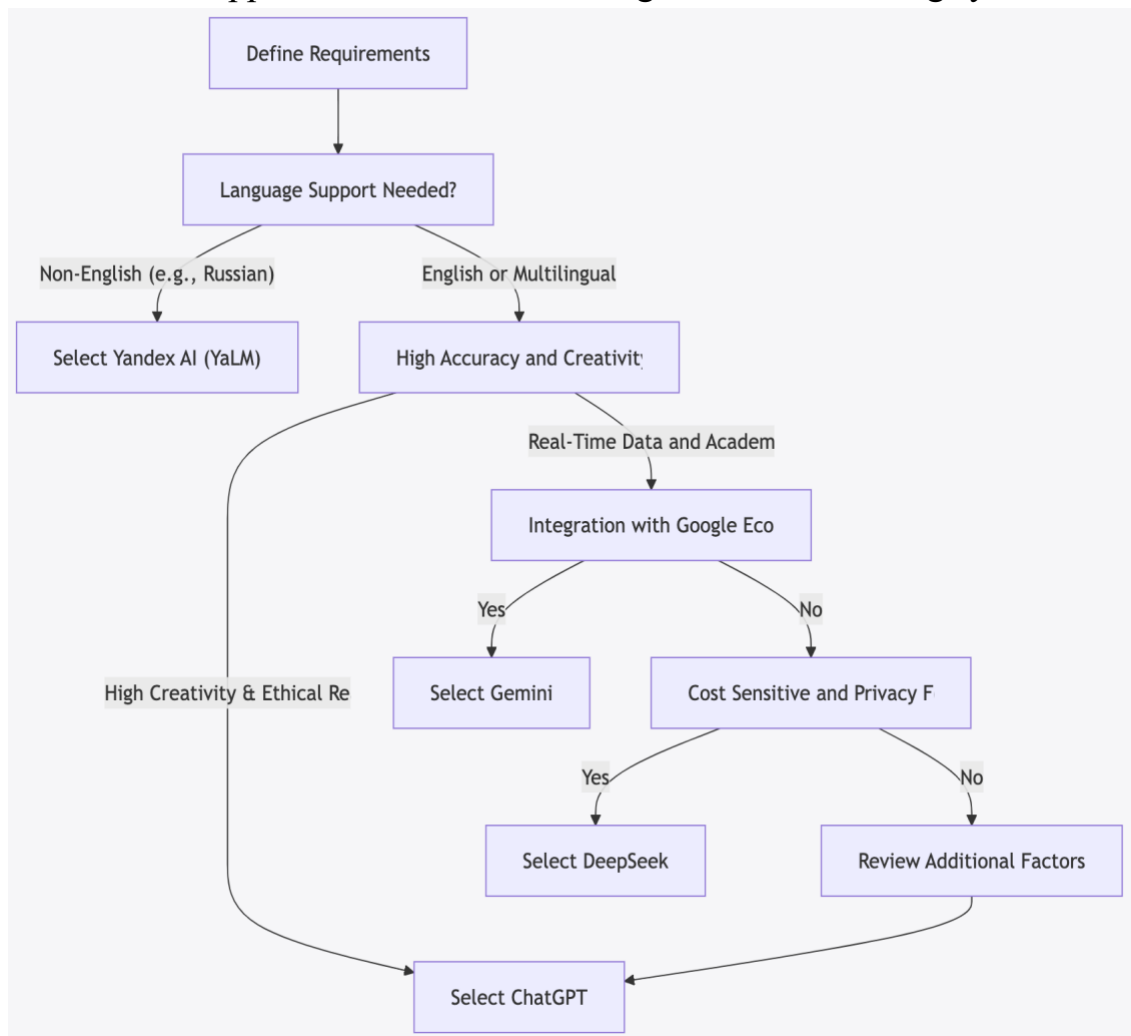
In total, the key features of ChatGPT, Yandex AI, Gemini and DeepSeek, highlighting their main benefits, drawbacks and fields of application are summarized in Table 1.

**Table 1.** Comparative Summary Table

AI tool	Advantages	Disadvantages	Application areas
ChatGPT	<ul style="list-style-type: none"> <li>-High human-like accuracy</li> <li>-Strong creative writing and ethical reasoning</li> <li>-Versatile in education, healthcare, and business</li> </ul>	<ul style="list-style-type: none"> <li>-May produce incorrect or biased output</li> <li>- Privacy concerns</li> <li>- Knowledge cutoff limits updates</li> </ul>	<ul style="list-style-type: none"> <li>- Education (tutoring, content creation)</li> <li>- Healthcare (diagnostic support)</li> <li>- Marketing &amp; coding</li> </ul>
Yandex AI (YaLM)	<ul style="list-style-type: none"> <li>- Optimized for Russian language</li> <li>- Cost-efficient and works on less powerful hardware</li> <li>- Integrated with Yandex products</li> </ul>	<ul style="list-style-type: none"> <li>- Limited language support</li> <li>- Less suited for complex tasks compared to larger models</li> <li>- Fewer developer resources</li> </ul>	<ul style="list-style-type: none"> <li>- Search and virtual assistance</li> <li>- Customer service chatbots</li> <li>- Russian media content generation</li> </ul>
Gemini	<ul style="list-style-type: none"> <li>- Real-time data access</li> <li>- Seamless integration</li> </ul>	<ul style="list-style-type: none"> <li>- Lower diagnostic accuracy compared to</li> </ul>	<ul style="list-style-type: none"> <li>-Academic research and</li> </ul>

	with Google Workspace - Enterprise-grade security	ChatGPT - Less creative output - Ethical reasoning limitations	writing - Business productivity tools - Limited healthcare support
DeepSeek	- Open-source and low cost - Offline deployment enhances privacy - Efficient Mixture of Experts architecture	- Insufficient real-world performance data in critical fields - Some privacy policy concerns	- Healthcare (clinical decision support) - Cost-sensitive educational deployments - AI research

In addition, we provide a flowchart that illustrates the process of selecting an artificial intelligence tool based on specific requirements such as language support, cost-effectiveness, application domain, and integration with existing systems below.



**Figure 1. Process Flow for AI Tool Selection**

Selecting the right AI tool is, for the most part, a matter of what the user needs. If you need a person to assist with writing stories, or to study or teach, you could do a lot worse than ChatGPT, which is excellent at producing clear explanations and content. Yandex AI would be more suitable if the job is related to a Russian-speaking audience or the Russian market. It's specifically designed for Russian, and it excels in that regard. Gemini is best for those who are in the Google ecosystem with tools such as Gmail, Docs, or Sheets and require the most current information. It's tightly integrated with Google services and can provide answers with up-to-date data from Google Search. DeepSeek is ideal for those people who are looking to save money (basic package available) or need to protect their data. It is low-cost, and it can operate without the internet, which can be critical in sensitive locations like hospitals, research labs, or government offices.

Instead, when we consider different types of AI tools, we see that each one is strong in different ways and for different reasons. ChatGPT is the right choice when a task requires a good deal of creativity, say to tell stories or produce content, and when it's essential to pay attention to what is right and wrong. It is of the greatest assistance to us in education, in writing, and in ethical decision-making. Here, a particular strength of Yandex AI shines: the Russian language. It works very accurately for Russian, and is also a less expensive choice for businesses or individuals conducting business in a Russian-speaking territory. This is useful in local markets where cost and language are a factor. Gemini is strong when folks need perfect information that's constantly fresh. It's able to live-search the web, and it works great in stuff like Google Docs, Gmail, and Sheets. This helps people work faster and smarter. However, Gemini is not as good at creative tasks, like storytelling or writing imaginative content. DeepSeek is another strong option because it is open-source, which means people can use and change it freely. It can also work offline, so it does not need an internet connection. This is very helpful in places where privacy is important, like hospitals or research centers. It is also a low-cost solution, so it's useful for schools, small businesses, and other places that need to save money and protect sensitive information.

## **1.2. AI in chemistry education**

In the age of advanced science and technology, students and teachers are using artificial intelligence positively in the field of education and benefiting from it. As a chemistry teacher, I researched articles on the use of artificial intelligence in chemistry education.

Russian scientists Vladimir L. Kodkin, Ekaterina V. Artemeva(2024) described the progress and uses of neural networks, especially ChatGPT, in different areas like chemistry education and research in the article "ChatGPT: Application in Chemistry

Education and Challenges”. It explains the advantages of artificial intelligence and ChatGPT, like their ability to handle large amounts of data, make personalized learning systems, and give suggestions for solving problems. It shows real-life examples of how ChatGPT can help improve learning in chemistry, like making tests, generating multiple-choice questions, studying chemistry topics, and writing scientific ideas in chemistry. It ends by discussing the exciting possibilities of AI to solve complex problems that might be beyond human abilities, opening doors for more research and collaboration between humans and intelligent machines.

From Ohio State University, scientist Ted M. Clark (2023) studied how well an artificial intelligence chatbot called ChatGPT could answer chemistry exam questions in “Investigating the Use of an Artificial Intelligence Chatbot with General Chemistry Exam Questions” article. They tested ChatGPT on exams from two chemistry classes. ChatGPT did a good job at understanding the main ideas in questions that had multiple-choice answers, even when there were lots of chemical symbols. However, it only got 44% of the answers right, which was lower than what most students got. When it came to questions where students had to write out their answers, ChatGPT was good at understanding language, especially for questions that needed general knowledge. But it had trouble with questions that needed specific skills, especially if those skills were taught in lectures. Even though some of its wrong answers sounded convincing, ChatGPT isn't yet good at giving reliable answers or explanations for many exam questions. However, it could be useful for making assignments where students work on improving ChatGPT's answers.

Researchers Margarida Figueiredo and M. Lurdes Esteves (2014) from Portugal looked into how lab classes in Chemistry learning are taught using artificial intelligence in “Lab Classes in Chemistry Learning an Artificial Intelligence View” article. They studied 702 students in 10th grade in Portuguese Secondary Schools. They used a method called k-Means clustering to group the data into different segments. Also used decision trees to create models explaining the segmentation. The findings revealed that most students think doing experiments is very important for learning Chemistry. They also found that how much importance students place on research in chemistry learning depends a lot on how much they participate in lab work.

In the article "Artificial intelligence in chemistry" by N.A.B Gray (1988), it's mentioned that they review the methods and tools of artificial intelligence. Chemical problems are seen as good areas to test new artificial intelligence methods. Artificial intelligence techniques can be used to solve real-life chemistry problems. In chemistry, these techniques can help with tasks that need chemical knowledge, as well as tasks where chemistry is just one part of the problem. Programs that use encoded chemical knowledge can help solve problems like figuring out chemical structures, planning

synthesis, and designing experiments. Robotics methods that use artificial intelligence and expert systems can make chemical instruments work better. Systems that understand human language could make it easier to deal with chemical information, and artificial intelligence techniques could make computer-based teaching even better.

In their paper titled "Artificial Intelligence Generative Tools and Conceptual Knowledge in Problem Solving in Chemistry" Wajeeh Daher, Hussam Diab, and Anwar Rayan (2023) discuss how artificial intelligence (AI) has become increasingly important in education and problem-solving. They focus on the use of ChatGPT, an AI tool, to support learning in chemistry. The researchers examined the problems ChatGPT has when trying to understand and answer chemistry questions related to the topic called *Introduction to Material Science*. To examine this, the researchers used a method developed by Holme and others. This method includes different categories such as transfer (how knowledge is used in new situations), depth (how detailed the understanding is), predict/explain (how well the model can explain or guess outcomes), problem solving, and translate (changing information into different forms, like from words to diagrams). The researchers fed ChatGPT thirty chemistry problems on *Introduction to Material Science* and tasked it to answer them. After verifying the answers, they discovered that ChatGPT had struggled tremendously to grasp the concepts behind the questions. It particularly had difficulty in two areas: the extent to which it used representations (symbols, graphs, or diagrams) and the depth of its understanding of topics. Because it had problems using representations correctly, it did not reach generalization and transfer its knowledge to solve problems in new ways.

In their paper "Evaluating Academic Answers Generated Using ChatGPT", American academics Suzanne Fergus, Michelle Botha and Mehrnush Ostovar (2023) explore the use of technology in education, especially in the context of the COVID-19 pandemic. They focused on the Chat Generative Pre-Trained Transformer (ChatGPT), an artificial intelligence technology that generates answers to conversation questions based on user prompts. This study set out to find out how ChatGPT responds to chemistry assessment answers and the implications this may have for learning and assessment. The researchers analyzed two chemistry modules in the first and second year of a pharmacy program, comparing answers given by ChatGPT to marks in an end-of-year exam. They found ChatGPT excelled in responding to questions needing knowledge and comprehension, but had difficulties when answering questions needing the application of knowledge and understanding non-textual information. The paper also quality-checks ChatGPT responses and finds fraud risk low. Furthermore, the researchers also hypothesize that conversations about ChatGPT might catalyze conversations around academic integrity and the design of assessment, just as educational conversations have occurred around COVID-19.

Wajeeh Daher, Hussam Diab, and Anwar Rayan(2024) explored the growing use of generative artificial intelligence tools in different fields, especially education, in their article "Generative Artificial Intelligence in Chemistry Problem Solving across Versions and Languages". They focused on ChatGPT versions 3.5 and 4.0, which are known for their ability to understand multiple languages, aiding both students and teachers. Their research aimed to compare the performance of these versions, specifically looking at how accurately ChatGPT answered questions in English compared to Arabic. They tested ChatGPT with 39 chemistry problems typically found in 6th-7th grade curriculum, consisting of both open-ended and multiple-choice questions. They categorized each response as accurate, partially accurate, or inaccurate. Their analysis revealed significant improvements in version 4.0, especially in handling Arabic. However, despite these enhancements, the study found that ChatGPT's responses in English were consistently more accurate than those in Arabic. To address this issue, they proposed either including more Arabic data during training or implementing a method where questions are first translated from Arabic to English, answered, and then translated back into Arabic for users. This approach leverages the higher accuracy of English responses to benefit Arabic-speaking users and potentially improves outcomes for users in other languages, especially in science education contexts like chemistry.

The next research work is Mary E. Emenike and Bright U. Emenike(2023) article "Was This Title Generated by ChatGPT? Considerations for Artificial Intelligence Text-Generation Software Programs for Chemists and Chemistry Educators". They discuss the concerns raised by the release of ChatGPT, a free text-based system developed by OpenAI, especially regarding honesty in academics and how students are evaluated across different levels of education. However, they believe that these systems can have a bigger impact beyond just teaching and learning chemistry. In their study, the researchers said that AI systems such as ChatGPT can benefit students, teachers, and college administrators for teaching, research, and professional activities. They look at a variety of ways students and teachers might use ChatGPT, and they flag both the benefits and challenges, as well as fairness and accessibility concerns. For instance, students might use ChatGPT to receive aid in answering homework questions, writing up lab reports, conducting research, studying, and finding tutors. At the same time, teachers can use it not just for making assessment tools-there are many features for that-preparing class resources, carrying out research, and managing their professional life-their career and speaking-as well. They seek to stimulate discussions on how to use artificial intelligence technology in the best and most responsible ways.

The same thing in the article "Proposal to integrate artificial intelligence tools in chemistry teaching in higher education" by the authors R. Perezzan, S. Montalvo-

Quirós, R. Rama-Ballesteros, D. Herráez-Aguilar(2024). In today's education, Artificial Intelligence (AI) is becoming increasingly important, especially in Natural Sciences like Chemistry. This study proposes using AI tools such as ChatGPT 3.5, Microsoft Bing, and Wolfram Alpha to enhance learning in Chemistry. These tools help students understand complex concepts and improve their ability to learn independently. Students can ask different types of questions, problems from AI and make communication. They can discuss answers with the teacher to develop their critical thinking skills. These pathways will help to improve knowledge in chemistry education, but students will meet academic and professional challenges. Through the integration of AI into the learning process, education nowadays is becoming more contemporary and individualized. AI tools do more interactive and effective learning by allowing students to explore concepts independently and creatively. In sum, integrating AI into education opens up new possibilities for learning and preparation for all aspects of life.

The article “Limitations and Potential of ChatGPT in Chemistry Teacher's Work” by Tu Nguyen and Yanika Sirichokcharoenkun (2023) looks at the strengths and weaknesses of using ChatGPT in chemistry teaching. ChatGPT is a chatbot that communicates using language and is used in education, including chemistry. The authors discuss its problems, such as difficulty understanding context, inability to recognize images, and lack of specialized training data. They also point out benefits like personalized help, improving students' critical thinking, and promoting independent learning. However, they mention some negative points, including mistakes in accuracy and ethical concerns. Finally, the authors suggest ways teachers and developers can better use ChatGPT in chemistry teaching while dealing with these issues.

In the article "AI-Professional Development Model for Chemistry Teacher: Artificial Intelligence in Chemistry Education," Bekir Yildirim and Ahmet Tayfur Akcan (2024) looked at how to train chemistry teachers to use artificial intelligence. Their study made a Professional Development Model (PDM) to help teachers get better at using artificial intelligence. The research included 17 chemistry teachers and used a qualitative case study method. The team got data from documents, teacher interviews, and tips from AI tools. The documents showed that teachers need to know AI basics, such as knowledge, tools, skills, ethics, attitudes, and AI understanding, to use AI well in classes. It also said teachers need special skills for teaching chemistry using AI. When checking AI tool tips from ChatGPT and Gemini, the researchers saw they matched the document data. Many teachers said they had trouble understanding AI and making lessons with AI. They suggested using AI for topics like organic chemistry, lab

experiments, and reactions. From all the data, the researchers made a 10-step plan to help chemistry teachers use AI in their work.

A. O. Alli (2025) looked at how artificial intelligence (AI) can be used to help students learn chemistry and improve it. In the 2025 paper under the title: "A Review of the Potential of Artificial Intelligence (AI) in Chemistry Education," Alli, A. O. sampled numerous scientific books, peer-reviewed articles, and reputable online references to collect information on the potential areas where AI tools may be employed in chemistry education. The research involved the use of tools like intelligent tutoring systems, virtual laboratories, and adaptive learning platforms, which can mitigate the monotonicity in the classroom environment and prepare more effective and interesting chemistry lectures for the students. AI might help students learn faster and more deeply by giving them personalized learning experiences and by explaining abstract chemical concepts in straightforward terms, the researchers concluded in their study. It also helps teach problem-solving by providing immediate feedback and prompting. But the writer also cites some obstacles. These issues can range from ethical considerations to the risk of over-reliance on technology and can extend to the necessity of adequate teacher training for effective deployment of AI tools. The research emphasizes that although AI can be used as a valuable resource in education, it will never replace teachers entirely. Finally, the paper proposes that AI is most successful when it complements rather than replaces traditional teaching. With the blend of the two, students can have better support as well as exposure to today's technology, which therefore facilitates a better and more balanced chemistry education practiced among students.

Authors studied how ChatGPT can help teachers in their daily work and improve the teaching and learning process. In the article "ChatGPT for Teachers: Practical Examples for Utilizing Artificial Intelligence for Educational Purposes", Himel Mondal et al. (2023) looked at real examples of how teachers can use ChatGPT in education. The authors explained that ChatGPT can act like a virtual assistant or tutor. It can help both teachers and students in different ways. For example, teachers can use ChatGPT to create presentation slides, write essay or multiple-choice questions, make viva (oral exam) questions, and even answer students' questions. It can also help teachers make learning materials that match students' levels of understanding, check their answers, create case study examples, plan lessons more easily, and design blended learning content for both online and classroom use. The article explains that ChatGPT can save teachers a lot of time, especially when they are busy with many school tasks. However, the writers also say that ChatGPT is not always correct - it can sometimes make mistakes or give wrong scientific information. So, teachers should always check

and edit the content before using it in class. In conclusion, the article says ChatGPT is a useful and smart tool, but it must be used carefully.

Indah Sari, Ratna Farwati, and Geby Ayu Fadhilah (2024) study how students in chemistry education use artificial intelligence (AI), like ChatGPT, while learning. In their article “Transforming the Chemistry Education Curriculum: Students’ Creativity Using AI in Lectures,” they look at students from two universities in Palembang and Serang to check their skills with AI in learning. They asked 76 students to take part in the study. The results showed that many students used AI to help them understand chemistry, write papers, and do experiments. This shows that AI can be a helpful and important tool in chemistry classes. However, students also shared some problems. These included high cost, little knowledge about AI, weak access to tools, and technical troubles. These problems stop students from using AI fully. Because of this, the writers suggested adding AI into chemistry courses through workshops, training, and simple guides to help students learn easily. The study also advises giving more technical help, better AI access, and updating lessons to match new tech. It also found student views on AI differed by gender and class level. These differences affect how AI should be used in learning. The authors believe AI can improve chemistry education with proper support.

In their study called *"Investigating the Effect of Artificial Intelligence on Chemistry and Physics Students' Achievement and Conceptual Change in Heat Change in SSS2 in Rivers State"* Zudonu et al.(2024) examined the influence of Artificial Intelligence (AI) on the academic performance and conceptual understanding of heat change among senior secondary school (SSS2) Chemistry and Physics students in Rivers State, Nigeria. The researchers employed a quasi-experimental design involving non-randomized pretest and posttest control groups. A total of 160 students from sixteen secondary schools in the Rivers West Education Zone were purposively selected to participate. To address five research questions, the authors applied descriptive statistics such as mean, standard deviation, percentages, and visual tools (column and pie charts), while five hypotheses were tested using independent sample t-tests at a 0.05 significance level. The tools used for assessment were the Chemistry Achievement Test (CAT), Physics Achievement Test (PAT), and Conceptual Understanding Test (CUT), all designed to check student learning and changes in their understanding. Also, planned interviews were held with students who had lessons using AI to understand their thoughts and learning experiences. Instrument reliability was verified using the test-retest method alongside Kuder-Richardson Formula-21, yielding a reliability coefficient of 0.78. Findings revealed that while Physics students scored slightly higher than Chemistry students in both pretests and posttests (14.40 vs. 12.45 and 46.35 vs. 44.63, respectively), the difference was not statistically significant.

However, when comparing AI-based instruction with conventional teaching methods, the results demonstrated a statistically significant advantage for students exposed to AI, both in academic achievement and conceptual understanding. Moreover, the feedback gathered through interviews indicated generally favorable attitudes among students towards AI-supported learning. The study highlights the transformative potential of AI in science education by promoting deeper understanding, improved academic performance, and student engagement. The writers finish by suggesting more long-term research to study the lasting effects of AI on student learning and support wider use of AI-based teaching in Chemistry and Physics classes.

In the article “Capabilities, Benefits, and Role of ChatGPT in Chemistry Teaching and Learning in Vietnamese High Schools”, Thuy Nguyen Linh Cao and Phong Nguyen(2023) investigate the potential of ChatGPT as a supportive tool in Vietnamese high school chemistry education. As a large language model developed by OpenAI, ChatGPT has attracted growing interest for its educational applications, particularly in enhancing both teaching and learning processes. Authors discuss ChatGPT’s effectiveness in chemistry education, analyzing its strengths and limitations. Notably, ChatGPT is recognized for its ability to support personalized learning, stimulate critical thinking, and aid in the explanation of complex scientific concepts. The conclusion reinforces these findings, emphasizing that ChatGPT’s conversational and AI-driven features allow it to deliver tailored feedback and foster deeper student engagement. The authors argue that, if applied responsibly and strategically, ChatGPT can transform traditional chemistry instruction by promoting innovative, student-centered learning environments. Ultimately, the integration of ChatGPT in Vietnamese chemistry classrooms represents a promising step toward modernizing science education and equipping students with essential skills for the future.

Summarizing the above information, the impact of artificial intelligence in teaching chemistry was discussed and literature research was conducted. The purpose of this study is to search for the use of AI in chemistry education. The research question is “what are the main advantages and disadvantages of using AI tools in teaching chemistry?”.

**Table 2.** Advantages and disadvantages of using AI tool in chemistry education

<b>Authors &amp; Year</b>	<b>Advantages</b>	<b>Disadvantages</b>
<b>Kodkin &amp; Artemeva (2024)</b> <i>“ChatGPT: Application in</i>	- Capable of handling large datasets. - Offers personalized learning systems. - Provides problem-solving	- Challenges not explicitly detailed in the abstract, but implied: still evolving

<p><i>Chemistry Education and Challenges”</i></p>	<p>suggestions.</p> <ul style="list-style-type: none"> <li>- Helps generate tests, multiple-choice questions, study topics, and scientific ideas.</li> <li>- Potential to solve complex problems beyond human capability, fostering new research and collaboration.</li> </ul>	<p>technology with potential limitations (e.g., accuracy, context understanding).</p>
<p><b>Clark (2023)</b> <i>“Investigating the Use of an Artificial Intelligence Chatbot with General Chemistry Exam Questions”</i></p>	<ul style="list-style-type: none"> <li>- Performs relatively well on multiple-choice questions (understands question text, chemical symbols).</li> <li>- Potentially useful for creating assignments where students improve the AI’s responses.</li> </ul>	<ul style="list-style-type: none"> <li>- Overall accuracy of only 44%, lower than students’ average.</li> <li>- Struggles with deeper skills needed for certain lecture-taught content.</li> <li>- Gives incorrect yet plausible-sounding explanations.</li> </ul>
<p><b>Figueiredo &amp; Esteves (2014)</b> <i>“Lab Classes in Chemistry Learning an Artificial Intelligence View”</i></p>	<ul style="list-style-type: none"> <li>- Demonstrates how AI methods (e.g., k-Means clustering, decision trees) can be used in educational research.</li> <li>- Confirms that most students value hands-on experiments for learning Chemistry.</li> <li>- Shows the role of lab participation in enhancing students’ perception of research importance.</li> </ul>	<ul style="list-style-type: none"> <li>- The abstract does not discuss direct drawbacks of AI for instruction but focuses on data analysis. Potential limitations could be related to data-driven segmentation not capturing all individual learning nuances.</li> </ul>
<p><b>Gray (1988)</b> <i>“Artificial intelligence in chemistry”</i></p>	<ul style="list-style-type: none"> <li>- AI techniques can solve real-life chemistry problems (e.g., structure elucidation, synthesis planning, experiment design).</li> <li>- Robotics and expert systems can enhance instrumentation and processes.</li> <li>- Natural language processing can improve handling of chemical information.</li> </ul>	<ul style="list-style-type: none"> <li>- Not explicitly stated in the abstract but possible issues include the complexity of encoding chemical knowledge and ensuring AI covers diverse chemistry domains effectively.</li> </ul>

	<ul style="list-style-type: none"> <li>- Potential for improving computer-based teaching in chemistry.</li> </ul>	
<p><b>Daher, Diab, &amp; Rayan (2023)</b>  <i>“Artificial Intelligence Generative Tools and Conceptual Knowledge in Problem Solving in Chemistry”</i></p>	<ul style="list-style-type: none"> <li>- Demonstrates ChatGPT’s potential to support learning and problem-solving in chemistry.</li> <li>- Can assist with some steps in transferring knowledge (e.g., general knowledge).</li> </ul>	<ul style="list-style-type: none"> <li>- Struggles with deeper conceptual understanding (representations, depth).</li> <li>- Insufficient ability to transfer detailed conceptual knowledge effectively.</li> </ul>
<p><b>Fergus, Botha, &amp; Ostovar (2023)</b>  <i>“Evaluating Academic Answers Generated Using ChatGPT”</i></p>	<ul style="list-style-type: none"> <li>- Performs well on questions requiring straightforward knowledge and understanding.</li> <li>- Poses low risk of cheating according to their assessment.</li> <li>- Encourages discussions on academic integrity and better assessment design.</li> </ul>	<ul style="list-style-type: none"> <li>- Limited in answering questions that require application of knowledge and interpretation of non-textual (visual) information.</li> <li>- Quality can vary, sometimes lacking depth or contextual relevance.</li> </ul>
<p><b>Daher, Diab, &amp; Rayan (2024)</b>  <i>“Generative Artificial Intelligence in Chemistry Problem Solving across Versions and Languages”</i></p>	<ul style="list-style-type: none"> <li>- Shows improved performance in ChatGPT version 4.0 over 3.5.</li> <li>- Offers multilingual support (English and Arabic), beneficial for diverse learners.</li> <li>- Useful for different question formats (open-ended and multiple-choice).</li> </ul>	<ul style="list-style-type: none"> <li>- Accuracy in Arabic still lags behind English.</li> <li>- May require translation-based workarounds (Arabic → English → Arabic) to improve precision.</li> <li>- Indicates possible broader challenges in languages with less training data.</li> </ul>
<p><b>Emenike &amp; Emenike (2023)</b>  <i>“Was This Title Generated by”</i></p>	<ul style="list-style-type: none"> <li>- AI systems can aid students, teachers, and school administrators in teaching, research, and professional activities.</li> </ul>	<ul style="list-style-type: none"> <li>- Raises concerns about academic integrity and honest assessments.</li> <li>- Issues of fairness and</li> </ul>

<p><i>ChatGPT? Considerations for Artificial Intelligence Text-Generation Software Programs for Chemists and Chemistry Educators”</i></p>	<ul style="list-style-type: none"> <li>- Helps with writing assignments, lab reports, research proposals, and tutoring.</li> <li>- Teachers can create assessments, class materials, and manage professional tasks.</li> </ul>	<p>accessibility arise for different student populations.</p> <ul style="list-style-type: none"> <li>- Overreliance on AI could impede genuine skill development if not monitored.</li> </ul>
<p><b>Perezan et al. (2024)</b>  <i>“Proposal to integrate artificial intelligence tools in chemistry teaching in higher education”</i></p>	<ul style="list-style-type: none"> <li>- AI tools (ChatGPT 3.5, Microsoft Bing, Wolfram Alpha) can enhance understanding of complex concepts.</li> <li>- Fosters independent learning and critical thinking via Q&amp;A with an AI tool.</li> <li>- Makes education more interactive and adaptable to current demands.</li> <li>- Prepares students for academic and professional challenges in modern contexts.</li> </ul>	<ul style="list-style-type: none"> <li>- Not explicitly detailed in the abstract, but potential concerns include ensuring students do not become overly dependent on AI, handling misinformation, and integrating AI-based approaches responsibly.</li> </ul>
<p><b>Nguyen &amp; Sirichokcharoenkun (2023)</b>  <i>“The Limitations and Potential of ChatGPT in Chemistry Education”</i></p>	<ul style="list-style-type: none"> <li>- Provides personalized help and supports self-paced learning.</li> <li>- Encourages critical thinking by offering alternative viewpoints or clarifications.</li> <li>- Popular and user-friendly in educational settings.</li> </ul>	<ul style="list-style-type: none"> <li>- Struggles with context understanding and lacks the ability to interpret visuals or symbolic representations.</li> <li>- May have limited specialized training data for advanced topics.</li> <li>- Accuracy issues and ethical concerns (e.g., potential misinformation).</li> </ul>
<p><b>Bekir Yildirim and Ahmet Tayfur</b></p>	<ul style="list-style-type: none"> <li>- Supports individualized and permanent learning</li> <li>- Helps visualize complex chemistry</li> </ul>	<ul style="list-style-type: none"> <li>- Teachers feel unprepared to integrate AI into lesson plans</li> </ul>

<p><b>Akcan(2024)</b>  <i>"AI-Professional Development Model for Chemistry Teacher: Artificial Intelligence in Chemistry Education"</i></p>	<p>concepts (e.g., molecules, reactions)</p> <ul style="list-style-type: none"> <li>- Assists in solving and analyzing chemical problems</li> <li>- Useful in lesson planning, experimentation, and student assessment</li> <li>- Can predict learning outcomes and personalize instruction</li> </ul>	<ul style="list-style-type: none"> <li>- Insufficient AI training programs tailored for chemistry educators</li> <li>- Integration of AI into chemistry curricula is still limited</li> </ul>
<p><b>A.O. Alli (2025)</b>  <i>"A Review of the Potential of Artificial Intelligence (AI) in Enhancing Chemistry Education"</i></p>	<ul style="list-style-type: none"> <li>- Gamification and interactivity increase student interest and active participation</li> <li>- AI-powered simulations help visualize complex chemical reactions and structures</li> <li>- Provides immediate feedback on problem-solving and conceptual errors</li> <li>- AI tailors instruction to individual student needs, pace, and style</li> </ul>	<ul style="list-style-type: none"> <li>- Many educators lack sufficient training to integrate AI effectively</li> <li>- Less direct teacher-student interaction could impact communication and emotional support</li> <li>- Implementing and maintaining AI tools can be financially demanding</li> </ul>
<p><b>Himel Mondal et al. (2023)</b>  <i>"ChatGPT for Teachers: Practical Examples for Utilizing Artificial Intelligence for Educational Purposes"</i></p>	<ul style="list-style-type: none"> <li>- Helps teachers generate lesson plans, slides, quizzes, viva questions, and assignments quickly</li> <li>- Creates tailored content for students with different comprehension levels</li> <li>- Can be used for essay questions, MCQs, case vignettes, and blended learning materials</li> <li>- Reduces the need for expensive materials or extensive textbooks</li> </ul>	<ul style="list-style-type: none"> <li>- Using AI may reduce the originality of teacher-created materials</li> <li>- Potential for text similarity, which may affect academic integrity</li> <li>- AI cannot provide emotional support or human guidance to students</li> </ul>
<p><b>Ratna Farwati, Indah Sari, and Geby Ayu Fadhilah(2024)</b></p>	<ul style="list-style-type: none"> <li>- AI (especially ChatGPT) helps students grasp difficult chemistry concepts via personalized explanations</li> </ul>	<ul style="list-style-type: none"> <li>- Some AI tools (especially premium versions) are costly for students</li> </ul>

<p><i>"Transforming the Chemistry Education Curriculum: Students' Creativity Using AI in Lectures"</i></p>	<ul style="list-style-type: none"> <li>- Students complete assignments, write reports, and analyze data more efficiently</li> <li>- Encourages active learning through dialogue, feedback, and problem-solving</li> </ul>	<ul style="list-style-type: none"> <li>- Some students struggle to operate AI tools due to interface or language issues</li> <li>- Dependence on AI could hinder deeper, independent thinking</li> </ul>
<p><b>Zudonu et al.(2024)</b> <i>"Investigating the Effect of Artificial Intelligence on Chemistry and Physics Students' Achievement and Conceptual Change in Heat Change in SSS2 in Rivers State"</i></p>	<ul style="list-style-type: none"> <li>- AI helps to improve academic achievement than using traditional methods</li> <li>- AI adapted to individual student needs, strengths, and weaknesses, enhancing the learning experience</li> <li>- Reduced workload through automation of tasks like test creation, grading, and content delivery</li> <li>- AI tools encouraged engagement through simulations, interactive content, and virtual assistance</li> </ul>	<ul style="list-style-type: none"> <li>- Without guidance, students might use AI to shortcut tasks rather than learn (e.g., copying answers)</li> <li>- Implementing AI-based education requires financial investment in tools, software, and maintenance</li> <li>- Over-reliance on AI might reduce students' natural learning effort and critical thinking</li> </ul>
<p><b>Thuy Nguyen Linh Cao and Phong Nguyen(2023)</b> <i>"Capabilities, Benefits, and Role of ChatGPT in Chemistry Teaching and Learning in Vietnamese High Schools"</i></p>	<ul style="list-style-type: none"> <li>- Provides real-time explanations and feedback on chemistry questions</li> <li>- Encourages students to think analytically, explore multiple solutions, and ask deeper questions</li> <li>- Enables anytime, anywhere learning, especially beneficial in remote or under-resourced areas</li> <li>- Helps break down complex topics into manageable parts with step-by-step guidance</li> </ul>	<ul style="list-style-type: none"> <li>- ChatGPT's performance on chemistry exams is lower than human students; struggles with complex application questions</li> <li>- May produce incomplete, incorrect, or out-of-context explanations</li> <li>- Responses need to be verified; cannot fully replace expert human instruction</li> </ul>

There are many benefits associated with the integration of AI products within education. These include the app's efficiency for big data and its capability to personalize learning. AI can help in generating quizzes, tests, and more study resources, for the benefit of teachers and students alike. It can also help improve students' conceptual understanding and problem-solving abilities. Furthermore, it supports self-paced learning and promotes independent study. Students can ask questions and receive immediate responses, which helps them to better understand chemistry. Those tools provide personal support to the extent they explain topics in a manner attuned to each student's level. They can demystify hard chemistry ideas with easy-to-understand examples and step-by-step guidance. Equally importantly, it also generates rich discussions around academic integrity and prompts updates to assessment design. AI can save teachers and students time, and learning can occur anywhere and at any time. But there are also some drawbacks. AI-based tools can also generate errors especially for context-rich, visual, and advanced chemistry items. Students could also come to overly depend on AI, sparking concerns about academic integrity. Professional practice and niche chemistry topics can come at a cost in terms of performance. AI gets things wrong or gives out wrong, confusing, or off-topic answers every now and then, particularly in hard chemistry subject matter. And some students, perhaps, would rely too much on AI and not bother to try to learn on their own. Most teachers and students do not know how to use AI well, and not all schools have the internet or devices required. Furthermore, language issues can occur, particularly in non-English settings. Last but not least, ethical problems such as data use, fairness, and open access are still pressing issues in adopting AI in education.

Overall, while AI (including ChatGPT) offers significant advantages for enhancing chemistry teaching and learning such as personalized support, resource generation, and modernized education practices educators and institutions must address its drawbacks by implementing thoughtful pedagogy, ethical guidelines, and robust evaluation methods.

## **2. METHODOLOGY**

### **2.1. Research Design**

This study aims to explore the effectiveness and usability, advantages and disadvantages of artificial intelligence (AI) tools in chemistry education and to identify the ethical considerations, like fair access, academic integrity, and authorship related to their use. The research will focus on how AI tools like ChatGPT, Gemini, Yandex AI, and DeepSeek help students understand chemistry concepts, solve problems, and prepare for exams. At the same time, the research will examine the risks of using AI, such as giving wrong answers, reducing student effort, and problems with academic honesty. The study will also look at how teachers use these tools to solve problems, explain difficult topics, and give feedback. The research employs a mixed-methods approach, combining both quantitative (structured questionnaires) and qualitative (open-ended questions and expert evaluations) methodologies.

### **2.2. Participants and Sampling Criteria**

For Research Questions 1 and 2, a total of 108 participants took part in the survey. All participants were students (bachelor's and master's level) from Suleyman Demirel University (SDU), Abai Kazakh National Pedagogical University (KazNPU), and Kazakh State Women's Pedagogical University (KazSWPU) studying to become chemistry teachers. For Research Question 3, the evaluation was conducted by five independent experts, representing different academic and practical levels:

- one PhD student
- one master's student
- one undergraduate student
- one high school student (11th grade)
- one chemistry teacher in high school.

### **2.3. Data Collection**

All data were collected via the Google Forms platform. Survey links were distributed through university Telegram and WhatsApp groups. Participation was fully voluntary. Respondents provided informed consent after reviewing the research objectives and privacy terms.

#### **2.3.1. Survey for Research Question 1**

In the questionnaire, 9 closed-ended questions were asked to determine the effectiveness and potential use of artificial intelligence tools in chemistry

education. Closed-ended questions were rated using a 3-point Likert scale: 1-Disagree, 2-Partially agree, 3-Fully agree.

### **2.3.2. Survey for Research Question 2**

In the questionnaire, 7 closed-ended questions were asked regarding the reliability of artificial intelligence tools in chemistry education, their usefulness for students, academic integrity (such as modifying AI-generated responses and using AI during class), and ethical considerations. Closed-ended questions were rated using a 3-point Likert scale: 1-Disagree, 2-Partially agree, 3-Fully agree. In addition, 2 open-ended questions were included to explore students' fairness when using AI for completing assignments and, if they were teachers, how they would address the issue of AI use in education.

### **2.3.3. Solving Problems Using ChatGPT (Research Question 3)**

During the survey, 2 open-ended questions were asked about the advantages and disadvantages of artificial intelligence. The most common responses were collected, categorized, and coded. In the second stage of the research, 5 real questions from the Unified National Testing (UNT) of Kazakhstan were selected. These questions were taken from official materials published by the National Testing Center (NTC). The questions were divided by levels by Blum taxonomy: 2 questions–Level A, 2 questions–Level B, 1 question–Level C. The selected questions covered the following chemistry topics:

- Atomic structure and composition - including two level-A questions
- Factors affecting reaction rate- including one level-B question
- Factors affecting chemical equilibrium- including one level-B question
- Electrolysis- including one level-C question

The questions were translated into English and entered into ChatGPT using the following prompt: “Solve the question step by step and provide the final answer.” ChatGPT’s responses and solution steps were saved into one file. Then, five independent experts evaluated each response using the following criteria:

1. Accuracy of translation
2. Clarity of the solution process
3. Relevance to the chemistry context
4. Correctness of the final answer

## **2.4. Data analysis**

The data collected during the research included expert ratings and written comments. A mixed-method approach was used, combining both quantitative and qualitative analysis.

### **For quantitative analysis:**

- The Mean calculation method was used to compare the scores for each evaluation criterion.
- The Kendall's *W* coefficient was used to measure the level of agreement among the experts. A total of 5 expert chemistry teachers took part in the evaluation. Each of them has different teaching experience but is familiar with the national curriculum and the UNT exam format. All evaluations were done independently. A value between 0.7 and 1 shows a high level of agreement between raters.

### **For qualitative analysis:**

- The open comments provided by the experts (about strengths and weaknesses) were analyzed using content analysis.
- Repeated topics and common ideas were grouped and coded into key categories.

## **2.5. Ethical Considerations**

All questions used in the study were taken from public sources. No personal information was collected. The experts gave their informed consent to participate in the study, and their evaluations were kept anonymous.

### 3. RESULTS

#### 3.1. RQ1 The effectiveness and usability of AI integrated chemistry education

A literature review was used to carry out the research question that found the effectiveness AI-integrated chemistry education. About 16 works of foreign and domestic scientists written on AI in chemistry education were studied. These articles were about the AI in chemistry education, the advantages and disadvantages of AI and impact to the study . Research works of scientists were taken from the Google Scholar platform. Keywords such as AI, chemistry education and AI in chemistry education were used to search for the article. The scientific articles were shown in the following table (Table 3). Each activity with frequency and percentage calculation as analysis method.

**Table 3.** Presentation name of the searched articles

Article	Number
Vladimir L. Kodkin & Ekaterina V. Artemeva(2024); Ted M. Clark(2023); Margarida Figueiredo & M. Lurdes Esteves(2014); N.A.B Gray (1988); Wajeeh Daher, Hussam Diab & Anwar Rayan (2023); Suzanne Fergus, Michelle Botha & Mehrnoosh Ostovar(2023); Wajeeh Daher, Hussam Diab & Anwar Rayan(2024); Mary E. Emenike & Bright U. Emenike(2023); R. Perezzan, S. Montalvo-Quirós, R. Rama-Ballesteros & D. Herráez-Aguilar(2024); Thu Nguyen and Yanika Sirichokcharoenkun(2023); Bekir Yildirim &Ahmet Tayfur Akcan(2024); A.O.Alli(2025); Himel Mondal et al.(2023); Ratna Farwati et al.(2024); Zudonu et al.(2024); Thuy Nguyen Linh Cao & Phong Nguyen(2023)	16 articles

After extensive data collection, it was determined that artificial intelligence can be applied to 4 types of activities in chemistry education. The results of the research are shown in Table 4. The table provides a list of types of activity that can be created by artificial intelligence and research articles, frequency and percentage in the literature. The most common activities identified were answering questions and solving problems(81%). This indicates that students use artificial intelligence to assist with learning chemistry through these methods. The least common activities included laboratory work(44%) and making scientific predictions(31%), such as writing

research papers. These tasks tend to be more interesting to college and university students.

**Table 4.** Activities used in AI integrated Instruction

Activity name	Article (author, Year)	Frequency	Percentage
Creating and solving open-ended and multiple-choice questions	Vladimir L. Kodkin, Ekaterina V. Artem'eva(2024); Ted M. Clark(2023); Wajeeh Daher, Hussam Diab, and Anwar Rayan(2024); Mary E. Emenike and Bright U. Emenike(2023); Bekir Yildirim and Ahmet Tayfur Akcan(2024); A.O.Alli(2025); Himel Mondal et al.(2023); Thuy Nguyen Linh Cao & Phong Nguyen(2023); Zudonu et al.(2024)	9	56%
For Studying Chemistry (answering questions, solving problems)	Vladimir L. Kodkin, Ekaterina V. Artem'eva(2024); N.A.B Gray (1988); Wajeeh Daher, Hussam Diab, and Anwar Rayan (2023); Suzanne Fergus, Michelle Botha, and Mehrnoosh Ostovar(2023); Mary E. Emenike and Bright U. Emenike(2023); R. Perezzan, S. Montalvo-Quirós, R. Rama-Ballesteros, D. Herráez-Aguilar(2024); Thu Nguyen and Yanika Sirichokcharoenkun(2023); Bekir Yildirim and Ahmet Tayfur Akcan(2024); A.O.Alli(2025); Himel Mondal et al.(2023); Ratna Farwati et al.(2024); Thuy Nguyen Linh Cao & Phong Nguyen(2023); Zudonu et al.(2024)	13	81%
For Scientific Assumptions	Vladimir L. Kodkin, Ekaterina V. Artem'eva(2024); Mary E. Emenike and Bright U. Emenike(2023); A.O.Alli(2025); Ratna Farwati et al.(2024); Thuy Nguyen Linh Cao &	5	31%

	Phong Nguyen(2023); Zudonu et al.(2024)		
Laboratory works	Margarida Figueiredo et al. (2014); Mary E. Emenike and Bright U. Emenike(2023); A.O.Alli(2025); Himel Mondal et al.(2023); Ratna Farwati et al.(2024); Thuy Nguyen Linh Cao & Phong Nguyen(2023); Zudonu et al.(2024)	7	44%

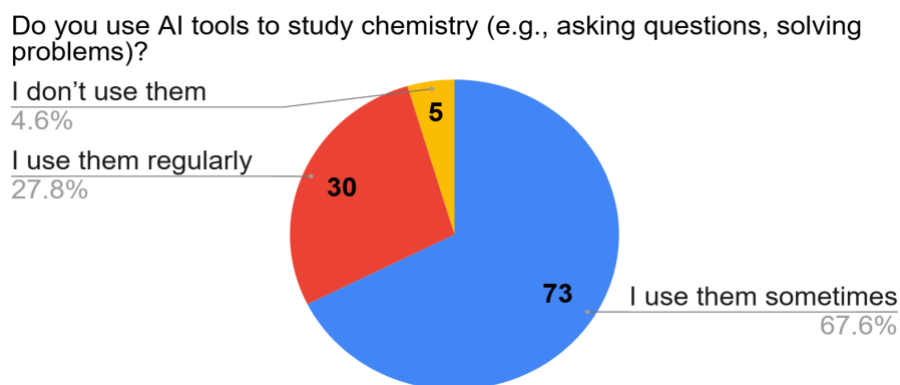
Secondly, for the usability of AI-integrated chemistry education, the survey was used to find out the uses of AI-integrated tools in chemistry education. Table 5 shows which AI tools are used most often by the participants. The results clearly show that ChatGPT is the most popular tool, with 99.1% of users choosing it. This means almost all participants prefer ChatGPT over other tools. It is likely that ChatGPT is easy to use, gives helpful answers, and supports many topics, especially in education. In comparison, only 0.9% of users use DeepSeek, and no one uses Yandex AI or Gemini. This shows that many people may not know about these other tools or do not find them useful. The table suggests that most users are comfortable with one tool, and they do not try other options. This may also show a lack of awareness or access to different AI tools. To improve this, teachers and schools can introduce students to other AI tools and explain their benefits. This can help students choose the best tool for their needs and become more independent in their learning.

**Table 5.** Most commonly used AI tools.

<b>Question</b>	<b>ChatGPT</b>	<b>DeepSeek</b>	<b>Yandex AI</b>	<b>Gemini</b>
Which AI tools do you use most often?	99.1%	0.9%	0%	0%

The Figure 2. shows how often students use AI tools to study chemistry, such as asking questions or solving problems. Most students, 67.6%, said they use AI tools sometimes. This means many students are familiar with AI but may not use it every day. 27.8% of students said they use AI tools regularly, which shows that more than a quarter of the group use these tools often to help with chemistry. Only 4.6% of students said they do not use AI tools at all. This means almost all students have tried or used AI tools in their learning. These results show that AI tools are becoming an important part of chemistry education. However, not everyone uses them regularly. Schools and

teachers can help students learn how to use AI tools more effectively and confidently. This can support their learning and problem-solving skills in chemistry.



**Figure 2.** Frequency of AI tool usage in chemistry studies

The table 6. shows students' opinions about using AI tools to learn chemistry. Most students believe that AI is helpful in their learning process. 54.6% of students partly agree and 43.5% fully agree that AI tools are useful for asking questions and solving problems in chemistry. This means that a total of 98.1% of students find AI helpful in some way. Only 1.9% disagree, which is a very small group. These results show that AI tools are widely accepted and used by students to improve their understanding of chemistry. The high percentage of agreement shows that students see AI as a useful support in their education, especially when it comes to solving difficult tasks or learning new topics. This also means that AI tools can play an important role in modern science education.

The second question reflects students' opinions about whether AI tools increase their interest in chemistry. More than half of the students (51.9%) said they partly agree, and 30.6% said they fully agree. It means that a total of 82.5% of students feel that AI helps them become more interested in chemistry, at least to some extent. However, 17.6% of students disagree, which is a higher percentage compared to the previous question. This shows that while many students find AI tools interesting and motivating, some students still do not feel more engaged in the subject. Overall, the results suggest that AI tools can increase interest for most students, but they may not work the same way for everyone. Teachers may need to guide students on how to use AI in a fun and educational way to keep them more interested in chemistry.

Third question from table represents how students feel about the effect of AI tools on their test or assignment results. Most students gave a positive answer. 50.9% of students said they partly agree, and 39.8% said they fully agree that their results improved after using AI tools. This means that a total of over 90% noticed some benefit. Only 9.3% of students disagree, which is a small group. These results suggest

that AI tools can help students do better in their chemistry tasks. The tools may give clearer explanations, help with problem-solving, or allow students to practice more. It shows that AI has a positive impact not only on learning but also on academic performance.

The following question provides information about students' opinions on how quickly they learn chemistry when using AI tools. A large number of students gave a positive answer. 49.1% fully agree and 43.5% partly agree that AI helps them learn faster. This means that a total of 92.6% believe AI tools support quicker learning. Only 7.4% disagree, which is a small group. These results show that most students feel AI gives fast explanations, clear examples, and quick help with questions. This saves time and makes learning more effective. The high agreement also shows that AI tools can be very useful in helping students study at their own speed and understand chemistry in a shorter time.

The next question shows what students think about the role of AI tools in helping them study chemistry independently. Most students answered positively. 48.1% partly agree, and 39.8% fully agree that AI tools helped them become more independent learners. Only 12% disagree, which means a small number of students did not see improvement in this area. The high percentage of agreement shows that many students feel more confident to study alone when they use AI tools. This could be because AI is always available to explain, check answers, or give support without needing a teacher.

The last question represents students' opinions about whether AI can replace some parts of traditional teaching. Most students are not fully sure. 60.2% partly agree, meaning they think AI can help in some areas, but not completely replace teachers. Only 16.7% fully agree, which is a small group. On the other hand, 23.1% disagree, showing that many students believe traditional teaching is still very important. These results suggest that students see AI tools as a good support, but not as a full substitute for human teachers. AI can help explain topics, answer questions, and support learning, but students still value personal interaction, guidance, and explanations from real teachers. This shows that the best way to use AI in education is to combine it with traditional teaching methods. These outcomes suggest that AI tools can support students not just in learning content, but also in building good study habits and self-discipline.

**Table 6.** The usability and Effectiveness of AI Tools

	<b>Question</b>	<b>Disagree</b>	<b>Partly agree</b>	<b>Fully agree</b>
1	In your opinion, is using AI tools helpful for learning chemistry (e.g., asking questions, solving problems)?	<b>1.9%</b>	<b>54.6%</b>	<b>43.5%</b>
2	Do AI tools increase your interest in chemistry?	<b>17.6%</b>	<b>51.9%</b>	<b>30.6%</b>
3	My test or assignment results improved after using AI tools	<b>9.3%</b>	<b>50.9%</b>	<b>39.8%</b>
4	I learn chemistry faster when I use AI tools	<b>7.4%</b>	<b>43.5%</b>	<b>49.1%</b>
5	Using AI tools improved my ability to study chemistry independently	<b>12%</b>	<b>48.1%</b>	<b>39.8%</b>
6	Do you think AI can replace some parts of traditional teaching?	<b>23.1%</b>	<b>60.2%</b>	<b>16.7%</b>

The table 7. presents data on the main purposes for which students use AI tools in their studies, based on responses from 108 participants. The most common reason is to understand difficult topics, with 90 students (83.3%) choosing this option. This suggests that AI tools are especially valued as a learning support when students face challenges with complex or unclear material. These tools are likely helping learners by providing simplified explanations, examples, or step-by-step solutions that may not always be available in traditional classroom settings. The second most frequent purpose is to ask theory-related questions, reported by 67 students (62%). This indicates that many learners use AI as a resource for checking definitions, concepts, and theoretical explanations when they want immediate answers without waiting for teacher feedback. Another big way students use AI is to get ready for exams, chosen by 61 students (56.5%). This means more than half of the students use AI tools to review lessons, practice problems, or ask questions before tests. AI seems to act like a study helper during exam time, helping students prepare better. Also, 59 students (54.6%) said they use AI to look for test answers, which is close to the number using it for studying. This shows a pattern of using AI to check answers or learn how to solve test questions. While useful, this also brings up the need for honest and careful use. Overall, the data

shows that students use AI tools not only to deepen their understanding of content but also as a practical aid in assessments and exam readiness. These findings highlight how AI is becoming an essential part of students' learning strategies. However, it also suggests that students may benefit from guidance on how to use AI tools effectively and ethically, ensuring they develop critical thinking and problem-solving skills alongside the support provided by technology.

**Table 7.** Main Purposes for Using AI Tools in Chemistry Learning

Purposes	Number	Percentage
For exam preparation	61	56.5%
To understand difficult topics	90	83.3%
To find test answers	59	54.6%
To ask theory-related questions	67	62%

To sum up, when reviewing the literature, it was found that most people use artificial intelligence in studying and learning chemistry (answering questions, solving problems), while the rest use it in creating open and closed questions and problems, writing laboratory works, and making scientific predictions. The most frequently used artificial intelligence tool by students was ChatGPT. This indicates that ChatGPT is more popular and widely used among students than AI tools such as Yandex AI, Gemini, and DeepSeek. The fact that the total percentage of those who “sometimes use” and “always use” AI tools when studying chemistry is 95.4% indicates that AI is important for students in their learning process. The usefulness of artificial intelligence tools in chemistry education and the increase of interest in the subject in the process of studying at the university will lead to an increase in exam results. Thanks to AI tools, students develop their own learning skills and learn discipline. Since artificial intelligence tools cannot completely replace traditional teaching methods, offline lectures and practical classes should be continued. 83.3% of students use AI tools to understand and learn difficult topics, 62% use them to ask for answers to theoretical questions, 56.5% use them to prepare for exams, and 54.6% use them to find answers to test questions. This shows that AI tools act as a free virtual online tutor.

### **3.2. RQ2 Ethical and privacy concerns**

The survey, which collected among students, was used to find out the ethical considerations of AI-integrated tools in chemistry education.

The first question from table 8 asks, “In your opinion, are AI tools reliable for studying chemistry?” A very small share of students only 7.4% disagree, so nearly everyone places at least some trust in AI. The clear majority, 81.5%, only *partly* agree,

which shows that most learners see AI as helpful but recognize it can still make mistakes or give surface-level explanations. Just 11.1% *fully* agree, meaning total confidence is rare. Taken together, these figures suggest students treat AI as a useful assistant rather than an unquestionable authority and will likely cross-check AI answers with textbooks or teacher feedback. For teachers, this means AI can be safely integrated into lessons, but students still need reminders to verify complex information. Improving AI accuracy and including more chemistry-specific examples could raise trust over time. Overall, the data tell us that reliability concerns exist, yet they are small enough for AI to play a positive role in chemistry learning.

The next question “Do you think using AI tools in education is fair?” reveals near-universal acceptance of AI’s fairness. No student selected “disagree,” which means everyone sees at least some level of justice in using AI. Even so, 69.2% only *partly* agree, while 30.8% *fully* agree. The large “partly agree” share hints that, although students like AI’s on-demand help, they still worry about gaps in digital skills, internet speed, or device quality that could create hidden inequalities. Teachers can reduce these concerns by offering training sessions, loaner devices, and offline AI resources. Because all students perceive some degree of fairness, educators can confidently use AI in class but they should keep working to ensure truly equal access. The data therefore encourage action: reinforcing support structures may move more students from partial to full agreement in the future.

A third question “Do you think using AI tools breaks academic integrity?”—elicits mixed feelings. Roughly 26.9% of students disagree, believing AI can be used ethically. The majority, 61.5 %, *partly* agree, admitting there are risks such as plagiarism or unapproved assistance if guidelines are unclear. A smaller 11.5% *fully* agree, viewing AI use as a direct threat to honesty. Since most respondents fall into the middle, they likely see a “gray area” where AI can be either helpful or harmful, depending on how it is used. Teachers should respond by defining clear rules: when AI assistance is allowed, how to cite it, and which tasks must be completed unaided. Classroom discussions on acceptable versus unacceptable AI use can reduce uncertainty. With firm guidance, the majority may shift toward the “disagree” column, seeing AI as safe when used responsibly.

The fourth question “Do you think AI tools are equally accessible to all students?” – receives a strongly positive answer. An impressive 65.4% *fully* agree that AI is accessible to everyone, and 26.9% *partly* agree, while only 7.7% disagree. These results suggest that stable internet and usable devices are common among respondents, yet roughly one-third see at least minor barriers. Those students might rely on shared devices, limited data plans, or slow connections. Schools can help by providing on-campus Wi-Fi, loaner laptops, or downloadable AI resources to ensure nobody is left

behind. Because most students already feel AI access is strong, targeted support for the small disadvantaged group could quickly close the remaining gap.

The fifth question “Should teachers limit or control AI usage in learning?” shows opinions are split. About 34.6% disagree with any limits, wanting full freedom to explore AI. The largest portion, 42.3%, *partly* agree, supporting some guidance but not strict restrictions. Another 23.1% *fully* agree that teachers should impose controls, perhaps to prevent over-reliance or misuse during exams. This diversity suggests students appreciate autonomy yet recognize the need for balanced guardrails. A sensible class policy might allow AI for practice but ban it during assessments, or require reflective notes explaining how AI helped. Such flexible rules would satisfy most students while teaching responsible usage.

The sixth question “Do you rewrite AI-generated answers in your own words before submitting them?” highlights strong ethical awareness. No student chose “disagree,” meaning everyone paraphrases AI content at least some of the time. More than half, 53.8%, *fully* agree they always rewrite, and 46.2% *partly* agree, suggesting they do so in most cases. This widespread rewriting shows students understand plagiarism risks and want to demonstrate real comprehension. Teachers can reinforce this habit by grading for originality and asking students to explain how they modified AI text. Maintaining this practice will strengthen academic integrity and deepen learning.

Finally, the seventh question “Do you present an AI-generated answer as your own original work?” uncovers remaining challenges. Half the students (50%) disagree, stating they never claim AI content as their own. Yet 42.3% *partly* agree, implying occasional or uncertain misuse, and 7.7% *fully* agree, openly admitting to submitting AI work unchanged. These figures show most learners aim for honesty, but a sizeable minority still needs clearer rules. Teachers should provide explicit guidelines on when and how to acknowledge AI assistance, perhaps requiring short citations such as “Generated with ChatGPT, adapted by the student.” Regular discussions about intellectual honesty can move students in the “partly agree” and “fully agree” categories toward full transparency. In sum, the table reveals that students value AI for chemistry learning but also want clear boundaries to ensure fairness, reliability, and academic integrity.

**Table 8.** Survey Findings on Ethical and privacy concerns

Question	Disagree	Partly agree	Fully agree
In your opinion, are AI tools reliable for studying chemistry?	7.4%	81.5%	11.1%
Do you think using AI tools in education is fair?	0%	69.2%	30.8%
Do you think using AI tools breaks academic integrity?	26.9%	61.5%	11.5%
Do you think AI tools are equally accessible to all students?	7.7%	26.9%	65.4%
Should teachers limit or control AI usage in learning?	34.6%	42.3%	23.1%
Do you rewrite AI-generated answers in your own words before submitting them?	0%	46.2%	53.8%
Do you present an AI-generated answer as your own original work?	50%	42.3%	7.7%

Let's take a look at the students' open questions related to ethical considerations. The first question asked whether they considered themselves fair users of artificial intelligence (AI) tools when solving problems or completing assignments. Analysis of the answers revealed two clear positions. Approximately 64% of the respondents stated that they do regard their use of AI as fair. For them, AI functions as a *digital tutor*: it clarifies theory, presents step-by-step solutions, suggests alternative methods, and saves time. These students emphasise that they do not simply copy the final answer; instead, they aim to understand the underlying algorithm, verify the output, correct any errors, and then apply the learned method to similar tasks independently. They also note that AI encourages creative thinking by offering multiple approaches to a problem. Hence, they view AI as a legitimate aid that supports learning without violating academic integrity. By contrast, roughly 36% of students do not consider themselves fair users. Their shared concerns centre on three points. First, they argue that AI solves tasks on the student's behalf, so the outcome cannot be claimed as personal effort.

Second, they highlight the high likelihood of errors in AI-generated solutions and fear that over-reliance may weaken critical thinking and create unhealthy dependence on technology. Third, they believe that obtaining grades through such “shortcut” methods contradicts academic honesty because it demands neither genuine effort nor deep comprehension. Some respondents report using AI only for initial ideas or guidance while still solving the problem on their own, yet they confess feelings of guilt; others prefer to avoid AI entirely, convinced that it hinders independent intellectual growth. These findings indicate that the question of fairness lies not in the technology itself but in *how* and *why* it is used. When AI is employed to deepen understanding, correct mistakes, and foster creative inquiry, it is perceived as fair and beneficial. When it merely replaces individual effort, it becomes ethically questionable and may impede genuine learning. Accordingly, educational institutions should not ban AI outright but develop methodological frameworks that channel its use toward critical thinking and self-development. Training sessions on responsible AI usage, updated academic integrity guidelines, and explicit instruction to review and rephrase AI-generated content in one’s own words are essential. Such measures can transform AI into an effective and ethical learning tool, enhancing rather than diminishing students’ authentic knowledge.

Second open question about how students would regulate the use of artificial-intelligence tools if they were instructors. The students’ answers about regulating AI use fall into several broad positions that together outline a balanced, guidance-first policy. A little over one-third of respondents favour a guided-use model: they would let learners consult AI tools to clarify theory, find ideas, or draft solutions, but only after explicit training that shows how to verify output, spot errors, and rewrite answers in their own words. Roughly one quarter prefer a verification regime in which instructors run submissions through AI-detectors, cap the percentage of AI-generated text (for example, 10–15%), and require students to defend their work orally before marks are awarded. About one-fifth recommend open access outside high-stakes settings, allowing AI for homework or project research while restricting it during timed examinations; they see AI as a legitimate study aid when reflection time is available but insist on independent mastery for graded tests. A smaller group urge caution or conditional permission, arguing that AI often produces mistakes and can undermine genuine effort, so it should be used only for specific gaps in understanding after students seek human help. Finally, a small minority adopt a *laissez-faire* stance, trusting students to judge credibility much as they already do with internet searches. Across these viewpoints, several common themes emerge. Nearly every student calls for explicit lessons on academic honesty and responsible AI practice, even when they support broad access. Many also propose “explain-back” checks short viva questions

or reflections so students must prove they truly understand any AI-assisted solution. Error awareness recurs: respondents want instructors to demonstrate typical AI mistakes in class, ensuring that trust in the technology is tempered by critical thinking. Task design is another shared concern; some suggest using creative or data-interpretation assignments that generic language models cannot complete convincingly, thereby discouraging copy-and-paste habits.

Taken together, the answers point to a coherent policy. Instructors should open each course with practical workshops on prompt crafting, source checking, paraphrasing, and ethical citation of AI output. They can permit AI freely for exploratory or low-stakes tasks pre-reading, brainstorming, or rough homework drafts while monitoring or prohibiting it in high-stakes assessments. Technical safeguards such as detector reports, oral defences, or brief reflection notes can verify that the final reasoning belongs to the student, not the model. Finally, assessments themselves should invite personal insight and creativity, reducing the temptation to rely on AI alone. This blended strategy respects the majority view that guidance beats prohibition, addresses the minority's ethical concerns, and channels AI toward deeper learning rather than shortcuts.

### **3.3. RQ3 Advantages and disadvantages of using AI tools**

To determine the advantages and disadvantages of using AI tools, interventions with ChatGPT by experts, and survey open questions were conducted.

The chemistry tasks created by ChatGPT were evaluated by experts using four main criteria: accuracy of translation, clarity of the solution, relevance to the chemistry context, and correctness of the final answer. The experts evaluated each category using a three-point scale. A score of 1 indicated disagreement, meaning the expert did not support the quality or appropriateness of that aspect. A score of 2 reflected partial agreement, suggesting that the expert found the element somewhat acceptable but identified areas for improvement. A score of 3 represented full agreement, meaning the expert was fully satisfied with the quality and considered that category to be well executed. This rating system allowed for a structured and consistent assessment of each question across multiple criteria. These evaluations were analyzed using mean statistics. For all criteria, the mean values were considered. Also the Kendall's W agreement coefficient was calculated to measure the consistency of the experts' ratings. This method helps to show, in numbers, how similar the experts' evaluations are for the same topic.

**Table 9.** The value of Kendall's W

<b>Kendall's W Value</b>	<b>Interpretation</b>
0	No agreement
0.1 – 0.3	Weak agreement
0.3 – 0.6	Moderate agreement
0.6 – 0.8	High agreement
0.8 – 1.0	Very high / Full agreement

Table 10 presents the evaluation results of the first question, which is from topic “the composition and structure of the atom”, based on the feedback from five experts. Each expert rated these categories on a scale of 1 to 3, where 1 represented “disagree,” 2 represented “partially agree,” and 3 represented “fully agree.” The first category, accuracy of translation, received a mean score of 2.4. This is the lowest among the four categories and indicates a degree of inconsistency or imprecision in the translation of the original task into English. Three out of five experts assigned a score of 2, suggesting that the wording was understandable but not entirely accurate in scientific or terminological terms. This result highlights the importance of using precise chemical terminology when translating educational content. A moderately low mean such as 2.4 implies that certain key terms might have been translated too literally or lacked the expected clarity for an English-speaking academic audience. Therefore, revision is recommended to improve alignment with international standards in chemistry education. The second category, clarity of the solution, achieved a higher average score of 2.8, indicating a generally strong performance. Four of the five experts gave the highest rating of 3, and only one gave a 2, implying that the step-by-step explanation of the problem was mostly coherent and logical. However, the one lower rating suggests that at least one expert encountered minor issues in understanding the sequence or logic of the solution. This could be due to a lack of transitional explanations, unclear assumptions, or insufficient breakdown of complex concepts. To further improve clarity, it may be helpful to include brief justifications for each step or to simplify certain calculations where possible. In terms of relevance to the chemistry context, the question received a perfect mean score of 3, reflecting unanimous agreement among all experts that the task was scientifically valid and appropriately grounded in chemistry content. This result confirms that the topic, concepts, and procedures used in the question are aligned with core chemistry principles and are suitable for assessment in an educational context. It also indicates that the question is well-integrated with the curriculum and successfully reflects the kinds of analytical thinking expected of students at this level. The fourth category, correctness of the final answer, also received a perfect mean score of 3. All five experts agreed that the final

solution provided was mathematically and scientifically accurate. This reflects a strong command of the subject matter and suggests that the final result was not only correct but also convincingly derived. High confidence in the correctness of the final answer enhances the overall reliability of the question and reduces the likelihood of misinterpretation or confusion on the part of students.

Taken together, these results show that the first question was well-received in terms of scientific accuracy and relevance. However, the relatively lower score in the translation category points to an area that requires attention. Ensuring that subject-specific terminology is used appropriately and that language is both grammatically and semantically precise is essential when preparing assessment materials for students. Furthermore, while the clarity of the solution was generally rated highly, even small improvements in structure or explanation could further raise the quality of the task.

**Table 10.** Points and means of first question.

<b>Expert</b>	<b>Accuracy of translation</b>	<b>Clarity of the solution</b>	<b>Relevance to the chemistry context</b>	<b>Correctness of the final answer</b>
1	2	3	3	3
2	3	3	3	3
3	2	3	3	3
4	2	2	3	3
5	3	3	3	3
Mean	2.4	2.8	3	3

Kendall's W value is calculated by calculating the sum of the ranks given in Table 11. The Kendall's W value of  $\approx 0.19$  reveals that the panel of five experts showed only a weak, statistically-insignificant level of concordance when judging the four quality dimensions of the atomic-structure question. In practical terms, the experts shared a broad sense that the task is scientifically sound and contextually appropriate evidenced by the low rank totals for "Relevance" and "Correctness" yet they diverged markedly in their views on language precision and explanatory depth. Such a pattern often arises when reviewers hold different implicit standards for what counts as an acceptable translation or a sufficiently transparent solution path. One expert, for example, penalised the clarity category more harshly than the others, while several flagged the wording as the principal weakness, driving the translation rank total to its conspicuously high value of 16.5. A low W carries two main implications for item quality. First, it undermines the reliability of the expert-review process: if judges do not rank aspects in a consistent order, their collective endorsement (or rejection)

becomes a shaky basis for high-stakes test decisions. Second, it signals that certain rubric descriptors especially those governing linguistic accuracy and step-by-step explanation are either too broad or interpreted idiosyncratically. Before the question is released to students, the translation should be tightened by adopting standard IUPAC terminology, removing literal wording artefacts, and ensuring that every chemical symbol and superscript follows international notation. Simultaneously, the worked solution ought to be re-structured into numbered steps that state key assumptions, justify intermediate calculations, and link each algebraic move to the underlying atomic-composition concept.

Finally, a short calibration session could bring the experts closer to a shared scoring anchor. By rating two or three exemplar items together and explicitly negotiating what merits “fully agree,” “partly agree,” or “disagree,” the reviewers would reduce idiosyncratic drift and yield a higher Kendall’s *W* in subsequent rounds. Achieving even moderate concordance ( $W \approx 0.4\text{--}0.5$ ) would greatly improve confidence that the task is not only chemically valid but presented so consistently that different subject-matter specialists reach the same qualitative verdict.

**Table 11.** Scores of each expert in within-expert ranks for first question.

<b>Expert</b>	<b>Accuracy of translation</b>	<b>Clarity of the solution</b>	<b>Relevance to the chemistry context</b>	<b>Correctness of the final answer</b>
1	4	2	2	2
2	2.5	2.5	2.5	2.5
3	4	2	2	2
4	3.5	3.5	1.5	1.5
5	2.5	2.5	2.5	2.5
Total Rank ( <i>R<sub>j</sub></i> )	16.5	12.5	10.5	10.5

The second question, which is based on the topic “Composition and Structure of the Atom”, received consistently high evaluations, which presented in Table 12, from all five experts. The most notable outcome is that three of the four categories received a perfect average score of 3, indicating unanimous agreement among all experts that the question meets high standards in those areas. The accuracy of translation achieved a flawless mean score of 3, with every expert rating it as “fully agree.” This suggests that the translation was linguistically precise, semantically clear, and scientifically appropriate. It reflects a strong command of both language and subject-specific terminology, ensuring that the English version of the question accurately conveyed the

original intent without introducing ambiguity or distortion. Similarly, the relevance to the chemistry context and correctness of the final answer both received perfect scores (3). These results confirm that the question is highly aligned with core chemistry content and was correctly solved using appropriate scientific methods. The experts' unanimous approval in these categories indicates that the question is both valid and reliable as a tool for assessing chemistry knowledge. The only slight variation was observed in the clarity of the solution, which received an average score of 2.8. Four out of five experts rated the solution as fully clear (score of 3), while one expert gave a score of 2, suggesting that a minor issue may have been present in the logical flow or presentation of the solution. This small drop in average does not significantly detract from the overall quality of the question but implies that a brief review or clarification of one part of the explanation could be beneficial.

In conclusion, the second question was rated very highly across all criteria, indicating that it is well-constructed, scientifically valid, and clearly communicated. Only minor refinement in the clarity of the solution could further enhance the quality. These results demonstrate that the question is suitable for use in a chemistry learning or assessment context without major revision.

**Table 12.** Points and means of second question.

<b>Expert</b>	<b>Accuracy of translation</b>	<b>Clarity of the solution</b>	<b>Relevance to the chemistry context</b>	<b>Correctness of the final answer</b>
1	3	3	3	3
2	3	3	3	3
3	3	3	3	3
4	3	3	3	3
5	3	2	3	3
Mean	3	2.8	3	3

The Kendall's W value, which are calculated by the information from table 13, of roughly 0.02 tells an unusual but reassuring story about the expert review of the second question. In strict statistical terms, a W that low signals "no concordance," yet, in this case, the near-zero coefficient does not spring from conflict among the judges; it stems from their overwhelming unanimity. Four of the five reviewers marked every one of the four quality dimensions translation accuracy, solution clarity, contextual relevance, and answer correctness with the highest possible score. The fifth expert agreed on three of the four criteria and lowered only *Clarity* from "fully agree" to "partially agree." Because Kendall's W operates on *rank differences*, when nearly all

scores are identical the algorithm finds nothing to rank, so the concordance statistic collapses toward zero. The result therefore confirms that the panel sees the item as uniformly strong: the language is precise, the explanation sufficiently transparent for most readers, the chemistry squarely on target, and the numerical answer correct. The sole hint of improvement lies in that one lower clarity mark, suggesting a quick editorial pass perhaps adding a line-by-line rationale or refining a diagram label would make the solution flawlessly clear to every reviewer. Beyond that minor polish, the question appears classroom-ready, and the low W should be interpreted not as a weakness in expert agreement but as a mathematical artifact of their near-perfect consensus.

**Table 13.** Scores of each expert in within-expert ranks for second question.

<b>Expert</b>	<b>Accuracy of translation</b>	<b>Clarity of the solution</b>	<b>Relevance to the chemistry context</b>	<b>Correctness of the final answer</b>
1	2.5	2.5	2.5	2.5
2	2.5	2.5	2.5	2.5
3	2.5	2.5	2.5	2.5
4	2.5	2.5	2.5	2.5
5	2	4	2	2
Total Rank (R <sub>j</sub> )	12	14	12	12

The third question, which is based on the topic “factors affecting reaction rate”, accepted varied evaluations from the five experts, revealing a number of areas for improvement despite a scientifically correct final answer. The accuracy of translation obtained a mean score of 2.6, which indicates that the language used in the English version of the question was mostly acceptable but not perfect. Three of the five experts gave a score of 3, showing confidence in the translation quality, while the remaining two experts assigned a 2, suggesting there may have been some inconsistencies or ambiguity in the terminology or phrasing. In the context of chemistry, the precise use of technical terms is essential for clarity and academic integrity, so even minor linguistic inaccuracies can affect comprehension. These results highlight the need to carefully review the translated version of the question to ensure it adheres to standard scientific vocabulary and avoids any semantic confusion. The clarity of the solution emerged as the weakest aspect of the question, receiving a uniform score of 2 from all five experts. This consistent rating indicates a shared concern among reviewers regarding the presentation and logical flow of the solution. Although the steps may

have been technically correct, they were not articulated in a way that was fully accessible or easy to follow for learners. Potential issues may include insufficient explanation of why certain steps were taken, lack of intermediate reasoning, or abrupt transitions between stages of the solution. Given that clarity is essential for students to grasp chemical problem-solving strategies, this score strongly suggests the need to revise the solution by breaking it into more detailed steps, using clear justifications, and guiding the learner through the reasoning process. The relevance to the chemistry context was rated with a mean score of 2.4, implying that while the question was generally on-topic, it may not have fully captured the complexity or depth expected from a task dealing with reaction rate factors. For instance, the question might not have clearly addressed core variables such as temperature, concentration, surface area, catalyst presence, or particle collisions—all of which are fundamental to understanding rate changes in chemical reactions. It is also possible that the example chosen for the question lacked a real-world context or failed to emphasize the causal relationships that affect reaction speed. Experts may have felt that the question needed refinement to better align with curriculum goals and to more effectively engage students in critical chemical thinking. Despite the above concerns, the correctness of the final answer received a perfect score of 3, with full agreement among all experts. This result confirms that the question, while limited in presentation and depth, ultimately led to a scientifically valid and accurate conclusion. The correct final answer enhances the credibility of the question and suggests that its underlying conceptual framework is sound.

In summary, the third question demonstrates strong content accuracy but falls short in terms of solution clarity and contextual integration. While the translation was generally reliable, the solution would benefit from improved structure and clearer explanation.

**Table 14.** Points and means of third question.

<b>Expert</b>	<b>Accuracy of translation</b>	<b>Clarity of the solution</b>	<b>Relevance to the chemistry context</b>	<b>Correctness of the final answer</b>
1	3	2	2	3
2	3	2	3	3
3	3	2	2	3
4	2	2	2	3
5	2	2	3	3
Mean	2.6	2	2.4	3

The Kendall's  $W$  value, which are calculated by the information from table 15, of  $\approx 0.42$  indicates that the five experts reached only a moderate level of agreement when judging the question on *factors affecting reaction rate*. They converged most strongly on the correctness of the final answer and its chemical relevance both aspects were consistently rated “fully agree,” suggesting that the underlying kinetics principles (temperature, concentration, surface area, catalyst effects) are scientifically accurate and framed at an appropriate curricular depth. In contrast, opinions diverged on two fronts. First, clarity of the solution drew the widest spread of ranks: while several reviewers found the explanation acceptable, one judged it weak, pointing to gaps in the step-by-step logic or insufficient linkage between kinetic theory and the numerical reasoning. Second, translation accuracy divided the panel; some experts perceived imprecise terminology or awkward phrasing (for example, inconsistent use of “collision frequency” versus “collision rate”) that could confuse students.

These discrepancies matter pedagogically. A reaction-rate problem is often the first time learners must integrate qualitative particle-collision ideas with quantitative rate data; any ambiguity in wording or calculation flow therefore risks mis-teaching a foundational concept. The moderate  $W$  highlights that, before classroom deployment or summative assessment, the solution should be re-structured into clearly numbered steps each citing the specific rate factor being applied while the English wording should be tightened to align with IUPAC kinetics language. In addition, a brief reviewer-calibration exercise, where experts jointly score exemplar items and discuss what distinguishes “partly” from “fully” acceptable clarity, would likely push  $W$  closer to the strong-agreement zone ( $\geq 0.60$ ). Such refinements would not only raise statistical reliability but also ensure that students receive an unambiguous, pedagogically sound treatment of how temperature, concentration, surface area, and catalysts influence reaction speed.

**Table 15.** Scores of each expert in within-expert ranks for third question.

<b>Expert</b>	<b>Accuracy of translation</b>	<b>Clarity of the solution</b>	<b>Relevance to the chemistry context</b>	<b>Correctness of the final answer</b>
1	1.5	3.5	3.5	1.5
2	2	4	2	2
3	1.5	3.5	3.5	1.5
4	3	3	3	1
5	3.5	3.5	1.5	1.5
Total Rank ( $R_j$ )	11.5	17.5	13.5	7.5

The fourth question, which addresses the topic “factors affecting chemical equilibrium” and is classified as level B according to Bloom’s taxonomy, obtained generally low evaluations from the expert panel across most categories, as shown in Table 16. The accuracy of translation received the lowest average score of 1.2, indicating a critical weakness in the linguistic quality of the question. Four out of five experts rated it as 1 (disagree), suggesting that the English version of the question may have been unclear, misleading, or improperly conveyed key chemistry terminology. Such poor translation accuracy can significantly hinder student understanding and can distort the intended meaning of the original content. To address this, the wording should be carefully revised by aligning it with standard scientific language and by ensuring that all core terms are correctly used in English. The clarity of the solution scored slightly higher, with a mean of 2.2, yet this still signals a notable lack of explanation or logical structure. Although one expert gave a score of 3, indicating full agreement, the majority of the experts rated it as 2, which reflects only partial clarity. This suggests that while some steps in the solution may have been understandable, the overall process lacked consistency or detail. For a question at level B of Bloom’s taxonomy-where students are expected to apply knowledge-it is essential that the solution be logically sequenced and supported with clear reasoning. Therefore, additional elaboration and restructuring of the explanation would be necessary to raise the clarity of this response. In contrast, the relevance to the chemistry context received a relatively high mean score of 2.8, which implies that the question itself was well-aligned with the core concepts of chemical equilibrium. All experts except one rated this category as 3, indicating that the problem was appropriately related to the intended topic. This is a positive indication that the question was conceptually placed within the correct domain of chemistry and touches on meaningful aspects of equilibrium such as Le Chatelier’s principle or changes in pressure, concentration, or temperature. However, high contextual relevance alone cannot compensate for poor clarity and translation. The most significant concern lies in the correctness of the final answer, which received a mean score of 1. Every expert gave the lowest possible score, indicating that the answer was scientifically incorrect. This is a critical flaw, especially in the context of educational assessments, as students rely on correct solutions not only to check their understanding but also to learn proper problem-solving techniques. A wrong answer invalidates the educational purpose of the task and can mislead students. The solution needs to be re-examined carefully, verified by multiple sources or professionals, and corrected to reflect the accepted scientific explanation. In summary, although the fourth question demonstrates good alignment with the chemistry curriculum in terms of topic relevance, its poor translation quality, partially unclear explanation, and a completely incorrect final answer severely undermine its

educational value. To make this question appropriate for use at Bloom’s level B, it is essential to fully revise the translation, reconstruct the solution with greater clarity, and most importantly, correct the scientific content of the final answer. Only after these revisions can the question be considered reliable and pedagogically sound.

**Table 16.** Points and means of fourth question.

<b>Expert</b>	<b>Accuracy of translation</b>	<b>Clarity of the solution</b>	<b>Relevance to the chemistry context</b>	<b>Correctness of the final answer</b>
1	1	2	3	1
2	1	3	3	1
3	1	2	3	1
4	1	2	3	1
5	2	2	2	1
Mean	1.2	2.2	2.8	1

The strong Kendall’s  $W$  of  $\approx 0.70$  tells us that the expert panel spoke with one voice: they valued the equilibrium task’s chemical *relevance* but unanimously condemned both the *translation* and the *correctness* of the final answer. The fact that translation accuracy received the worst possible score from four of five reviewers suggests that key terms linked to Le Châtelier’s principle such as “forward reaction,” “shift to the right,” “partial pressure,” or “endothermic vs exothermic” were rendered imprecisely or even reversed. When a machine-generated translation (for example, an initial draft produced by ChatGPT) swaps cause and effect “increase” for “decrease,” “products” for “reactants,” or mis-labels a catalyst’s role the error ripples through the calculation: the algebra may be executed flawlessly, yet it is applied to the wrong conceptual premise. That scenario appears to have unfolded here: all experts recognised the numerical work as *methodical* but judged the final equilibrium prediction incorrect, implying that a mistranslated statement (e.g., mis-identifying the stress applied to the system) led the solution down an erroneous path. Pedagogically, this tight consensus is invaluable feedback. First, the stem should be back-translated and compared line-by-line with the original to locate semantic shifts introduced by the language model. Employing a bilingual chemist or a vetted terminology bank can prevent pitfalls such as mistranslating “constant-volume vessel” or misplacing the heat term in an equilibrium expression. Second, once the wording is corrected, the solution steps must be rerun under the proper chemical assumptions; often the numeric machinery ( $K_c$ ,  $Q$ , direction of shift) is sound, and a single corrected parameter will yield the right answer. Finally, a quick recalibration review should confirm that, with

accurate terminology, experts now agree on both translation and correctness. Because the panel has already demonstrated high concordance, such targeted repairs are likely to convert their shared criticism into shared endorsement, resulting in a reliable and linguistically robust item on the *factors affecting chemical equilibrium*.

**Table 17.** Scores of each expert in within-expert ranks for fourth question

<b>Expert</b>	<b>Accuracy of translation</b>	<b>Clarity of the solution</b>	<b>Relevance to the chemistry context</b>	<b>Correctness of the final answer</b>
1	3.5	2	1	3.5
2	3.5	1.5	1.5	3.5
3	3.5	2	1	3.5
4	3.5	2	1	3.5
5	2	2	2	4
Total Rank (R <sub>j</sub> )	16	9.5	6.5	18

The fifth question, which focuses on the topic “Electrolysis” and is classified as Level C according to Bloom’s taxonomy, was evaluated as a practice-based question that requires numerical problem-solving. This type of task is expected to test a student’s ability to analyze data, apply formulas, and carry out calculations in a chemistry context. However, based on the scores presented in Table 13, the expert evaluations highlight several concerns about the formulation and presentation of the question, particularly regarding translation and clarity. The accuracy of translation received a low mean score of 1.6, suggesting that the language used in the English version was problematic. Three of the five experts gave the lowest score of 1, indicating strong disagreement with the quality of the translation. This implies that the text may have contained terminology errors, awkward phrasing, or unclear scientific expressions. Since Level C tasks involve critical thinking and precise interpretation of quantitative data, any confusion in wording can severely hinder a student’s ability to understand and solve the problem. Therefore, it is essential that the question be reviewed and rewritten with a focus on chemical accuracy, standardized units, and proper use of technical language. The clarity of the solution also scored 1.6, reflecting a similar concern. Three experts gave the lowest score of 1, indicating that the solution was not clearly presented. At Level C, students are expected to carry out multi-step problem solving, often involving molar relationships, Faraday’s laws, and current/time calculations. If the solution lacks logical structure, fails to explain each calculation step, or skips necessary reasoning, students may struggle to follow and learn from it.

The feedback suggests that this particular solution needs restructuring to ensure that it guides students step-by-step through the problem, clearly explaining what is being calculated and why. In contrast, the relevance to the chemistry context received a mean score of 2.4, which indicates that the question was generally appropriate for the topic of electrolysis but may not have fully captured its complexity. The majority of experts rated this category as a 2, with two assigning a 3. This suggests that while the question included content related to electrolysis, it might have lacked depth, real-world application, or diversity in the variables involved. For Level C tasks, it is important that the question not only include numbers but also reflect deeper understanding, such as how different parameters influence the electrolysis process in industrial or experimental settings. Despite these issues, the correctness of the final answer was rated very positively, with a perfect mean score of 3. All five experts fully agreed that the final result was scientifically accurate. This indicates that, although the translation and explanation require improvement, the underlying chemistry concepts and calculations are correct and reliable.

In summary, while the fifth question demonstrates scientific accuracy in the final result, it suffers from significant weaknesses in translation and solution clarity. As a Level C question involving calculations, these elements are especially important for helping students build problem-solving skills. To improve the educational value of this task, it is recommended to revise the language, reorganize the solution into detailed steps, and slightly enrich the contextual framing to better align with the analytical thinking expected at this cognitive level.

**Table 18.** Points and means of fifth question.

<b>Expert</b>	<b>Accuracy of translation</b>	<b>Clarity of the solution</b>	<b>Relevance to the chemistry context</b>	<b>Correctness of the final answer</b>
1	1	1	2	3
2	3	2	3	3
3	1	1	2	3
4	1	1	2	3
5	2	3	3	3
Mean	1.6	1.6	2.4	3

The Kendall's W value of about 0.55 shows that the five experts have a moderate but real agreement on this C-level question about electrolysis. They all think the task is chemically correct: it names the anode as the place where oxidation removes electrons and the cathode as the place where reduction gains electrons. The question

also asks students to work in a solution, so they must choose the ion that is discharged first, balance the half-equations, count the moles of electrons, and use simple stoichiometry to find the theoretical yield of product made at each electrode. Because this is Bloom's C level, learners are expected to take the data for time and current, turn it into moles of electrons, change that into moles of product, and finally change moles into grams or litres. Even so, the experts agree that two parts need work. First, the translation is weak. Some words are unclear or mixed up, for example writing "reduction at the anode" or using different symbols for the same ion. Second, the clarity of the solution is only "partly OK." The worked answer jumps too quickly from the half-reactions to the final mass, so one can lose the logic in the middle. To fix this, the wording should follow standard terms every time anode (oxidation), cathode (reduction), ions, moles, yield. The solution should be rewritten in five clear steps: (1) write and balance each half-reaction, (2) show the total electrons, (3) turn electrons into moles of product, (4) convert moles to the correct units (grams or litres), and (5) state the final theoretical yield and, if given, compare with the real yield. With these small changes, the translation will be smooth, the explanation will be easy to follow, and future reviewers would likely give higher and even more similar scores, raising Kendall's W and making the exercise clearer for students.

**Table 19.** *Scores of each expert in within-expert ranks for fifth question*

<b>Expert</b>	<b>Accuracy of translation</b>	<b>Clarity of the solution</b>	<b>Relevance to the chemistry context</b>	<b>Correctness of the final answer</b>
1	3.5	3.5	2	1
2	2	4	2	2
3	3.5	3.5	2	1
4	3.5	3.5	2	1
5	4	2	2	2
Total Rank (R <sub>j</sub> )	16.5	16.5	10	7

### **Survey open questions tables and graphs**

Content analysis of open-ended responses collected from chemistry students on the topic "Advantages of Artificial Intelligence tools" was conducted in several stages. First, we collected all the statements into a text document, corrected minor spelling errors, and left the original form unchanged in order to preserve the meaning. Then, we read each sentence individually and converted the main semantic units in it into short keyword codes. In this way, several codes were assigned to one response, resulting in

a total of more than two hundred codes. In the next step, eight major themes (categories) were identified by grouping similar codes; their frequency was then measured in percentage terms.

The results of the analysis show that the main value of artificial intelligence in the eyes of students is speed and time saving. About half ( $\approx 49\%$  of the total) say that “I save time when using artificial intelligence”, “I get answers to my questions quickly” or “it works quickly”, and evaluate AI as an “express reference” for theoretical questions and calculations in chemistry. This priority clearly reflects the demand of modern students for urgent information. There are three factors behind this. First, chemistry is a subject full of formulas and reactions, so most students want to get concrete information “here and now”. Second, since the number of tasks in the credit-module system is large and time-intensive, quick reference help reduces time pressure and stress. Third, the rate of information absorption of Generation Z is naturally high. Therefore, the lower the AI’s response latency, the more the tool will consolidate its “first choice” status. What does this mean for the teacher? If he wants to focus the student’s attention more on theoretical issues, he should first present the content in a concise form and then use a strategy of deepening it (the “inverted AI-pyramid” approach).

In second place are the advantages of explaining complex topics in simple language. Every fourth student ( $\approx 27\%$ ) mentioned statements such as “explains in a simple and easy way”, “teach difficult topics in a simple way” or “explains visually”. Although this inclusive style of explanation makes learning easier, two things should not be forgotten. First, teachers and students should not rely too much on academic jargon; on the other hand, terminological precision may be lost if “simplification” is excessive. Therefore, it is beneficial to return to scientific vocabulary after a “short-form explanation”. We also noticed that 27% of students think that AI is not just a source of information, but also a free tutor who can help them at any time and place. In addition, artificial intelligence can be a good helper in improving the level of students with low academic knowledge in certain subjects.

The third important block is information search and filtering ( $\approx 20\%$ ). Students say that they do not waste time by retrieving the necessary information from a variety of materials on the Internet with one query. Here, AI plays the role of a “smart filter” and reduces information overload. At the same time, it can be observed that students are tired of the excessive amount of information on the Internet. From an educational and methodological point of view, this is an opportunity to strengthen the digital literacy component. If students are given prompts that teach them to indicate the source, year, and author of data obtained from artificial intelligence when researching information on their own, confidence increases and academic integrity is strengthened.

Among the answers given by students, comments such as "Report, theory, test - I ask all at the same place" are found quite often ( $\approx 8\%$ ). Platform models that turn AI into a universal "chemistry hub" and connect everything from balancing reaction equations to theoretical questions and ways to solve problems increase cognitive efficiency.

Only 6.5% of students indicated that "artificial intelligence adapts to my level." That is, students often use AI as a ready answer, but do not fully use its adaptive functions. As a solution to this problem, students should be taught to use adaptive functions with prompts such as: "Explain this topic to me at the level of 1st-year students," "Imagine that you are a chemistry teacher teaching at a university. Teach me how to find the hydrogen number of weak acids and bases." In addition, students should be trained in the use of AI and shown the level-adapting and feedback functions of ChatGPT. The accessibility factor ( $\approx 5\%$ ) ("I can look at it at any time") and the ability to help prepare for exams ( $\approx 5\%$ ) are also mentioned less, but their practical benefits are high. AI's formative assessment tools, such as test creation and schedule creation, were also appreciated. A student who uploads lecture slides to ChatGPT and teaches them to create short thematic tests or question-and-answer cards based on them will undoubtedly further strengthen their knowledge and get high scores in exams. Therefore, if university faculties begin to introduce additional test-generating bots and automatic feedback into the formative assessment ecosystem through AI, the quality of students' education will undoubtedly improve. Only 1.5% of students said that AI "increases my critical thinking." This low figure indicates two things: students use artificial intelligence to get ready-made answers; AI's potential as a tool for debate, hypothesis testing, and reflection has not yet been fully realized. That is, it is rarely used as an idea generator or "critical discussion partner." All students should not just learn and memorize answers to questions through AI, but also delve deeper into their underlying meaning, logic, and questions like "why," "how," and "what."

Overall, students' feedback indicated that they were highly dependent on AI's ability to accelerate learning and facilitate explanation. They expected a tool that would quickly solve problems and simplify complex concepts. However, it seems that deeper features such as adaptive learning, feedback, or level-based learning were not widely introduced. Therefore, if educational institutions demonstrated the level-based learning, error analysis, and "collaborative reasoning" capabilities of ChatGPT through practical training, students would perceive it as a learning and research partner, not just an "answer machine". Based on these data, several recommendations arise. First, it is necessary to hold more frequent seminars and trainings that teach students how to use AI correctly and effectively, write prompts at different levels. Second, universities should create question-and-answer bots related to their subjects together with their students and conduct formative assessment. Third, if tasks that develop the ability to

adapt and think critically (for example, “find logical errors in the answer”) are introduced, the intellectual potential of AI will be revealed more fully. Fourth, it would be much more useful if subject teachers at universities identified and discussed AI errors with students.

In conclusion, the main advantage of artificial intelligence for chemistry students is that it saves time and conveys complex concepts in an easy-to-understand language. These two properties make the learning process much faster and more accessible. Now, based on these basic values, systematically implementing ways to turn AI into a tutor, critical discussion partner, and formative assessment tool is an effective way to raise the quality of learning to the next level.

**Table 20.** Advantages of AI tools.

<b>№</b>	<b>Category (theme)</b>	<b>Code groups</b>	<b>Percentage %</b>
1	Time & efficiency	time, fast answer, one tool	48.8 %
2	Clarity & plain language	simple lang, visual, breakdown	26.8 %
3	Quick information/answer retrieval	info search, internet filter	20.3 %
4	Integrated multifunctionality	multi func, rich resource	8.1 %
5	Individual learning & adaptation	self study, level adapt	6.5 %
6	Accessibility (24/7)	anytime, accessible	4.9 %
7	Exam & assignment support	exam support, quiz, task	4.9 %
8	Creativity / critical-thinking perspective	idea critic, motivate	1.6 %

Above, we learned the opinions of students about "advantages of artificial intelligence tools", now it's time to talk about "disadvantages of artificial intelligence tools". We collected all the opinions in a document, corrected minor errors, read each sentence separately, and turned them into several short keyword codes according to their meaning. The first and most important category was "Accuracy and reliability problems". More than two-thirds of the students (73%) said that AI often gives wrong answers, makes mistakes because they do not understand what is given in the problems, and incorrectly balances reaction equations. Several respondents gave examples of receiving two different, even contradictory, answers to the same question. As a result of such errors, there were also those who prepared for the exam using AI, saying that they "got a low score on the exam". This trend clearly shows that even if AI is used as a "first option", additional checks are definitely necessary.

The second large block was grouped around the topic of “Addiction and cognitive decline”. Approximately every third student (39%) admitted that they rely on AI and rarely think for themselves, search for information, do additional research, or review other literature. The comments “it causes laziness,” “it inhibits critical thinking,” and “it reduces strategic thinking” are particularly common. Students also realize that the disadvantage here is not the AI itself, but the consequence of its excessive use: there were those who said “there is no disadvantage if you can use it correctly.” Over time, this can lead to a student losing the ability to solve unexpected problems encountered in a laboratory setting.

The third category was called “Inability to solve complex problems and specialized content in chemistry.” It was noted that 25% of students made mistakes in ChatGPT C-level (complex) problems, reaction kinetics and balancing equations. “Even the paid version of ChatGPT balances the reaction incorrectly,” it was noted that the paid version also has errors. In such a case, if the student does not know the topic well, he may copy the wrong solution directly on the exam and lose points. The fourth category is “Unavailable features”. Several students (7%) said that the fact that AI tools, such as ChatGPT, have full functionality and the latest models are only available in paid versions is an obstacle for the average user. The high number of errors in the free version was also mentioned in this category. The fifth, relatively small group (4%) belongs to the problem of “Linguistic and incomplete segments”. There are few complaints about the poor quality of answers in Kazakh, lack of technical or terminological accuracy. This indicates that the artificial intelligence database needs to be expanded with information in Kazakh.

13% of students complained that "there are not enough real examples" and "only one solution is shown." This shows that students expect better quality from AI. If difficult chemistry problems do not show different ways to solve them or results depending on conditions, it can confuse the student. The last, but most notable, block is “Socio-ethical and professional implications.” The argument that the widespread use of AI will lead to job losses in the future (mentioned by several respondents), as well as the argument regarding the human factor that “there are not enough live tutors,” fell into this category.

As the analysis shows, the main concern for students is false information. This combines many of the commonly mentioned problems and leads to other issues: believing in incorrect facts increases dependence, weakens critical thinking, and limits the ability to solve complex problems. Over time, people may start to doubt the knowledge they have received. Therefore, using AI in education is not possible without creating a culture of “double-checking.” The second most important issue is the problem of dependence: students may try to find an easy way right away and ignore

textbooks or lab experiments. As a result, they may lose the ability to think critically and search for deeper answers to theoretical questions.

As a solution to the problems mentioned above:(1) students should check the results given by AI using different books and sources, and include a “fact-checking” step; (2) they should first try to solve complex problems in the traditional way, and only then compare the answer with AI; (3) before asking ChatGPT about theory or problems, students should spend 5–10 minutes trying to solve them on their own; (4) improve localization and use specialized models for each subject (like chem-GPT for chemistry); (5) explain the difference between paid and free features, and help students develop critical thinking during training. These actions can reduce the problems of AI and help use it safely and effectively in the learning process.

**Table 21.** Disadvantages of AI tools.

<b>№</b>	<b>Category</b>	<b>Code groups</b>	<b>Perc.(%)</b>
1	Accuracy and reliability problems	Incorrect answers, false information	73%
2	Addiction and cognitive decline	Lazy, think less, crit_down	39%
3	Inability to solve complex problems and specialized content in chemistry	Incorrect methods of solving, false balanced equation	25%
4	Unavailable features	High price, paid version	7%
5	Linguistic and incomplete segments	Language limit	4%
6	Not enough real examples	Less real examples	13%
7	Socio-ethical and professional implications	Job risk	4%

## 4. DISCUSSION

The first research question of the study focused on the effectiveness and usability of AI integrated chemistry education to evaluate the usefulness and impact of AI tools in the learning process. Literature review and a structured questionnaire was used. From literature review, it was identified that artificial intelligence (AI) is applied in four main types of activities within chemistry education. These categories are clearly outlined in Table 4, which presents the types of activities supported by AI, relevant research sources, and the frequency and percentage of their appearance in the academic literature. The findings indicate that the most common application of AI in chemistry education is for answering questions and solving problems, accounting for 70% of the reported cases. This suggests that students primarily rely on AI tools to support their understanding of chemistry concepts and to assist in problem-solving tasks. Such usage reflects the practical benefits AI offers in reinforcing learning and promoting independent study. In contrast, the least frequently reported activities were related to laboratory work and scientific prediction, each representing 20% of the cases. These include tasks such as writing laboratory reports or generating research-based hypotheses. Although less commonly practiced, these activities appear to be of particular interest to students at the college and university level, where there is greater emphasis on research and critical thinking. The lower percentage may also indicate the current limitations of AI tools in handling more complex or hands-on tasks, or it may reflect a lack of integration of AI in these areas within the current educational systems. Overall, these findings highlight not only the growing role of AI in supporting student learning in chemistry but also the variation in how it is applied depending on the nature of the task.

Also after the literature review in the second part aimed to investigate the effectiveness and usability of AI-integrated tools in chemistry education, a survey was administered to identify which AI tools students commonly use in their learning process. As shown in Table 5, the results indicate that ChatGPT is overwhelmingly the most utilized AI tool, with 99.1% of participants reporting its use. This striking majority suggests that students strongly prefer ChatGPT, likely due to its user-friendly interface, responsiveness, and ability to address a broad range of educational topics, including complex chemistry concepts. This finding agrees with Vladimir L. Kodkin and Ekaterina V. Artemeva(2024), Ted M. Clark(2023), Wajeeh Daher et al(2023), Suzanne Fergus et al(2023) , Wajeeh Daher et al (2024), Mary E. Emenike and Bright U. Emenike(2023), R. Perezzan et al(2024), Thu Nguyen and Yanika Sirichokcharoenkun(2023), Bekir Yildirim and Ahmet Tayfur Akcan(2024), Himel Mondal, Gujaram Marndi, Joshil Kumar Behera, and Shaikat Mondal(2023), Thuy Nguyen Linh Cao and Phong Nguyen(2023), who in their articles used the ChatGPT

as the most usable AI tools for observation. In contrast, usage of other AI tools such as DeepSeek (0.9%) and none for Yandex AI or Gemini was almost negligible. These findings imply either a lack of awareness or a perception that these alternative tools offer less utility or relevance in the context of chemistry education. The data highlights a potential over-reliance on a single AI tool, pointing to a need for broader exposure and training on various AI applications.

Further insights into students' perceptions of AI tools are provided in Table 6. The findings reveal a generally positive attitude toward the integration of AI in chemistry learning. A significant proportion of students acknowledged the helpfulness of AI tools in their learning process, with 54.6% partly agreeing and 43.5% fully agreeing that such tools are beneficial for asking questions and solving chemistry problems. With a combined total of 98.1% expressing some level of agreement, it is clear that AI tools have gained wide acceptance among learners. This finding agrees with that of Vladimir L. Kodkin and Ekaterina V. Artemeva (2024), who studied real-life examples of how ChatGPT can help improve learning in chemistry. Furthermore, they reported positive sides of AI tools for studying chemistry topics, and writing scientific ideas in chemistry. The minimal disagreement (1.9%) further reinforces the view that AI is seen as a valuable supplement to traditional learning methods, particularly in assisting students with problem-solving and comprehension of challenging material.

Regarding the influence of AI tools on students' interest in chemistry, responses were also generally positive, though slightly more varied. A total of 82.5% of students (51.9% partly agreeing and 30.6% fully agreeing) reported that AI tools increased their interest in the subject. However, 17.6% disagreed, indicating that for a minority of students, AI tools did not enhance engagement. This variation suggests that while AI can be a motivating factor for many, its effectiveness in stimulating interest may depend on individual learning styles and the ways in which these tools are integrated into the curriculum. To address this, educators could design more interactive or personalized AI-based learning activities that appeal to a broader spectrum of learners.

Students' perceptions of the impact of AI on their academic performance were also assessed. Over 90% of respondents believed that their results had improved as a result of using AI tools (50.9% partly agreeing, 39.8% fully agreeing). Only 9.3% of students disagreed. These results suggest that AI tools not only support learning but may also contribute directly to improved academic outcomes. The likely reasons include clearer explanations, increased access to practice resources, and enhanced problem-solving support—all of which can positively influence performance in tests and assignments. This finding agrees with that of Mary E. Emenike and Bright U. Emenike(2023), who explained the positive sides of AI tools for students. They

reported that AI helps with writing assignments, lab reports, research proposals, and tutoring.

Similarly, the majority of students perceived that AI tools contributed to faster learning. Specifically, 49.1% fully agreed and 43.5% partly agreed that AI helped them learn more quickly, amounting to a total of 92.6% positive responses. This reflects the role of AI in delivering immediate feedback, concise explanations, and real-time assistance, which may accelerate the learning process and enable students to work more efficiently. The finding highlights the potential of AI tools to support personalized pacing and adaptive learning, allowing students to engage with content at a speed suited to their individual needs. The responses also shed light on students' growing independence as learners when using AI. A combined 87.9% of students (48.1% partly agree, 39.8% fully agree) indicated that AI tools helped them become more autonomous in their studies. This is a particularly important finding, as it suggests that AI may play a role in fostering self-directed learning habits. The ability to receive continuous support, explanations, and feedback from AI tools without relying entirely on teachers may help students build confidence and resilience in their learning journey. However, when asked whether AI could replace aspects of traditional teaching, students expressed more cautious views. While 60.2% partly agreed and 16.7% fully agreed that AI could replace some teaching functions, 23.1% disagreed. This distribution indicates that while students recognize the supportive role of AI, they still value the human elements of teaching, such as mentorship, emotional connection, and context-specific feedback. These results underline the necessity of a blended learning approach, where AI is used to complement rather than substitute traditional teaching methods.

Table 7 provides additional context by identifying the main purposes for which students use AI tools. The most commonly reported use was for understanding difficult topics, cited by 83.3% of participants. This emphasizes the perceived value of AI in clarifying complex concepts, often through simplified explanations, visual aids, or interactive problem-solving features. The second most frequent use, reported by 62% of students, was for asking theory-related questions. This reflects the immediacy and convenience of AI tools as sources of theoretical clarification, reducing dependency on teachers for basic inquiries. This findings agree with that of Daher, Diab, and Rayan (2024), who reported the usefulness of AI for different question formats like open-ended and multiple-choice. A notable 56.5% of students also reported using AI for exam preparation, and 54.6% indicated using it to find answers to test questions. This findings agree with that of Sridhar et al. (2022), who explained that Chatbots can act as virtual tutors, providing instant feedback, answering questions, and guiding students throughout the learning process. While these practices suggest that students rely

heavily on AI during high-stakes assessments, they also raise concerns about the potential for misuse or over-dependence. This underlines the need for academic integrity education and for guiding students on responsible and ethical AI use. Encouraging students to engage in reflective learning, rather than simply seeking answers, is essential to ensure that AI tools enhance, rather than replace, critical thinking and problem-solving skills.

The second research question explored the ethical and privacy concerns that arise when students use AI tools in chemistry education. The results, derived from both closed- and open-ended survey responses, provide a detailed view of students' perceptions, behaviors, and ethical considerations. Overall, students appear to have a pragmatic and moderately cautious approach toward AI, recognizing its usefulness while remaining aware of its limitations and risks.

According to the data (Table 8), a large majority of students (81.5%) partly agreed that AI tools are reliable, while only 11.1% fully agreed. This cautious optimism reflects the broader academic consensus observed in the literature. For instance, Clark (2023) and Fergus et al. (2023) noted that while ChatGPT can answer simple conceptual questions correctly, it often fails to handle tasks requiring higher-order thinking or interpretation of visual data. Similarly, Daher et al. (2024) pointed out that ChatGPT tends to produce more accurate responses in English than in Arabic, indicating language-based reliability concerns. Thus, the student skepticism reflected in your data aligns well with published findings.

Fairness perceptions were largely positive: 69.2% partly agreed and 30.8% fully agreed that AI use in education is fair. This suggests a general belief in equitable access, though open-ended responses revealed concerns about unequal access to devices, internet speed, or premium AI versions. This resonates with concerns raised by Emenike & Emenike (2023), who discussed the ethical dilemma of accessibility and fairness when AI tools are embedded into learning systems. These findings highlight that while most students view AI as a democratizing force in education, systemic inequalities can still impact its fair usage.

One of the most debated topics was academic integrity. While 26.9% of students disagreed that AI use breaks integrity, 61.5% partly agreed and 11.5% fully agreed that it could compromise honest academic conduct. This middle-ground response illustrates an ongoing tension between AI as a learning aid and its potential misuse for cheating. This ambiguity is also echoed in the literature: Fergus et al. (2023) emphasized that while ChatGPT itself does not increase cheating risks directly, the lack of clear guidelines and overreliance could foster unethical practices. The open-ended responses support this view, as students themselves call for clearer boundaries and educational policies that define what constitutes acceptable AI use.

The accessibility of AI tools was perceived as high: 65.4% fully agreed and 26.9% partly agreed that all students have equal access. However, literature and qualitative responses challenge this view. Some students noted that premium AI versions offer better accuracy and features, which may create a two-tiered access system. Akgun & Greenhow (2022) similarly cautioned that access to AI tools could be uneven across socioeconomic lines, reinforcing the need for equitable implementation strategies by educators and institutions.

Opinions about teacher control over AI use were split: 34.6% opposed restrictions, 42.3% supported moderate control, and 23.1% favored strict limits. Open responses suggest that students value guidance more than prohibition. They want structured training on how to use AI responsibly an approach supported by Bekir Yildirim and Ahmet Tayfur Akcan (2024), who advocate for teacher professional development in AI literacy to set ethical norms and foster responsible usage among students. These responses collectively suggest that rather than imposing bans, educators should adopt a policy of guided autonomy supported by transparency, dialogue, and scaffolding.

Encouragingly, students show a strong inclination toward ethical use: 53.8% always paraphrase AI-generated content, and none admitted to submitting answers verbatim without at least some modification. This practice aligns with findings by Daher et al. (2023), who noted that ChatGPT is best used when students refine or critique its output rather than use it blindly. However, 7.7% of students admitted to submitting unedited AI-generated work, and 42.3% only partially rephrase, suggesting that a notable minority still needs ethical literacy training.

The qualitative responses further enrich these results. Students who viewed their use of AI as fair stressed that they verify results, study the solution process, and use the tool as a learning partner. In contrast, others admitted to using AI to shortcut effort, acknowledging that this undermines their own understanding. This duality echoes the ethical paradox raised in Alli (2025), who warned that overdependence on AI without critical engagement can impair long-term learning. The responses reinforce the importance of promoting a culture of reflection, responsible use, and critical thinking.

To sum up, while most students acknowledge the benefits of AI in chemistry education and strive to use it ethically, gaps in translation accuracy, digital access, and clear usage policies remain. These concerns are consistent with themes in the reviewed literature and suggest that integrating AI tools into education requires not only technological readiness but also ethical and pedagogical infrastructure. Educators are encouraged to adopt a blended strategy-combining AI tools with training, dialogue, and integrity-focused policies to ensure that AI enhances learning without compromising fairness or academic honesty.

The third research question investigated the main advantages and disadvantages of using AI tools in teaching chemistry. This was examined through a combination of student surveys (both open- and closed-ended questions) and expert evaluation of ChatGPT's ability to solve national chemistry test items. This dual approach allowed for a holistic understanding of how AI tools are perceived and how they actually perform in real educational contexts.

From the content analysis of student responses about the advantages of AI tools, the most frequently cited advantage (reported by 48.8%) was speed and efficiency. Students emphasized how AI, especially ChatGPT, helps them receive fast, direct answers, which saves significant time when studying for exams or doing assignments. This aligns with Alli (2025), who reported that AI tutoring systems can help accelerate learning by delivering immediate, personalized feedback. The second major benefit (26.8%) was the clarity and simplicity of AI-generated explanations. Students appreciated how complex chemistry topics were broken down into more understandable components. This finding supports those by Nguyen & Sirichokcharoenkun (2023) and Emenike & Emenike (2023), who emphasized that AI tools like ChatGPT are effective in facilitating comprehension by providing user-friendly and accessible content. Another advantage (20.3%) involved information retrieval. Students used AI to filter and summarize large amounts of content quickly. In line with Daher, Diab, & Rayan (2024), this suggests AI serves not only as a tutor but also as an intelligent search and synthesis assistant. Students also cited multifunctionality, exam support, 24/7 accessibility, and individualized adaptation, though these were mentioned less frequently. However, the low percentage (1.6%) of students who believed AI stimulated their critical thinking indicates that AI's role is still largely passive and underutilized in promoting higher-order skills—a concern echoed by Fergus et al. (2023).

From the content analysis of student responses about the disadvantages of AI tools, the most prominent concern, expressed by 73% of students, was accuracy and reliability. Many noted that AI tools sometimes provided incorrect answers, misleading explanations, or contradictory outputs. This confirms prior findings by Clark (2023) and Daher et al. (2023), who documented similar issues with AI's inability to handle visual data, complex reactions, or symbolic representation accurately. The problem becomes more critical during exam preparation, where incorrect solutions can directly lead to lower academic performance. A second major concern (39%) was over-reliance on AI, leading to cognitive decline and a drop in self-directed learning. This aligns with ethical warnings by Akgun & Greenhow (2022) and the observations of Emenike & Emenike (2023), who cautioned that AI may unintentionally foster superficial learning unless used under pedagogical supervision. Students also noted AI's inability to handle

complex problems (25%), especially those requiring abstract thinking, mathematical calculation, or multi-step logic. This criticism is consistent with the findings of Daher et al. (2023, 2024), who reported ChatGPT's struggles with deep conceptual chemistry knowledge and representation-based problem solving. Other disadvantages included limited access to premium features (7%), linguistic issues in Kazakh (4%), and a lack of real-world examples (13%). These findings suggest that while AI is perceived as helpful, it still lacks localization and contextual adaptation in many educational environments.

To validate student perceptions with actual data, five experts evaluated ChatGPT's responses to five real chemistry test questions using four criteria: translation accuracy, clarity of solution, chemistry relevance, and final correctness.

The results were mixed:

- Translation and clarity were frequently rated lower (mean scores as low as 1.2 and 1.6), confirming concerns about linguistic quality and logical organization. Experts agreed that while answers were often mathematically correct, the language used could be confusing or scientifically imprecise.
- Correctness of final answers was rated highly in most tasks (scores of 3 in four out of five cases), showing that ChatGPT can compute accurate results-provided the prompt is clear and the question well-structured.
- However, Kendall's *W* coefficients often revealed low agreement among experts, especially for subjective aspects like clarity or translation, highlighting inconsistency in how experts judged AI performance.

These findings echo earlier research by Clark (2023) and Fergus et al. (2023), who found that ChatGPT often produced plausible but flawed explanations or missed important contextual cues. They also support Daher et al.'s (2024) recommendation that AI performance can be significantly improved by translating non-English content into English first, and only then processing it with AI.

Together, the survey data and expert review suggest that AI tools-especially ChatGPT-offer clear benefits in speed, accessibility, and support for individual learning, but require careful oversight due to errors, superficial reasoning, and ethical risks. These findings are echoed in the literature and underscore the need for AI-integrated pedagogy, rather than unstructured adoption. Teachers and educational institutions should play an active role in guiding the responsible use of AI tools. This includes offering targeted training that helps students understand how to use AI ethically and effectively, as well as promoting habits of critical thinking by encouraging them to verify AI-generated information rather than accepting it at face value. In addition, instructors should design assignments that go beyond simple fact retrieval and instead require students to engage in reasoning, problem-solving, and

deeper analysis. To further enhance the accuracy and relevance of AI assistance, it is also recommended to integrate subject-specific models-such as Chem-GPT for chemistry-into the educational process, as these tools are better equipped to handle domain-specific content.

In conclusion, while students find AI tools effective for practice and exploration, both their perceptions and expert evaluations point to clear limitations that must be addressed through better instructional design, transparency, and ethical policies.

## 5. CONCLUSION

This dissertation research was devoted to a comprehensive study and identification of the main challenges and opportunities encountered by students and educators in Kazakhstan when integrating Artificial Intelligence (AI) tools, specifically ChatGPT, into the chemistry education process. The aim of the work was not only to identify these challenges and advantages but also to analyze them based on student perceptions, expert evaluations, ethical considerations, and actual effectiveness in enhancing learning outcomes. The conducted research allowed us to draw several key conclusions.

Research Question 1 evaluated the effectiveness and usability of AI-integrated chemistry education. The findings indicated a strong preference among students for ChatGPT due to its efficiency in delivering rapid and clear explanations, simplifying complex chemical concepts, and promoting independent learning. AI was mainly employed for solving chemistry problems, clarifying theoretical content, and aiding exam preparation, significantly enhancing students' engagement and academic performance. Despite these benefits, the use of AI in laboratory settings and scientific prediction remains limited, highlighting potential areas for further research and development. Additionally, students highly valued the multifunctional aspect of AI, as it provided a single platform for diverse educational needs, ranging from theoretical inquiries to practical exercises.

Research Question 2 explored the ethical and privacy concerns related to AI use in education. Students generally recognized AI as fair and accessible but expressed concerns regarding academic integrity, potential over-reliance, and unequal access to premium features. They called for clear guidelines and explicit training to mitigate these ethical risks effectively. Furthermore, students noted cognitive decline and increased dependence on AI, potentially undermining critical thinking skills. The research underscores the importance of institutional policies, training sessions, and strategies that encourage responsible AI use and independent verification of AI-generated information, ensuring that learners develop critical evaluation skills alongside AI competencies.

Research Question 3 examined the advantages and disadvantages of AI tools through student feedback and expert evaluation. Key advantages identified included speed, efficiency, accessibility, simplified explanations, and adaptive learning capabilities. However, significant limitations were also noted concerning the accuracy, reliability, and clarity of AI-generated solutions, particularly for complex problems. Expert evaluations further confirmed issues such as translation inaccuracies, logical inconsistencies, and limited linguistic functionality in Kazakh, underscoring the need for improvements in these areas. Additionally, students raised concerns about

inadequate real-world examples provided by AI and limitations in handling specialized chemistry content accurately.

Overall, this research highlights that while AI tools like ChatGPT offer substantial benefits for chemistry education, careful oversight, structured integration, and ethical considerations are essential. Recommendations for future practice include developing specialized chemistry-focused AI applications, improving linguistic localization, conducting regular training for critical thinking enhancement, and establishing comprehensive frameworks that balance AI assistance with traditional educational methodologies. These strategic actions will ensure AI serves as an effective and responsible educational resource, supporting both students and educators in the chemistry learning process.

## SHORTCOMINGS AND LIMITATIONS

Several limitations should be considered when interpreting the results of the dissertation, as they may affect the results and conclusions of the study.

The first limitation is the number of respondents and their level of education during the quantitative analysis stage. The quantitative survey involved 108 chemistry students from only three universities in Kazakhstan (SDU, KazNPU, KazSWPU). While this is sufficient for exploratory research, our sample may not reflect the varying technological access, digital literacy levels, or educational levels of chemistry students across the country. Therefore, the findings should be interpreted as indicative rather than broadly representative. It is important to remember that since the survey was conducted among students only, the impact of AI tools on schoolchildren is likely to be different.

Secondly, although the study references several AI tools, the empirical evaluations and expert assessments are focused on ChatGPT. The conclusions drawn from ChatGPT may not generalize to other AI platforms such as Gemini, DeepSeek, or Yandex AI. It is important to note that they present different opportunities and limitations in educational contexts.

Thirdly, a central component of the study was based on the evaluation of AI-generated chemistry answers, especially those generated using ChatGPT. While the expert groups confirmed that most of the answers were technically correct, there were persistent concerns about the accuracy of the translation, clarity of the solution, and conceptual coherence, especially for complex or high-order chemistry problems. In many cases, the AI provided scientifically inaccurate or misleading solutions. This not only limited its pedagogical value, but also posed a risk if students used it uncritically.

Further, the study found a significant deficiency in the performance of AI tools such as ChatGPT in the Kazakh language. This linguistic limitation significantly affects accessibility and usability for Kazakh-speaking students and highlights the broader challenges in localizing AI for low-resource languages. Consequently, the educational effectiveness of AI cannot be ensured equally in all linguistic contexts in Kazakhstan.

Most of the data on AI use, effectiveness, and ethical considerations in the study was collected from self-reported surveys. These responses may be biased, particularly regarding academic integrity and societal concerns about AI misuse. Students may report instances of inappropriate use of AI tools or overestimate ethical caution.

Also, the qualitative evaluation of AI, including ChatGPT, on chemistry questions translated from Kazakh to English and provided solutions was conducted by only five individuals with varying academic backgrounds (from high school students to PhDs and teachers). While there were individuals with diverse perceptions and

levels, the small sample groups may limit the reliability and depth of the level of expertise.

Finally, the study did not track the actual academic performance of students before and after using AI tools. Therefore, while many students report improved interest and understanding in surveys, there is no objective, concrete evidence to support that AI integration has significantly improved learning outcomes over time.

## **REFERENCES**

Akgun, S., & Greenhow, C. (2022). *Artificial intelligence in education: Addressing ethical challenges in K-12 settings*. *AI and Ethics*, 2(3), 431-440.

Alejandro Barredo Arrieta, Natalia Díaz-Rodríguez (2020). *Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI*. *Information Fusion*, volume 58, 82-115.

Clark, T. M. (2023). *Investigating the use of an artificial intelligence chatbot with general chemistry exam questions*. *Journal of Chemical Education*, 100(5), 1905-1916.

Daher, W., Diab, H., & Rayan, A. (2023). *Artificial intelligence generative tools and conceptual knowledge in problem solving in chemistry*. *Information*, 14(7), 409.

Daher, W., Diab, H., & Rayan, A. (2024). *Generative Artificial Intelligence in Chemistry Problem Solving across Versions and Languages*.

Emenike, M. E., & Emenike, B. U. (2023). *Was this title generated by ChatGPT? Considerations for artificial intelligence text-generation software programs for chemists and chemistry educators*. *Journal of Chemical Education*, 100(4), 1413-1418.

Fergus, S., Botha, M., & Ostovar, M. (2023). *Evaluating academic answers generated using ChatGPT*. *Journal of Chemical Education*, 100(4), 1672-1675.

Figueiredo, M., Esteves, M. L., Neves, J., & Vicente, H. (2014). *Lab classes in chemistry learning an artificial intelligence view*. In *International Joint Conference SOCO'14-CISIS'14- ICEUTE'14: Bilbao, Spain, June 25th-27th, 2014, Proceedings* (pp. 565-575). Springer International Publishing.

Gray, N. A. B. (1988). *Artificial intelligence in chemistry*. *Analytica chimica acta*, 210, 9-32

Kodkin, V. L., & Artem'eva, E. V. (2024). *ChatGPT: Application in Chemistry Education and Challenges*. *Journal of Computer and Communications*, 12(03), 196-206.

Nguyen, T., & Sirichokcharoenkun, Y. (2023). *The Limitations and Potential of ChatGPT in Chemistry Education*.

Perezan, R., Montalvo-Quirós, S., Rama-Ballesteros, R., & Herráez-Aguilar, D. (2024). *A PROPOSAL TO INTEGRATE ARTIFICIAL INTELLIGENCE TOOLS IN CHEMISTRY TEACHING IN HIGHER EDUCATION*. In *INTED2024 Proceedings* (pp. 3435-3441). IATED.

Yildirim, B., & Akcan, A. T. (2024). *AI-professional development model for chemistry teacher: Artificial intelligence in chemistry education*. *Journal of Education in Science, Environment and Health (JESEH)*, 10(4), 161-182.

Alli, A.O. 2025. *A Review of the Potential of Artificial Intelligence (AI) in Enhancing Chemistry Education*. *International Journal of Recent Innovations in Academic Research*, 9(2): 31-43.

Mondal, Himel; Marndi, Gujaram1; Behera, Joshil Kumar2; Mondal, Shaikat3. *ChatGPT for Teachers: Practical Examples for Utilizing Artificial Intelligence for Educational Purposes*. *Indian Journal of Vascular and Endovascular Surgery* 10(3):p 200-205, Jul–Sep 2023.

Russell, S.J., Norvig, P. (2016). *Artificial Intelligence: A Modern Approach*. Pearson Education Limited, Malaysia.

West, D. M. (2018). *The future of work: Robots, AI, and automation*. Brookings Institution Press.

Farwati, R., Sari, I., & Fadhilah, G. (2024). *Transforming the Chemistry Education Curriculum: Students' Creativity Using AI in Lectures*. *Hydrogen: Jurnal Kependidikan Kimia*, 12(4), 687-696.

Iorliam, Aamo & Ingio, Joseph. (2024). *A Comparative Analysis of Generative Artificial Intelligence Tools for Natural Language Processing*. *Journal of Computing Theories and Applications*. Volume 2. 10.62411/jcta.9447.

Gupta, R., Hamid, A. M., Jhaveri, M., Patel, N., & Suthar, P. P. (2024). *Comparative Evaluation of AI Models Such as ChatGPT 3.5, ChatGPT 4.0, and Google Gemini in Neuroradiology Diagnostics*. *Cureus*, 16(8), e67766. <https://doi.org/10.7759/cureus.67766>

Chu, Hillary & Pasion, Emily & Yeh, Stephanie & Chu, Gary. (2024). *Assessing the Ethical and Professional Capabilities of AI: A Study of ChatGPT and Google Gemini versus PReview (Situational Judgement Test) for Medical Student Applicant*. *Journal of Clinical Question*. 1. 82-88. [10.69854/jcq.2024.0011](https://doi.org/10.69854/jcq.2024.0011).

Gill, Sukhpal Singh & Xu, Minxian & Patros, Panos & Wu, Huaming & Kaur, Rupinder & Kaur, Kamalpreet & Fuller, Stephanie & Singh, Manmeet & Arora, Priyansh & Parlikad, Ajith Kumar & Stankovski, Vlado & Abraham, Ajith & Ghosh, Soumya & Lutfiyya, Hanan & Kanhere, Salil & Bahsoon, Rami & Rana, Omer & Sakellariou, Rizos & Buyya, Rajkumar. (2023). *Transformative Effects of ChatGPT on Modern Education: Emerging Era of AI Chatbots*. *Internet of Things and Cyber-Physical Systems*. Volume 4, 2024, Pages 19-23, ISSN 2667-3452, <https://doi.org/10.1016/j.iotcps.2023.06.002>.

Fraiwan, Mohammad & Khasawneh, Natheer. (2023). *A Review of ChatGPT Applications in Education, Marketing, Software Engineering, and Healthcare: Benefits, Drawbacks, and Research Directions*. [10.48550/arXiv.2305.00237](https://arxiv.org/abs/2305.00237).

Baidoo-Anu, David and Owusu Ansah, Leticia, *Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning (January 25, 2023)*. Available at SSRN: <http://dx.doi.org/10.2139/ssrn.4337484>

M. Shukla, I. Goyal, B. Gupta, J. Sharma, "A Comparative Study of ChatGPT, Gemini, and Perplexity", *International Journal of Innovative Research in Engineering & Management (ijircst)*, Vol-12, Issue-4, Page No-10-15, 2024. Available from: <https://doi.org/10.55524/ijircst.2024.12.4.2>

Temsah, A., Alhasan, K., Altamimi, I., Jamal, A., Al-Eyadhy, A., Malki, K. H., & Temsah, M. H. (2025). *DeepSeek in Healthcare: Revealing Opportunities and Steering Challenges of a New Open-Source Artificial Intelligence Frontier*. *Cureus*, 17(2), e79221. <https://doi.org/10.7759/cureus.79221>

Okaiyeto S A, Bai J W, Wang J, Mujumdar Arun S, Xiao H W. *Success of DeepSeek and potential benefits of free access to AI for global-scale use*. *Int J Agric & Biol Eng*, 2025; 18(1):304–306.

Zudonu, O. C., Osiah, C. U., Ogbu, M. O., Osu, A. D., Ndukwu, D. E., Afolabi, B. A., John, J. P., & Nkisa, O. D. (n.d.). Investigating the effect of Artificial Intelligence on Chemistry and Physics students' achievement and conceptual change in heat change in SSS2 in Rivers State. *International Journal of Chemistry and Chemical Processes E-ISSN 2545-5265 P-ISSN 2695-1916, Vol 10. No. 3 2024* [www.iiardjournals.org](http://www.iiardjournals.org)

Nguyen, Thuy & Ai, & Trần, Vũ & Nguyen, Phuong & Nguyen, Phong & Cao, Linh. (2023). *Capabilities, Benefits, and Role of ChatGPT in Chemistry Teaching and Learning in Vietnamese High Schools*. 10.35542/osf.io/4wt6q.

# APPENDICES

## Appendix 1. Survey questions

**Қай университетте оқисыз? / Which university do you study at?**

- СДУ / SDU
- ҚазҰПУ / KazNPU
- ҚазҰҚПУ / KazNWTTU

**1. Қай жасанды интеллект түрлерін жиі қолданасыз? / Which AI tools do you use most often?**

- ChatGPT
- Gemini
- DeepSeek
- Yandex AI

**2. Жасанды интеллектінің құралдарын химияны үйренуде (сұрақтар сұрау, есептер шығаруда) қолданасыз ба? / Do you use AI tools to study chemistry (e.g., asking questions, solving problems)?**

- Қолданбаймын / I don't use them
- Жартылай қолданамын - I use them sometimes
- Толықтай қолданамын - I use them regularly

**3. Сіздің ойыңызша жасанды интеллектінің құралдарын химия үйренуде қолдану пайдалы ма? / In your opinion, is using AI tools helpful for learning chemistry (e.g., asking questions, solving problems)?**

- Келіспеймін - I disagree
- Жартылай келісемін - I partly agree
- Толықтай келісемін - I fully agree

**4. Сіздің ойыңызша жасанды интеллектінің құралдары сенімді ме? / In your opinion, are AI tools reliable for studying chemistry?**

- Келіспеймін - I disagree
- Жартылай келісемін - I partly agree
- Толықтай келісемін - I fully agree

**5. Жасанды интеллект құралдары сіздің химияға деген қызығушылығыңызды арттырады ма? / Do AI tools increase your interest in chemistry?**

- Келіспеймін - I disagree
- Жартылай келісемін - I partly agree
- Толықтай келісемін - I fully agree

**6. ЖИ құралдарын қолданғанда химияны жылдамырақ меңгеремін / I learn chemistry faster when I use AI tools**

- Келіспеймін - I disagree
- Жартылай келісемін - I partly agree
- Толықтай келісемін - I fully agree

**7. ЖИ қолданғаннан кейін тест немесе тапсырмаларда нәтижелерім жақсарды/ My test or assignment results improved after using AI tools**

- Келіспеймін - I disagree
- Жартылай келісемін - I partly agree
- Толықтай келісемін - I fully agree

**8. ЖИ құралдарын жиі қандай мақсатта қолданасыз? / For which purposes do you usually use AI tools?**

- Емтиханға дайындалуға - For exam preparation
- Түсінбеген тақырыптарды үйренуге - To understand difficult topics
- Тесттердің жауабын анықтауға - To find test answers
- Теориялық сұрақтардың жауабын анықтауға - To ask theory-related questions

**9. ЖИ көмегімен химияны оқу барысында өз бетіммен үйрену қабілетім артты / Using AI tools improved my ability to study chemistry independently**

- Келіспеймін - I disagree
- Жартылай келісемін - I partly agree
- Толықтай келісемін - I fully agree

**10. Сіз ЖИ дәстүрлі оқытудың кейбір бөліктерін алмастыра алады деп ойлайсыз ба? / Do you think AI can replace some parts of traditional teaching?**

- Келіспеймін - I disagree
- Жартылай келісемін - I partly agree
- Толықтай келісемін - I fully agree

**11. Сіз үшін ЖИ құралдарын қолдану әділ деп ойлайсыз ба? / Do you think using AI tools in education is fair?**

- Келіспеймін - I disagree
- Жартылай келісемін - I partly agree
- Толықтай келісемін - I fully agree

**12. ЖИ құралдарын пайдалану академиялық адалдықты бұзады деп ойлайсыз ба? / Do you think using AI tools breaks academic integrity?**

- Келіспеймін - I disagree
- Жартылай келісемін - I partly agree
- Толықтай келісемін - I fully agree

**13. Студенттердің бәріне ЖИ құралдары тең қолжетімді деп ойлайсызба? / Do you think AI tools are equally accessible to all students?**

- Келіспеймін - I disagree
- Жартылай келісемін - I partly agree

- Толықтай келісемін - I fully agree

**14. Мұғалімдер ЖИ қолдануды шектеуі керек деп ойлайсыз ба? / Should teachers limit or control AI usage in learning?**

- Келіспеймін - I disagree

- Жартылай келісемін - I partly agree

- Толықтай келісемін - I fully agree

**15. ЖИ шығарған жауаптарды өз сөзіңізбен қайта жазып, содан кейін тапсырасыз ба? Do you rewrite AI-generated answers in your own words before submitting them?**

- Келіспеймін - I disagree

- Жартылай келісемін - I partly agree

- Толықтай келісемін - I fully agree

**16. ЖИ шығарған жауапты өзіңіздің түпнұсқа жауабыңыз ретінде қабылдайсыз ба? Do you present an AI-generated answer as your own original work?**

- Келіспеймін - I disagree

- Жартылай келісемін - I partly agree

- Толықтай келісемін - I fully agree

**17. ЖИ арқылы есеп шығарғанда немесе тапсырма орындағанда өзіңді әділ пайдаланушы ретінде сезінесіз бе? Неге? / Do you feel fair when using AI for solving problems or assignments? Why?**

**18. Егер сіз оқытушы болсаңыз, студенттердің ЖИ қолдануын қалай реттейсіз? / If you were a teacher, how would you manage student use of AI tools?**

**19. Жасанды интеллект құралдарын қолданудың сіз байқаған артықшылықтары қандай? / What are the advantages of using AI tools in learning chemistry, based on your experience?**

**20. Жасанды интеллект құралдарын қолданудың қандай кемшіліктері бар деп ойлайсыз? / What do you think are the disadvantages or limitations of using AI tools in learning chemistry?**

## Appendix 2. Questions from UNT in Kazakh language

1. V топ бейметалдарының сыртқы энергетикалық деңгейдегі электрондық конфигурациясы
  - A)  $ns^2np^3$
  - B)  $ns^2np^5$
  - C)  $ns^2np^4$
  - D)  $ns^2np^2$
2. Mn және Cr атомдарының нейтрондарының қосындысы
  - A) 98
  - B) 58
  - C) 88
  - D) 78
3. Аммиактың катализатор қатысынсыз тотығу реакциясында оттектің концентрациясын 3 есе арттырғандағы реакция жылдамдығы
  - A) 27 есе артады
  - B) өзгермейді
  - C) 243 есе артады
  - D) 27 есе кемиді
4.  $C_2H_{4(g)} + H_2O_{(g)} \rightleftharpoons C_2H_5OH_{(g)}$   $\Delta H < 0$  тепе-теңдікті этанолдың айырылу бағытына ығыстыру үшін қажетті жағдай
  - A) Температураны төмендету
  - B) Қысымды арттыру
  - C) Этилен концентрациясын төмендету
  - D) Этанол концентрациясын төмендету
5. Мыс(II) нитратының ерітіндісі электролизге ұшырағанда 3,36 л(қ.ж.) газ түзілді. Өнім шығымы 80% болса электролизге ұшыраған тұз массасы
  - A) 70.5 г
  - B) 85.0 г
  - C) 65.5 г
  - D) 25.0 г

### Appendix 3. Questions from UNT and their solutions by ChatGPT in English language

1. The electron configuration of the outer energy level of group V nonmetals is:

- A)  $ns^2np^3$
- B)  $ns^2np^5$
- C)  $ns^2np^4$
- D)  $ns^2np^2$

#### Step 1: Understand Group V nonmetals

In the periodic table, **Group V** nonmetals (modern notation: **Group 15**) include:

- **Nitrogen (N)**
- **Phosphorus (P)**
- **Arsenic (As)** (metalloid, but still included in the group)
- etc.

These elements are in **Group 15**, meaning they have **5 electrons in their outermost shell**.

---

#### Step 2: General electron configuration of p-block elements

p-block elements have the outer shell electron configuration:

- $ns^2 np^1\text{--}np^6$

Group 15 elements have **5 valence electrons**, which fill the **s and p orbitals** as follows:

- **2 electrons in the s-orbital**
- **3 electrons in the p-orbital**

This gives:

- $ns^2 np^3$
- 

#### Step 3: Match with the options

- A)  $ns^2 np^3$   Correct
- B)  $ns^2 np^5$  → Group 17 (halogens)
- C)  $ns^2 np^4$  → Group 16 (oxygen group)
- D)  $ns^2 np^2$  → Group 14 (carbon group)

2. The total number of neutrons in Mn and Cr atoms

- A) 98
- B) 58
- C) 88
- D) 78

**Step 1: Get the atomic numbers and mass numbers from the Periodic Table**

- **Cr (Chromium)**
  - Atomic number = 24
  - Mass number (most common isotope)  $\approx$  52
- **Mn (Manganese)**
  - Atomic number = 25
  - Mass number (most common isotope)  $\approx$  55

---

**Step 2: Calculate the number of neutrons in each**

The formula to calculate neutrons:

$$\text{Neutrons} = \text{Mass Number} - \text{Atomic Number}$$

- **Cr:**

$$52 - 24 = 28 \text{ neutrons}$$

- **Mn:**

$$55 - 25 = 30 \text{ neutrons}$$

---

**Step 3: Add them together**

$$28(\text{Cr}) + 30(\text{Mn}) = 58 \text{ neutrons}$$

---

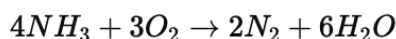
**✔ Final Answer: B) 58**

3. The reaction rate in the oxidation of ammonia without a catalyst when the oxygen concentration is increased 3 times

- A) Increases 27 times
- B) Does not change
- C) Increases 243 times
- D) Decreases 27 times

**Step 1: Understand the reaction**

The reaction is:



This is the **oxidation of ammonia** (without a catalyst).

---

**Step 2: Use the rate law (experimental)**

From experiments, it's known that for the **oxidation of ammonia without a catalyst**, the **rate law** is approximately:

$$\text{Rate} \propto [NH_3]^2 [O_2]^3$$

This means:

- Reaction rate depends **quadratically** on ammonia concentration.
  - Reaction rate depends **cubically** on oxygen concentration.
- 

**Step 3: Analyze what happens when  $[O_2]$  increases 3 times**

We are told:

$$[O_2] \text{ increases 3 times} \Rightarrow [O_2]_{\text{new}} = 3[O_2]$$

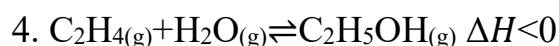
Using the rate law:

$$\text{New rate} \propto [NH_3]^2 \cdot (3[O_2])^3 = [NH_3]^2 \cdot 27[O_2]^3$$

$$\text{New rate} = 27 \times \text{Original rate}$$

---

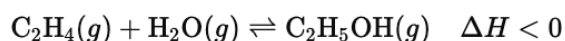
✅ **Final Answer: A) Increases 27 times**



To shift the equilibrium toward the formation of ethanol, the required condition is:

- A) Lower the temperature
- B) Increase the pressure
- C) Decrease the concentration of ethylene
- D) Decrease the concentration of ethanol

**Given Reaction:**



This tells us:

- It's an **exothermic reaction** (releases heat).
- All substances are in the **gas phase**.

We are asked: **What condition favors the formation of ethanol ( $\text{C}_2\text{H}_5\text{OH}$ )?**

---

### Step 1: Apply Le Chatelier's Principle

This principle says that if a stress is applied to a system at equilibrium, the system will **shift to oppose** the change.

We will now evaluate each option.

---

#### Option A) Lower the temperature

- Since the reaction is **exothermic** ( $\Delta H < 0$ ), lowering temperature **removes heat**.
  - The system will shift **toward the product side** (ethanol) to generate more heat.
  - ✓ **This will increase ethanol production.**
- 

#### Option B) Increase the pressure

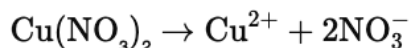
- Count the moles of gas:
  - Left side: 1 mol  $\text{C}_2\text{H}_4$  + 1 mol  $\text{H}_2\text{O}$  = 2 mol
  - Right side: 1 mol  $\text{C}_2\text{H}_5\text{OH}$  = 1 mol
- Increasing pressure favors the side with **fewer gas molecules** → **shifts right**
- ✓ **This also favors ethanol formation.**

5. When an aqueous solution of copper (II) nitrate undergoes electrolysis, 3.36 L (at STP) of gas is released. If the product yield is 80%, what is the mass of the salt subjected to electrolysis?

- A) 70.5 g
- B) 85.0 g
- C) 65.5 g
- D) 25.0 g

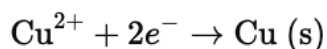
### Step 1: Understand the Electrolysis of Copper(II) Nitrate

$\text{Cu}(\text{NO}_3)_2$  in aqueous solution dissociates into:

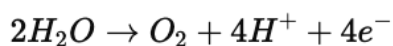


During electrolysis:

- At the **cathode**, copper metal is deposited:



- At the **anode**, **oxygen gas** is released (from water oxidation):



So the gas released is  $\text{O}_2$ .

---

### Step 2: Use the gas volume to find moles of $\text{O}_2$

We're told:

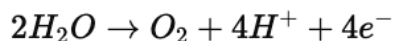
- Volume of gas = 3.36 L (at STP)
- **STP**: 1 mol gas = 22.4 L

So:

$$n_{\text{O}_2} = \frac{3.36}{22.4} = 0.15 \text{ mol O}_2$$

### Step 3: Use stoichiometry of the anode reaction

Anode reaction:



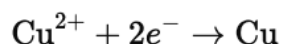
From this:

- 1 mol  $O_2 \leftrightarrow 4$  mol electrons

So:

$$0.15 \text{ mol } O_2 \Rightarrow 0.15 \times 4 = 0.6 \text{ mol electrons}$$

These 0.6 mol electrons also pass through the **cathode**, where  $Cu^{2+}$  is reduced:



So:

- 2 mol electrons deposit 1 mol Cu
- 0.6 mol electrons  $\rightarrow \frac{0.6}{2} = 0.3$  mol Cu

### Step 4: Find theoretical mass of $Cu(NO_3)_2$ required

We need 0.3 mol  $Cu(NO_3)_2$  to produce 0.3 mol Cu.

$$\begin{aligned} \text{Molar mass of } Cu(NO_3)_2 &= \\ &= 63.5 \text{ (Cu)} + 2 \times (14 + 3 \times 16) \\ &= 63.5 + 2 \times 62 = 63.5 + 124 = \mathbf{187.5 \text{ g/mol}} \end{aligned}$$

So:

$$\text{Theoretical mass} = 0.3 \times 187.5 = 56.25 \text{ g}$$

---

### Step 5: Adjust for 80% yield

Only 80% of the salt was effectively used, so the actual amount used is:

$$\text{Actual mass} = \frac{56.25}{0.80} = 70.31 \text{ g} \approx 70.5 \text{ g}$$

---

 **Final Answer: A) 70.5 g**