

# **J E S E H**

**Journal of Education in Science,  
Environment and Health**

**Volume: 10   Issue: 4   Year: 2024**

**ISSN: 2149-214X**

## EDITORIAL BOARD

### Editors

**Valarie L. Akerson**- Indiana University, U.S.A

**Seyit Ahmet Kiray**, Necmettin Erbakan University, Turkiye

### Section Editors

**Manuel Fernandez** - Universidad Europea de Madrid, Spain

**Mustafa Sami Topcu** - Yildiz Technical University, Turkiye

### Editorial Board

**Angelia Reid-Griffin**- University of North Carolina, United States

**Ching-San Lai**- National Taipei University of Education, Taiwan

**Ingo Eilks** - University of Bremen, Germany

**Jennifer Wilhelm**- University of Kentucky, United States

**Lloyd Mataka**-Lewis-Clark State College, United States

**Manuel Fernandez** - Universidad Europea de Madrid, Spain

**Osman Çardak** - Necmettin Erbakan University

**P.N. Iwuanyanwu**-University of the Western Cape, S.Africa

**Sinan Erten**, Hacettepe University, Turkiye

**Steven Sexton**-College of Education,University of Otago,New Zealand

**V. Ferreira Pinto**, Universidade Estadual do Norte Fluminense Darcy

Ribeiro (UENF), Brazil

**Zalpa Ayoubi**- Lebanese University, Lebanon

**William W. COBERN** - Western Michigan University, U.S.A.

**Ilkka Ratinen**, University of Jyväskylä, Finland

**Iwona Bodys-Cupak**-Jagiellonian University, Poland

**Kamisah Osman**- National University of Malaysia, Malaysia

**Luecha Ladachart**- University of Phayao, Thailand

**Mustafa Sami Topcu**, Yildiz Technical University, Turkiye

**Patrice Potvin**- Université du Québec à Montréal, Canada

**Sandra Abegglen**- London Metropolitan University, England

**Sofie Gärdebjer**, Chalmers University of Technology, Sweden

**Tammy R. McKeown**- Virginia Commonwealth University, U.S.A.

**Wan Ng**- University of Technology Sydney, Australia

**Ying-Chih Chen**, Arizona State University, United States

## Journal of Education in Science, Environment and Health (JESEH)

The Journal of Education in Science, Environment and Health (JESEH) is a peer-reviewed and online free journal. The JESEH is published quarterly in January, April, July and October. The language of the journal is English only. As an open access journal, Journal of Education in Science, Environment and Health (JESEH) does not charge article submission or processing fees. JESEH is a non-profit journal and publication is completely free of charge.

The JESEH welcomes any research papers on education in science, environment and health using techniques from and applications in any technical knowledge domain: original theoretical works, literature reviews, research reports, social issues, psychological issues, curricula, learning environments, book reviews, and review articles. The articles should be original, unpublished, and not in consideration for publication elsewhere at the time of submission to the JESEH.

### Abstracting/ Indexing

Journal of Education in Science, Environment and Health (JESEH) is indexed by following abstracting and indexing services: Education Resources Information Center (ERIC), DOAJ, Wilson Education Index

### Submissions

All submissions should be in electronic (.Doc or .Docx) format. Submissions in PDF and other non-editable formats are not acceptable. Manuscripts can be submitted through the journal website. All manuscripts should use the latest APA style. The manuscript template for formatting is available on the journal website.

### Contact Info

Journal of Education in Science, Environment and Health (JESEH)

Email: [jesehoffice@gmail.com](mailto:jesehoffice@gmail.com)

Web : [www.jeseh.net](http://www.jeseh.net)

**CONTENTS****AI-Professional Development Model for Chemistry Teacher: Artificial Intelligence in Chemistry Education .....161-182***Bekir Yildirim, Ahmet Tayfur Akcan***Bibliometric Analysis of Studies on the Artificial Intelligence in Science Education with VOSviewer .....183-195***Hayriye Nevin Genc, Nuriye Kocak***Using Artificial Intelligence to Analyze Nature of Science Themes: Einstein and Eddington Documentary Film Example .....196-207***Yusuf Kartal, Munise Seckin-Kapucu***Investigation of STEM Supported Engineering and Design Experiences of Students Participating in Design and Innovation Workshop in the Context of Situational Interest Theory ..... 208-223***Ibrahim Benek, Ismail Donmez, Sahin Idil***Use of Multi-Tier Concept Diagnostic Tests in Biology Education: A Systematic Review of the Literature ..... 224-244***Tugce Duran, Musa Dikmenli***Pharmacy Undergraduates' Epistemic Cognition and Medication Beliefs as Predictors of Conceptual Learning and Academic Progress .....245-256***Ilona Sodervik, Leena Hanski, Eero Laakkonen, Nina Katajavuori*

# AI-Professional Development Model for Chemistry Teacher: Artificial Intelligence in Chemistry Education

Bekir Yildirim, Ahmet Tayfur Akcan

## Article Info

### Article History

Published:  
01 October 2024

Received:  
08 March 2024

Accepted:  
21 August 2024

### Keywords

Artificial intelligence  
Chemistry education  
Teacher Education  
ChatGPT  
Gemini

## Abstract

This study aimed to propose a Professional Development Model (PDM) for chemistry teachers to enhance their professional development in Artificial Intelligence (AI). The research group consisted of 17 chemistry teachers. The study was designed using a particular case study suitable for qualitative research methods. Document review, teacher interviews, and AI opinions were utilized to create the model. Data were analyzed using inductive content analysis. The document analysis emphasized the teachers' knowledge of various topics, such as AI knowledge, AI tools, AI skills, AI ethics, AI attitudes, and AI literacy, to enable them to incorporate AI into their lessons. It was also highlighted that teachers should acquire domain-specific knowledge, skills, and competencies in the areas where artificial intelligence will be integrated. When examining the recommendations of artificial intelligence (ChatGPT and Gemini), it was found that they addressed similar content to the information included in the document analysis. Furthermore, chemistry teachers stated their deficiencies in AI literacy, AI competencies, and developing AI lesson plans. They also stated that AI applications could be included in various subjects such as organic chemistry, chemical experiments, and chemical reactions. Following the analysis of documents and teacher and AI opinions, a 10-step PDM has been proposed to enhance chemistry teachers' professional development in AI.

## Introduction

Scientific and technological advancements impact many areas of our lives, from health to the economy (Yoo & Yi, 2022). One of the areas experiencing change is education. Countries are integrating technology into their education systems (Lu et al., 2022). Technology enriches education, making learning more effective and lasting. Enriching educational environments with technological advancements has also paved the way for the effective implementation of the education process (Yildirim, 2023). Indeed, many countries have utilized various technological advancements such as augmented reality, machine learning, and virtual reality to enrich education. One of the technological advancements is the AI. AI systems, created by humans to fulfill specific tasks and constantly improve themselves with the knowledge they acquire, mimic humans (Kuleto et al., 2022).

With the emergence of AI tools, the expectations of countries from individuals have also changed. The most critical impact of this change started with ChatGPT and Gemini (Arman & Lamiyar, 2023). With the introduction of artificial intelligence into our lives, many artificial intelligence tools have entered our lives rapidly. ChatGPT is one of the AIs used in this process. Along with ChatGPT, different artificial intelligence tools, such as Gemini, have also been developed and started to be used (Rane et al., 2024). With these changes, countries have started integrating AI into educational environments to raise AI-literate individuals (Yau et al., 2023). Similarly, many institutions and researchers have conducted and continue to conduct studies on integrating AI into educational environments. As AI is a relatively new concept and has yet to be fully understood, there is no consensus on how AI should be integrated into the educational process (Sanusi et al., 2022). Although there is no consensus on the full integration of AI, there are efforts to integrate AI into the educational process. For example, the governments of China and the United States have initiated projects to integrate AI into school curricula by 2028 (Chiu & Chai, 2020). Similarly, Canada and South Korea are investing in AI for K-12 students (Asia Pacific Foundation of Canada Korea AI Ethics Association, 2021). These projects of the countries aim to raise AI literate individuals. AI literacy means that an individual knows AI, follows developments in AI, has AI skills, and applies AI-related knowledge in their lives. Therefore, countries emphasize AI literacy in students, necessitating its integration into education.

Teachers play a crucial role in education and are primarily responsible for integrating and implementing AI (Luckin et al., 2022). Teachers have primary responsibility for students' learning outcomes as they play a vital

role in motivating and preparing students for education (Gökdere & Çepni, 2003). Similarly, teachers are crucial in developing students' AI literacy (Casal-Otero et al., 2023). The vital role of teachers in the educational process underlines their critical role in implementing AI in classrooms (Long & Magerko, 2020). For example, South Korea has launched efforts to train 5000 teachers in AI literacy by 2024.

The high-quality professional development of teachers in AI is directly linked to students' achievement of high academic standards in AI. For instance, teachers' science and mathematics proficiency correlates with students' achievements in these fields (Hibpsman, 2007). Similarly, teachers' competencies in AI are directly related to students' achievement of desired qualifications in AI. Moreover, a teacher must know mathematical concepts and rules to teach mathematics effectively (Eckman et al., 2016). Teachers must also have sufficient knowledge and expertise in AI to effectively teach AI in their classrooms (Salas-Pilco et al., 2022). Therefore, there is a need for specially designed, robust pedagogical training in AI to ensure that teachers have the necessary knowledge and skills in AI because studies emphasize that teachers need more knowledge about AI (Lindner & Romeike, 2019).

Teachers' insufficient knowledge of AI leads to low performance in implementing AI in their classrooms (Park et al., 2023). Therefore, teachers must have sufficient knowledge of AI to effectively implement AI practices in their classrooms (Chounta et al., 2022). This is because teachers' mastery of the subjects they teach directly affects their self-efficacy in the teaching and learning process (Stohlmann et al., 2012). Self-efficacy refers to individuals' beliefs about what they can or cannot do in situations based on their knowledge, skills, and abilities (Bandura, 1997).

Self-efficacy beliefs are one of the significant factors influencing teachers' behaviors and activities in the classroom, classroom management, and instructional decisions, affecting the effectiveness and efficiency of teaching (Serin & Bayraktar, 2015). Self-efficacy belief motivates individuals to take action and guides them to achieve desired outcomes (Bandura, 1986). Therefore, the efforts exerted by teachers during the teaching-learning process significantly impact their beliefs about what they can achieve in the classroom (Bandura, 1997). Thus, teachers' low self-efficacy in AI can hinder effective teaching (Çelik, 2023). Indeed, the lack of sufficient AI knowledge among teachers can also lead to negative attitudes toward implementing AI in their classrooms (Park et al., 2023). Hence, teachers' knowledge, skills, and self-confidence are directly related to creating an effective learning environment for AI (Bandura, 1993). Therefore, it is necessary to develop high-quality AI PD programs for teachers to implement AI in classrooms (Park et al., 2023).

When the literature is examined, there needs to be more topics covered in AI PD programs for teachers and how to integrate these topics into their fields. In addition, more AI program units are required for classroom practice (Chiu et al., 2021). Moreover, few studies emphasize how to conduct subject-specific training processes on AI for teachers (Muslim al-Zyoud, 2020). Therefore, organizing AI training to meet the needs of teachers will facilitate the integration of AI into the educational process. Thus, in this study, a PDM for chemistry teachers was proposed. This is because, when the literature is examined, it is understood that there are insufficient PD programs for chemistry teachers related to AI. Despite insufficient programs, studies show the connection between AI and chemistry education and its use in chemistry teaching (Emenike & Emenike, 2023; Clark, 2023). Moreover, AI applications are accepted and used as a powerful tool in chemistry education (Choudhary et al., 2022). The fact that artificial intelligence is a powerful tool facilitates its application in chemistry (Owczarek, 2022).

The dominant role of AI in chemistry research and education (Choudhary et al., 2022) facilitates its connection to chemistry education topics. For example, AI helps to solve general chemistry problems (Daher et al., 2023). In addition, AI tools are used for visualizing molecules in chemistry education, conducting measurement and evaluation processes for chemistry courses, and preparing chemistry lesson plans (Clark, 2023; Emenike & Emenike, 2023; Joshi & Kumar, 2021). In addition, AI is used in various application processes such as molecular property prediction, drug discovery, nanotechnology, wastewater treatment, retrosynthesis, and reaction results in chemistry education (Clark, 2023).

In chemistry education, AI can be an assessment tool to provide feedback on students' achievements and make inferences about students (Emenike & Emenike, 2023). Although there are studies on using AI in chemistry education, its integration into chemistry education is not at the desired level (Deng et al., 2023). Therefore, there are calls to increase the use of AI in chemistry education due to the limited studies of AI in chemistry education (Yuriev et al., 2023). In addition, one of the reasons for the limited integration of AI in chemistry education is the insufficient knowledge of chemistry teachers in this field. Therefore, it is essential to provide chemistry teachers professional development on AI.

In this study, a PDM is proposed to enhance the professional development of chemistry teachers in AI. This model aims to provide professional development opportunities tailored explicitly to chemistry teachers in the field of AI. In this context, the research problem was defined as follows: “How should a professional development program be designed to enhance the professional development of chemistry teachers?”. This study sought answers to the following subquestions:

RQ1: Which AI-related content has been mentioned in artificial intelligence studies on teachers?

RQ2: What are chemistry teachers' views on AI?

RQ3: How does artificial intelligence propose a model for teachers' PD related to AI?

RQ4: What are the opinions of experts and artificial intelligence about the model developed for teacher training?

## Method

### Research Design

A PDM was proposed to enhance chemistry teachers' PD in AI within the study's scope. To achieve this objective, the study was designed using a unique case study approach suitable for qualitative research methods. In case study research, various qualitative data collection tools are used simultaneously. This approach aims to provide a detailed and comprehensive understanding of a situation within its real-life context (Yin, 2017). Additionally, diversifying the data collection methods ensures that qualitative data mutually reinforce each other.

### Study Group

Participants were recruited using purposive convenience sampling, a non-probability sampling method. This approach is advantageous as it enables researchers to focus their resources on specific subgroups or segments of the population that are of particular interest, making it both time- and cost-effective (Balci, 2016). Accordingly, the study group consisted of 17 chemistry teachers. Table 1 shows their demographic characteristics.

Table 1. Demographic characteristics

Theme	Category	Codes	f
Sociodemographic characteristics	Gender	Woman	13
		Man	4
	Work experience (year)	1-10	10
		11-18	4
		≥ 21	3
	School type	Public	14
		Private	3
	Branch	Chemistry teachers	113

### Model Development Process

The AI PDM for chemistry teachers was developed according to the ADDIE instructional design model (Dick & Carey, 1996). The ADDIE design model consists of five stages: Analysis, design, development, implementation, and evaluation. This model emerges as an instructional design model suitable for all types of instruction due to its essential nature, its inclusion of the components of other instructional design models, and its coverage of the basic definitions of instruction. For this reason, artificial intelligence professional development model for chemistry teachers was preferred. The development process of the AI PDM for Chemistry Teachers is detailed below.

### Analysis Phase

In the AI PDM developed for chemistry teachers, (1) needs analysis, (2) participant analysis, (3) identification of the target group, (4) problem definition, and (5) task and instruction analysis stages were used. When the literature is examined, the lack of a study that emphasizes the content required for a professional development program aiming to develop chemistry teachers' artificial intelligence skills and the process of delivering such a

program reveals the importance of this study. Accordingly, a needs analysis was conducted to facilitate the development of chemistry teachers' artificial intelligence skills. Following the needs analysis, the research started with the question, "How should AI education for chemistry teachers be conducted?". To address this question, the instructional analysis process was initiated. In the process of instructional analysis, emphasis was placed on determining the content that should be taught to facilitate the PD of chemistry teachers on AI. Within the scope of the study, all these processes were integrated with a comprehensive perspective by combining the information obtained from the literature, teacher opinions, and inquiries about AI.

### **Design Phase**

The teaching process for chemistry teachers was defined in the design phase of the AI PDM. Literature documents were analyzed to formulate the teaching process. For the documents used at this stage, the following stages were taken into consideration: (1) accessing the documents, (2) confirming the authenticity of the documents (these documents were analyzed both from the journals in which they were published and from the indexed pages), (3) understanding the documents, (4) analyzing the data, and (5) using the data (Yıldırım & Şimşek, 2011). After the literature review, AI's articles on teacher education were included in the process. Verification of the authenticity of the studies used in determining the teaching stages and content of AI is essential for the validity of the model. Therefore, the documents used in the study were analyzed through Web of Science, Scopus, Google Scholar, and Eric databases. The authenticity of the articles obtained through research was confirmed. Then, after the authenticity of the documents was confirmed, the stages and content were determined.

### **Development Phase**

Following the document reviews, the emerged model underwent further refinement based on teacher feedback and the insights of AI tools. During this phase, the design was adjusted according to the feedback received from teachers, experts, and AI tools.

### **Implementation Phase**

During this phase, a pilot study of the educational program designed for chemistry teachers was conducted with three teachers. This enabled testing the effectiveness of the proposed model.

### **Evaluation Phase**

The final version of the developed model was evaluated after the pilot implementation in this phase.

### **Data Collection Tools**

#### *AI Interview Guide*

This study utilized a semi-structured interview protocol to elicit chemistry teachers' perspectives on AI. The interview protocol was developed following a literature review, during which a pool of questions was generated. After creating the question pool, the interview protocol was evaluated by two different experts. One of the experts had publications on science education, while the other had publications on AI education. Based on the experts' feedback, adjustments were made to the interview protocol. Following these adjustments, a pilot study was conducted, and the final version of the interview protocol was established. As a result of these stages, the interview protocol comprised ten questions (See. Appendix).

#### *Documents*

The documents used in the study were investigated through the Web of Science, Scopus, Google Scholar, and Eric databases. The words "artificial intelligence," "teacher education," "PD," and "model" were searched for studies. In this direction, 19 studies were reached. Artificial intelligence studies conducted with teachers

between 2019 and 2024 were analyzed. As a result of the research, the journals in which the accessed articles were published and their originality were confirmed. Following the confirmation of the documents, the instructional stages and content of the model were examined and determined.

### Data Analysis

The qualitative data obtained were analyzed using the inductive content analysis method. The data were collected through semi-structured interviews and documents. The analysis and interpretation of the data were conducted in four stages. These stages involved processing the data from documents, coding the data, organizing the codes, identifying the findings, and interpreting them (Yıldırım & Şimşek, 2012). Initially, two experts conducted separate analyses of the transcripts, independently identifying and elaborating on themes and codes. Following this, the themes and codes were structured and presented by the research inquiries. To guarantee coder consistency, two distinct professionals conducted individual coding sessions. The interrater reliability stood at 83%, signifying the reliability of the coding process. Then, the originality of the documents obtained was examined and confirmed. Similarly, two separate experts were involved in document reviews and conducted the coding. Along with coding qualitative data, the coding reliability was 82%. In determining the reliability value, changes were made to some codes. For example, 'Lesson plans' was changed to 'AI-lesson plans. Similarly, pedagogical knowledge was considered a province in 'AI pedagogical knowledge.' Subsequently, this code was treated separately as 'AI pedagogical knowledge and pedagogical' knowledge. Another example is related to the 'need' code. The 'needs' code was changed to 'Identification of needs. On the other hand, the study did not analyze Cohen's kappa value, because while calculating the Cohen Kappa value, not all of the codes are considered. For this reason, the reliability of all the codes was examined to reveal all of the codes in the article.

### Reliability and Validity

Experts developed the codes and themes to minimize researcher bias and bolster internal validity. For data reliability, codes were devised based on input from an academic with publications in science education and one expert working in AI. These experts provided insights on the themes and codes, modifying non-common ones to avoid subjective judgments. Subsequently, two experts with publications in science education and AI reviewed and finalized the codes and themes. Interrater reliability was calculated using the formula  $[\text{Reliability} = (113) / (113+25)*100]$  suggested by Miles et al. (2014). Qualitative and quantitative approaches were employed to diversify data sources, significantly enhancing the validity and reliability of research findings. Finally, the relationships between the themes were identified, findings were defined and tabulated, and quotations about the codes were included to ensure internal reliability and eliminate bias. A similar process was conducted for the analysis of document reviews.

### Findings

Themes, categories, and codes were presented in tables related to the research questions. Direct quotations were used to provide an accurate and coherent representation of participants' views and to facilitate readers' analysis and interpretation of the findings.

*RQ1: Which AI-related content has been mentioned in AI studies on teachers?*

In the first research question, studies on AI-related documents were examined, and the codes and contents obtained from the analyzed studies are presented in Table 2.

Table 2. Contents topics related to AI

No	Studies	Study	Title	Content Related to AI and Supporting Elements
1.	Nyaaba & Zhai, (2024)	Generative AI professional development needs for teacher education	AI knowledge AI tools AI ethical AI skills	
2.	Nazaretsky et. al.	Teachers' trust in AI-powered	AI knowledge	

	(2022)	educational technology and a Professional development program to improve it	AI tools AI attitude PK
3.	Çelik (2023)	Towards intelligence -TPACK: An empirical study on teachers' Professional knowledge to ethically integrate artificial intelligence (AI) based tools into education	AI technological knowledge AI pedagogical knowledge AI knowledge AI tools AI ethical PCK TPCAK
4.	Kuleto et. al. (2022)	The Attitudes of K-12 Schools' Teachers in Serbia towards the potential of Artificial Intelligence	AI knowledge AI tools AI lesson plans
5.	Lee et. al. (2024)	Artificial Intelligence (AI) Education for teaching in schools and their AI teacher training programs	AI skills AI tools AI knowledge AI pedagogical knowledge Practical knowledge AI attitude AI lesson plans
6.	Lindner & Romeike (2019)	Teachers' perspectives on artificial intelligence	AI knowledge AI attitude AI tools AI ethical Special content knowledge TPACK PK Practice examples, lack of adequate teaching material
7.	Chiu & Chai (2020)	Sustainable curriculum planning for artificial intelligence education: a self-determination theory perspective	AI-lesson plans AI knowledge AI tools AI ethical AI skills AI attitude Special content knowledge Pedagogical knowledge Limited resources Guided teacher Guide learned Teacher needs
8.	Sala-Pilco et. al. (2022)	Artificial intelligence and learning analytics in teacher education: A systematic review	AI ethical AI tools AI knowledge AI skills Special content knowledge Pedagogical knowledge Technical supports
9.	Zhou (2020)	Analysis of the professional development of teachers in the era of Artificial intelligence	AI skills AI knowledge AI Literacy Technological Knowledge
10.	Wardant et al. (2024)	Artificial intelligence in education: Mathematics teachers' perspectives, practices, and challenges	AI skills AI knowledge AI ethical AI pedagogical knowledge Pedagogical knowledge Special content knowledge Technical support

11.	Su et al. (2023)	Artificial intelligence (AI) literacy in early childhood education: the challenges and opportunities	AI-lesson plan AI literacy AI Attitude AI skills AI tools AI ethical Pedagogical knowledge TPACK Teacher/learning guided.
12.	Mondal et. al. (2023)	ChatGPT for Teachers: Practical examples for utilizing artificial intelligence for educational purposes	AI lesson plan PK AI knowledge Technical supports Resources
13.	Lin et. al. (2022)	Teachers' perceptions of teaching sustainable artificial intelligence: A design frame perspective	AI knowledge AI lesson plan AI practices AI ethical Pedagogical knowledge Special content knowledge Special context knowledge TPACK AI resources
14.	Yang (2022)	Artificial intelligence education for young children: Why, what, and how in curriculum design and implementation	AI literacy AI ethical AI attitude AI knowledge AI tools Learning goals Learning evaluation Teacher guided AI related activities AI activities
15.	Kitcharoen et al. (2024)	Enhancing teachers' AI competencies through Artificial Intelligence of things professional development training	AI literacy AI knowledge AI tools AI skills AI attitude Technological knowledge
16.	Kim & Kim (2022)	Teachers' perceptions of using an artificial intelligence based educational tool for scientific writing	AI attitude AI knowledge AI pedagogical knowledge AI tools
17.	Çelik et. al. (2022)	The promises and challenges of artificial intelligence for teachers: A Systematic review of research	AI skills AI knowledge AI attitude AI tools AI lesson plan AI ethical Pedagogical knowledge Special content knowledge Identification of needs Teacher needs Technical supports Teaching materials Using AI tools guided
18.	Kim et al. (2021)	Why and what to teach: AI curriculum for elementary school	AI knowledge AI skills AI attitude

		AI literacy Pedagogical knowledge AI lesson plan
		Guided teacher
19.	Park et. al. (2023)	Integrating artificial intelligence into science lessons: teachers' experiences and views
		AI knowledge AI skills AI literacy AI-lesson plans AI technological knowledge AI pedagogical knowledge AI tools pedagogical knowledge Special content knowledge Teaching materials Resources activities

The results showed that AI studies on teachers focus on various topics such as AI knowledge, AI attitude, AI skills, AI ethics, AI tools, AI literacy, AI lesson plans, TPACK, PCK, domain-specific knowledge, and pedagogical knowledge. In addition, the documents highlight the need for various supporting elements, including teacher and student guides, AI tool usage guides, expert support, technical assistance, and others. These supporting elements are essential to make the educational process more effective.

#### *RQ2: What are chemistry teachers' views on AI?*

The first research question of the second subproblem addressed chemistry teachers' views on AI competencies (See Table 3). The results showed that most chemistry teachers need to be more adequately prepared regarding their AI competencies.

Table 3. Views on the AI competencies

Theme	Code
AI competencies	No (n=16) Yes (n=1)

The second research question of the second subproblem focused on chemistry teachers' perspectives regarding areas of deficiency in their AI competencies (See Table 4). The results showed that teachers mainly express feelings of inadequacy in their knowledge of AI, AI tools, and AI applications.

Table 4. Teachers' views on the areas of deficiency regarding AI competencies

Theme	Code
The areas of deficiency regarding AI competencies	AI knowledge (n=7) AI tools (n=5) AI applications (n=5) AI lesson plans (n=4) AI ethics (n=3) AI curriculum (n=3) Integration of AI into the lesson (n=2)

\* Participants have provided multiple responses

The third research question of the second subproblem addressed participants' views of AI literacy. Table 5 shows the results. The results indicate that most chemistry teachers feel inadequate in AI literacy. On the other hand, another teacher feels competent. The fourth research question of the second subproblem addressed participants' views on developing AI-based chemistry lesson plans. Table 6 shows the results.

Table 5. Views on AI literacy

Theme	Code
AI literacy	No (n=16) Yes (n=1)

Table 6. Views on developing AI-based chemistry lesson plans

Theme	Code
Developing AI-based chemistry lesson plans	No (n=14) Yes (n=3)

The results showed that most teachers stated they did not feel adequate in preparing AI-based chemistry lesson plans. When asked about the reasons for their inability to prepare an AI-based chemistry lesson plan, they provided similar responses as those in Table 4. The fifth research question of the second subproblem addressed participants' concerns about the ethical dimension of using AI in chemistry classes. Table 7 shows the results.

Table 7. Views on concerns about AI's ethical

Theme	Code
AI Ethical	Yes (n=15) No (n=2)

The results showed that most chemistry teachers believe using AI in chemistry classes may raise ethical concerns. When asked about these ethical concerns, chemistry teachers provided their perspectives. The sixth research question of the second subproblem addressed chemistry teachers' views on the ethical status of using AI in chemistry classes (See Table 8).

Table 8. Views on the ethical status of using AI

Theme	Code
The ethical status of using AI	Assignment/Research tasks (n=6) Loss of personal data (n=4) Lack of privacy (n=2) Using others' information (n=2) Misleading information (n=1)

The results showed that chemistry teachers have emphasized that using AI in chemistry classes may pose ethical issues such as assigning homework/research, losing personal data, and lacking privacy. The seventh research question of the second subproblem addressed participants' views of the benefits of using AI in chemistry classes for students. Table 9 shows the results. The findings show that most chemistry teachers believe integrating artificial intelligence into chemistry courses benefits students.

Table 9. Participants' views of the benefits of using AI in chemistry classes for students

Theme	Code
Benefits of using AI in chemistry classes for students	Yes (n=16) Not sure (n=1)

Participants who found AI useful in chemistry classes for students were asked what kind of benefits they thought AI had. Table 10 shows the results. The results showed that chemistry teachers believe integrating AI into lessons will enhance students' AI literacy, facilitate lasting learning, and encourage individualized learning.

Table 10. Participants' views of the benefits of using AI for students

Theme	Code
Benefits of using AI	AI literacy (n=10) Permanent learning (n=6) Individualized learning (n=5) Effective learning (n=4) AI skills (n=3) Creating their content (n=2) Feedback about lesson (n=1)

\* Participants have provided multiple responses

The ninth research question of the second subproblem addressed the opinions of chemistry teachers on the connections between AI and chemistry topics (See Table 11).

Table 11. Views on the connection between AI and chemistry topics

Theme	Code
The connection between AI and Chemistry topics	Organic chemistry (n=6)
	Chemical reactions (n=5)
	Molecules and atoms (n=5)
	Chemistry experiments (n=4)
	Properties of matter (n=3)
	Periodic systems (n=2)
	Nanotechnology (n=1)

\* Participants have provided multiple responses

The results show that teachers highlighted the connection between AI and chemistry, especially in organic chemistry, chemical reactions, molecules, and atoms. The tenth research question of the second subproblem addressed chemistry teachers' opinions on how AI could be customized for chemistry lessons (See Table 12). The results showed that chemistry teachers indicated that AI was used to customize chemistry concepts in various ways, such as visualizing, solving chemistry problems, analyzing chemical reactions, and conducting chemistry experiments.

Table 12. Views on customization of AI for chemistry lessons

Theme	Code
Customization of AI for chemistry lessons	Visualization of chemistry topics (n=12)
	Solving chemistry problems (n=9)
	Analyzing chemical reactions (n=9)
	Conducting chemistry experiments (n=8)
	Measurement and evaluation (n=5)
	Personalization of chemistry lessons (n=2)
	Predicting students' future in chemistry courses (n=1)
	Preparing appropriate lesson plans for chemistry lessons (n=1)

### RQ3: How does artificial intelligence propose a teacher PD model related to AI?

For this research question, AI was asked two questions. In the first question, artificial intelligences were asked to suggest a professional development program for teachers. The professional development program proposed by the artificial intelligence and its content are presented in Figures 1, 2, 3, and 4.

<p>Title: Professional Development Program Model for Artificial Intelligence Education for Teachers</p> <p>Introduction: As artificial intelligence (AI) continues to permeate various aspects of society, integrating AI education into the curriculum becomes imperative to prepare students for the future workforce. Equipping teachers with the knowledge and skills to teach AI concepts effectively is essential. This document outlines a model for a professional development program tailored to teachers aiming to incorporate AI education into their teaching practices.</p> <p>Program Goals:</p> <ol style="list-style-type: none"> <li>1. Familiarize teachers with fundamental concepts and applications of artificial intelligence.</li> <li>2. Provide hands-on experience with AI tools, platforms, and programming languages.</li> <li>3. Enable teachers to integrate AI education across different subject areas and grade levels.</li> <li>4. Foster critical thinking, problem-solving, and ethical considerations related to AI.</li> <li>5. Support teachers in designing engaging AI-related activities and projects for their students.</li> </ol> <p>Program Components:</p> <ol style="list-style-type: none"> <li>1. Foundations of AI: <ul style="list-style-type: none"> <li>• Introduction to key concepts, terminology, and principles of artificial intelligence.</li> <li>• Exploration of AI applications in various fields such as healthcare, finance, education, and entertainment.</li> <li>• Discussion of ethical considerations, bias, and societal implications of AI technologies.</li> </ul> </li> <li>2. Hands-on Workshops and Training: <ul style="list-style-type: none"> <li>• Practical sessions on using AI tools, platforms, and programming languages (e.g., TensorFlow, Python, AI platforms like Google AIY).</li> <li>• Guided tutorials and projects to help teachers gain proficiency in AI-related skills.</li> <li>• Collaborative problem-solving activities and hackathons to apply AI concepts in real-world scenarios.</li> </ul> </li> <li>3. Curriculum Integration: <ul style="list-style-type: none"> <li>• Strategies for integrating AI education into existing curriculum frameworks and lesson plans.</li> <li>• Development of interdisciplinary AI projects that connect with subjects like mathematics, science, social studies, and language arts.</li> <li>• Exploration of resources, educational apps, and online platforms for teaching AI across different grade levels.</li> </ul> </li> <li>4. Pedagogical Approaches: <ul style="list-style-type: none"> <li>• Instructional methods and best practices for teaching AI concepts effectively, including inquiry-based learning, project-based learning, and flipped classrooms.</li> <li>• Differentiation strategies to meet the diverse learning needs of students when teaching AI.</li> <li>• Assessment techniques for evaluating students' understanding of AI concepts and skills.</li> </ul> </li> <li>5. Collaboration and Networking: <ul style="list-style-type: none"> <li>• Opportunities for teachers to collaborate with AI experts, industry professionals, and researchers.</li> <li>• Participation in online communities, webinars, and workshops focused on AI education.</li> <li>• Establishing local or regional networks of educators interested in advancing AI education in schools.</li> </ul> </li> <li>6. Ongoing Support and Resources: <ul style="list-style-type: none"> <li>• Provision of ongoing support through mentorship, coaching, and online forums.</li> <li>• Access to curated resources, lesson plans, and teaching materials for AI education.</li> <li>• Professional development grants or funding opportunities to support teachers in implementing AI initiatives in their classrooms.</li> </ul> </li> <li>7. Evaluation and Feedback: <ul style="list-style-type: none"> <li>• Regular assessment of teachers' progress and proficiency in AI concepts and skills.</li> <li>• Feedback mechanisms to gather input from teachers, students, and stakeholders on the effectiveness of the program.</li> <li>• Continuous improvement based on evaluation findings to enhance the relevance and impact of the professional development program.</li> </ul> </li> </ol> <p>Conclusion: Integrating AI education into schools requires a concerted effort to equip teachers with the necessary knowledge, skills, and resources. By implementing the professional development program model outlined above, educational institutions can empower teachers to effectively teach AI concepts and prepare students for the opportunities and challenges of the AI-driven world. (Ctrl) AI concepts and prepare</p>	
--	--

Figure 1. PDM proposed by ChatGPT for teachers

<p><b>AI for Teachers Professional Development Program Model</b></p> <p><b>Program Goals</b></p> <ul style="list-style-type: none"> <li>To help teachers understand the basics of AI and its potential impact on education.</li> <li>To provide teachers with the skills and knowledge they need to use AI effectively in the classroom.</li> <li>To create a community of teachers who are committed to using AI to improve student learning.</li> </ul> <p><b>Program Structure</b></p> <p>The program will be delivered in a blended format, with online modules and in-person workshops. The online modules will provide teachers with the foundational knowledge they need to understand AI. The in-person workshops will provide teachers with the opportunity to learn how to use AI in the classroom and to collaborate with other teachers.</p> <p><b>Program Content</b></p> <p>The program will cover the following topics:</p> <ul style="list-style-type: none"> <li>What is AI?</li> <li>The different types of AI</li> <li>The potential impact of AI on education</li> <li>How to use AI in the classroom</li> <li>Ethical considerations for using AI in education.</li> </ul> <p><b>Program Assessment</b></p> <p>Teachers will be assessed on their understanding of the content and their ability to use AI in the classroom. Assessment will be conducted through a combination of online quizzes, in-person workshop activities, and a final project.</p> <p><b>Program Dissemination</b></p> <p>The program will be disseminated through a variety of channels, including:</p> <ul style="list-style-type: none"> <li>A website</li> <li>A social media campaign</li> <li>A series of webinars</li> <li>A national conference</li> </ul> <p><b>Program Sustainability</b></p> <p>The program will be sustained through a combination of funding sources, including:</p> <ul style="list-style-type: none"> <li>Government grants</li> <li>Corporate sponsorships</li> <li>Private donations</li> </ul> <p><b>Program Impact</b></p> <p>The program is expected to have a positive impact on teachers, students, and the education system as a whole. Teachers will gain the skills and knowledge they need to use AI effectively in the classroom. Students will benefit from learning with AI-powered tools and resources. The education system will become more innovative and efficient as a result of the program.</p> <p><b>Conclusion</b></p> <p>The AI for Teachers Professional Development Program Model is a comprehensive and innovative program that will help teachers understand and use AI to improve student learning. The program is designed to be accessible to teachers of all levels of experience and expertise. The program is also designed to be sustainable, so that it can continue to provide benefits to teachers and students for years to come.</p> <p><b>Here are some additional resources that you may find helpful:</b></p> <ul style="list-style-type: none"> <li>The AI Index: <a href="https://aiindex.org/">https://aiindex.org/</a></li> <li>The Partnership on AI: <a href="https://www.partnershiponai.org/">https://www.partnershiponai.org/</a></li> <li>The AI for Education Initiative:</li> </ul>
---

Figure 2. PDM Proposed by Gemini for teachers

When the PD programs related to artificial intelligence suggested by AIs are examined, it is understood that they propose similar models. However, the model offered by ChatGPT is more detailed than Gemini's. The topics AI tools express are identical to those found in the literature. ChatGPT and Gemini were also asked to suggest an AI PDM for chemistry teachers, as shown in Figures 3 and 4.

<p><b>Title: Professional Development Program Model for Artificial Intelligence Education for Chemistry Teachers</b></p> <p><b>Introduction:</b> As artificial intelligence (AI) continues to revolutionize various fields, including chemistry, it becomes essential for chemistry teachers to stay updated with AI applications relevant to their subject. Integrating AI education into chemistry classrooms can enhance teaching methodologies, facilitate research, and deepen students' understanding of chemical concepts. This document presents a model for a professional development program specifically designed for chemistry teachers interested in integrating AI education into their teaching practices.</p> <p><b>Program Goals:</b></p> <ol style="list-style-type: none"> <li>Equip chemistry teachers with a foundational understanding of artificial intelligence concepts and their applications in chemistry.</li> <li>Provide hands-on experience with AI tools and algorithms relevant to chemical data analysis, modeling, and simulations.</li> <li>Enable teachers to develop AI-enhanced learning experiences and experiments tailored to chemistry curriculum objectives.</li> <li>Foster collaboration and knowledge-sharing among chemistry educators interested in AI integration.</li> <li>Promote ethical considerations and responsible use of AI technologies in chemistry education.</li> </ol> <p><b>Program Components:</b></p> <ol style="list-style-type: none"> <li><b>Fundamentals of AI in Chemistry:</b> <ul style="list-style-type: none"> <li>Overview of artificial intelligence concepts, machine learning algorithms, and data analysis techniques relevant to chemistry.</li> <li>Exploration of AI applications in chemical research, drug discovery, materials science, and computational chemistry.</li> <li>Discussion of ethical considerations, data privacy, and bias in AI-driven chemical research.</li> </ul> </li> <li><b>AI Tools and Techniques for Chemical Data Analysis:</b> <ul style="list-style-type: none"> <li>Hands-on training in using AI tools and software packages for chemical data preprocessing, feature selection, and predictive modeling.</li> <li>Practical sessions on machine learning algorithms commonly used in chemistry, such as neural networks, decision trees, and support vector machines.</li> <li>Case studies and real-world examples demonstrating the application of AI techniques in chemical data analysis and interpretation.</li> </ul> </li> <li><b>Integration of AI into Chemistry Curriculum:</b> <ul style="list-style-type: none"> <li>Strategies for integrating AI concepts and tools into chemistry lesson plans, laboratory experiments, and research projects.</li> <li>Development of AI-enhanced learning activities, simulations, and virtual experiments that align with chemistry curriculum standards.</li> <li>Collaborative planning and sharing of AI-integrated teaching resources among participants.</li> </ul> </li> <li><b>Designing AI-Driven Chemistry Experiments:</b> <ul style="list-style-type: none"> <li>Guidance on designing and conducting chemistry experiments enhanced by AI technologies, such as automated data collection and analysis systems.</li> <li>Exploration of AI-driven approaches to chemical synthesis, reaction optimization, and materials discovery.</li> <li>Incorporation of AI-based modeling and simulation tools for predicting chemical properties, reactions, and molecular structures.</li> </ul> </li> <li><b>Community Building and Networking:</b> <ul style="list-style-type: none"> <li>Opportunities for chemistry teachers to connect with AI experts, computational chemists, and industry professionals.</li> <li>Establishment of online forums, discussion groups, and professional networks focused on AI in chemistry education.</li> <li>Collaboration on research projects, curriculum development initiatives, and grant opportunities related to AI integration.</li> </ul> </li> <li><b>Ethical Considerations and Best Practices:</b> <ul style="list-style-type: none"> <li>Discussion of ethical guidelines, responsible conduct of research, and fair use of AI technologies in chemistry education.</li> <li>Exploration of case studies and scenarios highlighting ethical dilemmas and considerations in AI-driven chemical research.</li> <li>Development of strategies for teaching students about the ethical implications of AI applications in chemistry.</li> </ul> </li> <li><b>Evaluation and Reflection:</b> <ul style="list-style-type: none"> <li>Assessment of teachers' learning outcomes and proficiency in AI concepts and techniques through formative and summative evaluations.</li> <li>Reflection on the impact of AI integration on teaching practices, student engagement, and learning outcomes in chemistry classrooms.</li> <li>Continuous feedback and program improvement based on participant input and evaluation findings.</li> </ul> </li> </ol> <p><b>Conclusion:</b> By equipping chemistry teachers with the knowledge, skills, and resources to integrate artificial intelligence into their teaching practices, this professional development program aims to enhance the quality and relevance of chemistry education in the AI era. Through collaboration, experimentation, and ethical considerations, chemistry educators can leverage AI technologies to inspire curiosity, facilitate discovery, and prepare students for careers in chemistry and related fields.</p>
--

Figure 3. PDM proposed by ChatGPT for chemistry teachers

<p><b>AI for Chemistry Teachers Professional Development Program Model</b></p> <p><b>Program Goals</b></p> <ul style="list-style-type: none"> <li>To help chemistry teachers understand the basics of AI and its potential impact on chemistry education.</li> <li>To provide chemistry teachers with the skills and knowledge they need to use AI effectively in the chemistry classroom.</li> <li>To create a community of chemistry teachers who are committed to using AI to improve student learning in chemistry.</li> </ul> <p><b>Program Structure</b></p> <p>The program will be delivered in a blended format, with online modules and in-person workshops. The online modules will provide chemistry teachers with the foundational knowledge they need to understand AI. The in-person workshops will provide chemistry teachers with the opportunity to learn how to use AI in the chemistry classroom and to collaborate with other chemistry teachers.</p> <p><b>Program Content</b></p> <p>The program will cover the following topics:</p> <ul style="list-style-type: none"> <li>What is AI?</li> <li>The different types of AI</li> <li>The potential impact of AI on chemistry education</li> <li>How to use AI in the chemistry classroom</li> <li>Ethical considerations for using AI in chemistry education.</li> </ul> <p><b>Program Assessment</b></p> <p>Chemistry teachers will be assessed on their understanding of the content and their ability to use AI in the chemistry classroom. Assessment will be conducted through a combination of online quizzes, in-person workshop activities, and a final project.</p> <p><b>Program Dissemination</b></p> <p>The program will be disseminated through a variety of channels, including:</p> <ul style="list-style-type: none"> <li>A website</li> <li>A social media campaigns.</li> <li>A series of webinars</li> <li>A national conference</li> </ul> <p><b>Program Sustainability</b></p> <p>The program will be sustained through a combination of funding sources, including:</p> <ul style="list-style-type: none"> <li>Government grants</li> <li>Corporate sponsorships</li> <li>Private donations</li> </ul> <p><b>Program Impact</b></p> <p>The AI for Chemistry Teachers Professional Development Program Model is a comprehensive and innovative program that will help chemistry teachers understand and use AI to improve student learning in chemistry. The program is designed to be accessible to chemistry teachers of all levels of experience and expertise. The program is also designed to be sustainable, so that it can continue to provide benefits to chemistry teachers and students for years to come.</p> <p><b>Here are some additional resources that you may find helpful:</b></p> <ul style="list-style-type: none"> <li>The AI Index: <a href="https://aindex.org/">https://aindex.org/</a></li> <li>The Partnership on AI: <a href="https://www.partnershiponai.org/">https://www.partnershiponai.org/</a></li> <li>The AI for Education Initiative:</li> </ul> <p><b>Specific Examples of AI Applications in Chemistry Education</b></p> <ul style="list-style-type: none"> <li>AI-powered tutoring systems can provide students with personalized help and support in learning chemistry concepts.</li> <li>AI-powered virtual labs can allow students to conduct chemistry experiments in a safe and controlled environment.</li> <li>AI-powered assessment tools can provide teachers with real-time feedback on student learning.</li> </ul> <p><b>Conclusion</b></p> <p>The AI for Chemistry Teachers Professional Development Program Model is a comprehensive and innovative program that will help chemistry teachers understand and use AI to improve student learning in chemistry. The program is designed to be accessible to chemistry teachers of all levels of experience and expertise. The program is also designed to be sustainable, so that it can continue to provide benefits to chemistry teachers and students for years to come.</p>
--

Figure 4. PDM proposed by Gemini for chemistry teachers

When examining the PD programs suggested by ChatGPT and Gemini, it is evident that they emphasize the teaching of AI knowledge, pedagogical skills, hands-on activities and training, assessment and feedback, and support elements to enhance teachers' professional development. Additionally, it is noted that for a specific PD program tailored to chemistry teachers, in addition to the features proposed in the general PD program, it is necessary to address the integration of chemistry with AI and ethical issues. Furthermore, as part of this research question, ChatGPT and Gemini were questioned about the competencies teachers should possess in AI. The competencies suggested by ChatGPT and Gemini are presented in Figures 5, 6, 7, and 8.

<p><b>Title: Essential Competencies for Teachers in Artificial Intelligence Education</b></p> <p><b>Introduction:</b> In an era where artificial intelligence (AI) is increasingly integrated into various aspects of society, it is essential for educators to possess specific competencies to effectively teach AI concepts to students. This document outlines the key competencies that teachers should have in the field of AI education.</p> <ol style="list-style-type: none"> <li><b>Foundational Knowledge of Artificial Intelligence:</b> <ul style="list-style-type: none"> <li>Understanding of basic AI concepts, terminology, and principles.</li> <li>Awareness of the history, development, and current trends in AI technologies.</li> <li>Knowledge of key subfields of AI, including machine learning, neural networks, natural language processing, and robotics.</li> </ul> </li> <li><b>Pedagogical Skills:</b> <ul style="list-style-type: none"> <li>Ability to design and deliver engaging lessons that effectively communicate AI concepts to students of varying ages and backgrounds.</li> <li>Skill in integrating AI education into existing curriculum frameworks across different subject areas.</li> <li>Competence in selecting appropriate instructional strategies, resources, and assessment methods for teaching AI.</li> </ul> </li> <li><b>Proficiency in AI Tools and Technologies:</b> <ul style="list-style-type: none"> <li>Familiarity with AI tools, software, and platforms commonly used for data analysis, modeling, and programming (e.g., TensorFlow, Python, AI platforms like Google AIY).</li> <li>Ability to navigate and utilize online resources, educational apps, and AI-related learning materials for teaching purposes.</li> <li>Experience in leveraging AI tools to enhance teaching, learning, and research in their respective subject areas.</li> </ul> </li> <li><b>Critical Thinking and Problem-Solving:</b> <ul style="list-style-type: none"> <li>Capacity to foster critical thinking skills and inquiry-based learning in students through AI-related activities and projects.</li> <li>Skill in guiding students through problem-solving tasks and real-world applications of AI concepts.</li> <li>Ability to encourage exploration, experimentation, and creativity in students' engagement with AI technologies.</li> </ul> </li> <li><b>Ethical and Social Awareness:</b> <ul style="list-style-type: none"> <li>Understanding of ethical considerations, biases, and societal implications associated with AI technologies.</li> <li>Commitment to promoting responsible use of AI and fostering discussions on ethical dilemmas in AI development and deployment.</li> <li>Sensitivity to cultural, social, and ethical issues related to AI and its impact on individuals and communities.</li> </ul> </li> <li><b>Collaboration and Networking:</b> <ul style="list-style-type: none"> <li>Capacity to collaborate with colleagues, AI experts, industry professionals, and community stakeholders to enhance AI education initiatives.</li> <li>Skill in participating in professional learning networks, communities of practice, and collaborative projects focused on AI integration in education.</li> <li>Ability to leverage external partnerships and resources to enrich students' learning experiences in AI.</li> </ul> </li> <li><b>Continuous Learning and Adaptability:</b> <ul style="list-style-type: none"> <li>Commitment to ongoing professional development and staying updated with advancements in AI technologies and educational practices.</li> <li>Openness to experimenting with new teaching methods, technologies, and pedagogical approaches in AI education.</li> <li>Flexibility in adapting instructional strategies and resources to meet the evolving needs of students and the changing landscape of AI.</li> </ul> </li> </ol> <p><b>Conclusion:</b> Teachers play a pivotal role in preparing students for the AI-driven future by equipping them with the necessary knowledge, skills, and ethical awareness. By possessing the competencies outlined above, educators can effectively engage students in AI education, empower them to harness the potential of AI technologies, and foster responsible and ethical use of AI in society.</p>
---

Figure 5. The views of ChatGPT regarding the AI competencies

<p><b>AI Competencies for Teachers</b></p> <p>The rapid advancement of Artificial Intelligence (AI) has significant implications for the education sector. As AI becomes increasingly integrated into teaching and learning, it is crucial for teachers to possess the necessary competencies to effectively utilize this technology.</p> <p>Here are some key AI competencies that teachers should develop:</p> <ol style="list-style-type: none"> <li><b>1. Foundational Knowledge of AI:</b> <ul style="list-style-type: none"> <li>Understand the basics of AI, including its history, different types, and potential impact on education.</li> <li>Be familiar with the ethical considerations surrounding the use of AI in education.</li> </ul> </li> <li><b>2. AI Pedagogy:</b> <ul style="list-style-type: none"> <li>Understand how AI can be used to enhance student learning and engagement.</li> <li>Identify and select appropriate AI tools and resources for specific teaching and learning objectives.</li> <li>Effectively integrate AI into lesson plans and curriculum design.</li> </ul> </li> <li><b>3. AI Assessment:</b> <ul style="list-style-type: none"> <li>Understand how AI can be used to assess student learning and progress.</li> <li>Use AI-powered assessment tools to provide personalized feedback to students.</li> </ul> </li> <li><b>4. AI Technical Skills:</b> <ul style="list-style-type: none"> <li>Be able to use basic AI tools and applications.</li> <li>Have a basic understanding of AI programming languages and frameworks.</li> <li>Troubleshoot common technical issues related to AI in the classroom.</li> </ul> </li> <li><b>5. AI-related Learning Skills:</b> <ul style="list-style-type: none"> <li>Be able to learn and adapt to new AI technologies and tools.</li> <li>Encourage students to develop critical thinking and problem-solving skills related to AI.</li> <li>Promote responsible and ethical use of AI in the classroom.</li> </ul> </li> <li><b>6. Collaboration and Communication:</b> <ul style="list-style-type: none"> <li>Collaborate with other teachers, experts, and stakeholders to effectively integrate AI into the curriculum.</li> <li>Communicate effectively with students, parents, and the community about the benefits and risks of AI in education.</li> </ul> </li> </ol> <p>It is important to note that these competencies are not exhaustive and will continue to evolve as AI technology advances. However, by developing these core competencies, teachers can be well-positioned to leverage the power of AI to create a more personalized, engaging, and effective learning experience for all students.</p> <p>Here are some additional resources that may be helpful for teachers who want to develop their AI competencies:</p> <ul style="list-style-type: none"> <li>The AI Index: <a href="https://aiindex.org/">https://aiindex.org/</a></li> <li>The Partnership on AI: <a href="https://www.partnershiponai.org/">https://www.partnershiponai.org/</a></li> <li>The AI for Education Initiative:</li> </ul> <p>By developing their AI competencies, teachers can ensure that they are prepared to use this powerful technology to its full potential in the classroom.</p>
--

Figure 6. The views of Gemini regarding the AI competencies

ChatGPT has asked ChatGPT for the qualifications that chemistry teachers need regarding AI. Below is the response provided by ChatGPT.

<p>Title: Essential Competencies for Chemistry Teachers in Artificial Intelligence Education</p> <p>Introduction: As the integration of artificial intelligence (AI) into various scientific fields, including chemistry, becomes increasingly prevalent, it is imperative for chemistry teachers to possess specific competencies to effectively incorporate AI concepts into their teaching. This document outlines the key competencies that chemistry teachers should have in the realm of AI education.</p> <ol style="list-style-type: none"> <li><b>1. Profound Understanding of Chemistry Concepts:</b> <ul style="list-style-type: none"> <li>Mastery of fundamental principles, theories, and concepts in chemistry, including atomic structure, chemical bonding, kinetics, thermodynamics, and organic chemistry.</li> <li>In-depth knowledge of key chemical processes, reactions, and phenomena across different branches of chemistry.</li> </ul> </li> <li><b>2. Foundation in Artificial Intelligence:</b> <ul style="list-style-type: none"> <li>Familiarity with basic AI concepts, terminology, and applications relevant to chemistry, such as machine learning algorithms for chemical data analysis and modeling.</li> <li>Understanding of AI-driven tools and techniques used in chemical research, drug discovery, materials science, and computational chemistry.</li> </ul> </li> <li><b>3. Integration of AI into Chemistry Curriculum:</b> <ul style="list-style-type: none"> <li>Ability to identify opportunities for integrating AI concepts and tools into chemistry curriculum objectives, lesson plans, and laboratory experiments.</li> <li>Skill in designing AI-enhanced learning experiences and projects that align with chemistry curriculum standards and foster student engagement and inquiry.</li> </ul> </li> <li><b>4. Computational Skills:</b> <ul style="list-style-type: none"> <li>Proficiency in using computational tools and software for chemical data analysis, modeling, and simulation, such as molecular modeling software, quantum chemistry packages, and statistical analysis tools.</li> <li>Experience in programming languages commonly used in chemistry and AI, such as Python, MATLAB, and R, for implementing AI algorithms and analyzing chemical data.</li> </ul> </li> <li><b>5. Critical Thinking and Problem-Solving:</b> <ul style="list-style-type: none"> <li>Capacity to promote critical thinking skills and scientific inquiry among students through AI-driven problem-solving tasks and real-world applications in chemistry.</li> <li>Skill in guiding students through the analysis, interpretation, and evaluation of AI-generated chemical data and predictions.</li> </ul> </li> <li><b>6. Ethical Considerations in AI-Driven Chemistry:</b> <ul style="list-style-type: none"> <li>Awareness of ethical considerations, data privacy issues, and societal implications associated with the use of AI technologies in chemistry research and education.</li> <li>Commitment to promoting responsible conduct of research, integrity, and ethical decision-making in AI-driven chemical investigations.</li> </ul> </li> <li><b>7. Collaboration and Professional Development:</b> <ul style="list-style-type: none"> <li>Ability to collaborate with AI experts, computational chemists, and interdisciplinary teams to enhance AI education initiatives and advance chemical research.</li> <li>Commitment to continuous professional development, staying updated with advancements in AI technologies and their applications in chemistry, and participating in relevant workshops, conferences, and professional networks.</li> </ul> </li> </ol> <p>Conclusion: Chemistry teachers play a pivotal role in preparing the next generation of scientists and innovators to navigate the intersection of chemistry and artificial intelligence. By possessing the competencies outlined above, educators can effectively integrate AI concepts into their chemistry classrooms, inspire students' curiosity and creativity, and empower them to leverage AI technologies for solving complex chemical problems and addressing societal challenges.</p>
---

Figure 7. ChatGPT's views on the qualifications chemistry teachers need to have regarding AI

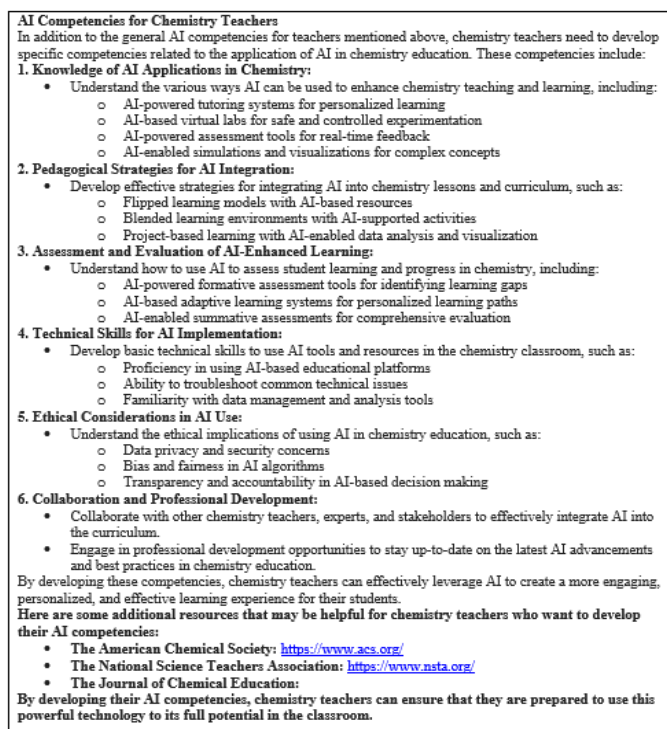


Figure 8. Gemini's views on the qualifications chemistry teachers need to have regarding AI

When Figures 5, 6, 7, and 8 are examined, the titles suggested by ChatGPT and Gemini for PD programs are similar. Unlike PD programs, however, they focus on various AI-related skills, such as computational thinking, collaboration, communication, and learning skills.

*RQ4: What are the opinions of experts and artificial intelligence about the model developed for teacher training?*

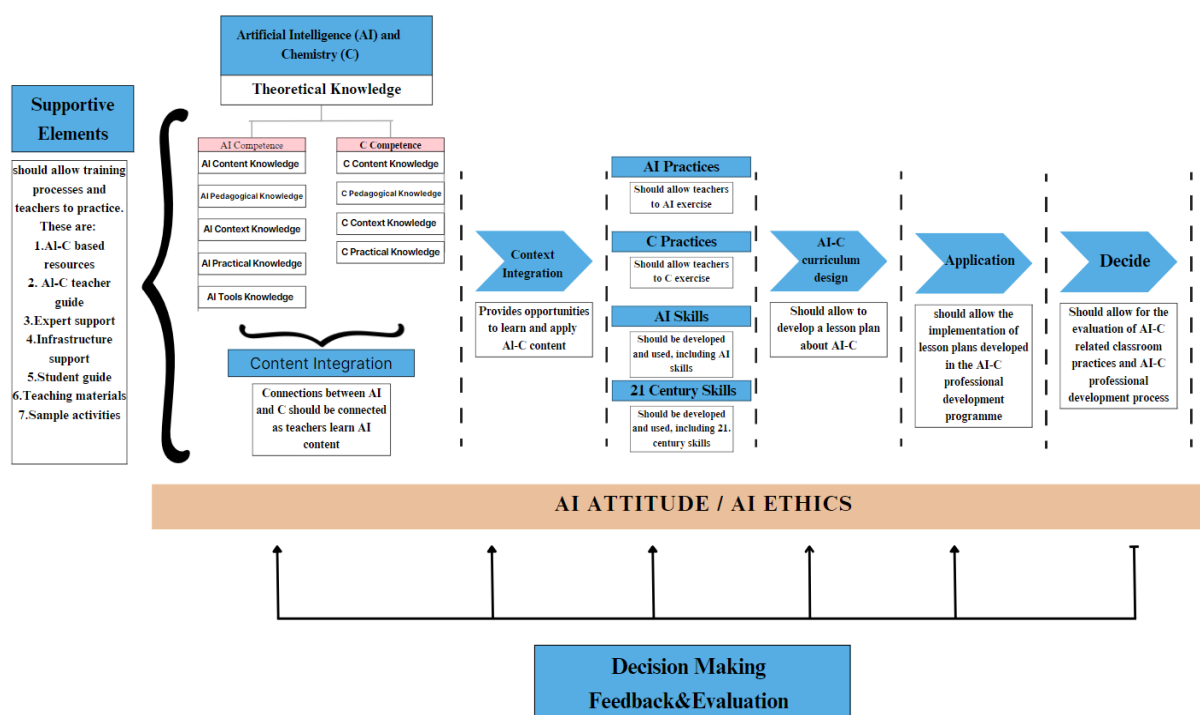


Figure 9. AI PDM for chemistry teachers

A PDM for chemistry teachers was developed based on document reviews, model proposals from AI, and teacher opinions. The model's first stage addresses teacher needs, while the second stage focuses on program objectives. The model presented below was formulated per identified needs and goals, considering input from experts and AI (Figure 9). AI was asked whether the model developed had any areas for improvement. ChatGPT was first asked about the model (Figure 9).

Here are my thoughts on the developed model:
The model encompasses several strengths:
<ol style="list-style-type: none"> <li>1. Needs assessment for teachers: The initial phase involves identifying the specific needs of teachers, ensuring that the subsequent training is tailored to address these needs effectively.</li> <li>2. Goal setting: By establishing clear objectives, the model provides direction and purpose to the educational process, ensuring that the outcomes are aligned with the intended goals.</li> <li>3. Comprehensive content development: The inclusion of various aspects such as content knowledge, pedagogy, practical skills, tools, and contextual understanding ensures a well-rounded approach to content development, enhancing its effectiveness.</li> <li>4. Integration of AI into daily life: Emphasizing the connection between AI content and everyday life through contextual understanding helps make the learning experience more relevant and relatable for teachers and students alike.</li> <li>5. Focus on practical application and skill development: The model emphasizes not only the theoretical aspects but also the practical application of AI, along with the acquisition of specific skills and competencies, including 21st-century skills.</li> <li>6. Systematic program design: The structured approach to designing AI programs, including the development of tailored lesson plans and program units, ensures coherence and alignment with educational objectives.</li> <li>7. Feedback and evaluation: Incorporating feedback mechanisms and evaluation processes enables continuous improvement and ensures that decisions are informed by evidence and assessment data.</li> </ol>
However, there are some areas where the model could be further enhanced:
<ol style="list-style-type: none"> <li>1. Clarity on supportive elements: While the model mentions supportive elements such as equipment, resources, and expert support, providing more specific details on how these elements will be provided and utilized would enhance clarity and implementation.</li> <li>2. Emphasis on AI ethics and attitudes: While the model acknowledges the importance of considering AI ethics and attitudes throughout the process, further emphasis and guidance on how to integrate these aspects into training and curriculum development would be beneficial.</li> </ol>
Overall, the model provides a robust framework for integrating AI into education, but further refinement and clarity in certain areas would strengthen its effectiveness and implementation.

Figure 10. ChatGPT's views on the model

<b>Detailed Thoughts on the Developed Model</b>
<b>Strengths of the Model:</b>
<ul style="list-style-type: none"> <li>• <b>Comprehensiveness:</b> The model comprehensively addresses the integration of AI into the teacher training program. Everything from the basics of AI to its integration into a specific field like chemistry has been considered.</li> <li>• <b>Flexibility:</b> The model offers flexibility to be adapted to different fields and different levels of teachers.</li> <li>• <b>Sustainability:</b> The supporting elements of the model will contribute to the long-term sustainability of the program.</li> <li>• <b>Up-to-dateness:</b> The model is designed to be updated considering the ever-evolving nature of AI.</li> <li>• <b>Feedback and Evaluation:</b> There are feedback and evaluation mechanisms at each stage of the model.</li> <li>• <b>Supporting Elements:</b> Many supporting elements such as equipment, resources, expert, technical support, teacher and student guides, and sample activities will increase the success of the model.</li> <li>• <b>Ethics and Attitude:</b> Considering AI ethics and attitude throughout the process is an important step.</li> </ul>
<b>Areas for Improvement:</b>
<ul style="list-style-type: none"> <li>• <b>Pilot Implementation:</b> A pilot implementation can be conducted to test the model's effectiveness in real life.</li> <li>• <b>Stakeholder Involvement:</b> Teachers, experts, and other stakeholders can be involved more actively in the development and implementation of the model.</li> <li>• <b>Implementation Phase:</b> It would be beneficial to plan the implementation phase in more detail and provide guidance to teachers.</li> <li>• <b>Concrete Examples:</b> Concrete examples and case studies can be added for each stage of the model to ensure better understanding.</li> </ul>
<b>Detailed Review of the Model:</b>
<ol style="list-style-type: none"> <li>1. <b>Needs Identification, Goal Setting, and Content Process:</b> <ul style="list-style-type: none"> <li>• It is important to plan these stages in detail and determine concrete outputs for each stage.</li> <li>• Examples of how AI can be used in different fields and at different levels can be added to the content process.</li> </ul> </li> <li>2. <b>Contextual Information and Connection to Daily Life:</b> <ul style="list-style-type: none"> <li>• Concrete examples and activities can be provided on how teachers can connect AI content to daily life.</li> </ul> </li> <li>3. <b>AI Applications, Field-Specific Applications, and Skill Acquisition:</b> <ul style="list-style-type: none"> <li>• Sample lesson plans and program units can be developed for AI applications in different fields and levels.</li> <li>• Concrete activities and assessment tools can be provided for the acquisition of AI skills and 21st century skills.</li> </ul> </li> <li>4. <b>AI Program Design:</b> <ul style="list-style-type: none"> <li>• A detailed guide can be prepared for the development of AI and field-specific lesson plans and program units.</li> </ul> </li> <li>5. <b>Implementation and Feedback:</b> <ul style="list-style-type: none"> <li>• Mentoring and coaching support can be provided to teachers during the implementation phase.</li> <li>• Feedback and evaluation mechanisms can be further detailed.</li> </ul> </li> <li>6. <b>Decision Making and Supporting Elements:</b> <ul style="list-style-type: none"> <li>• A clear framework can be created for the decisions to be made at each stage of the model.</li> <li>• Detailed information can be provided on how and when the supporting elements will be used.</li> </ul> </li> </ol>
<b>Overall Evaluation:</b>
<p>The developed AI teacher training program model is very comprehensive and detailed. The model has many strengths and potential. I believe that the model can be made even more effective and sustainable with some of the development suggestions mentioned above.</p> <p><b>Congratulations on your model! I am sure it will help teachers discover the power of AI and use it to improve their students' learning.</b></p> <p><b>Note:</b> Please feel free to ask me any questions or feedback you have about your model.</p> <p><b>Additional:</b></p> <ul style="list-style-type: none"> <li>• Preparing a visual presentation of your model will help it to be better understood and disseminated.</li> <li>• Translating your model into different languages can make it accessible to a wider range of teachers worldwide.</li> </ul> <p><b>I hope this detailed evaluation is helpful for you.</b></p>

Figure 11. Gemini's views on the model

The results showed that ChatGPT highlighted seven strengths of the developed model (Figure 10). Additionally, ChatGPT emphasized the need for (1) clearly defined support elements in the model and (2) comprehensive addressing of ethical and attitudinal aspects, along with strategies for cultivating such attitudes among teachers. In this context, the model's support elements were elucidated with explicit references to the literature. Subsequently, Gemini's perspectives on the model were obtained, and Gemini's responses are presented in Figure 11.

Gemini highlighted that the model developed for teachers has many strengths. Furthermore, Gemini emphasized that this model will provide robust learning experiences for teachers in AI. However, Gemini suggests that the model needs improvement in terms of (1) piloting, (2) stakeholder involvement, (3) implementation phase, and (4) providing concrete examples. In this context, based on the perspectives of both ChatGPT and Gemini, improvements were made to the model, sample content for chemistry courses was developed, and pilot implementations were conducted for chemistry teachers. Expert opinions regarding the model were also consulted within the scope of the study. Experts highlighted that the model has many strengths and will enhance teachers' competencies in AI. Additionally, experts emphasized that (1) the steps should be explained in detail, (2) examples related to the field should be provided, (3) pilot studies should be conducted, and (4) the decision-making mechanism should be clarified.

## Discussion and Conclusion

The first research question of the study focused on analyzing studies related to artificial intelligence (AI) for teachers. The review results identified codes such as AI knowledge, AI attitude, AI skills, AI ethics, AI tools, AI literacy, AI lesson plans, TPACK, domain-specific knowledge, and pedagogical knowledge. When the identified codes are analyzed, it is understood that concepts are highlighted in teacher education literature. For example, Zhou (2020) focused on AI skills, AI knowledge, technology knowledge, and AI literacy. Similarly, Lee et al. (2024) highlighted AI skills, AI tools, AI knowledge, AI pedagogical knowledge, practical knowledge, AI attitude, and AI lesson plans. Many studies in the literature also emphasize similar issues related to AI (Lee et al., 2023; Lin et al., 2022; Mondal et al., 2023; Nazaretsky et al., 2022; Nyaaba & Zhai, 2024; Su et al., 2023; Wardant et al., 2024; Yang, 2022).

The second result obtained within the scope of the first research question is related to the supportive elements necessary for successfully implementing the AI process. It is emphasized that these supportive elements generally include teacher and student guides for AI, AI tool usage guides, expert support, technical support, and various other supports. Additionally, it is noted that sample lesson plans and activities should be available for teachers to implement AI in their classrooms. For example, Su et al. (2023) emphasized the importance of teacher and student guides. Similarly, Mondal et al. (2023) emphasized the importance of technical support and AI resources in the AI process. Linder and Romeike (2019) mentioned needing more application examples and sufficient materials in the AI process. In addition, many other studies emphasize the importance of similar supporting elements (Chiu & Chai, 2020; Salas-Pilco et al., 2022). Our result is consistent with the literature (Kim & Kim, 2022; Kitcharoen et al., 2024)).

Within the scope of the second sub-question, chemistry teachers' views on their competencies in AI were investigated. Most chemistry teachers feel they need to be more competent in AI. Moreover, chemistry teachers must improve AI knowledge, tools, applications, lesson planning, ethics, and course integration. In addition to this, teachers also feel they need to improve AI literacy. Similarly, they think they need to improve in developing AI-based lesson plans.

Furthermore, chemistry teachers have expressed ethical concerns about various aspects of artificial intelligence use, including assigning homework and research, personal data loss, privacy issues, using others' information, and the possibility of artificial intelligence misguidance. For instance, Park et al. (2023) study indicated that teachers must be more competent in artificial intelligence. Another study emphasized that insufficient support and resources related to artificial intelligence affect teachers' in-class applications (Lindner & Romeike, 2019). Çelik et al. (2022) study also highlighted that teachers want to use artificial intelligence in their classrooms. Our results are consistent with the literature.

Another result obtained within the scope of the study's second research question is examining teachers' views on the effects of using AI on students in chemistry classes. Teachers stated that using AI in chemistry classes provides various advantages, such as developing students' AI literacy, ensuring permanent learning, supporting individual learning, and promoting effective learning. These results are consistent with studies in the literature

(Kim et al., 2021). Additionally, Ng et al.'s (2022) study emphasized the importance of students' having AI literacy. Similarly, many countries support students' AI literacy by including AI in K-12 schools.

Another result obtained within the scope of the second research question is the teachers' views on the connections between AI and chemistry subjects. Teachers stated that AI can be associated with organic chemistry, chemical reactions, molecules and atoms, chemistry experiments, the periodic table, and nanotechnology. In addition, the teachers believed that AI could help visualize chemistry topics, solve chemistry problems, analyze chemical reactions, conduct chemistry experiments, evaluate and predict students' future in chemistry, and develop lesson plans. For example, Yuriev et al. (2023) highlighted the importance of artificial intelligence in chemistry education. Sarioğlu (2023) highlighted that AI tools could be used to visualize molecules in measurement and evaluation processes in chemistry courses and the course planning process. Choudhary et al. (2022) stated that AI could be used in molecular property prediction, drug discovery, molecular design, nanotechnology, wastewater treatment, retrosynthesis, and reaction results. Clark (2023) highlighted its potential use in general chemistry laboratories and organic chemistry courses. The findings of this study are consistent with the conclusions of the literature.

Within the scope of the third problem of the research, the views of artificial intelligence on PD programs were examined. In addition to the information obtained from the document reviews, artificial intelligence stated that there is a need for information about AI, pedagogical knowledge, practical activities, and feedback to support teachers' PD programs. They also highlighted the need to emphasize integrating chemistry with AI, attitudes, and ethical issues in a PD program specific to chemistry teachers. The information provided by AI is in line with the documentation in the literature.

The fourth research question of the study explores the perspectives of experts and artificial intelligence on the model developed for chemistry teachers. ChatGPT highlighted both the strengths and weaknesses of the developed model for AI education, stressing the importance of clarity in support elements, ethics, and attitudes. Similarly, Gemini noted the model's strengths in supporting teachers' PD, emphasizing the significance of pilot implementation, providing support with concrete examples, the implementation phase, and conducting a pilot study. Experts also underscored the model's strengths in supporting teachers' PD. Additionally, they highlighted the necessity of explaining the model's steps, offering domain-specific examples, conducting pilot studies, and clarifying the decision-making mechanism. Based on the literature review, views on AI, and teacher views, a model was proposed to support chemistry teachers' professional development on AI.

#### AI PDM for Chemistry Teachers

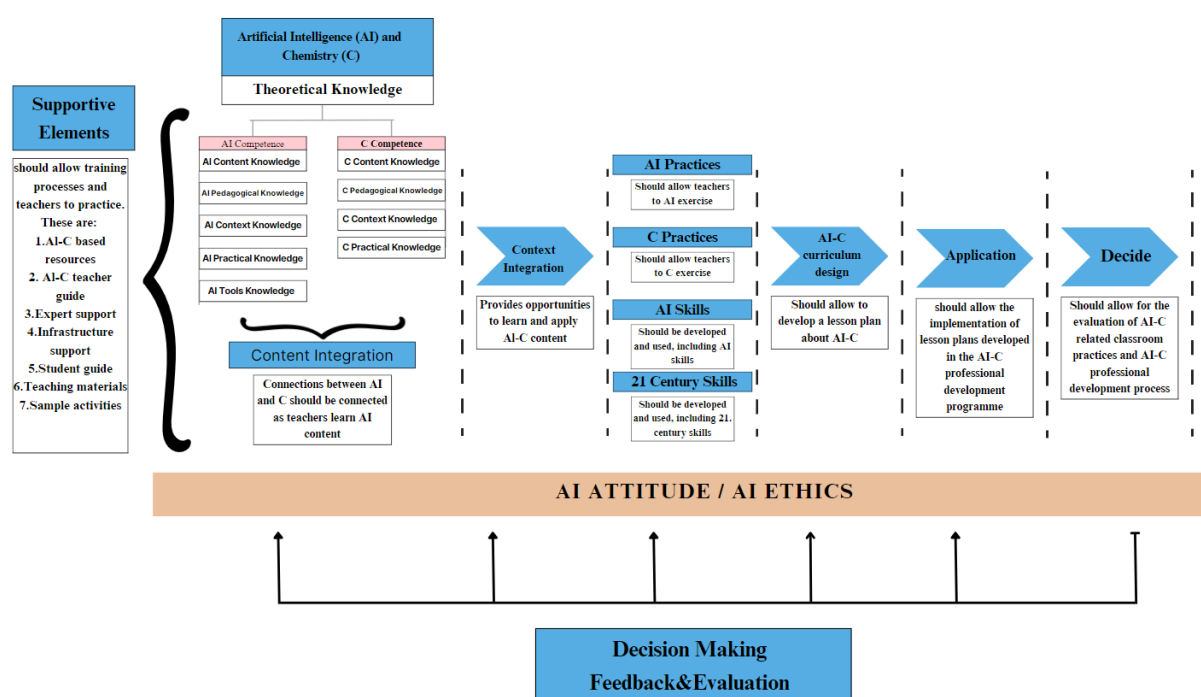


Figure 12. Proposed model for AI PDM for chemistry teachers

The stages and contents of the developed model are detailed below:

**Identification of Teacher Needs:** This stage involves identifying the expectations of chemistry teachers concerning AI to determine their training requirements about AI before the education process commences.

**Determination of Objectives:** This stage entails establishing the objectives expected to be accomplished in the AI professional development program for chemistry teachers.

**AI and Chemistry Theoretical Knowledge:** This stage involves providing teachers with knowledge about AI and the field of chemistry. It consists of two parts: (1) AI competencies and (2) chemistry-related competencies. AI competencies include AI-related domain, pedagogy, context, and application information that should be provided in teacher training. Competencies related to chemistry encompass domain-specific, pedagogical, contextual, and applied knowledge of chemistry. Essentially, this section forms the content knowledge.

**Context Integration:** This stage establishes the connection between AI training for chemistry teachers and the environment in which teachers provide education. In other words, it entails establishing the real-life connection of the AI-related knowledge acquired by chemistry teachers. Contextual knowledge is vital for AI training because it creates a connection between the field of chemistry and daily life. Additionally, it provides opportunities to learn AI content and apply it in real-life chemistry classes.

**AI and Chemistry Practices:** This stage enables the implementation of practices in real life that are relevant to AI and chemistry. These practices include information about the practical applications of AI in chemistry education in classroom settings.

**AI and 21st-Century Skills:** This stage encompasses the 21st-century skills that chemistry teachers need to acquire, along with AI skills.

**AI-Chemistry Program Design:** This involves developing lesson plans related to AI and chemistry. This stage is essential for including AI applications in chemistry education because it outlines when teachers will use AI in their lessons.

**Implementation Phase:** This is the stage where the lesson plans developed by chemistry teachers as part of the AI professional development program are implemented. During this stage, teachers can apply the knowledge they have acquired in their classrooms. Conducting in-class implementations is crucial for teachers to identify any shortcomings in their training. Successful implementation of in-class activities within the scope of their training indicates the program's effectiveness.

**Decision Stage:** This stage involves identifying any shortcomings in the professional development program provided to teachers. It aims to pinpoint which parts of the program are deficient or not functioning as intended, and improvements are made accordingly. Teachers' feedback is gathered, and the training process is evaluated. Regular input from chemistry teachers is solicited throughout this process, and assessments are conducted.

**AI Attitude and AI Ethics:** Although not explicitly outlined as stages in the developed model, AI attitude and AI ethics are two crucial factors influencing it. Improving teachers' attitudes towards AI is essential for integrating AI into classroom practices, affecting teachers' self-efficacy. AI ethics is a significant concern addressed by teachers and users, highlighting the importance of providing teachers with training to support their professional development.

**Supporting Elements:** These components are necessary for teachers to integrate AI practices into their classrooms. These include AI-based resources, AI teacher guides, AI subject matter expert support, AI infrastructure support, AI student guides, AI teaching materials, and sample activities. All these elements constitute the necessary support mechanisms for chemistry teachers to implement in their lessons.

## **Limitations of the Study**

The study had three limitations. First, the sample consisted only of chemistry teachers interested in AI education, so the results cannot be generalized. Second, the developed model is based on the views of ChatGPT and Gemini artificial intelligence. Third, the model presented in this study focuses on providing professional

development in artificial intelligence, especially for chemistry teachers. Also, the processes provided in this model are specific to chemistry teachers.

## Suggestions

The model developed for chemistry teachers in this study could be adapted for use by teachers in other subject areas. Longitudinal studies can be conducted to evaluate the long-term effects of this model on teachers. Training programs based on this model can increase teachers' competencies in artificial intelligence. Furthermore, the model proposed in this study is expected to provide a basic framework for future research in the field of AI.

## Ethical Considerations

The data for this study were collected voluntarily from chemistry teachers. I am grateful to all the teachers who participated in this process.

## Scientific Ethics Declaration

The authors declare that the scientific ethical and legal responsibility of this article published in JESEH journal belongs to the authors.

## References

- Arman, M., & Lamiyar, U. R. (2023). Exploring the implication of ChatGPT AI for business: Efficiency and challenges. *Applied Quantitative Analysis*, 3(2), 46-67.
- Asia Pacific Foundation of Canada Korea AI Ethics Association, (2021). *Talent for the future: AI education for K-12 in Canada and South Korea*. Asia Pacific Foundation of Canada. <https://www.asiapacific.ca/sites/default/files/publication-pdf/AI%2012%20Education%20Report%20FINAL.pdf>
- Balcı, A. (2016). *Research methods, techniques, and principles in social sciences*. Pegem.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. Freeman.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall.
- Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, 28(2), 117-148.
- Casal-Otero, L., Catala, A., Fernández-Morante, C., Taboada, M., Cebreiro, B., & Barro, S. (2023). AI literacy in K-12: A systematic literature review. *International Journal of STEM Education*, 10(1), 29. <https://doi.org/10.1186/s40594-023-00418-7>
- Chiu T.K.F., & Chai, C.S. (2020). Sustainable curriculum planning for artificial intelligence education: A self-determination theory perspective. *Sustainability*, 12(14):5568. <https://doi.org/10.3390/su12145568>
- Chiu, T. K. F., Meng, H., Chai, C. S., Yeung, Y., King, I., & Wong, S. (2021). Creation and evaluation of a pre-tertiary Artificial Intelligence (AI) curriculum. *IEEE Transactions on Education*, 30-39. <https://doi.org/10.1109/TE.2021.3085878>
- Choudhary, N., Bharti, R., & Sharma, R. (2022). Role of artificial intelligence in chemistry. *Materials Today: Proceedings*, 48, 1527-1533. <https://doi.org/10.1016/j.matpr.2021.09.428>
- Chounta, I.A., Bardone, E., Raudsep, A., & Pedaaste, M. (2022). Exploring teachers' perceptions of artificial intelligence as a tool to support their practice in Estonian K-12 education. *International Journal of Artificial Intelligence in Education*, 32, 725-755 (2022). <https://doi.org/10.1007/s40593-021-00243-5>
- 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. <https://doi.org/10.1021/acs.jchemed.3c00027>
- Çelik, İ. (2023). Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. *Computers in Human Behavior*, 138, 1-12. <https://doi.org/10.1016/j.chb.2022.107468>

- Çelik, I., Dindar, M., Muukkonen, H., & Järvelä, S. (2022). The promises and challenges of artificial intelligence for teachers: a systematic review of research. *TechTrends*, 66, 616–630. <https://doi.org/10.1007/s11528-022-00715-y>
- Daher, W., Diab, H., & Rayan, A. (2023). Artificial intelligence generative tools and conceptual knowledge in problem solving in chemistry. *Information*, 14(7), 409. <https://doi.org/10.3390/info14070409>
- Deng, J. M., Lalani, Z., McDermid, L.A., & Szozda, A. R., (2023). Using generative artificial intelligence in chemistry education research: Prioritizing ethical use and accessibility. *Journal of Chemical Education*. <https://doi.org/10.26434/chemrxiv-2023-24zfi>
- Dick, W., & Carey, L. (1996). *The systematic design of instruction*. Harper Collins College Publishers.
- Eckman, E. W., Williams, M. A. and Silver-Thorn, M. B. (2016). An integrated model for STEM teacher preparation: The value of a teaching cooperative educational experience. *Journal of STEM Teacher Education*, 51(1), 71-82.
- 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. <https://doi.org/10.1021/acs.jchemed.3c00063>
- Gökdere, M., & Çepni, S. (2003). The role of the teacher in values education given to gifted children. *Journal of Values Education*, 1(2), 93-107.
- Hibpshman, T. L. (2007). *Analysis of transcript data for mathematics and science teachers*. Education Professional Standards Board.
- Joshi, R.P., & Kumar, N. (2021). Artificial intelligence for autonomous molecular design: A perspective. *Molecules*, 26(22), 6761. <https://doi.org/10.3390/molecules26226761>
- Kitcharoen, P., Howimanporn, S., & Chookaew, S. (2024). Enhancing teachers' AI competencies through artificial intelligence of things professional development training. *International Journal of Interactive Mobile Technologies (IJIM)*, 18(2), 4–15. <https://doi.org/10.3991/ijim.v18i02.46613>
- Kim, N.J., & Kim, M.K. (2022). Teacher's perceptions of using an artificial intelligence-based educational tool for scientific writing. *Frontiers in Education*, 7, 755914. <https://doi.org/10.3389/feduc.2022.755914>
- Kim, S., Jang, Y., Kim, W., Choi, S., Jung, H., Kim, S., & Kim, H. (2021). Why and what to teach: AI curriculum for elementary school. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(17), 15569-15576. <https://doi.org/10.1609/aaai.v35i17.17833>
- Kuleto, V., Ilić, M.P., Bucea-Manea-Toniş, R., Ciocodeică, D-F., Mihălcescu, H., & Mindrescu, V. (2022). The attitudes of K–12 schools' teachers in Serbia towards the potential of artificial intelligence. *Sustainability*, 14(14), 8636. <https://doi.org/10.3390/su14148636>
- Lee, Y. J., Davis, R. O., & Ryu, J. (2024). Korean in-service teachers' perceptions of implementing artificial intelligence (ai) education for teaching in schools and their ai teacher training programs. *International Journal of Information and Education Technology*, 14(2), 214-219. <https://doi.org/10.18178/ijiet.2024.14.2.2042>
- Lin, X.-F., Chen, L., Chan, K.K., Peng, S., Chen, X., Xie, S., Liu, J., & Hu, Q. (2022). Teachers' perceptions of teaching sustainable artificial intelligence: A design frame perspective. *Sustainability*, 14, 7811. <https://doi.org/10.3390/su14137811>
- Lindner, A., & Romeike, R. (2019). Teachers' perspectives on artificial intelligence. In *ISSEP 2019 - 12th International Conference on Informatics in Schools: Situation, evaluation and perspectives*, local proceedings 2019 (pp. 22–29).
- Long, D., & Magerko, B. (2020). What is AI literacy?. Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-16). Association for Computing Machinery. <https://doi.org/10.1145/3313831.3376727>
- Lu, S. Y., Lo, C. C., & Syu, J. Y. (2022). Project-based learning oriented STEAM: The case of micro-bit paper-cutting lamp. *International Journal of Technology and Design Education*, 32(5), 2553-2575. <https://doi.org/10.1007/s10798-021-09714-1>
- Luckin, R., George, K., & Cukurova, M. (2022). *AI for school teachers*. CRC Press. <https://doi.org/10.1201/9781003193173>
- Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Sage.
- Muslim al-Zyoud, H. M., (2020). The role of artificial intelligence in teacher professional development. *Universal Journal of Educational Research*, 8(11B), 6263 - 6272. <https://doi.org/10.13189/ujer.2020.082265>
- Mondal, H., Marndi, G., Behera, J. K., & Mondal, S. (2023). ChatGPT for teachers: Practical examples for utilizing artificial intelligence for educational purposes. *Indian Journal of Vascular and Endovascular Surgery*, 10(3), 200-205. [https://doi.org/10.4103/ijves.ijves\\_37\\_23](https://doi.org/10.4103/ijves.ijves_37_23)

- Nazaretsky, T., Ariely, M., Çukurova, M., & Alexandron, G. (2022). Teachers' trust in AI-powered educational technology and a professional development program to improve it. *British Journal of Educational Technology*, 53(4), 914-931. <https://doi.org/10.1111/bjet.13232>
- Ng, D. T. K., Luo, W., Chan, H. M. Y., & Chu, S. K. W. (2022). Using digital story writing as a pedagogy to develop AI literacy among primary students. *Computers and Education: Artificial Intelligence*, 3, 100054. <https://doi.org/10.1016/j.caeai.2022.100054>
- Nyaaba, M., & Zhai, X. (2024). Generative ai professional development needs for teacher educators. *Journal of AI*. <http://dx.doi.org/10.2139/ssrn.4628126>
- Owczarek, D. (2022). *Artificial intelligence in the chemicals: How AI benefits the chemical industry*. Nexcode. <https://nexocode.com/blog/posts/artificial-intelligence-in-the-chemicals-how-ai-benefits-the-chemical-industry/>
- Park, J., Teo, T.W., Teo, A., Chang, J., Huang, J.S., & Koo, S. (2023). Integrating artificial intelligence into science lessons: teachers' experiences and views. *International Journal of STEM Education*, 10, 61. <https://doi.org/10.1186/s40594-023-00454-3>
- Rane, N.L., Choudhary, S.P., & Rane, J. (2024). Gemini or ChatGPT? capability, performance, and selection of cutting-edge generative artificial intelligence (AI) in business management. *Studies in Economics and Business Relations*, 5(1), 40-50.
- Salas-Pilco, S.Z., Xiao, K., & Hu, X. (2022). Artificial intelligence and learning analytics in teacher education: A systematic review. *Education Sciences*, 12(8), 569. <https://doi.org/10.3390/educsci12080569>
- Sanusi, I.T., Olaleye, S.A., Agbo, F.J., & Chiu, T.K.F. (2022). The role of learners' competencies in artificial intelligence education. *Computers and Education: Artificial Intelligence*, 3, 1-10. <https://doi.org/10.1016/j.caeai.2022.100098>
- Serin, M. K., & Bayraktar, Ş. (2015). According to the locus of control status of prospective primary school teachers' science the examination of efficacy beliefs towards teaching. *Kafkas University Institute of Social Sciences Journal*, 16, 51-71.
- Stohlmann, M., Tamara J.M., & Roehring, G.H. (2012). Considerations for teaching integrated STEM education. *Journal of Pre-College Engineering Education Research (J-PEER)*, 2(1), 28-34.
- Su, J., Ng, D.T.K., & Chu, S.K.W. (2023). Artificial intelligence (AI) literacy in early childhood education: the challenges and opportunities. *Computers and Education: Artificial Intelligence*, 4, 1-14. <https://doi.org/10.1016/j.caeai.2023.100124>
- Wardat, Y., Tashtoush, M. A., Alali, R., & Saleh, S. (2024). Artificial intelligence in education: mathematics teachers' perspectives, practices and challenges. *Iraqi Journal for Computer Science and Mathematics*, 5(1), 60-77. <https://doi.org/10.52866/ijcsm.2024.05.01.004>
- Yang, W. (2022). Artificial Intelligence education for young children: Why, what, and how in curriculum design and implementation. *Computers and Education: Artificial Intelligence*, 3. <https://doi.org/10.1016/j.caeai.2022.100061>
- Yıldırım, A., & Şimşek, H. (2011). *Qualitative research methods in social sciences*. Seçkin Publications.
- Yıldırım, B. (2023). Flipped TRIZ-STEM: Enhancing teacher training through innovative pedagogy?. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-023-12242-1>
- Yin, R. K. (2017). *Case study research and applications: Design and methods*. SAGE.
- Yau, K.W., CHAI, C.S., Chiu, T.K., Meng, H., King, I., & Yam, Y. (2023). A phenomenographic approach on teacher conceptions of teaching Artificial Intelligence (AI) in K-12 schools. *Education and Information Technologies*, 28, 1041–1064 (2023). <https://doi.org/10.1007/s10639-022-11161-x>
- Yoo, I., & Yi, C. G. (2022). Economic innovation caused by digital transformation and impact on social systems. *Sustainability*, 14(5), 2600. <https://doi.org/10.3390/su14052600>
- Yuriev, E., Wink, D.J., & Holme, T.A. (2023). Virtual special issue call for papers: Investigating the uses and impacts of generative artificial intelligence in chemistry education. *Journal of Chemical Education*, 100(9), 3168-3170. <https://doi.org/10.1021/acs.jchemed.3c00829>
- Zhou, S. (2020). Analysis of the Professional Development of teachers in the era of artificial intelligence. *Advances in Education*, 10(3), 339-345. <https://doi.org/10.12677/ae.2020.103056>

---

### Author(s) Information

**Bekir Yıldırım**

PhD, Associate Professor, Department of Mathematics and Science Education, Faculty of Education, Muş Alparslan University, Muş 49250, Türkiye  
 Contact Email: [bekir58bekir@gmail.com](mailto:bekir58bekir@gmail.com)  
 ORCID iD: <https://orcid.org/0000-0002-5374-4025>

**Ahmet Tayfur Akcan**

PhD, Associate Professor, Department of International Trade, Faculty of Applied Sciences, Necmettin Erbakan University, Konya 42100, Türkiye  
 ORCID iD: <https://orcid.org/0000-0001-8210-7327>

---

## Appendix: AI Interview Guide

Dear Teachers,

This interview guide has been designed to gather your perspectives on AI. Your input will be utilized solely for research purposes, and the collective responses will be evaluated. To ensure the reliability of this study, it is crucial that you candidly express your genuine opinions. Respond to each question; your participation is integral to this scientific inquiry. Your answers will be utilized exclusively for this research and will not be disclosed to others. Thank you for your valuable contribution to this study.

1. What do teachers think about the AI competencies?
2. What do teachers think about the areas of deficiency regarding AI competencies?
3. What do teachers think about AI literacy?
4. What do teachers think about developing AI-based chemistry lesson plans?
5. What do teachers think about the ethical concerns about AI?
6. What do teachers think about the ethical status of using AI?
7. What do teachers think about the benefits of using AI in chemistry classes for students?
8. What do teachers think about using AI for students?
9. What do teachers think about the connection between AI and Chemistry topics?
10. What do teachers think about customizing AI for chemistry lessons?

## Bibliometric Analysis of Studies on the Artificial Intelligence in Science Education with VOSviewer

Hayriye Nevin Genc, Nuriye Kocak

### Article Info

#### Article History

Published:  
01 October 2024

Received:  
14 May 2024

Accepted:  
21 August 2024

#### Keywords

Artificial intelligence,  
Bibliometric analysis,  
Science education,  
VOSviewer

### Abstract

Understanding the trends and developments in artificial intelligence research in science education, which has rapidly advanced in recent years, is crucial for technological innovations and applications in education. Therefore, this study examines research on artificial intelligence in science education conducted between 2019 and 2023 (the last five years) through bibliometric analysis, utilizing the Web of Science database. VOSviewer was used for the analysis. As a result of scanning the Web of Science database under these criteria, 867 studies were identified. Analyzing the distribution of these 867 publications by year reveals a concentration in 2023 and 2022. In terms of publication distribution by country, the leading contributors were the USA, China and Australia. The most prolific authors were identified as D. Gasevic, Z. Xiaoming and S. S. Oyelere. Regarding the number of documents and citations in journals, "Computers & Education" ranked first. The most cited used keywords were "Machine Learning" and "Artificial Intelligence", followed by "Learning Analytics", "Data Science" and "Higher Education". The findings illuminate recent research on artificial intelligence in science education. This study is expected to assist researchers in identifying trends within the field and to provide guidance for future studies.

### Introduction

Artificial intelligence (AI) can be defined as the development and application of systems that are capable of thinking, learning, and making decisions in a manner similar to human intelligence (Fogel, 2022). It aims to imitate some aspects of human intelligence by utilising computer science, psychology, philosophy and other disciplines. Nowadays, AI finds widespread use in areas such as healthcare, finance, transportation, education, and entertainment (Mukhamediev et al., 2022).

The relationship between AI and educational systems is becoming increasingly significant (Chen et al., 2022). AI stands out as a technology with the potential to transform educational processes (Jaiswal & Arun, 2021). For instance, AI can create personalized educational programs tailored to each individual's learning style and pace by analyzing student data. Therefore, it helps students learn more effectively (Arslan, 2020; Chassignol et al., 2018). Educational institutions can leverage AI-based systems to collect and analyze large datasets. This advancement enables teachers and administrators to gain a better understanding of student performance and adjust educational strategies accordingly (Buluş & Elmas, 2024). Additionally, AI facilitates the development of innovative tools such as virtual teachers and intelligent learning platforms. These tools enable students to learn in an interactive environment. Consequently, AI exerts a substantial impact on educational systems, making learning processes more efficient and effective (Khosravi, 2022; Schiff, 2021; Ulukök Yıldırım & Sönmez, 2024; Xuesong et al., 2021).

In recent years, AI has increasingly been utilized in science education. A large number of studies have explored the application of AI in various aspects of science education, including teaching, learning, assessment, and curriculum development (Akhmadiyeva et al., 2023; Al Darayseh, 2023; AlKanaan, 2022; Holmes & Tuomi, 2022; Kalogiannakis et al., 2021; Swiecki et al., 2022; Xu & Ouyang, 2022). AI is being applied in diverse ways to enrich students' learning experiences in science classes and to enhance the educational processes of teachers. AI-based simulation software helps students to better understand complex science concepts. For example, in science laboratory courses, AI applications enable students to transfer their theoretical knowledge to practice by conducting virtual experiments (Park et al., 2023). Interactive applications and games that use AI can make the teaching of scientific concepts more engaging and comprehensible (Dimitriadou & Lanitis, 2023). AI also provides virtual teachers that assist students in question answering and guiding their learning processes. Through these applications, students can receive teacher support whenever needed (Chen et al., 2020). In conclusion, AI is effectively utilized in science education across areas such as simulation, interactive learning,

data analysis, and virtual teaching assistants. These applications help students better understand scientific concepts and assist teachers in improving their educational practices (Almasri, 2024).

Due to the increasing number of studies on artificial intelligence, it has become essential to provide a general overview of the research conducted in this area. Bibliometric analysis is the process of quantitatively examining and evaluating scientific publications and citations (Gutiérrez-Salcedo et al., 2018; Ulukök Yıldırım, 2024). It is used to determine the status and trends of scientific activities in a specific field, identify gaps and opportunities in research areas, assess and compare publication performance, visualize collaboration networks among researchers, institutions, and countries, and map scientific fields while tracking their structural changes (Donthu et al., 2021). Bibliometric analysis also examines the number of publications, such as articles, books, and conference papers, published in a particular field or journal, the number of citations, the relationships between related publications, and the frequency and relationships of keywords used in these publications (Akhavan et al., 2016). As a result, bibliometric analysis provides important data for science policy and management by examining the quantitative aspects of scientific communication.

In the field of science education, bibliometric analysis serves as a significant research tool. This analysis reveals innovations and patterns in the field by examining the number of publications related to science education, number of citations, and research trends (Arici et al., 2019). By examining collaborations and interactions among researchers and countries, it uncovers international research networks in science education. Additionally, it is used to identify which topics in science education receive more attention and which research methods are preferred (Comarú et al., 2021). So, these analyses contribute to the development of educational policies and the adoption of innovative approaches in science education.

A bibliometric analysis of studies related to AI can be found in the literature (Al Husaeni et al., 2022; Guo et al., 2024; Kaban, 2023; Khosravi et al., 2023; Lin & Yu, 2024; López-Chila et al., 2024; Oliński et al., 2024; Pradana et al., 2023; Radu et al., 2024; Shang, 2024). For instance, Lin and Yu (2024) reviewed the existing literature on AI chatbots from an educational perspective and aimed to address research gaps. Guo et al. (2024), conducted a bibliographic analysis encompassing 6843 publications over the past decade to identify trends in AI research within the field of education and to understand its development. Khosravi et al. (2023) focused specifically on ChatGPT, performing a bibliometric analysis of the scientific literature concerning chatbots. Their study examined sources, countries, author impact, and keywords, concluding that ChatGPT represents the latest trend in the field of chatbots.

Kaban (2023) examined articles on AI in education and employed bibliometric mapping methods to reveal trends in the field of AI in education across various variables. The study presented results related to the most cited publications, trending topics, thematic maps of keywords, and co-occurrence networks. López-Chila et al. (2024) analyzed the current state of AI in higher education to provide a basis for future research. They performed a bibliometric analysis using the Scopus database for the period between 2017 and 2023. Radu et al. (2024) identified the emerging trends, challenges and new opportunities as a result of AI and Competency Based Education. Al Husaeni et al. (2022) examined the application and use of AI chatbots in the field of education. They conducted a data search in the Scopus database using the keywords ‘chatbot’ and ‘education’ for the research period from 2007 to 2024. Shang (2024) reviewed the existing research on ChatGPT in education using bibliometric analysis methods. The data were collected from English-language studies by searching for the terms ‘ChatGPT’ and ‘Education’. A total of 385 documents were analyzed in this study. Pradana et al. (2023) presented a review of existing research on the use of OpenAI’s ChatGPT in education, employing both bibliometric analysis and a systematic literature review. Oliński et al. (2024) conducted a bibliometric analysis of ChatGPT, an AI tool proficient in text analysis within the social sciences. By utilizing data obtained from the Scopus database, a comprehensive selection of 814 related publications was made and subsequently analyzed with VOSviewer to investigate co-citations, keyword occurrences, and patterns of international collaboration.

Unlike these studies, the aim of this research is to examine the studies published between 2019 and 2023 (the last five years) regarding the use of AI in science education in detail under headings such as year, author, citation, journal, country, and keywords, in order to reveal relationships. Thus, it is expected to provide researchers with an overview of the current situation and developments in the field and to identify new research opportunities. This study aims to answer the research questions listed below:

1. What is the distribution of relevant publications by year?
2. Which countries are the most prolific in terms of publication output?
3. Who are the most prolific and influential authors in the field?

4. Which authors are most frequently co-cited?
5. What are the most cited journals in the field?
6. What are the distributions and trends of keywords used in the literature?

## Method

This study conducts a bibliometric analysis to investigate the current state of international publications in journals indexed in the Web of Science (WoS) database regarding the use of AI in science education. Bibliometric analysis is an approach that quantitatively measures certain indicators by analyzing citations of published works, inter-author relationships, keywords, theoretical and practical topics in a specific field using various statistical techniques. This method allows for monitoring and evaluating the development and advancements within the discipline (Ulukök Yıldırım & Sönmez, 2024). A successful bibliometric study can lay a solid foundation for innovative and meaningful progress in a field, thereby providing researchers with a comprehensive overview, identifying knowledge gaps, generating new research ideas, and positioning their planned contributions within the existing literature.

## Purpose and Limitations of the Study

Due to the fact that the current year is not yet completed, publications from 2024 have not been included in the study. The determination of the data to be used in the study as the last five years, the utilization of the WoS database, the application of VOSviewer software for bibliometric analyses, and the focus on selected titles for network mapping in the analyses constitute the limitations of this research.

## Data Collection Process

While there are many databases indexing educational research, one of the most widely used databases for various analyses, including bibliometric analyses, is the WoS, which is one of the world's most important scientific citation search and analytical information platforms. For this reason, WoS has been utilized in this study due to its provision of a comprehensive data set across different disciplines, includes of the most influential journals and publications in the field, holds the distinction of being one of the oldest and most widely used databases, provides daily updates, and is compatible with VOSviewer, a widely used program in bibliometric studies, that enables file downloads in the desired format. (Hu et al., 2020; Li, Kazak & Kazak, 2023; Rollins, & Yan, 2018). The study includes only articles due to their status as the most common and representative type of scientific publication, their bibliometric indicators, inclusion of original research findings, comparability, and the ability to be analyzed using bibliometric methods (Atmaca Aksoy, 2024). As of September 3, 2024, a total of 867 studies were identified in the WoS database depended on the search criteria presented in Table 1.

Table 1. Article selection process

Parameter	Details
Database	WoS
Keywords	("artificial intelligence" OR "AI" OR "AIED" OR "machine learning" OR "intelligent tutoring system" OR "expert system" OR "recommended system" OR "recommendation system" OR "feedback system" OR "personalized learning" OR "adaptive learning" OR "prediction system" OR "student model" OR "learner model" OR "data mining" OR "learning analytics" OR "prediction model" OR "automated evaluation" OR "automated assessment" OR "robot" OR "virtual agent" OR "algorithm" OR "machine intelligence" OR "intelligent support" OR "intelligent system" OR "deep learning" OR "AI education") and ("science" or "science education")
Research Area	"Education and Educational Research" OR "Education Scientific Disciplines" OR "Education Special"
Publication Type	Article
Publication Years	2019-2023
Indexes	SCI-EXPANDED, SSCI, A&HCI, ESCI
Languages	English
Date	03 September 2024

## Data Analysis

In this study, the bibliographic data of the 867 documents obtained from WoS were imported into VOSviewer (version 1.6.20), a software that allows for the creation and detailed examination of bibliometric maps. The imported format includes publication year, language, journal, title, author, institution, keywords, document type, abstract, and citation count. VOSviewer is a software tool designed for creating, exploring, and visualizing maps derived from bibliometric network data (Van Eck & Waltman, 2010). It can be employed to construct networks consisting of scientific publications, scientific journals, researchers, research institutions, countries, keywords, or terms. The elements in these networks can be connected through co-authorship, co-publication, citation, bibliographic links or co-citation links (Van Eck & Waltman, 2022). In this context, the study identified year, country, journal, citation, co-citation, and keywords. Before each analysis, the relevant data were carefully examined, and essential data cleaning processes were performed. This included the creation of "thesaurus files" for author, journal, and institution names that were written in different languages and scripts, as well as for identical or closely related terms.

## Findings

Under this heading, findings related to distribution of publications by year and country, the most influential authors, journals, and the most used keywords are presented.

### Distribution of the Number of Studies (2019-2023)

Figure 1 presents the distribution of articles obtained from the WoS database, illustrating the number of publications from 2019 to 2023. The data indicates a clear trend in the annual output of scholarly works, revealing fluctuations in publication rates across these years. This temporal analysis offers significant insights into the changing landscape of research activity within the defined timeframe.

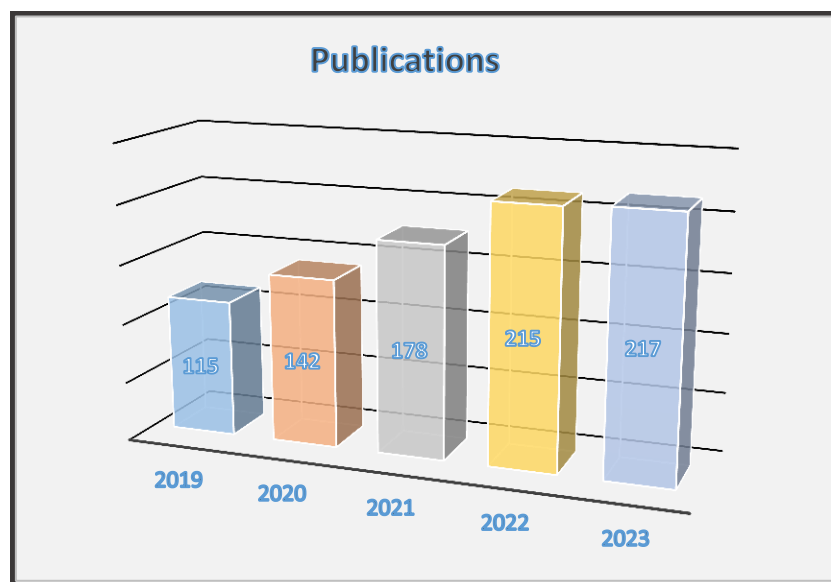


Figure 1. Distribution of publications by year

As can be seen from the Figure 1, studies on AI in science education have increased day by day. In 2023, the number of publications reached the highest level.

### A Distribution of Publications by Country

Table 2 presents the geographical distribution of publications. In analyzing the most productive countries in scientific research, a minimum threshold of 10 documents and at least 1 citation was established. This criterion ensures the inclusion of countries with a substantial impact on the scholarly output, facilitating a clearer understanding of global research dynamics.

Table 2. A distribution of publications by ten top countries

Rank	Country/ Region	Number of publications	Citation	TLS
1	USA	273	2896	104
2	China	93	873	43
3	Australia	52	870	19
4	Spain	51	581	11
5	Taiwan	49	577	12
6	Germany	44	534	47
7	Türkiye	40	290	13
8	England	34	561	10
9	Canada	32	543	23
10	Finland	29	618	17

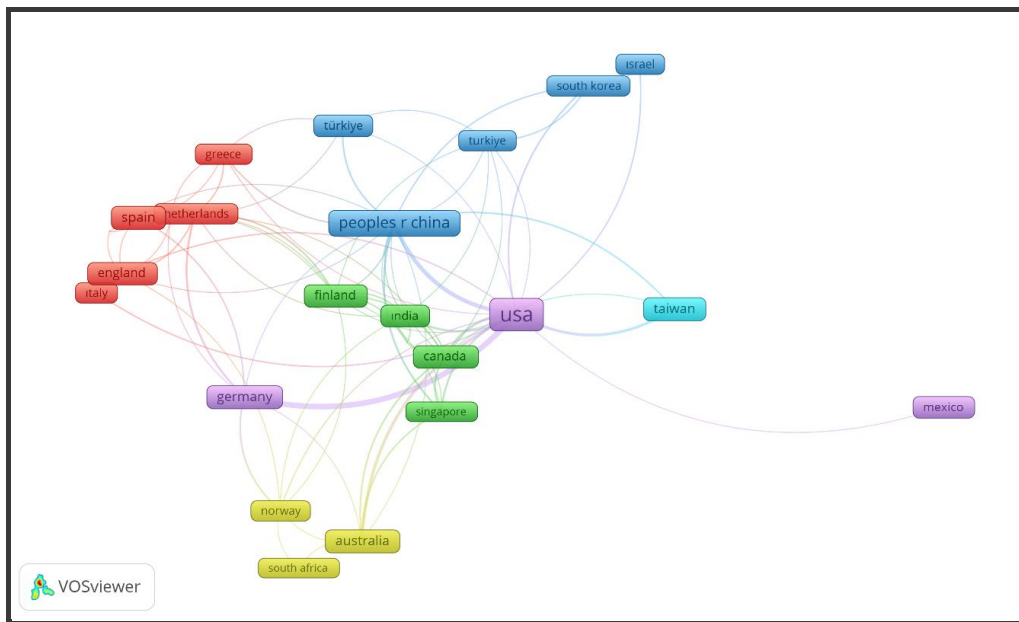


Figure 2. Global collaboration through documents published by country

An analysis was conducted using 24 observation units that exhibit relationships (Figure 2). Six clusters, 76 link and 202 total link strength were identified. The countries with the highest number of citations are the USA (2896 citations), China (873 citations), and Australia (870 citations). In terms of the number of publications, the ranking is as follows: USA (273 publications), China (93 publications), and Australia (52 publications).

### Most Influential Researchers in terms of the Number of Citations

A citation network map was created based on the criteria of at least three publications and a minimum of one citation to identify citation networks among authors (see Table 3 and Figure 3).

Table 3. Ranking of the most influential researchers by citation

Rank	Author	Number of publications	Citation	TLS
1	D. Gasevic	7	244	1
2	Z. Xiaoming	10	204	87
3	S. S. Oyelere	7	188	11
4	X. Wanli	7	157	28
5	H-S. Lee	4	156	26
6	C. Xie	7	141	35
7	F. J. Agbo	3	137	6
8	G. Chen	4	127	24
9	G. Zhu	5	111	7
10	I. T. Sanusi	3	98	0

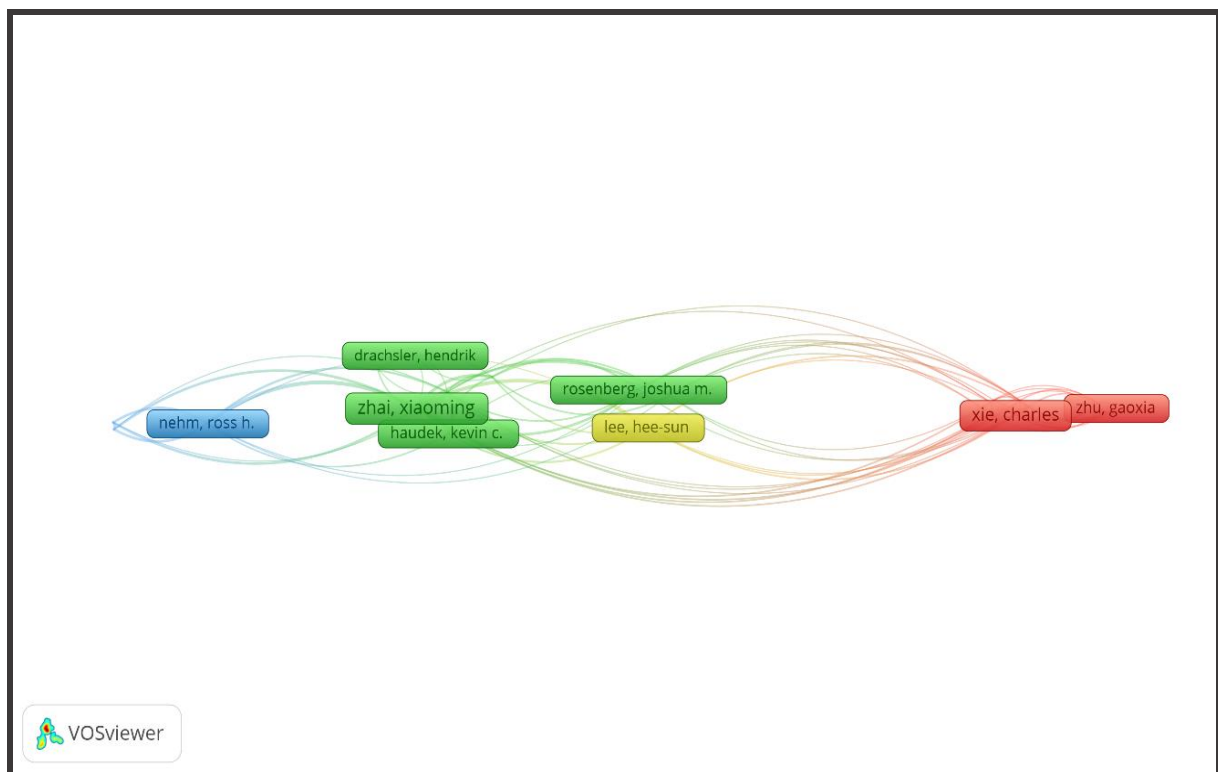


Figure 3. Author citation network

In an analysis conducted on 58 interconnected units, a total of four clusters, 86 link, and total link strength of 237 were identified. The most cited authors are Dragan Gasevic with 244 citations, Zhai Xiaoming with 204 citations, and Solomon Sunday Oyelere with 188 citations. Additionally, 'co-citation' was selected as the type of analysis, with 'cited authors' designated as the analysis unit within the VOSviewer program. A threshold value of 35 was established to minimize clutter in the data visualization. The resulting map is presented in Figure 4.

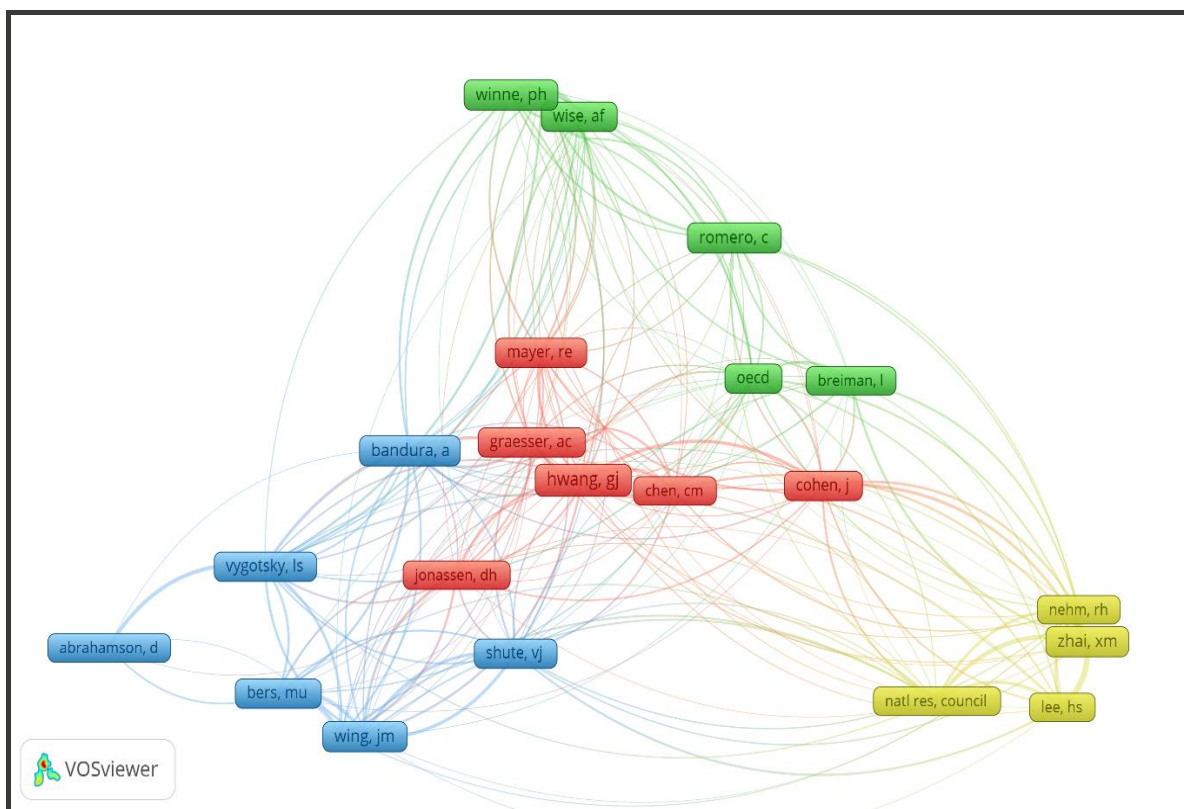


Figure 4. Co-author citation network

Table 4. Ranking of the most influential researchers by co-citation

Rank	Author	Co-Citation	TLS
1	G-J. Hwang	86	139
2	X. M. Zhai	72	284
3	P. H. Winne	67	125
4	J. M. Wing	66	188
5	A. Bandura	65	110
6	C. Romero	65	70
7	D. Gasevic	59	196
8	A. F. Wise	56	142
9	M. U. Bers	55	118
10	OECD	55	99

When the map in Figure 4 and Table 4 are examined, it is seen that there are four different colored clusters related to the common referenced authors. G-J. Hwang is in the center of the names in the red cluster, at the center of the yellow cluster is X. M. Zhai, at the center of the blue cluster is J.M. Wing and at the center of the green cluster is P.H. Winne. G-J. Hwang (86 co-citations), X. M. Zhai (72 co-citations), P. H. Winne (67 co-citations), J. M. Wing (66 co-citations) and A. Bandura (65 co-citations) are highly cited authors.

### The Most Influential Journals in Terms of the Number of Citations

Journals with a minimum of ten publications were included in the analysis. A citation analysis was conducted to identify the most influential publications in the field. The results indicated that 21 out of 232 journals published ten or more studies on the topic (Figure 5). Table 5 presents the top ten most influential journals.

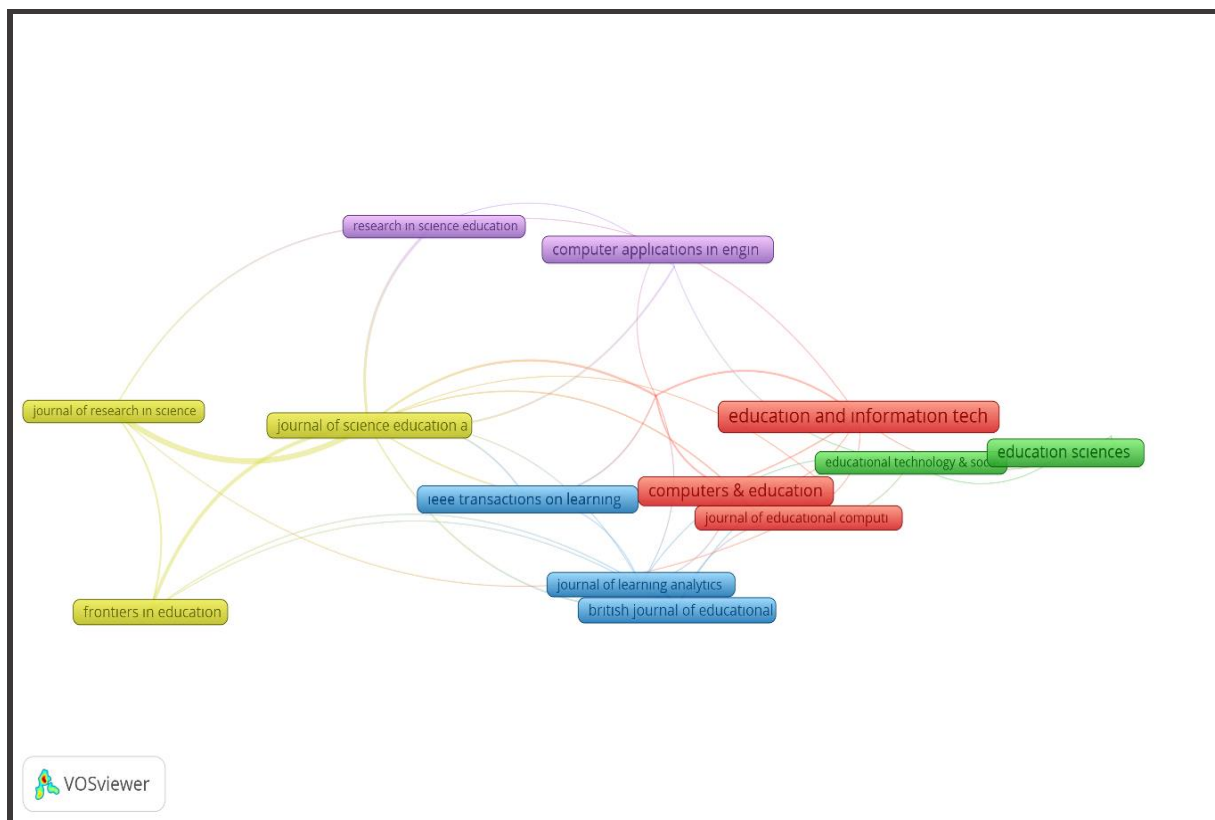


Figure 5. Journal network

According to the findings in Table 5, the journals "Computers & Education" journal is in first place with 90 articles. "Education and Information Technologies" ranks second with 60 publications. It is followed by the "Education Sciences" with 34 articles. The journal with the highest TLS value is "Journal of Science Education and Technology". The most cited journal is "Computers & Education" with 1131 citations.

Table 5. The most influential journals

Rank	Journals	Number of publications	Citation	TLS
1	Computers & Education	90	1131	9
2	Education and Information Technologies	60	588	9
3	Journal of Science Education and Technology	23	535	61
4	British Journal of Educational Technology	22	387	10
5	Education Sciences	34	318	2
6	IEEE Transactions on Learning Technologies	27	266	5
7	Etr&D-Educational Technology Research and Development	19	243	14
8	Journal of Educational Computing Research	12	230	6
9	Computer Applications in Engineering Education	26	210	2
10	Interactive Learning Environments	19	209	6

### Keyword Analysis and Trending Topics

A total of 2661 keywords were used in 867 publications regarding AI in science education. The minimum threshold number of a keyword in VOSviewer is set to 15. As a result of the analysis, 20 keywords, 4 clusters, 101 link and 286 total link strength that met the usage criteria emerged. Table 6 shows the top ten most influential keywords.

Table 6. The most influential keywords

Rank	Keyword	Occurrences	TLS
1	Machine Learning	103	86
2	Artificial Intelligence	103	67
3	Learning Analytics	85	53
4	Data Science	53	70
5	Higher Education	44	41
6	e-Learning	36	26
7	Education	34	41
8	Computational Thinking	32	19
9	Computer Science Education	31	14
10	Data Science Applications in Education	29	5

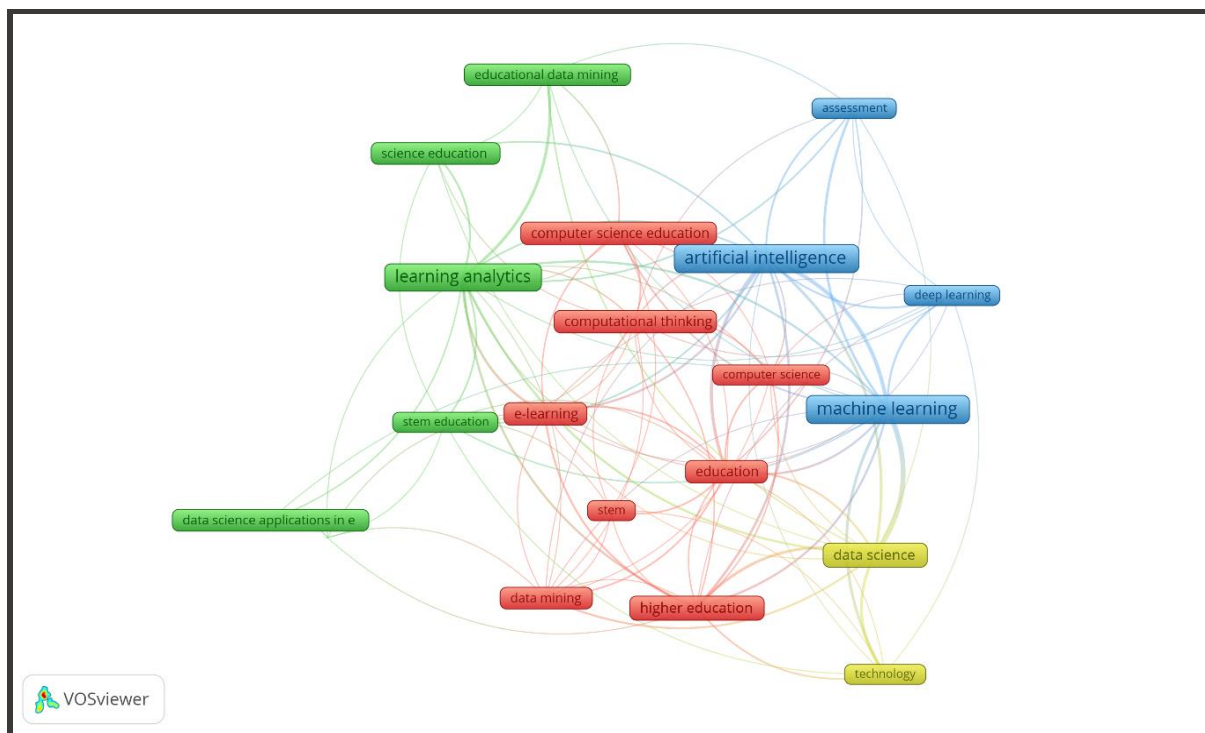


Figure 6. Keywords network

As shown in Figure 6, the first red cluster contains the words computational thinking, computer science, computer science education, data mining, e-learning, education, higher education and stem. The second cluster, colored green, includes adaptive learning, data science applications in education, educational data mining, learning analytics, science education and stem education. The words artificial intelligence, assessment, deep learning and machine learning are included in the blue cluster. The fourth cluster is yellow. The prominent keywords in this cluster are data science and technology. Machine learning, artificial intelligence, learning analytics are the most frequently used keywords.

## Conclusion and Discussion

Publications related to AI in science education were retrieved from WoS and analyzed in an objective and comprehensive manner. The study comprises articles published in English. A total of 867 articles indexed in the WoS database on the topic of AI in science education from 2019 to 2023 (the past five years) were contained in the bibliographic analysis.

The results of this study show that the total of AI studies in science education has increased significantly, especially in recent years. An examination of the distribution of 867 publications related to AI by publication year reveals that the highest output occurred in 2023 and 2022. Notably, approximately half of the articles (49.8%) were published in the last two years. The increase in the number of publications in recent years is attributed to the rapid advancement of technology and the growing importance of AI applications in education. It is anticipated that research in the field of AI within science education will continue to rise in the future (Jia, Sun, & Looi, 2024; Yılmaz, 2024). Talan (2021), examined the articles between the years 2001-2022 in the bibliometric analysis of studies on AI in education. publications. Talan concluded that the number of publications between 2001 and 2004 was low, but there was a notable increase in publications in the subsequent years, with studies published after 2015 accounting for 60% of the total publications. Kaban (2023), examined all studies related to AI in education scanned in the WoS database. As a result of bibliometric analysis, the number of publications started to climb after 2004, and there was a severe increase in 2019. Lopez-Chila et al. (2024) examined 870 articles sourced from the Scopus database related to AI in higher education from 2017 to 2023. Their study found that the number of publications recorded up to 2022 exhibited a continuous increase over the years, with a particularly pronounced rise following the release of ChatGPT, OpenAI's generative AI product, in November 2022. When examining the distribution of studies by country, the USA, China and Australia emerge as notable contributors. The literature indicates that similar findings have been reported, with the USA leading in publications on educational AI (Chen, Xie & Hwang, 2020; Moreno-Guerrero et al., 2020; Song & Wang, 2020).

The top three journals that publish the most articles are "Computers & Education," "Education and Information Technologies" and "Education Sciences" with "Computers & Education" being the most cited journal. This journal is recognized as one of the leading scientific publications with a notable academic impact in the field of educational technology. It is indexed in numerous international databases, particularly the Social Sciences Citation Index (SSCI). Talan (2021) identified "Computers & Education" and the "International Journal of Emerging Technologies in Learning" as the most frequently published journals. Analyzing citation metrics, he found that "Computers & Education", "IEEE Transactions on Education" and the "Educational Technology Society" emerged as the most popular journals based on citations per article. Similarly, Moreno-Guerrero et al. (2020) examined the scientific development of educational AI in the WoS and also noted that "Computers & Education" ranked among the top publishing journal. Additionally, Chen, Xie and Hwang (2020) analyzed studies on AI in education across various dimensions, including grants, conferences, journals, software tools, institutions, and researchers. Their findings revealed that the majority of published research in this area consisted of conference papers, with "Computers & Education" and the "International Journal of Engineering Education" standing out in terms of scientific publications. Likewise, Kaban (2023) concluded that the highest number of articles on AI in education were published in "Education and Information Technologies" and "Computers & Education".

Analyzing the keywords used by authors in publications is a vital method for identifying trending topics and providing insights for researchers engaged in related work (Song et al., 2019). Keyword analysis facilitates a swift determination of the topic and focus of a given publication. Upon examining the keywords, it was found that "machine learning" "artificial intelligence" and "learning analytics" are the most frequently used terms. In the study conducted by Chen, Xie, and Hwang (2020), the most commonly employed keywords related to AI in education were identified as "education", "machine learning", "robotics", "artificial intelligence", "deep learning", "system" and "educational robotics". Guo et al. (2024) performed a bibliometric analysis of Artificial

Intelligence in Education (AIED) to explore the current status and key research trends over the past decade. They categorized the results into three aspects: AIED technology, applications, and subject domain outcomes. Jia, Sun, & Looi (2024) investigated the trends and research focuses of AI in the early stages of education. Their findings highlighted the most prominent keywords and their associated themes in AI in science education, such as science education, robotics, artificial intelligence, and machine learning. In the bibliometric study conducted by Talan (2021) in the field of AI in education, it was observed that keywords such as artificial intelligence, intelligent tutoring systems, machine learning, deep learning, and higher education are positioned at the center of the keyword network map.

When examining authors who have conducted research in the field of AI in science education, it is evident that the most cited authors Dragan Gasevic, Zhai Xiaoming and Solomon Sunday Oyelere. It can be stated that these authors have made important contributions to the field of AI in science education and shaped its direction. It is noteworthy that the number of publications and citations for these three authors are almost proportional. Additionally, they rank among the top three in terms of total link strength. Upon reviewing Dragan Gasevic's works, it is evident that his highly cited and valuable article (2020), "Vision, Challenges, Roles, and Research Issues of AI in Education" addresses the critical roles of AI in education and proposes ten research topics, thereby providing a valuable resource for those entering this field. Zhai Xiaoming has conducted studies across various domains, including education, science education, and artificial intelligence, contributing to the field by delineating the current state of artificial intelligence, key research topics, and educational outcomes over the past decade. His work offers researchers a comprehensive overview of this promising area, providing insights for future research topics and directions. Solomon Sunday Oyelere's research emphasizes the importance of utilizing AI in education and offers recommendations. He highlights the crucial role teachers play in encouraging the next generation of students to engage in AI-supported learning and preparing them for human-AI collaboration in the future.

This bibliometric analysis of AI in science education has revealed a significant increase in the scientific literature regarding the use of AI technologies in this field in recent years. The studies have become increasingly diverse and detailed, indicating that AI technologies are enhancing teaching and learning experiences. However, this field is still relatively new, and understanding the existing research can provide insights into how to optimize the integration of AI into science education. By examining the results of this bibliometric study, researchers will gain information about the most active and collaborative authors in the last five years, the most cited articles, the most frequently used keywords, and the journals that have published the highest number of articles in this area. Equipped with this information, researchers can plan their studies accordingly, significantly contributing to the advancement and development of AI in science education.

In future studies, alongside WoS, other databases such as Scopus and Google Scholar could also be systematically reviewed to expand the search data. Collaborations could be established with prominent authors based on citation and publication counts, and planned research could focus on less commonly used keywords in addition to the highlighted ones, contributing to the field through new studies.

## Scientific Ethics Declaration

The authors declare that the scientific ethical and legal responsibility of this article published in JESEH journal belongs to the authors.

## Acknowledgements or Notes

On behalf of authors, the corresponding author states that there is no conflict of interest

## References

- Akhavan, P., Ebrahim, N. A., Fetrati, M. A., & Pezeshkan, A. (2016). Major trends in knowledge management research: A bibliometric study. *Scientometrics*, 107, 1249–1264. <https://doi.org/10.1007/s11192-016-1938-x>

- Akhmadieva, R. S., Udina, N. N., Kosheleva, Y. P., Zhdanov, S. P., Timofeeva, M. O., & Budkevich, R. L. (2023). Artificial intelligence in science education: A bibliometric review. *Contemporary Educational Technology*, 15(4), ep460. <https://doi.org/10.30935/cedtech/13587>
- Al Darayseh, A. (2023). Acceptance of artificial intelligence in teaching science: Science teachers' perspective. *Computers and Education: Artificial Intelligence*, 4, 100132. <https://doi.org/10.1016/j.caeai.2023.100132>
- Al Husaeni, D., Haristiani, N., Wahyudin, W., & Rasim, R. (2022). Chatbot artificial intelligence as educational tools in science and engineering education: A literature review and bibliometric mapping analysis with its advantages and disadvantages. *ASEAN Journal of Science and Engineering*, 4(1), 93-118. <https://doi.org/10.17509/ajse.v4i1.67429>
- AlKanaan, H. M. N. (2022). Awareness regarding the implication of artificial intelligence in science education among pre-service science teachers. *International Journal of Instruction*, 15(3), 895-912. <https://doi.org/10.29333/iji.2022.15348a>
- Almasri, F. (2024). Exploring the impact of artificial intelligence in teaching and learning of science: A systematic review of empirical research. *Res Sci Educ*, 54, 977-997. <https://doi.org/10.1007/s11165-024-10176-3>
- Arici, F., Yildirim, P., Caliklar, Ş., & Yilmaz, R. M. (2019). Research trends in the use of augmented reality in science education: Content and bibliometric mapping analysis. *Computers & Education*, 142, 103647. <https://doi.org/10.1016/j.compedu.2019.103647>.
- Arslan, K. (2020). Artificial intelligence and applications in education. *Batı Anadolu Eğitim Bilimleri Dergisi*, 11(1), 71-88.
- Atmaca-Aksoy, A. C. (2024). Using technology in science education: A bibliometric analysis. *Journal of Education in Science, Environment and Health (JESEH)*, 10(3), 230-244. <https://doi.org/10.55549/jeseh.730>
- Buluş, B., & Elmas, R. (2024). Yapay zeka uygulamalarının kimya eğitiminde kullanımı alternatif araçlar. *Türkiye Kimya Dernegi Dergisi Kısım C: Kimya Eğitimi*, 9(1), 01-28. <https://doi.org/10.37995/jotesc.1366999>
- Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A. (2018). Artificial Intelligence trends in education: A narrative overview. *Procedia Computer Science*, 136, 16-24. <https://doi.org/10.1016/j.procs.2018.08.233>.
- Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE*, 8, 75264-75278. <https://doi.org/10.1109/ACCESS.2020.2988510>
- Chen, X., Xie, H., & Hwang, G. J. (2020). A multi-perspective study on artificial intelligence in education: Grants, conferences, journals, software tools, institutions, and researchers. *Computers and Education: Artificial Intelligence*, 1, 100005. <https://doi.org/10.1016/j.caeai.2020.100005>.
- Chen, X., Zou, D., Xie, H., Cheng, G., & Liu, C. (2022). Two decades of artificial intelligence in education: Contributors, collaborations, research topics, challenges, and future directions. *Educational Technology & Society*, 25(1), 28-47. <https://www.jstor.org/stable/48647028>
- Comarú, M. W., Lopes, R. M., Braga, L. A. M., Batista Mota, F., & Galvão, C. (2021). A bibliometric and descriptive analysis of inclusive education in science education. *Studies in Science Education*, 57(2), 241-263. <https://doi.org/10.1080/03057267.2021.1897930>
- Dimitriadou, E., & Lanitis, A. (2023). A critical evaluation, challenges, and future perspectives of using artificial intelligence and emerging technologies in smart classrooms. *Smart Learn Environ*, 10, 12. <https://doi.org/10.1186/s40561-023-00231-3>
- Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, 133, 285-296. <https://doi.org/10.1016/j.jbusres.2021.04.070>.
- Fogel, D. B. (2022). *Defining artificial intelligence*. Machine Learning and the City: Applications in Architecture and Urban Design, 91-120. <https://doi.org/10.1002/9781119815075.ch7>
- Guo, S., Zheng, Y., & Zhai, X. (2024). Artificial intelligence in education research during 2013-2023: A review based on bibliometric analysis. *Educ Inf Technol*. <https://doi.org/10.1007/s10639-024-12491-8>

- Gutiérrez-Salcedo, M., Martínez, M. Á., Moral-Munoz, J. A., Herrera-Viedma, E., & Cobo, M. J. (2018). Some bibliometric procedures for analyzing and evaluating research fields. *Appl Intell*, 48, 1275–1287. <https://doi.org/10.1007/s10489-017-1105-y>
- Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education*, 57(4), 542-570. <https://doi.org/10.1111/ejed.12533>
- Hu, G., Wang, L., Ni, R., & Liu, W. (2020). Which h-index? An exploration within the Web of Science. *Scientometrics*, 123(3), 1225-1233. <https://doi.org/10.1007/s11192-020-03425-5>
- Hwang, G.-J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 1, 100001. <https://doi.org/10.1016/j.caeai.2020.100001>
- Jaiswal, A., & Arun, C. J. (2021). Potential of artificial intelligence for transformation of the education system in India. *International Journal of Education and Development using Information and Communication Technology (IJEDICT)*, 17(1), 142-158.
- Jia, F., Sun, D., & Looi, Ck. (2024). Artificial intelligence in science education (2013–2023): Research trends in ten years. *J Sci Educ Technol*, 33, 94–117. <https://doi.org/10.1007/s10956-023-10077-6>
- Kaban, A. (2023). Artificial intelligence in education: A science mapping approach. *International Journal of Education in Mathematics, Science, and Technology (IJEMST)*, 11(4), 844-861. <https://doi.org/10.46328/ijemst.3368>
- Kalogiannakis, M., Papadakis, S., & Zourmpakis, A. I. (2021). Gamification in science education. A systematic review of the literature. *Education Sciences*, 11(1), 22. <https://doi.org/10.3390/educsci11010022>
- Kazak, M., & Kazak, G. (2023). Bibliometric analysis of health tourism studies with Vosviewer. *Journal of Healthcare Management and Leadership (JOHMAL)*, 1, 34-45. <https://doi.org/10.35345/johmal.1253318>
- Khosravi, H., Shafie, M. R., Hajiabadi, M., Raihan, A. S., & Ahmed, I. (2023). Chatbots and ChatGPT: A bibliometric analysis and systematic review of publications in Web of Science and Scopus databases. arXiv preprint arXiv:2304.05436.
- Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y.-S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., & Gašević, D. (2022). Explainable artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 3, 100074. <https://doi.org/10.1016/j.caeai.2022.100074>.
- Li, K., Rollins, J., & Yan, E. (2018). Web of Science use in published research and review papers 1997-2017: A selective, dynamic, cross-domain, content-based analysis. *Scientometrics*, 115, 1-20. <https://doi.org/10.1007/s11192-017-2622-5>
- Lin, Y., & Yu, Z. (2024). A bibliometric analysis of artificial intelligence chatbots in educational contexts. *Interactive Technology and Smart Education*, 21(2), 189-213. <https://doi.org/10.1108/ITSE-12-2022-0165>
- Lopez-Chila, R., Llerena-Izquierdo, J., Sumba-Nacipucha, N., & Cueva-Estrada, J. (2024). Artificial intelligence in higher education: An analysis of existing bibliometrics. *Educ Sci*, 14, 47. <https://doi.org/10.3390/educsci14010047>
- Moreno-Guerrero, A. J., López-Belmonte, J., Marín-Marín, J. A., & Soler-Costa, R. (2020). Scientific development of educational artificial intelligence in web of science. *Future Internet*, 12(8), 1-17. <https://doi.org/10.3390/fi12080124>
- Mukhamediev, R. I., Popova, Y., Kuchin, Y., Zaitseva, E., Kalimoldayev, A., Symagulov, A., Levashenko, V., Abdoldina, F., Gopejenko, V., & Yakunin, K. (2022). Review of artificial intelligence and machine learning technologies: Classification, restrictions, opportunities and challenges. *Mathematics*, 10, 2552. <https://doi.org/10.3390/math10152552>
- Oliński, M., Krukowski, K., & Sieciński, K. (2024). Bibliometric overview of ChatGPT: New perspectives in social sciences. *Publications*, 12, 9. <https://doi.org/10.3390/publications12010009>
- Park, J., Teo, T. W., Teo, A., Chang, J., Huang, J. S., & Koo, S. (2023). Integrating artificial intelligence into science lessons: Teachers' experiences and views. *IJ STEM Ed*, 10, 61. <https://doi.org/10.1186/s40594-023-00454-3>

- Pradana, M., Elisa, H. P., & Syarifuddin, S. (2023). Discussing ChatGPT in education: A literature review and bibliometric analysis. *Cogent Education*, 10(2), 2243134. <https://doi.org/10.1080/2331186X.2023.2243134>
- Radu, C., Ciocoiu, C. N., Veith, C., & Dobrea, R. C. (2024). Artificial intelligence and competency-based education: A bibliometric analysis. *Amfiteatru Economic*, 26(65), 220-240. <https://doi.org/10.24818/EA/2024/65/220>
- Schiff, D. (2021). Out of the laboratory and into the classroom: The future of artificial intelligence in education. *AI & Soc*, 36, 331-348. <https://doi.org/10.1007/s00146-020-01033-8>
- Shang, L. (2024). Bibliometric analysis of potential themes and trend development of ChatGPT in the field of education. *Advances in Educational Technology and Psychology*, 8, 49-55. <http://dx.doi.org/10.23977/aetp.2024.080308>.
- Song, P., & Wang, X. (2020). A bibliometric analysis of worldwide educational artificial intelligence research development in recent twenty years. *Asia Pacific Education Review*, 21(3), 473-486. <https://doi.org/10.1007/S12564-020-09640-2/FIGURES/3>
- Song, Y., Chen, X., Hao, T., Liu, Z., & Lan, Z. (2019). Exploring two decades of research on classroom dialogue by using bibliometric analysis. *Computers & Education*, 137, 12-31. <https://doi.org/10.1016/J.COMPEDU.2019.04.002>
- Swiecki, Z., Khosravi, H., Chen, G., Martinez-Maldonado, R., Lodge, J. M., Milligan, S., Selwyn, N., & Gašević, D. (2022). Assessment in the age of artificial intelligence. *Computers and Education: Artificial Intelligence*, 3, 100075. <https://doi.org/10.1016/j.caeai.2022.100075>
- Talan, T. (2021). Artificial intelligence in education: A bibliometric study. *International Journal of Research in Education and Science*, 7(3), 822-837. <https://doi.org/10.46328/ijres.2409>
- Ulukök Yıldırım, Ş., & Sönmez, D. (2024). A bibliometric look at eye tracking research in video-based learning. *Van Yüzyüncü Yıl University Journal of Education*, 21(2), 378-400. <https://doi.org/10.33711/yyuefd.1378898>
- Ulukök Yıldırım, Ş. (2024). Trends in planetarium research: A bibliometric analysis. *Millî Eğitim*, 53(241), 31-56. <https://doi.org/10.37669/milliegitim.1188601>
- Van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523-538. <https://doi.org/10.1007/s11192-009-0146-3>
- Van Eck, N. J., & Waltman, L. (2022). *VOSviewer Manual*. Universiteit Leiden.
- Xu, W., & Ouyang, F. (2022). The application of AI technologies in STEM education: A systematic review from 2011 to 2021. *International Journal of STEM Education*, 9(1), 1-20. <https://doi.org/10.1186/s40594-022-00377-5>
- Xuesong, Z., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., Liu, J.-B., Yuan, J., & Li, Y. (2021). A review of artificial intelligence (AI) in education from 2010 to 2020. *Complexity*, 18. <https://doi.org/10.1155/2021/8812542>
- Yılmaz, A. (2024). Strengths and weaknesses of teachers' use of artificial intelligence, transhumanism and creativity applications in science education. *International Journal of Eurasia Social Sciences (IJOESS)*, 15(55), 17-36. <http://dx.doi.org/10.35826/ijoess.4448>

---

### Authors Information

---

**Hayriye Nevin Genc**

Necmettin Erbakan University  
Department of Science Education, A. K. Education  
Faculty, Konya, Türkiye  
Contact e-mail: [nnevin85@hotmail.com](mailto:nnevin85@hotmail.com)  
ORCID iD <https://orcid.org/0000-0003-3240-0714>

**Nuriye Kocak**

Necmettin Erbakan University  
Department of Science Education, A. K. Education  
Faculty, Konya, Türkiye  
ORCID iD <https://orcid.org/0000-0002-0531-3538>

---

# Using Artificial Intelligence to Analyze Nature of Science Themes: Einstein and Eddington Documentary Film Example

Yusuf Kartal, Munise Seckin-Kapucu

## Article Info

### Article History

Published:  
01 October 2024

Received:  
10 March 2024

Accepted:  
29 July 2024

### Keywords

Nature of science  
Natural Language Processing  
Latent Dirichlet Allocation  
LDA  
Topic Modelling

## Abstract

Understanding the nature of science is an essential component of scientific literacy. In a technology and media-oriented environment, text-processing algorithms and various artificial learning approaches are crucial and continue to develop. Latent Dirichlet Allocation is a topic modeling algorithm that has been used frequently for many years to extract the main themes in many documents. This study examined the nature of science themes in the documentary film "Einstein and Eddington," which lasted 1 hour and 28 minutes, with the Latent Dirichlet Allocation topic extraction algorithm. First, the texts in the documentary film were fragmented into 30-second, 1-minute, and 3-minute periods to obtain three different datasets. Considering the literature on these datasets, experiments were carried out using the Latent Dirichlet Allocation algorithm to extract 5, 7, and 9 topics. The Latent Dirichlet Allocation algorithm developed with R programming language was used to analyze the data. In the analysis made by the computer, it is seen that science-related words such as science, scientist, theory, the universe, Eddington, think, speed, and answer stand out. While it was observed that it was difficult to distinguish the detected topics from each other on limited data, it was concluded that the dataset created with 30-second periods made more sensitive topic inferences. Despite the challenges posed by subjectivity regarding the nature of science, it is thought that computer-aided models can provide much more than information retrieval and advanced search. In this context, although it seems pretty difficult to extract the nature of science themes on limited data with the Latent Dirichlet Allocation algorithm, it is possible that the artificial learning models to be used in addition to the Latent Dirichlet Allocation algorithm can detect the nature of science themes.

## Introduction

21st-century individuals live in a technology and media-oriented environment. In this environment, individuals use many applications or devices containing advanced technology. Commonly used and developing technologies include augmented reality, virtual reality, mixed reality, mobile applications, web 3.0/4.0, cloud technology, simulation, social networks, educational and digital games, digital storytelling, artificial intelligence, online learning environments, wearable technology, QR Code applications, three-dimensional printing, etc. (Adams Becker et al., 2016; Johnson et al., 2015). These technologies are developing under the leadership of science and reaching individuals.

Societies are increasingly confronted with science and technology, which is its application, day by day. This situation arises from the constantly developing and changing structure of science. Understanding science is necessary to see the results of the application of science and to recognize scientific thought. This is an intellectual necessity for the intellectuals of our age (Yıldırım, 2010). Knowing science and its characteristics is also of critical importance for scientific literacy. According to Shen (1975), scientific literacy has three dimensions: practical, civil, and cultural. Practical scientific literacy focuses on solving the daily problems of life, civic scientific literacy focuses on issues related to the protection of the natural environment, and cultural scientific literacy focuses on knowing the ideas that represent outstanding cultural achievements in science. According to Hodson (2008), in order to understand scientific literacy, it is necessary to understand the nature of science (NOS), understand the basic theoretical frameworks of biology, chemistry, and physics, understand the complex relationships between science, technology, society, and environment, to be aware of contemporary applications of science, to have the ability to use science in solving daily problems. Having a personal opinion on controversial issues from a scientific and technological perspective and a basic understanding of global environmental issues is vital. It is advocated that the NOS should be taught in order to increase scientific literacy (Hand et al., 1999).

It is possible to define the NOS as the way of accessing knowledge or the values and beliefs inherent in developing scientific knowledge (Abd-El-Khalick et al., 1998; Lederman, 1992). Understanding the NOS is seen as an essential component of science literacy (Lederman, 2007) and is emphasized in many reform reports (American Association for the Advancement of Science [AAAS], 1989, 1993; National Science Teachers Association [NSTA], 1982; National Research Council [NRC], 1996).

McComas et al. (1998) explore how science operates, the dynamics of scientists as a social unit, societal influence on and reactions to scientific pursuits, and different facets of social sciences like philosophy of science, history of science, and sociology of science. They view it as a fusion of cognitive sciences like psychology, which delve into interpretations. According to Lederman (1992), the NOS is seen as the values and beliefs inherent in scientific knowledge. Although there is no single, universally accepted definition of the NOS, there is now a significant academic consensus on what aspects of the NOS should be taught in schools (Lederman et al., 2002; Smith et al., 1997; Smith & Scharman, 1999). As a result of scientific studies, it has been determined that some dimensions are necessary for teaching the NOS in formal education (McComas, 1998). These dimensions have been stated as follows. Scientific knowledge is not specific, is based on experiments, is subjective, is partly a product of human imagination and creativity, and is affected by the social and cultural environment. In addition, observation and inference and scientific theories and laws are different types of information (Lederman, 2007; Lederman et al., 2002; McComas, 1998; Smith & Scharman, 1999).

The dimensions of the NOS are not adequately understood by students, scientists, teachers, and teacher candidates (Irez, 2006; Ryan & Aikenhead, 1992; Dagher & Boujaoude, 2005). Many people have misconceptions about the NOS. The sources of these common misunderstandings are textbooks, teachers, and the media (McComas, 2003; Irez, 2006). In the literature, there are studies on films and film genres in the field of education (Başkalyoncu, 2017; Dark, 2005; Efthimiou & Llewellyn, 2006; 2007; Navakanesh et al., 2019; Piliouras et al., 2011; Yıldırım et al., 2015; Yılmaz, 2018). However, some studies examine documentary films, which are a means of transferring science, and their content in terms of the concepts of the NOS (Başkalyoncu, 2017; Seckin-Kapucu et al., 2015; Seckin-Kapucu, 2016; Seckin-Kapucu, 2023). When studies on the NOS are examined, studies on the Latent Dirichlet Allocation (LDA) model attract attention but are limited (Allen & Murdock, 2020; Wang et al., 2023). In a study by Allen and Murdock (2020), applying LDA topic modeling to questions of interest to historians and philosophers of science was discussed and exemplified through Charles Darwin's work on modeling reading and writing behavior. In a study conducted by Wang et al. (2022), prospective teachers' understandings of the NOS were examined with the LDA model.

LDA is a widely used topic modeling algorithm applied to textual data, which forms the basis of many topic extraction models (Li & McCallum, 2006, pp. 577-584). Topic modeling is a common approach used for text processing. It is frequently used to reveal the main themes in documents (Yang & Zhang, 2018). LDA is a statistical method used to identify abstract topics embedded in a set of documents (Blei et al., 2003; Wang et al., 2018). On the other hand, LDA is an effective method that stands out in topic modeling (Zhao et al., 2020). LDA discovers hidden themes in documents. It uses a generative probability model and Dirichlet distributions (Kaya & Gülbandır, 2022). According to the LDA algorithm, all words represent a topic to a certain extent. All documents contain these topics to a certain extent so that each document can be expressed as a mixture of more than one topic. In other words, in LDA analysis, documents are seen as a component of different topics (Calvo et al., 2018; Wang et al. 2023). LDA enables the detection of topics in the entire archive by calculating the probability of a topic being in a document (Kartal, 2017). There are various studies in the literature using the LDA model. When these studies are examined, Facebook, Twitter, Blog, etc. It is seen that social media contents such as these are analyzed (Güven et al., 2018; Lin & He, 2011; Pak & Paroubek, 2010; Roberts et al., 2012; Tur et al., 2013). Lin and He (2009) simultaneously extracted product features and emotional expressions from cinema reviews with the JointSentiment topic model, an LDA-based method they developed.

Many methods are used to determine which NOS theme documentary content belongs to. In this study, computer analysis of the NOS themes in the science content presented, primarily through the media, has been a matter of curiosity. In this study, the LDA algorithm, one of the leading topic modeling techniques, was preferred to detect the nature of the science theme of documentary content. In this context, the study on the film *Einstein and Eddington*, made by Seckin-Kapucu (2016), was requested to be analyzed with artificial learning and analysis methods regarding the elements of the NOS. In this context, the results obtained in the computer environment were examined using the LDA model. This study aimed to examine the NOS themes in the documentary film *"Einstein and Eddington"* according to the LDA Model and to extract the NOS themes in the documentary film. Within the scope of this general purpose, answers were sought to the following questions:

- What are the results of the analysis performed with LDA to extract topics on “Einstein and Eddington Documentary Film”?
- What are the results of the analysis performed with LDA at a dataset divided by 30-second, 1-minute, and 3-minute intervals?

## Method

### Model of the Research

In this study, LDA, a probability-based topic modeling method, was used. The model reveals the semantic structure of the text document by basic topic modeling with the words and weight values formed by the observed dataset (Güven et al., 2018). Topic modeling technology emerges as a new method for large-scale literature research (Yin & Yuan, 2022). As a natural language processing method, topic modeling can discover hidden, valuable, and relevant topics in large amounts of textual data (Jiang et al., 2016; Zhou et al., 2021). This model was preferred in this study because the dialogues in the Einstein and Eddington documentary film needed to be examined in terms of the NOS.

### Data Source

The documentary film "Einstein and Eddington," which is 1 hour and 28 minutes long, was used as the data source in the study. Einstein and Eddington is a 2008 British drama produced by Company Pictures and the BBC in association with Home Box Office (HBO). The film was directed by Philip Martin. The film tells the story of Einstein's theory of general relativity, his relationship with Eddington, and introducing this theory to the World against the background of the Great War and Eddington's eclipse observations. The conversations in the Einstein and Eddington documentary film constitute the text data to which the method used in this study is applied. These text data were organized in 3 different ways to obtain three different datasets to work on (Table 1). The periods for the datasets are 30 seconds, 1 minute, and 3 minutes. The numbers of records are 174, 87, and 29 depending on the period.

Table 1. Information on the period and number of records of datasets obtained from the texts of the Einstein and Eddington documentary film

	<b>Period</b>	<b>Record Count</b>
Dataset-1	30 seconds	174
Dataset-2	1 minute	87
Dataset-3	3 minutes	29

In the referenced documentary (Seckin-Kapucu, 2016), five themes of the NOS were emphasized. For this reason, LDA was used to extract five topics in this study. However, to expand this and measure the general topic extraction abilities of the LDA algorithm, a study was conducted to identify seven topics based on the themes stated by the common opinion of experts in literature (Lederman et al., 2002; Smith et al., 1997; Smith & Scharman, 1999). In addition, considering that different topics could be included, a study was carried out to identify nine topics.

### Process

In the study conducted on the documentary film Einstein and Eddington (Seckin-Kapucu, 2016), the themes of the NOS were directly examined by the researcher, and transcripts of the conversations in the film were obtained for these analyses. By examining these transcripts, the NOS themes were revealed. In order to measure the ability of LDA topic modeling to extract the NOS themes within the scope of this study, these transcripts obtained by Seckin-Kapucu (2016) were divided into three different formats, 30 seconds, 1 minute, and 3 minutes apart, and documents were created (Table 1). In this way, three datasets of different sizes containing the same data were obtained. The LDA topic modeling algorithm was run in 3 different iterations to extract 5, 7, and 9 topics from these datasets, referring to the NOS themes. In this way, nine different combinations were obtained on the same data. By examining the results obtained with the LDA algorithm, the theme of the NOS to which the keywords suggested by LDA might correspond was examined (Figure 1).

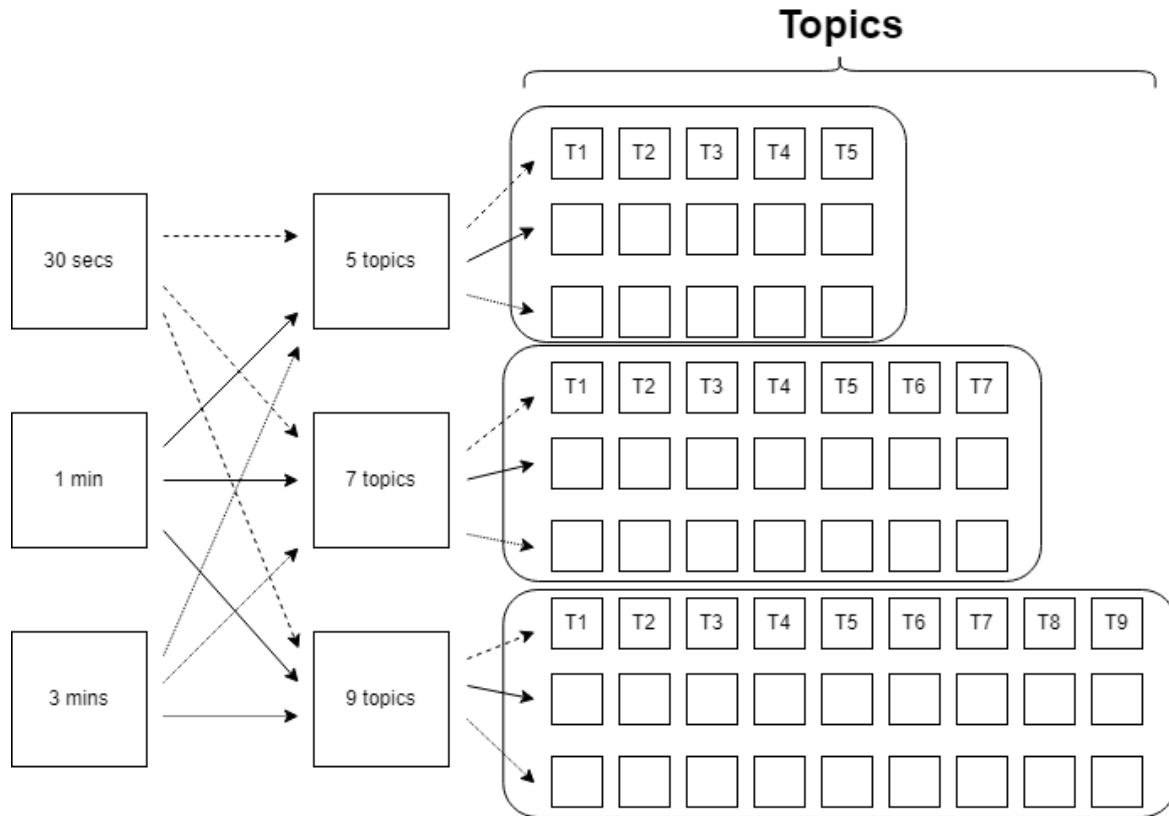


Figure 1. Process steps operated by LDA

### Data Analysis

R programming language was used to analyze the data of this study. R programming language is a statistical software development library with functions that work as open-source scripts (Hornik, 2012). This language, developed in the Department of Statistics at Auckland University, New Zealand, has found widespread use over time (Ihaka & Gentleman, 1996). Unlike other languages, R provides access to data through references, thanks to unique data structures, rather than providing direct access to data (Özdemir et al., 2010). Since it is open source, it has function libraries developed by many users from different fields. Since it works as a script, it is compatible with all operating systems (Çelik, 2018). With the R programming language, nine different processes were run separately for the LDA algorithm to extract five, seven, and nine topics on each dataset.

The LDA function provided by the R programming language was used. The parameters of the LDA function are given in Table 2. *Gibbs Sampling* is chosen as the method for the LDA function. Gibbs sampling is one of the Markov Chain Monte Carlo (MCMC) methods used to sample from high-dimensional probability distributions. Gibbs sampling is beneficial when it is difficult to test all variables of complex probability distributions simultaneously. The method works by sampling each variable sequentially and from a conditional probability distribution.

Table 2. LDA Parameters

Parameter Name	Parameter Value
Method	Gibbs
control.estimate.beta	TRUE
control.seed	2138
control.nstart	1L
control.best	TRUE
control.var.iter.max	20
control.var.iter.tol	1e-6
control.em.iter.max	1000
control.em.iter.tol	1e-3
control.initialize	random

## Results

It was determined that the film examined by Seckin-Kapucu touched on five themes related to the NOS. These are: Scientific knowledge is variable; Scientific knowledge involves logical, mathematical, and empirical inferences; Scientific knowledge is subjective; Imagination and creativity play an essential role in obtaining scientific knowledge; Scientific knowledge is affected by the social and cultural environment during its development and translation into practice. The NOS themes in the movie "Einstein and Eddington," which lasts 1 hour and 28 minutes, are explained by giving examples from the conversations in the movie (Table 3). When the words in each subject in the Datasets are examined, they have different frequency values. The resulting word sets are presented as word clouds so that they can be examined visually more clearly. The words in word clouds can be considered keywords belonging to a topic. Showing a word in a more extensive and bolder font means the subject emphasizes it more.

Table 3. Nature of science themes featured in the movie (Seckin-Kapucu, 2016)

Duration	Dialogue	Scientific knowledge
01:22:17,040 01:22:21,238 01:22:21,280 01:22:24,033	--> then the sun's gravitational field has shifted the stars' position --> and we have a new theory of gravity.	Scientific knowledge is reliable but is subject to change.
00:17:03,120 00:17:07,830 01:09:39,040 01:09:42,316 01:09:45,640 01:09:50,111 01:09:50,160 01:09:52,913	--> What use is science if it has no practical application? --> As we look at the eclipsed sun through the giant telescope, --> We'll take photographs of these stars during the five minutes of eclipse, --> and then compare them to photographs taken of the same stars at night.	Scientific knowledge includes logical, mathematical, or experimental inferences.
01:23:34,880 01:23:40,352	--> Einstein says that time is not the same for all of us...	Scientific knowledge is subjective.
01:07:37,160 01:07:39,276 01:07:43,280 01:07:45,794	--> Pick up the tablecloth --> Space. The tablecloth is space.	Human imagination and creativity have an important role in the acquisition of scientific knowledge.
01:06:19,600 01:06:21,079 01:06:24,840 01:06:28,913 01:06:28,960 01:06:31,030 01:06:33,520 01:06:35,954 01:06:36,000 01:06:39,470	--> Pass, please. --> Would you see that this gets to Cambridge? --> They won't allow me to post it. --> But you... you are... --> They would allow you to. Please.	Scientific knowledge is influenced by social and cultural factors at the developmental stage.

### Experimental Studies on Dataset-1

When the analysis results performed with LDA to extract five topics on Dataset-1 (Table 1) were examined, the word cloud related to the resulting word density is presented in Figure 2. When the word cloud for five topics in Dataset-1 is reviewed, it is seen that the words "science, max, theory, scientist, Eddington" stand out.

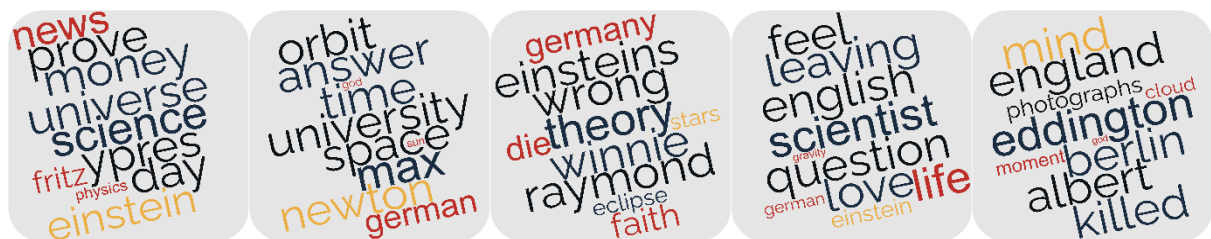


Figure 2. Word clouds of 5 topics extracted by LDA analysis on Dataset-1

When the analysis results performed with LDA to extract seven topics on Dataset-1 (Table 1) are examined, the word cloud related to the resulting word density is presented in Figure 3. As a result of the analysis made to extract seven topics in Dataset-1, the emphasis on the words "science, universe, max, theory, English, feel, Eddington" stands out.

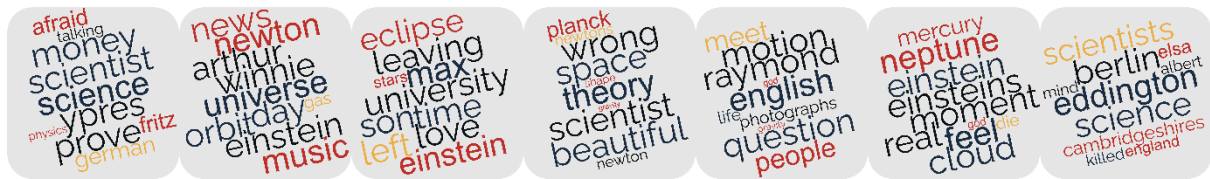


Figure 3. Word clouds of 7 topics extracted by LDA analysis on Dataset-1

When the analysis results performed with LDA to extract nine topics on Dataset-1 (Table 1) were examined, the word cloud related to the resulting word density is presented in Figure 4. In Dataset-1, the words "science, universe, space, theory, Winnie, max, Albert, question, Eddington" are included in the word density resulting from the analyses made to extract nine topics.



Figure 4. Word clouds of 9 topics extracted by LDA analysis on Dataset-1

When the analysis results performed with LDA to extract 5,7,9 topics on Dataset-1 are examined, it is seen that the most emphasized words are "science, max, theory, eddington". These words are followed by the word "universe".

### Experimental Studies on Dataset-2

When the analysis results made with LDA to extract five topics on Dataset-2 (Table 1) were examined, the word cloud related to the resulting word density is presented in Figure 5. As a result of the analysis made to extract five topics in Dataset-2, the words "theory, science, Eddington, Ypres, max" were highlighted.



Figure 5. Word clouds of 5 topics extracted by LDA analysis on Dataset-2

When the analysis results performed with LDA to extract seven topics on Dataset-2 (Table 1) were examined, the resulting word cloud regarding the word density is presented in Figure 6. As a result of the analyses made to

extract seven topics in Dataset-2, the words "Berlin, theory, Eddington, science, scientist, space, max" were seen.



Figure 6. Word clouds of 7 topics extracted by LDA analysis on Dataset-2

When the analysis results performed with LDA to extract nine topics on Dataset-2 (Table 1) were examined, the word cloud related to the resulting word density is presented in Figure 7. As a result of the analysis made to extract nine topics in Dataset-2, there is a density related to the words "max, Winnie, max, time, berlin, theory, science, Eddington, science."



Figure 7. Word clouds of 9 topics extracted by LDA analysis on Dataset-2

When the analysis results performed with LDA to extract 5,7,9 topics on Dataset-2 are examined, it is seen that the mainly highlighted words are "max, science, theory, eddington".

### Experimental Studies on Dataset-3

When the analysis results made with LDA extracting five topics on Dataset-3 (Table 1) were examined, the word cloud related to the resulting word density is presented in Figure 8. As a result of the analysis made to extract five topics in Dataset-3, a density related to the words "science, Eddington, science, space, theory" was formed.



Figure 8. Word clouds of 5 topics extracted by LDA analysis on Dataset-3

When the analysis results performed with LDA to extract seven topics on Dataset-3 (Table 1) were examined, the word cloud related to the resulting word density is presented in Figure 9. As a result of the analysis made to extract seven topics in Dataset-3, the words "time, max, science, theory, science, berlin, Eddington" stand out.



Figure 9. Word clouds of 7 topics extracted by LDA analysis on Dataset-3

When the analysis results performed with LDA to extract nine topics on Dataset-3 (Table 1) were examined, the word cloud related to the resulting word density is presented in Figure 10. As a result of the analysis made to extract nine topics in Dataset-3, the emphasis on the words "max, science, theory, Eddington, science, scientist, theory, space, berlin" stands out.



Figure 10. Word clouds of 9 topics extracted by LDA analysis on Dataset-3

When the analysis results performed with LDA to extract 5,7,9 topics on Dataset-3 are examined, it is seen that the most emphasized word is "science," and this word is followed by the words "theory, Eddington". When the prominent eliminations in the three data sets are examined, it is seen that the majority of the emphasis is on the words "science, theory, Eddington, and max."

## Discussion

When the word clouds that emerged because of the analysis to extract 5-7-9 topics in three datasets made on the Einstein and Eddington documentary film are examined, it is seen that words such as science, scientist, theory, universe, Eddington, think, speed, answer stand out. In addition to that, the LDA model appears as a method used to identify NOS themes by some other studies. In the study, published by Wang et al. in 2023, it was found that other issues other than the NOS were also identified. The 12 topics obtained in the relevant study were associated with eight NOS themes due to manual reviews. The research done by Wang et al. (2022) is similar to the current study. LDA topic modeling not only considers sentences containing terms, keywords, and phrases that contain a view of NOS but also analyzes the entire perspectives of participants (Wang et al., 2023). The words and topics that emerge with LDA are insufficient to determine the nature of science themes. Because the same terms, expressions, or keywords have different meanings in different contexts (Jaeger et al., 2019). If we consider the word "science" without context, the meaning is the same, but the meaning of the word "science" is different in different contexts. The use of the same word in different subjects provides a context for the word, and the meaning expressed is different (Wang et al., 2023). The words that appear in the analysis are words used in the nature of science themes but also words in the content of the documentary film. For this reason, it seems difficult to decide on the nature of science themes based on words.

Another result obtained in this study is that although the data size is the same, the results obtained with Dataset-1, which is handled in 30-second sections, reveal more sensitive subject sections in the context of keywords compared to Dataset-2 and Dataset-3, which have 1-minute and 3-minutes sections. When the studies conducted

with LDA are examined, large data sets are used (Gurcan & Çağıltay, 2019; Na, 2024; Yuan et al., 2020). Since the number of words to be studied in this study was limited, the data were divided according to certain time intervals, and datasets containing the same data but constructed in a different structure were created. The result of this study's division of data sets differs from other studies.

## Limitations

In this section, we will discuss the limitations and threats of our work. Although we have carefully established the research infrastructure, potential threats to the validity of our results and conclusions may still exist. In this section, we consider these threats and their possible impact on the results of our study.

LDA detects unknown issues within the documents given due to its structure and the results it proposes. In this research, the problems related to NOS are tried to be determined, and other issues and scopes may be gathered. This is among the expected threats.

LDA does not deal directly with the titles of the topics. It produces word sets that express a subject in a way that emphasizes the one that expresses it the most. In this case, the necessity of matching NOS topics with human intervention emerges. This situation is a common limitation in almost all of these similar studies.

Internal Validity. Although NOS topics are clear and limited, these topics may not have been mentioned in the documentary contents discussed in the research. In this case, choosing the number of topics 5-7-9 may have prevented more accurate determinations.

External Validity. When checking the results obtained with LDA, the results obtained from a previous study with subjective approaches were taken as references. Although the subjective analyses were carefully conducted, imprecision may be possible.

## Conclusion

This study aims to examine the NOS themes in the documentary film "Einstein and Eddington" according to the LDA Model and to extract the NOS themes in the documentary film. When the word clouds that emerged as a result of the analysis to extract 5-7-9 topics in three datasets made on the Einstein and Eddington documentary film are examined, it is seen that words such as science, scientist, theory, universe, Eddington, think, speed, answer stand out. However, when these words are reviewed to evoke the NOS themes, it is pretty challenging to identify the NOS themes in the documentary film under study. Topic extraction models enable topics other than the NOS themes in this documentary film to be revealed and even come to the fore more dominantly.

Another result obtained in this study is that although the data size is the same, the results obtained with Dataset-1, which is handled in 30-second sections, reveal more sensitive subject sections in the context of keywords compared to Dataset-2 and Dataset-3, which have 1-minute and 3-minutes sections. Although the total number of words is the same, more divisions may cause words to play a more distinctive role in the frequency of their inclusion in the sections. An ideal time interval may have formed more appropriate and unique phrases in this case.

On the other hand, it is predicted that topic extraction algorithms such as LDA can produce more meaningful word clouds on much larger datasets. Since large datasets are needed for artificial learning algorithms to work effectively, it can be said that it is not enough for the algorithm to produce successful results. Since topic extraction methods such as LDA produce results by extracting frequencies on the data without training, the detected phrases also cover topics other than the NOS themes.

In line with the increasing importance of scientific literacy, it is considered that the NOS themes is quite challenging to understand and contains a lot of subjectivity; artificial intelligence-supported systems need to identify these complex topics and make them easier to understand. The results obtained by this study are considered one of the first stage studies on detecting the NOS themes with artificial intelligence models. In this respect, it will guide artificial intelligence-supported studies on the NOS.

## Recommendations

For this reason, it is essential to use artificial learning models trained with the information in the datasets in future studies to focus directly on the NOS themes and make more accurate determinations. It is thought that very successful results can be obtained from analyzing the transcripts of many documentary films that will discuss the NOS with artificial learning models (deep learning, machine learning, artificial neural networks, etc.). For computer-assisted detection of the NOS themes, it is recommended to study artificial learning models trained with data rather than topic extraction algorithms such as LDA. In addition, data obtained from studies on the nature of science or science education can be analyzed using various artificial learning models. Analyses made using different artificial learning models can be compared with each other or with analyses made with computer programs.

## References

- Abd-El-Khalick, F., Bell, R. L., & Lederman, N. G. (1998). The nature of science and instructional practice: Making the unnatural natural. *Science Education*, 82(4), 417-436.
- Adams Becker, S., Freeman, A., Giesinger Hall, C., Cummins, M., & Yuhnke, B. (2016). *NMC/CoSN horizon report: 2016 K-12 edition*. The New Media Consortium
- Allen, C., & Murdock, J. (2020). LDA Topic Modeling: Contexts for the History & Philosophy of Science. Forthcoming in Ramsey, G., De Block, A.(Eds.) *The Dynamics of Science: Computational Frontiers in History and Philosophy of Science*. Pittsburgh University Press; Pittsburgh.
- American Association for the Advancement of Science. (1989). *Science for all Americans*. New York: Oxford University Press.
- American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. New York: Oxford University Press.
- Başkalyoncu, H. (2017). *The effects of documentary films of historical science in the teaching of the nature of science and the particulate nature of matter* [Unpublished master's dissertation]. University of Abant İzzet Baysal.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3(Jan), 993-1022. <https://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf?ref=https://githubhelp.com>
- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27-40.
- Calvo, H., Hernández-Castañeda, Á., & García-Flores, J. (2018). Author identification using latent Dirichlet Allocation. In: Gelbukh, A. (eds) *Computational Linguistics and Intelligent Text Processing*. CICLing 2017. Lecture Notes in Computer Science(), vol 10762. Springer, Cham. [https://doi.org/10.1007/978-3-319-77116-8\\_22](https://doi.org/10.1007/978-3-319-77116-8_22)
- Celik, U. (2018). BIST30 verilerinin R programlama dili ile zaman serisi analizi - Time series analysis of BIST30 data with R programming language. II. *International Symposium on Economics, Finance and Econometrics (ISEFE,18)*, Türkiye
- Dagher, Z. R. & Boujaoude, S. (2005). Students' perceptions of the nature of evolutionary theory. *Science Education*, 89(3), 378-391.
- Dark, M. (2005). *Using science fiction movies in introductory physics*. *The Physics Teacher*, 43, 463-465. <https://doi.org/10.1119/1.2060648>
- Efthimiou, C. J., & Llewellyn, R. A. (2006). Avatars of Hollywood in physical science. *The Physics Teacher*, 44, 28-33. <https://doi.org/10.1119/1.2150756>
- Efthimiou, C. J., & Llewellyn, R. A. (2007). Cinema, fermi problems, and general education. *Physics Education*, 42, 253-261. <https://doi.org/10.1088/0031-9120/42/3/003>
- Güven, Z. A., Erdoğan, R. T., Diri, B., & Çakaloğlu, T.(2018). *n-seviyeli gizli Dirichlet ayırımı ile Türkçe tivit duygularının sınıflandırılması Classification of Turkish tweet emotions by n-stage Latent Dirichlet Allocation*.
- Gurcan, F., & Cagiltay, N. E. (2019). Big data software engineering: Analysis of knowledge domains and skill sets using LDA-based topic modeling. *IEEE access*, 7, 82541-82552.
- Hand, B., Prain, V., Lawrence, C. & Yore, L. D. (1999). A writing in science framework designed to improve science literacy. *International Journal of Science Education*, 10, 1021-1036.
- Hodson, D. (2008). *Towards scientific literacy, a teachers' guide to the history, philosophy and sociology of science*. Sense Publishers.
- Hornik, K. (2012). The comprehensive R archive network. *Wiley Interdisciplinary Reviews: Computational Statistics*, 4(4), 394-398.

- Ihaka, R., & Gentleman, R. (1996). R: a language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 5(3), 299–314.
- Irez, S. (2006). Are we prepared?: An assessment of preservice science teacher educators' beliefs about nature of science. *Science Education*, 90(6), 1113–1143.
- Jaeger, S. R., Roigard, C. M., Jin, D., Vidal, L., & Ares, G. (2019). Valence, arousal and sentiment meanings of 33 facial emoji: Insights for the use of emoji in consumer research. *Food research international* (Ottawa, Ont.), 119, 895–907. <https://doi.org/10.1016/j.foodres.2018.10.074>
- Jiang, H., Qiang, M., & Lin, P. (2016). A topic modeling based bibliometric exploration of hydropower research. *Renewable & Sustainable Energy Reviews*, 57, 226–237. <https://doi.org/10.1016/j.rser.2015.12.194>
- Johnson, L., Adams Becker, S., Estrada, V., & Freeman, A. (2015). *NMC horizon report: 2015 higher education edition*. The New Media Consortium.
- Kartal, Y. (2017). *Topic modelling of TOJDE journal with LDA* [Unpublished master's dissertation]. Anadolu University.
- Kaya, A., & Gülbandılar, E. (2022). Comparison of subject modelling methods. *Eskişehir Turkish World Application and Research Centre Informatics Journal*, 3(2), 46-53.
- Lederman, N. G. (1992). Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of Research in Science Teaching*, 29(4), 331-359.
- Lederman, N. G. (2007). Nature of science: Past, present, and future. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 831–880). Mahwah, NJ: Lawrence Erlbaum Associates.
- Lederman, N. G., Abd-El-Khalick, F., Bell, R. L., & Schwartz, R. S. (2002). Views of nature of science questionnaire: Toward valid and meaningful assessment of learners' conceptions of nature of science. *Journal of Research in Science Teaching*, 39(6), 497-521.
- Li, W., & McCallum, A. (2006, June). Pachinko allocation: DAG-structured mixture models of topic correlations. In *Proceedings of the 23rd International Conference on Machine Learning* (pp. 577–584).
- Lin, C., & He, Y. (2009, November). Joint sentiment/topic model for sentiment analysis. In *Proceedings of the 18th ACM Conference on Information and Knowledge Management* (pp. 375–384).
- McComas, M. R. (2003). A textbook case of the nature of science: laws and theories in the science of biology. *International Journal of Science and Mathematics Education*, 1, 141–155.
- McComas, W. F. (1998). The principal elements of the nature of science: Dispelling the myths of science. In W. F. McComas (Ed.) *The nature of science in science education: Rationales and strategies*. (pp. 53–70). Kluwer Academic Publishers.
- Na, J. C., Kim, E. J., & Kim, J. Y. (2024). Unveiling metaverse social trends: Analysing big data regarding online sports news with LDA-based topic modelling. *Revista de Psicología del Deporte (Journal of Sport Psychology)*, 33(1), 115-125.
- National Research Council (1996). *National Science Education Standards*. National Academy Press: Washington, DC.
- National Science Teacher Association (1982). *Science-Technology-Society: Science education for the 1980s*. Washington, DC: NSTA.
- Navakanesh, B., Shah, A. A., & Prasanna, M. V. (2019). Earthquake education through the use of documentary movies. *Frontiers in Earth Science*, 7, 42. <https://doi.org/10.3389/feart.2019.00042>
- Özdemir, A. F., Yıldıztepe, E., & Binar, M. (2010). İstatistiksel yazılım geliştirme ortamı: R. XII. *Akademik Bilişim Konferansı*, 10-12.
- Pak, A., & Paroubek, P. (2010, May). Twitter as a corpus for sentiment analysis and opinion mining. In *LREc* (Vol. 10, No. 2010, pp. 1320-1326).
- Piliouras, P., Siakas, S., & Seroglou, F. (2011). Pupils produce their own narratives inspired by the history of science: Animation movies concerning the geocentric heliocentric debate. *Science & Education*, 20(7-8), 761-795. <https://link.springer.com/article/10.1007/s11191-010-9321-4>
- Roberts, K., Roach, M. A., Johnson, J., Guthrie, J., & Harabagiu, S. M. (2012, May). EmpaTweet: Annotating and detecting emotions on Twitter. In *Lrec* (Vol. 12, No. 12, pp. 3806-3813).
- Ryan, A. G., & Aikenhead, G. S. (1992). Students' preconceptions about the epistemology of science. *Science Education*, 76(6), 559–580.
- Seckin-Kapucu, M. (2023). Analyzing the documentary film "Marie Curie" in terms of the nature of science themes. *The Eurasia Proceedings of Educational and Social Sciences*, 31, 17-25.
- Seçkin-Kapucu, M. (2016). An examination of the documentary film "Einstein and Eddington" in terms of nature of science themes, philosophical movements, and concepts. *International Journal of Progressive Education*, 12(2), 34-46. [https://ijpe.inased.org/files/2/manuscript/manuscript\\_3/ijpe-3-manuscript-224451.pdf](https://ijpe.inased.org/files/2/manuscript/manuscript_3/ijpe-3-manuscript-224451.pdf)

- Seçkin-Kapucu, M., Cakmakci, G., & Aydogdu, C. (2015). The influence of documentary films on 8th grade students' views about nature of science. *Educational Sciences: Theory & Practice*, 15(3), 797-808, <https://doi.org/10.12738/estp.2015.3.2186>
- Shen, B. S. (1975). Views: Science Literacy: Public understanding of science is becoming vitally needed in developing and industrialized countries alike. *American Scientist*, 63(3), 265–268.
- Smith, M. U., Lederman, N. G., Bell, R. L., McComas, W. F., & Clough, M. P. (1997). How great is the disagreement about the nature of science: A response to Alters. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, 34(10), 1101–1103.
- Smith, U. M., & Scharmann, L. C. (1999). Defining versus describing the nature of science: A pragmatic analysis for classroom teachers and science educators. *Science Education*, 83(4), 493–509. [https://doi.org/10.1002/\(SICI\)1098-237X\(199907\)83:4<493::AID-SCE6>3.0.CO;2-U](https://doi.org/10.1002/(SICI)1098-237X(199907)83:4<493::AID-SCE6>3.0.CO;2-U)
- Tur, G., Celikyilmaz, A., & Hakkani-Tür, D. (2013, May). Latent semantic modeling for slot filling in conversational understanding. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 8307-8311). IEEE.
- Wang, M., Gao, S., Gui, W., Ye, J., & Mi, S. (2023). Investigation of pre-service teachers' conceptions of the nature of science based on the LDA model. *Science & Education*, 32(3), 589-615.
- Wang, W., Feng, Y., & Dai, W. (2018). Topic analysis of online reviews for two competitive products using latent Dirichlet allocation. *Electronic Commerce Research and Applications*, 29, 142-156.
- Yang, S., & Zhang, H. (2018). Text mining of Twitter data using a latent Dirichlet allocation topic model and sentiment analysis. *International Journal of Computer and Information Engineering*, 12(7), 525–529.
- Yıldırım, C. (2010). *Bilim felsefesi*. Remzi Kitapevi.
- Yıldırım, E. G., Köklükaya, A., & Selvi, M. (2015). Öğretim materyali olarak 3-İdiot filmi ile öğretmen adaylarının günlük hayatta fenin kullanımı ve eğitimde aile rolü üzerine görüşlerinin belirlenmesi. *Trakya Üniversitesi Eğitim Fakültesi Dergisi*, 5(2), 94-105. <https://dergipark.org>
- Yılmaz, M. (2018). Filmlerin öğretim materyali olarak kullanılması ve biyoloji eğitimindeki yansımaları. *İnformel Ortamlarda Araştırmalar Dergisi*, 3(2), 24-37. <https://dergipark.org.tr/en/pub/jrinen/issue/42184/427249>
- Yin, B., & Yuan, C. H. (2022). Detecting latent topics and trends in blended learning using LDA topic modeling. *Education and Information Technologies*, 27(9), 12689-12712.
- Yuan, L., Bin, J., Wei, Y., Huang, F., Hu, X., & Tan, M. (2020). Big data aspect-based opinion mining using the SLDA and HME-LDA models. *Wireless Communications and Mobile Computing*, 2020(1), 8869385.
- Zhao, L., Zhao, Q., & Wang, Y. (2020, October). Research on Chinese movie reviews based on latent Dirichlet Allocation Topic Model. In *2020 2nd International Conference on Machine Learning, Big Data and Business Intelligence (MLBDI)* (pp. 46–49). IEEE.
- Zhou, Z. P., Zhou, X. N., & Qian, L. F. (2021). Online public opinion analysis on infrastructure megaprojects: Toward an analytical framework. *Journal of Management in Engineering*, 37(1). [https://doi.org/10.1061/\(ASCE\)Me.1943-5479.0000874](https://doi.org/10.1061/(ASCE)Me.1943-5479.0000874)

---

### Author(s) Information

---

**Yusuf Kartal**

Eskisehir Osmangazi University  
Computer Engineering Dept. Eskişehir/Türkiye  
Contact e-mail: [ykartal@ogu.edu.tr](mailto:ykartal@ogu.edu.tr)  
ORCID iD: <https://orcid.org/0000-0002-0402-1701>

**Munise Seckin-Kapucu**

Eskisehir Osmangazi University  
Faculty of Education, Eskişehir/ Türkiye  
ORCID iD: <https://orcid.org/0000-0002-9202-2703>

---

# Investigation of STEM Supported Engineering and Design Experiences of Students Participating in “Design and Innovation Workshop” in the Context of Situational Interest Theory

Ibrahim Benek, Ismail Donmez, Sahin Idil

## Article Info

### Article History

Published:  
01 October 2024

Received:  
19 March 2024

Accepted:  
02 July 2024

### Keywords

STEM education  
Engineering and design  
Situational interest theory

## Abstract

This study aims to examine the engineering and design-supported STEM (Science, Technology, Engineering, and Mathematics) experiences of students who participated in the "Design and Innovation Workshop" organized at the Science and Art Center (SAC) during the semester break. The experiences and perceptions of 17 students (10 males and 7 females) were analyzed in depth in this study, which was conducted using a phenomenology design from qualitative research methods. Semi-structured interview form, student diaries, field notes and STEM product drawing form were used as data collection tools. The research findings showed that students' participation in design and engineering-based STEM activities enhanced their problem-solving skills, creativity and teamwork. The theory accompanying the research suggests that these activities increase students' interest and motivation in STEM disciplines. The research results suggest that design and innovation workshops can positively influence students' attitudes and skills towards STEM fields.

## Introduction

STEM (science, technology, engineering, and mathematics) education occupies a core position in education systems owing to its capacity to equip the youth with the skills that are relevant in the modern world (McDonald, 2016). Engineering and Design enhances the problem-solving abilities and imaginations of the learners (Daly et al., 2014). These disciplines offer students the opportunity to think critically and generate innovative solutions in the process of transforming theoretical knowledge into practice (Kelley & Knowles, 2016). In Turkey, Science and Art Centers (BİLSEM) play an important role in providing special educational opportunities for gifted children. In terms of physical resources and equipment, it is more favorable than other institutions. Science and Art Centers (BİLSEM) are also significant institutions providing special education programs that help gifted students maximize their potential. In 2022, the Ministry of National Education (MoNE) made an amendment, which brought in other students to be trained in these centers. The aim of the program is to blend scientific thinking and behaviors with aesthetic values through the education provided to students through workshops, and to enable them to develop as individuals who produce solutions, are productive and self-knowledgeable. Also, these trainings aim to help students discover their talents and creative potential early on use these talents at the highest level and acquire new skills (MoNE, 2022). In line with this goal, summer school workshops offer students hands-on learning experiences through engineering and design projects, and their academic and personal development has not yet been sufficiently examined in the literature. In the context of this study, the experiences and perceptions of students participating in design and innovation workshops in summer school workshops towards engineering and design-supported STEM practices were examined.

## Literature Review

### STEM Education

The importance of STEM education has increased worldwide with specific focus on intervariance and multifaceted approaches to development (Zhan et al., 2022). In STEM education, students practice the skills of the 21st century, in particular, complex problem solving, communicating, and working in collaboration (Wu & Anderson, 2015). Integration of STEM in the school curriculum has been beneficial to the knowledge and skills of the students (Robinson et al., 2014). Students are given projects that deal with engineering and design, which

makes it easier for them to put theory into practice. This approach makes the learning process more efficient and worthwhile. The implementation of the new information technology when teaching or learning STEM subjects, such as online communication technologies cause a better performance and attitude towards STEM subjects among students. (Fang et al., 2021). Literature supports the claim that the practice of doing STEM in relevant studies enhances learning outcomes of the students and it also provides solutions to other problems within the disciplines in question (Wahono et al., 2020). The need for integrated STEM education is primarily to develop the need for a STEM identity and also the ability of the teachers to deliver stimulating STEM content and curriculum in elementary school classrooms (Galanti & Holincheck, 2022). Collaboration within STEM in the activities and environments focuses on the design of their outcome to solve a problem and the glue that holds these components together (Juškevičienė et al., 2020). STEM education is positioned as a key driver in the development of critical thinking, creativity, communication skills, and leadership abilities to overcome complex practical challenges. There are various applications of STEM education in the literature. However, most researchers have argued that the engineering-oriented design process is one of the most useful practices for conducting STEM programs (Roberts & Cantu, 2012; Ting, 2016; Forawi, 2018).

### **Engineering and Design Supported STEM**

Engineering and design-enhanced STEM education is an approach that combines design and engineering principles to enable students to explore science, technology, engineering, and mathematics topics. The engineering design process (EDP) can be effectively used to enhance K-12 STEM courses and make STEM concepts more relatable and exciting for students by solving authentic problems through a student-centered approach (Billiar et al., 2014). When engineering design is integrated within STEM education, learners are faced with real-world problems and hence get to know how scientific and mathematical principles are used in practice (Uzun & Şen, 2023). In particular, integrative STEM approaches that have technology and teamwork components have been reported to enhance learning performance in engineering design practices and foster conceptual understanding and higher order thinking skills among learners (Fan & Yu, 2017). In the context of primary schooling, real-life engineering challenges are capable of greatly improving the ability of learners to implement the core principles of STEM. A task involving optical engineering, for example, went much beyond theoretical learning and allowed fifth-grade students to grasp the very practical application of designing, building and redesigning STEM concepts (King & English, 2016).

Studies have revealed that the engineering design process enhances the teaching and learning of STEM subjects tremendously. Schemes of engineering design process allow students to actively participate in activities, interact with each other, and find ways of resolving issues in a creative and systematic manner. This approach aids in deeper appreciation of STEM concepts and inculcates important capabilities like communication, teamwork and hardiness (Sudrajat et al., 2022). Also, the rationale for STEM education has been cited as the incorporation of engineering design into the pedagogical process as a means of enhancing student motivation and performance. Shahali et al. (2016) report that following an extensive integrated STEM education program, middle school students showed a marked increase in the interest for STEM subjects and careers. In the same vein, Fan and Yu (2017), found that students in high school who were taught a STEM engineering module did better in conceptual understanding, reasoning, and performance in engineering design projects than their education technology counterparts. Carr et al. (2012) made a case for the inclusion of engineering components within primary curricular learning standards while Carlson and Sullivan (2004) addressed the issue of design as a factor in interest building towards engineering within the K-12 engineering curriculum. This is why, as Winarno (2020) explains, this method boosts students' interest in STEM disciplines and equips them for the kind of careers which are multidisciplinary and creative in nature. Therefore, the following studies review and present research evidence about the benefits of infusing engineering design throughout the education system to the students' motivation and even performance.

### **Theoretical Framework**

"Situational Interest Theory" explains how individuals' interest in a particular topic or activity is shaped by current circumstances and interactions. "Situational Interest Theory," developed by Hidi and Renninger (2006), explains how situational interest triggered by certain conditions or activities can lead to the maintenance of individual interest in a topic. The Situational Interest Theory identifies four stages of learner interest development, which includes "triggered situational interest", "sustained situational interest", "emerging individual interest" and "well-developed individual interest" (Hidi & Renninger, 2006). In this theory, it is argued that situational interest plays a role of enhancing the mood of the learner which further enhances their

learning processes and learning outcomes (Kron et al. 2022). Additionally, it has been demonstrated that clarity affects situational interest, although its role in the growth of personal interest over time is still being researched (Tröbst et al., 2016). Related to the situational condition of interest, there are studies which argue that an activity's interest is not based solely on the activity, but also on the environment that surrounds the activity (Higgins et al., 2010).

In addition, another finding by Lo (2015) suggests that situational interest can play an important role in influencing students' learning and future behavior, pointing to the importance of situational factors in educational settings. Another study underlines the link between the quality of motivation in learning situations and the development of stable individual interest in a particular domain (Seidel et al., 2005). Conclusions endorse that, the educational environment, or situational factors, can also play a role in affecting an individual's level of interest and motivation in learning. According to Rotgans and Schmidt (2014), posit that a situation becomes interesting when there is a feeling of lacking adequate knowledge, and that level of interest can also go down when the problem becomes overly familiar. This study presents an opportunity to explore some implications of Situational Interest Theory with respect to explaining students' interest in the STEM fields. More specifically, workshops aim to change participants' motivation and interest in STEM areas by providing new and meaningful, interactive ways of learning. In this respect, it can suggest important ways how those involved in education and designing the curriculum by addressing questions of design and innovation workshops guided by Situational Relevance Theory can further enhance the motivation of the students.

### **Innovation and Contribution of the Research**

The objective of the Design and Innovation Workshop Summer School Program is to facilitate students with knowledge and skills necessary in the field of design and issues of innovation. In everyday practices, the basic approach of the program combines many different methods of teaching, while addressing the individual differences and the developmental levels of the students. This approach is meant to facilitate active engagement of students, where the reluctance to embrace innovation and design processes is averted (MoNE, 2022). The organization of the program involves informing the students about the project cycle process, designing as well as manufacturing any technological or innovative products. It aims to provide students with the skills to use innovation processes, to gain experience in three-dimensional solid modeling and design development, and to transform the ideas obtained through innovative thinking into practical applications. Through practical activities such as modeling exercises and the use of electronic components, the program enables students to increase their technological literacy while gaining experience in solid modeling and additive manufacturing techniques. The importance of this program lies in the fact that it guides students towards innovative thinking, creativity and problem-solving skills in parallel with the ever-evolving and changing technology in today's world. Giving students these skills at an early age increases their capacity to cope with the challenges they will face in the future and raises them as active and productive individuals of the information society. The Design and Innovation Workshop Summer School Program has the potential to contribute to both the academic and personal development of students, enabling them to take their place as active and respected individuals in the world of the future (MoNE, 2022). This research aims to fill the gaps in existing studies by examining the effects of design and innovation workshops on student experiences and perceptions, especially in special education centers such as BİLSEM. The research provides valuable findings for educators and policy makers by offering new perspectives on how such workshops affect students' interest and motivation in STEM fields. This study will contribute to the development of educational practices for students and help identify strategies to support their academic and personal development.

### **Aim**

The aim of this study is to examine the engineering and design-supported STEM experiences of the students who participated in the design and innovation workshop in the support and training course implemented in the Science and Art Center (BİLSEM) during the semester break. In line with this purpose, the following sub-questions were sought to be answered:

1. What are the students' experiences of engineering and design-supported STEM applications?
2. What are the students' perceptions of engineering and design-supported STEM practices?
3. How can students' perceptions and experiences towards engineering and design supported STEM implementation be evaluated in the context of Situational Interest Theory?

## Method

### Research Design

A phenomenological, qualitative research methodology was employed to determine students' views on a phenomenon. Phenomenology is used to illuminate phenomena that we are aware of but do not have an in-depth and detailed understanding (Yıldırım & Şimşek, 2013). Phenomenology is a research method that reveals the common meaning of a group of people's personal experiences about a particular phenomenon, fact or concept (Creswell, 2013). A phenomenological methodological approach enables data to be collected to explain the main features and significance of the experience (Merriam, 2013). In this study, students' perceptions and experiences regarding the phenomenon of engineering and design supported STEM application were examined.

### Participants

The participants of the study consisted of students who chose the "Design and Innovation Workshop" within the scope of the support and training course during the semester break at a Science and Art Center located in the east of Turkey in the 2022-2023 academic year. The study group was selected with the "easily accessible case sampling" method, which is one of the purposeful sampling methods. With this method, the researcher gains speed and practicality to the research because the researcher chooses a situation that is close and easy to access (Yıldırım & Şimşek, 2013). The study group consisted of a total of 17 students, 10 boys and 7 girls, attending different primary schools in the province. The ages of the participants ranged between 8 and 10 years old.

### Data Collection Tools

#### *Semi-structured Interview Form*

Semi-structured interview forms, especially in STEM education, allow researchers to delve deeply into the views of teachers, students or other stakeholders about STEM practices, pedagogical approaches or challenges in education. This method provides researchers with an important tool to understand participants' perspectives and gain the knowledge needed to improve educational practices (Kelley & Knowles, 2016). At the end of the implementation, semi-structured interviews were conducted with the students in the study group.

#### *Student Diaries*

Student diaries are used as an important tool in STEM education. Diaries enable students to reflect their experiences, personal observations and learning processes from classroom activities (Aktan & Budak, 2020). At the end of each activity, the students in the study group were asked to write their feelings and thoughts about the application in the form of a diary and the diaries were collected the next day.

#### *Field Notes*

Field notes are very important in STEM education and provide valuable information about the implementation of activities and the perspectives of teachers and students (Altan et al., 2018). Field notes play an important role in enriching special activities such as STEM camps and evaluating studies to improve students' STEM skills (Okulu et al., 2022). In this study, which was conducted according to the phenomenological design, the researcher/teacher acted as a participant observer, watched the groups as they carried out their work, established relationships with them, and took notes of an important event in the order of occurrence as soon as the practice was over.

#### *Document Analysis (STEM Product Drawing Form)*

Drawing in science education has been shown to have multiple benefits for children's learning and development. Research shows that drawing can increase memory retention, help to structure memories in stages and sequences, and provide guidance for recalling and articulating experiences (Bonilla-Sánchez et al., 2022). Drawing can also help children reveal their thinking on abstract or challenging topics, making it a valuable tool

for understanding children's perspectives on complex issues (Brechet et al., 2022). Examining students' drawings of STEM fields (Benek & Akçay, 2018) and employees in STEM fields (Dönmez, 2023) has been used in previous studies. This form was prepared to reveal what kind of tool/invention/product creation ideas students have in their minds for the future. In the form; "Can you draw a tool/invention/product that you want to make in your future life? Can you write the name of this product and what it does?". After the application was completed, the students were asked to fill in the form.

### Research Process

This study, which examines students' design and engineering-based STEM experiences and perceptions, was conducted during the semester break of the 2022-2023 academic year. The study was conducted within the scope of the Design and Innovation workshop opened within the scope of the support and training course opened within BİLSEMs. The study was conducted with students who participated in the Design and Innovation workshop during the semester break at a Science and Art Center in Van city center. In this context, the application was carried out for a total of 16 hours, eight hours in the first week and 8 hours in the second week.

Three separate design and engineering-based STEM activities were conducted during the implementation. The researcher/teacher informed the students about the activities before the actual implementation. In the first class hour, an introductory activity was conducted with the students and student groups were formed. In the second class hour, the groups were given information about the engineering design process. In the third and fourth class hours, a pilot application was carried out with a sample activity. In the light of the data obtained in the pilot application, it was decided that four class hours would be appropriate for each activity. The process followed in the study is presented in detail in the table below.

Table 1. Activities and date/time of realization

Events	Date	Hours
Pilot Application	26.01.2023	4
Energy Activity with Triangular Prism	27.01.2023	4
I Design My Seismograph Activity	02.02.2023	4
Mini Saw Making Activity	03.02.2023	4

During the implementation, three different STEM activities were carried out: "Energy with Triangular Prism", "I Design My Seismograph" and "Mini Saw Making". The activities were carried out by taking into account the outcomes of the Support Education Program (DEP) and Science Curriculum of BİLSEMs. The outcomes related to the activities are given in Table 2.

Table 2. Activities and related outcomes

Activities	Outcomes	Resource
Energy Activity with Triangular Prism	-Explains the concept of energy. -Understands that potential energy turns into kinetic energy.	-Science and Art Centers Science and Technology Supplementary Course Material -Support Education Program (SEP) Science Area Framework Plan
I Design My Seismograph Activity	-Explains the destructive natural events caused by natural processes.	-Science Curriculum (Grade 5)
Mini Saw Making Activity	-Makes his/her own design using an electric motor. -Creates his/her own design using the electric motor details the products. -Recognizes the circuit elements that make up the simple electrical circuit with their functions. -Battery, light bulb, cable and switch are introduced as circuit elements. -Establishes a working electrical circuit.	-Support Education Program (SEP) Science Area Framework Plan -Science Curriculum (Grade 4)

The activities were carried out following the "engineering design process" proposed by Cunningham (2009). The activities related to the design process are described below:

### Activities in the Engineering Design Process

*Ask:* In the first stage of engineering design, the problem situation was identified and students were encouraged to conduct research individually or in groups by searching various sources related to the problem. The groups determined the purpose of their work and planned what they would do for the solution. At the same time, they set criteria and limitations for the product they would make. In other words, they were provided to define the successful operation of the product to be developed and the obstacles in designing the product.

*Imagine:* At this stage, students brainstormed in groups and developed alternative solutions for the product. At the same time, it was ensured that they first made individual drawings about the design and then decided on a common drawing by bringing their individual drawings together. They chose the best solution among the alternative solutions they developed.

*Plan:* At this stage, students made plans about the design they were planning to make. They were made to clearly determine which materials they would use in the design process, where, why and how they would use the materials, and the reasons for choosing the materials. They also determined the conditions under which they would work and the sequence they would follow. As a result of the planning, the researcher/teacher gave the necessary materials to each group.

*Create:* After the planning, the groups created their products by following the necessary steps. They were told that they should pay attention to the predetermined criteria and limitations while creating their products. The products created by the groups were tested in front of the class and it was determined whether they met the criteria and limitations.

*Develop:* In the last stage, the groups further developed their products as a result of the tests. The products that did not meet the predetermined criteria and limitations were revised and retested after revision. Students used knowledge and skills from the fields of science, mathematics, engineering and technology. The association of the activities carried out by the students with STEM fields is given in Table 3.

Table 3. Association of activities with STEM fields

Activities	STEM Fields	Activities carried out
Energy Activity with Triangular Prism	Science	- Learning the definition of the concept of energy, the meaning of potential energy and kinetic energy, the transformation of potential energy into kinetic energy.
	Mathematics	- Learning the definition and properties of an equilateral triangular prism. - Creating a model of an equilateral triangular prism using angle and length measurements with the help of a ruler. - Making distance measurements of thrown balls.
	Engineering	- Creating a product by following the stages of engineering design.
	Technology	- Creating an audiovisual advertisement for the product they created using phones, tablets, computers, etc.
Designing My Seismograph Activity	Science	- Learning the negative effects of natural disasters and determining which precautions should be taken against natural disasters. - Learning what a seismograph is, where and for what purpose it is used and its working principles.
	Mathematics	- Creating the graph of straight and zigzag lines on the seismograph paper. - Interpreting the graph created.
	Engineering	- Creating a product by following the engineering design stages.
	Technology	- Producing audio-visual advertisements about the product they created using phones, tablets, computers, etc.
Mini Saw Making Activity	Science	- Identifying simple circuit components, knowing their functions, and using them to set up a simple electrical circuit.
	Mathematics	- Calculating the number of rotations and the speed of the saw. - Establishing the proportion between the number of teeth on the saw and the amount cut.
	Engineering	- Creating a product by following the engineering design stages.
	Technology	- Creating visual-audio advertisements related to the product they created using phones, tablets, computers, etc.

## Data Analysis

In the analysis of data, content analysis methodology was employed in a manner consistent with the nature of qualitative research methods. Initially, interview recordings, diaries, and field notes were transferred to digital format and readings were made from these. After the readings, coding was done, and related codes were grouped under certain themes. In the coding process, two independent researchers produced separate lists of codes, the obtained codes were compared, and consensus was reached by discussing the diverging parts of the coding. In the data analysis conducted by the researchers, coder reliability was calculated using the reliability calculation formula of Miles and Huberman (1994) and the reliability among coders was found to be 93%.

## Validity, Reliability and Ethics

In qualitative research methods, validity and reliability are ensured through methods such as credibility, transferability, dependability, and confirmability (Merriam, 2013). In this context, to enhance credibility in our study, expert opinion was sought in the preparation of the interview form, diversity was provided in data collection with the triangulation method, student opinions were given directly in data analysis, and participants' confirmation was sought for the data obtained. For transferability, detailed description was made, the obtained data were detailed, and direct quotations were included. Furthermore, purposeful sampling method was used in this context (Yıldırım & Şimşek, 2013).

To increase dependability; similar processes were followed in data collection, it was checked whether the obtained results were consistent with the collected data, and different experts were actively involved in coding. For confirmability in the research, the reached conclusions were continuously confirmed with the collected data. In this context, all data were transferred to a computer environment, and these data were revisited when necessary. Moreover, ethical principles were considered in the research. The study group was informed about the purpose of the research and how the data obtained in the research would be used. They were told that participation was voluntary. In addition, real names of the participants were not used in the study; instead, codes like Participant-1 (P1), Participant-2 (P2) ... Participant-17 (P17) were used.

## Findings

In this section, the findings of the research are discussed under the themes created in relation to the purpose of the study. As a result of the analysis of the data of the research, themes such as "benefit," "emotions," "learning new information," "activities," and "design making skill" were formed. Benefit based on interviews with students, students' diaries, and the teacher's field notes, the application was determined to be beneficial for students. The codes formed under this theme are provided as a word cloud in Figure 1.



Figure 1. Word cloud for the theme of benefit

As shown in Figure 1, based on the responses received from the participants in the study, seven (7) different codes have been created for the "benefit" theme. Under this theme, seven (7) students (P3, P4, P5, P6, P8, P10, P11) mentioned that the application helped them learn new information, three (3) students (P2, P6, P11) noticed

an improvement in their design skills, three (3) students (P7, P8, P10) developed innovative ideas/thoughts, two (2) students (P2, P4) felt an enhancement in their intelligence, one (1) student (P1) used their hands and fingers better, one (1) student (P1) observed an increase in their attention level, and one (1) student (P6) became more socialized. Additionally, eight (8) students (P1, P4, P5, P6, P7, P8, P9, P11) expressed their interest in participating in similar new workshops. Some of the students' opinions related to this theme are as follows:

*S1: "Thanks to this course, I used my hands and fingers better. I learned to be more careful and meticulous while making a design. I would like to participate again in such activities and courses because I really enjoy designing. I never get bored or tired while engaging in these activities."*

*S2: "My design skills improved. I realized my intelligence developed."*

*S3: "I learned new information. For example, I learned what a seismograph is, what it's used for, and where it's used."*

*S4: "I would like to participate again if this course is organized. Because it developed my intelligence and skills. Also, it was very fun."*

*S7: "I would like to participate again in this workshop because learning educational and innovative things is very beneficial."*

*S8: "It helped me develop better thinking for inventing new things. It allowed me to understand and make a design. I would apply and participate again if this workshop is organized because we spend quality time here. Instead of wasting time at home, I learn to invent here."*

*S11: "I learned new information. I learned how to design. Now, I too can design something like an engineer, an explorer, a scientist."*

Participants completed diaries on the days of the activities and wrote their thoughts and opinions in their diaries about learning new information, the usefulness of the activities they participated in, spending productive time during the application, and the development of their design skills. Some of the expressions written by students in their diaries are as follows:

*P8: "Hello diary. My name is P8. I registered for BILSEM during the semester break and joined the design and innovation workshop. Our teacher said we would be doing designs and activities related to STEM in this workshop. I really had fun today. Because I learned new things and had a good time with our friends. I'm glad I came here. Because we spend our time productively here. For example, today we learned what energy is, and how potential energy and kinetic energy are formed. Our teacher told us about some steps, and we completed our design by sticking to those steps. We used materials like skewers, electric tape, scissors, hard plastic cups, pens, rulers, metal tape, rubber. Our group finished the design nicely and it turned out great." (Date: 27.01.2023)*

*P13: "A seismograph is a device that measures earthquakes. I learned this today. I think it's a very nice thing. Imagine. There's an earthquake and you can measure its intensity. Our group tested our seismograph. The teacher was shaking the table, and we were slowly pulling the seismograph paper. It was like there was an earthquake. I enjoy learning new and different information like today." (Date: 27.01.2023)*

The usefulness of the activities for students is confirmed by the data obtained from the researcher/teacher's field notes. On 27.01.2023, the researcher/teacher noted: *"The necessary materials were provided to the students for the product to be made, and the students tried to make their designs in groups by following the engineering design stages. Students mentioned that they learned what an equilateral triangle prism is during the process of creating an equilateral triangle with skewers. During the testing of the design, they learned how flexibility, potential energy, and kinetic energy are formed, and found learning this enjoyable."* and on 03.02.2023: *"It was observed that students now design products more professionally using engineering design stages. Students who struggled during the 'Triangle Prism with Energy Activity' on the first day had no difficulty in today's activity. It was noticed that today's application helped students use their psychomotor skills better. In addition, it was found that more innovative and creative ideas emerged during the 'Ask' and 'Imagine' stages."* These field notes are considered under this theme and confirm the codes related to the "beneficial" theme.

## Emotions

The theme of emotions was formed based on the analysis of the opinions of the students in the study group and the field notes of the teacher. The codes under this theme are provided as a word cloud in Figure 2:



Figure 2. Word cloud for the theme of emotions

As depicted in Figure 2, based on the responses received from participants in the study, five (5) distinct codes have been formed under the "emotions" theme. Upon examining student responses under this theme, eight (8) students (P1, P3, P4, P5, P6, P7, P8, P10) stated they had a lot of fun and found the activities very entertaining during the application process, four (4) students (P3, P6, P9, P10) felt very excited and found the activities exciting, three (3) students (P8, P10, P11) enjoyed their time throughout the activities, one (1) student (P10) mentioned that their curiosity was piqued during this process, and one (1) student (P4) expressed an increased interest in the activities. Some of the students' opinions related to this theme are as follows:

*P1: "I had a great time. I found the activities very entertaining."*

*P2: "The activities were very nice. I made new friends during the activities, had good relations with our group members, and we could talk and chat with each other."*

*P3: "It was very fun. I was especially excited during the saw and energy activities. My excitement doubled in the seismograph activity."*

*P4: "The activities we did really caught my interest. I loved them. Throughout the process, my group mates and I had a lot of fun."*

*P9: "We did some activities in groups, and these group activities were exciting."*

*P10: "The activities we did were fun and exciting. Also, every day at home, I kept asking myself, 'I wonder what activity we will do tomorrow?' because I was very curious about the activity we would do."*

Upon reviewing the participants' diaries, it is observed that the majority wrote about how the application positively affected their emotions. Some of the feelings and thoughts written by students in their diaries are as follows:

*P3: "Hello diary. Today in the workshop, we designed a seismograph. First, we discussed with our group mates what we would do. Then we asked our teacher for the necessary materials and started making our design. We first cut the middle of the plastic cup to the size of a coin. We passed a felt-tip pen through the cut and fixed it with play-dough. Then we tied it to a shoebox with string. We obtained a long roll using A4 papers, ruler, scissors, and tape given by our teacher. After each group finished their design, they presented their seismograph. Each group tested their seismograph in turn. Ours was good. Zigzags formed on the paper. Those zigzags were the measure of the earthquake. In this process, I gained a lot of knowledge and had enjoyable times with our friends." (Date: 02.02.2023)*

*P17: "Dear diary, today at the Science and Art Center, we designed a seismograph to measure earthquakes. After nice work, we made our seismograph as a group. It turned out very well because we made it with team spirit and perseverance. The activity we did before was very nice. That's why I was excited about this activity. When I heard we were going to make a seismograph, I felt strange. Because it was the first time I heard this name. It turns out it's a very nice thing. We had a good time with our friends today. It was an enjoyable activity for me." (Date: 02.02.2023)*

The positive contribution of the application to the affective skills of the participant students is also confirmed by the field notes kept by the researcher/teacher. On 02.02.2023, the researcher/teacher noted: *"One of the most notable events today was some students coming to me in the corridor before the class started, asking me what activity we would do today. The students were curious about today's activity since they enjoyed the previous one. Their excitement was observed. Students had fun and spent a good time while designing and testing their seismographs."* This field note supports the codes formed under this theme.

## Learning New Information

Based on the responses received from the participants, the theme of learning new information has been formed. The codes under this theme are presented as a word cloud in Figure 3:

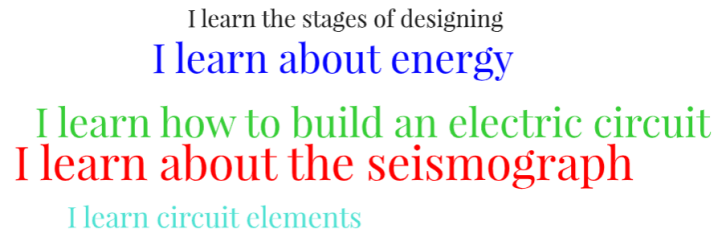


Figure 3. Word cloud for the theme of learning new information

As shown in Figure 3, in line with the participants' opinions in the study, the theme of "learning new information" has been created, and under this theme, five (5) distinct codes have been formed. Upon reviewing student opinions, it was found that all participating students reported learning what a seismograph is and its purpose, how to construct a simple electric circuit, and the properties of kinetic/potential energy. Additionally, six (6) students (P3, P4, P5, P6, P7, P8) mentioned learning the functions of electrical circuit components, and one (1) student (P4) reported learning the stages of making a design. Some responses from the participants are as follows:

*P1: "In the seismograph activity, I learned that a seismograph is a tool for measuring earthquakes. In the Triangle Prism activity, I learned about kinetic and potential energies. In the saw activity, I learned how to set up an electric circuit."*

*P2: "Even though it was in a beginner's fashion, thanks to the seismograph we made, I can measure earthquakes. I learned what potential and kinetic energy are. I learned how to set up an electric circuit."*

*P3: "I learned that a seismograph is a tool that measures earthquakes. Also, I learned the concepts of potential and kinetic energy. I learned the functions of electric circuit components."*

*P4: "I learned the stages of making a design. I learned all the details of a design. I learned the benefits and harms of the product we made."*

*P6: "I learned how earthquakes happen and what their effects are. I learned how to measure the magnitude of an earthquake and which instrument is used for it. Also, I learned how to set up an electric circuit using a switch, battery, DC motor, and wires. I learned that moving objects have kinetic energy and objects at height have potential energy."*

*P8: "I learned how and with which instrument earthquakes are measured. Besides, I learned what potential and kinetic energy are. Also, I learned the functions of electric circuit components and how to set up an electric circuit."*

*P9: "With the design we made, I learned that flexible objects store potential energy, and when these objects are released, that potential energy turns into kinetic energy. Besides, I learned how earthquakes are measured thanks to the seismograph activity."*

Students also mentioned learning new information in their diaries. Here are some expressions written by students in their diaries:

*P5: "Today at BİLSEM, we did a new activity. Thanks to this activity, I learned that a seismograph is an earthquake measuring device. Actually, we did some research about seismographs before making the design. As a result, we learned that seismographs have been used for a very long time. I had heard of this device for the first time. I brought the seismograph I made home. My brother was also very interested in it." (Date: 02.02.2023)*

*P6: "Dear diary, today in the workshop, we learned how potential energy is converted into kinetic energy. We did an application of this. It was a very exciting activity. The teacher gave us materials. First, we made a triangle prism out of skewers. Then we cut the cup and mounted it to our equilateral triangle prism with a rubber band. Then the bell rang, we played. When we came back to class, we started testing our designs. When we pulled back the cup, potential energy was building up, and when released, the ping pong ball gained speed and moved away. Here, kinetic energy was happening. After the class ended, we dispersed and went home." (Date: 27.01.2023)*

## Activities

At the end of the application, students participating in the study were asked to write down their favorite activities during the application, ranking them from 1 to 3.



Figure 4. Word cloud for the theme of activities

As shown in Figure 4, among the "activities" theme, participants in the study expressed enjoying the "mini saw making" activity the most (8 students). This activity was followed by "Energy Activity with Triangular Prism" (3 students) and "Designing My Seismograph" (1 student). The researcher/teacher noted on 03.02.2023, *"I think their favorite activity was the 'mini saw making'. They also liked the other two activities, but their enjoyment of this activity was evident in their eyes. I guess they liked setting up the electrical circuit and making the product work through a correctly assembled circuit."* This field note, upon analysis, confirms the codes within this theme are valid.

## Design Making Skill

Nine participants in the discussion stated that the application improved their design-making skills, positively influenced their desire and thoughts about creating an innovative tool/product/invention, prepared them for making new designs, enhanced their imagination and creativity, built their confidence towards designing a product, and accelerated the idea/thought of making a product. Their responses are as follows:

*P1: "Through the activities we did, I feel my ability to design a tool has improved. Now, I won't have any trouble doing my activities at home. Actually, I used to make designs at home, but the ones here are more professional. The designs we made here sparked new ideas in my mind. I started thinking about what kind of designs I could make."*

*P2: "I now have confidence in myself about making a tool, product, or invention. I've learned how a tool is made and which stages to follow. It increased my imagination, my ideas, and thoughts developed. My curiosity to make a discovery has grown. Now, making a new discovery seems more appealing to me."*

*P3: "Thanks to the activities we did, I think both my ability to create a new tool and my design-making skills have improved. My desire to make an invention has increased. Now, at home, I constantly think about what I can design or develop."*

*P5: "Now, I feel ready to use my intelligence more to create new things."*

*P8: "Thanks to what we learned in this workshop, I can think about and immediately design new things. It developed my idea and thought of making a new tool or design. I think most importantly, I can express my new ideas and thoughts more easily, and the person in front of me understands me better."*

The positive impact of the application on students' design-making skills is also evident from the field notes kept by the researcher/teacher. On 03.02.2023, the researcher noted: *"Concerning the saw to be made, students first identified the problem in groups and brainstormed to come up with the best solution. Then they made drawings related to their design and took note of the necessary materials. Afterward, they requested the necessary materials from me and presented their products, adhering to the other stages of engineering design. They fully utilized the engineering design stages they learned in both the pilot study and previous applications. Today, it was observed that the groups proposed different solutions, wrote more options, thought, and made decisions faster."* This note confirms the validity of the "design-making skill" theme.

## Participants' Drawings about a New Product

Participants were asked to make drawings of a tool/product that came to their mind after the application. The drawings made by the participants are as follows:

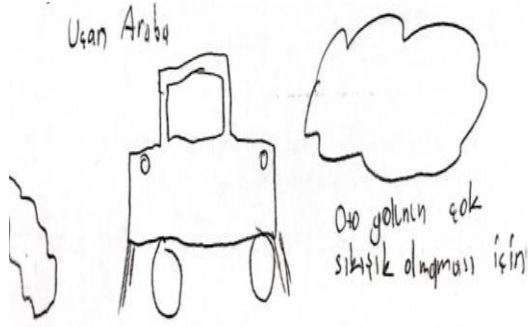


Figure 5. Drawing by P2 (flying car)

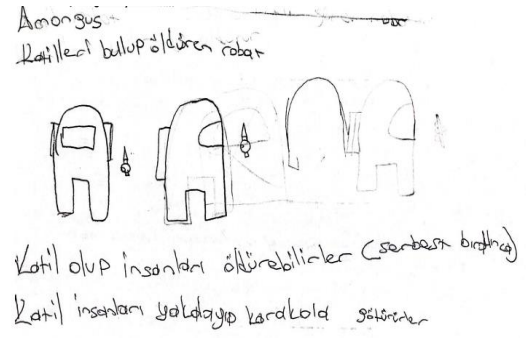


Figure 6. Drawing by P3 (A robot that finds and kills weeds)



Figure 7. Drawing by P4 (solar system detecting system)



Figure 8. Drawing by P5 (time machine)

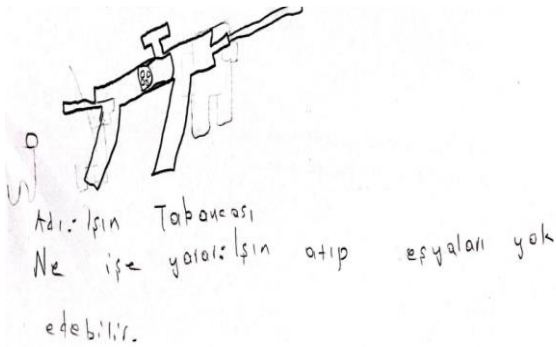


Figure 9. Drawing by P6 (ray gun)

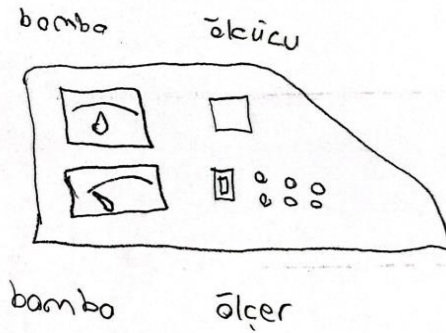


Figure 10. Drawing by P7 (bomb detector)

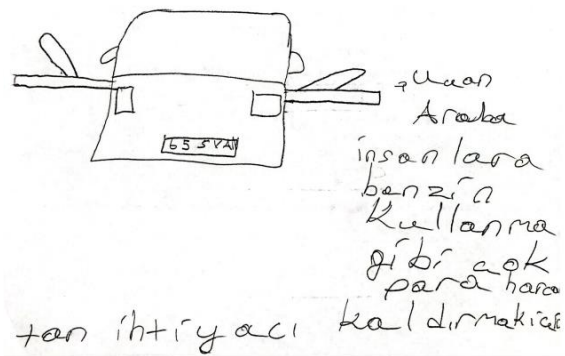


Figure 11. Drawing by P8 (flying car)

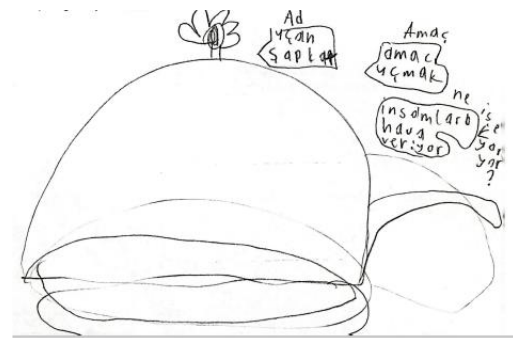


Figure 12. Drawing by P9 (flying hat)

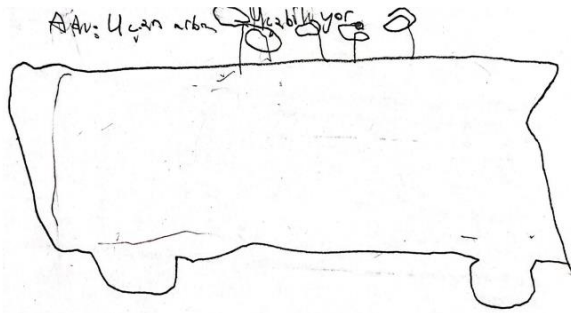


Figure 13. Drawing by P10 (flying car)

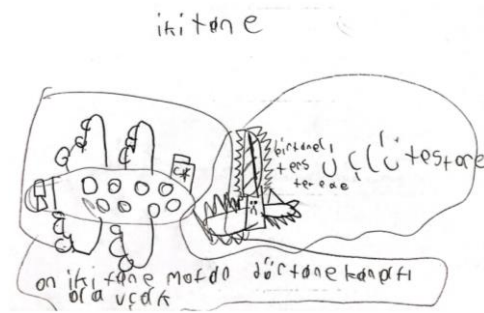


Figure 14. Drawing by P11 (airplane)

When analyzing the drawings made by the students, it was observed that they drew products named "Flying Car" (P2, P8, and P10), "A Robot that Finds and Kills Weeds" (P3), "Time Machine" (P5), "Solar System Detecting System" (P4), "Ray Gun" (P6), "Bomb Detector" (P7), "Flying Hat" (P9), and "Airplane" (P11). P2 suggested the idea of a flying car as a solution to the traffic problems in their area. P3 proposed a concept aiming at the quick apprehension of individuals committing crimes in society. P8, aware of the high economic value of fuels used to power cars, thought of escaping the high cost of fuels with the design of a flying car. P11 aimed at creating a new generation airplane. It's noteworthy that half of the students' drawings related to the concept of "flying". Overall, it can be concluded that students made drawings aimed at solving real-life problems. From this, it can be stated that the application positively influenced the students' creativity, innovation, and skills in solving societal issues.

## Discussion

The implications of the 'Design and Innovation Workshop' have been demonstrated in the existing literature on student engagement in engineering and design in STEM education. One of the implications of the study is that students enhance their problem-solving, creativity, and teamwork skills after engaging in engineering and design educational activities. This finding is consistent with earlier studies that support the notion of providing practical learning experiences for the students in the 21st century (Daly et al., 2014; Kelley and Knowles, 2016). Also, the propensity and willingness of participants particularly to engage in STEM subjects raises the need to implement design and innovation workshops in STEM education. This is particularly important in the literature as it proposes students in STEM fields should be given more interesting active forms of learning in order to make them stay in such fields (Wu and Anderson, 2015; Fang et al., 2021).

In a broader overall picture of the research conducted in STEM education, the outcome also indicates that conducting design and innovation workshops is not only meant to teach technical training, but also a way of enhancing the acceptable social behaviors and attitudes in STEM. This supports previous research that encourages the blending of engineering theories as well as design practices in the teaching of STEM subjects so as to increase the effectiveness of the learning (Fan and Yu, 2017; Wahono et al., 2020). The study was the finding that students' social skills and attention spans also improved significantly, workshop sessions were capable of enriching the educational experience towards all-rounded growth. The accounts of students engaged in the "Design and Innovation Workshop" were assessed through the prism of "Situational Interest Theory." The students' high levels of interest and motivation in the course of participation in the workshop activities are in line with the tenets of the theory which further posits that some conditions and interactions may foster individual interest.

The workshops were showcased to initiate, support and further develop the individual interests of the students in STEM fields which turned into a well nurtured personal interest. The Situational Interest Theory considers the development of learner interest to consist of four stages, namely: Triggered Situational Interest, Maintained Situational Interest, Emerging Individual Interest, and Well-Developed Individual Interest. The first stage of "triggered situational interest" was in design, and engineering-based STEM activities, which incited students upon the use of new and interactive design features, hence commencing their learning. The workshop activities enabled students to sustain their interest. While being in one of the "maintained situational interest" stages, the students kept and developed their interest, solving problems in a creative manner, and in a team. More so, during "Emerging Individual Interest" the students experienced an enhanced drive and interest in STEM fields. This could have resulted in students developing a need to delve more into the STEM areas after some time. However, there is no clear evidence on how the research helped students acquire "well-developed individual

interest” in the STEM areas of their choice. In addition to this, further studies may seek to investigate students’ levels of interest and motivation toward STEM aspects after undertaking design and innovation workshops over time.

## Conclusions

The study on students' interaction with engineering and design-based applications of STEM has shown that students acquired additional knowledge on topics such as energy, electric circuits, seismographs, and the engineering design spiral. This added knowledge helped in the students’ comprehension and the growth of the body of knowledge in the areas of science, Technology, engineering, and Mathematics. Among the activities that were carried out, “making mini saws” was highly rated by the learners. The different levels of enjoyment of the activities show the different interest areas and the ability of doing things concrete to encourage students. The applications sharpened students’ skills in making designs, developed their imagination and creativity, and made them more self-assured in developing new designs. Real life’s problems were the motivating force in students designing solutions. Real life problems were evident in the creative art work of the students. Drawings particularly associated with the idea of “flying” illustrated the diversity and inventiveness of students’ minds. The study indicates that STEM education delivered on the backdrop of design and engineering helps in improving the scientific knowledge and skills, creativity, and problem-solving capabilities of the students. Upon students’ attitudes’ assessment in engineering and design-related tasks, STEM was seen as useful to the students. They learned new things, improved in coming up with designs and in making them, and had new ideas. In addition, such activities promoted students’ socialization, enhanced attention and development of intelligence. The students enjoyed, were enthusiastic and had fun in the course of implementation which positive stems for attitude change in STEM education. When analyzing students' perceptions and experiences with the engineering and design-supported STEM application within the context of Situational Interest Theory, it suggests that applying this theory in the educational field within the context of STEM education can design interactive and meaningful learning environments to enhance students' interest and motivation. This provides guidance for educators and curriculum developers on designing learning experiences that will capture and sustain students' interest.

## Limitations and Future Research

The study has various limitations. The study’s sample size and the specific context of BİLSEM may interfere with the applicability of the findings. Later on, the research can deal with the issue of sample size and test such interventions in different settings and educational environments. Furthermore, longitudinal studies could investigate the impact of such workshop activities with regard to sight on the students' career options and educational pathways in ‘STEM’ after the workshops. It can also be studied how the design and innovation workshops can be used in other learning environments and how they can be combined for the students benefit.

## Scientific Ethics Declaration

The authors declare that the scientific ethical and legal responsibility of this article published in JESEH journal belongs to the authors.

## References

- Aktan, O. & Budak, Y. (2020). Evaluation of teachers' opinions about using team-assisted individualization technique in mathematics course in inclusive education practices. *Marmara Üniversitesi Atatürk Eğitim Fakültesi Eğitim Bilimleri Dergisi [Marmara University Atatürk Faculty of Education Journal of Educational Sciences]*, 53(53), 69-96. <https://doi.org/10.15285/maruaebed.732256>
- Altan, E. B., Öztürk, N., & Türkoğlu, A. Y. (2018). Socio-scientific issues as a context for stem education: a case study research with pre-service science teachers. *European Journal of Educational Research*, 7(4), 805-812. <https://doi.org/10.12973/eu-jer.7.4.805>
- Benek, İ., & Akçay, B. (2018). STEM in the my imaginary world! Investigation of student’s drawings in STEM field. *Journal of STEAM Education*, 1(2), 79-107.

- Billiar, K., Hubelbank, J., Oliva, T., & Camesano, T. (2014). Teaching STEM by design. *Advances in Engineering Education*, 4(1), 1-21.
- Bonilla-Sánchez, M. d. R., García-Flores, M. A., Méndez-Balbuena, I., Silva-González, J. G., & Ramírez-Arroyo, E. V. (2022). The benefits of role play in the development of drawing in preschool children. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.1010512>
- Brechet, C., Blanc, N., Mortier, A., & Rossi, S. (2022). Draw me a brain: the use of drawing as a tool to examine children's developing knowledge about the "black box". *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.951784>
- Carlson, L. E., & Sullivan, J. F. (2004). Exploiting design to inspire interest in engineering across the K-16 engineering curriculum. *International Journal of Engineering Education*, 20(3), 372-378.
- Carr, R. L., Bennett IV, L. D., & Strobel, J. (2012). Engineering in the K-12 STEM standards of the 50 US states: An analysis of presence and extent. *Journal of Engineering Education*, 101(3), 539-564. <https://doi.org/10.1002/j.2168-9830.2012.tb00061.x>
- Creswell, J. W. (2013). *Qualitative research methods*. Siyasal.
- Cunningham, C. M. (2009). Engineering is elementary. *The Bridge*, 30(3), 11-17.
- Daly, S. R., Mosyjowski, E.A., & Seifert, C.M. (2014). Teaching creativity in engineering courses. *Journal of Engineering Education*, 103, 417-449. <https://doi.org/10.1002/jee.20048>
- Dönmez, İ. (2023). Breaking gender stereotypes: How interacting with STEM professionals changed female students' perceptions. *Journal of Baltic Science Education*, 22(6), 974-990. <https://doi.org/10.33225/jbse/23.22.974>
- Fan, S., & Yu, K. (2017). How an integrative STEM curriculum can benefit students in engineering design practices. *International Journal of Technology and Design Education*, 27, 107-129. <http://dx.doi.org/10.1007/s10798-015-9328-x>
- Fang, M., Jandigulov, A., Snezhko, Z., Volkov, L., & Dudnik, O. (2021). New technologies in educational solutions in the field of STEM: The use of online communication services to manage teamwork in project-based learning activities. *International Journal of Emerging Technologies in Learning (IJET)*, 16(24), 4-22. <https://doi.org/10.3991/ijet.v16i24.25227>
- Forawi, S. (2018). Science, Technology, Engineering and Mathematics (STEM) education: Meaningful learning contexts and frameworks. *2018 International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE)*, 1-4.
- Galanti, T. M. & Holincheck, N. (2022). Beyond content and curriculum in elementary classrooms: Conceptualizing the cultivation of integrated STEM teacher identity. *International Journal of STEM Education*, 9(1). 1-10. <https://doi.org/10.1186/s40594-022-00358-8>
- Hidi, S. & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, 41(2), 111-127. [https://doi.org/10.1207/s15326985ep4102\\_4](https://doi.org/10.1207/s15326985ep4102_4)
- Higgins, E. T., Cesario, J., Hagiwara, N., Spiegel, S., & Pittman, T. S. (2010). Increasing or decreasing interest in activities: the role of regulatory fit. *Journal of Personality and Social Psychology*, 98(4), 559-572. <https://doi.org/10.1037/a0018833>
- Juškevičienė, A., Dagienė, V., & Dolgopolas, V. (2020). Integrated activities in STEM environment: Methodology and implementation practice. *Computer Applications in Engineering Education*, 29(1), 209-228. <https://doi.org/10.1002/cae.22324>
- Kelley, T. R., & Knowles, J. G. (2016). A conceptual framework for integrated STEM education. *International Journal of STEM education*, 3, 1-11. <https://doi.org/10.1186/s40594-016-0046-z>
- King, D., & English, L. D. (2016). Engineering design in the primary school: Applying STEM concepts to build an optical instrument. *International Journal of Science Education*, 38(18), 2762-2794. <https://doi.org/10.1080/09500693.2016.1262567>
- Kron, S., Sommerhoff, D., Achtner, M., Stürmer, K., Wecker, C., Siebeck, M., ... & Ufer, S. (2022). Cognitive and motivational person characteristics as predictors of diagnostic performance: combined effects on pre-service teachers' diagnostic task selection and accuracy. *Journal für Mathematik-Didaktik*, 43(1), 135-172. <https://doi.org/10.1007/s13138-022-00200-2>
- Lo, J. C. (2015). Developing participation through simulations: a multi-level analysis of situational interest on students' commitment to vote. *The Journal of Social Studies Research*, 39(4), 243-254. <https://doi.org/10.1016/j.jssr.2015.06.008>
- McDonald, C.V. (2016). STEM education: A review of the contribution of the disciplines of science, technology, engineering and mathematics. *Science Education International*, 27, (4), 530-569.
- Merriam, S. B. (2013). *Qualitative research: A guide to design and implementation* (Tr. Ed. S. Turan). Nobel
- Ministry of National Education [MoNE]. (2022). *Bilim ve sanat merkezleri yaz okulu destekleme ve yetiştirme kursu programı tasarım ve inovasyon*. [https://orgm.meb.gov.tr/meb\\_iys\\_dosyalar/2022\\_07/07180753\\_BYLSEM\\_Yaz\\_Okulu\\_TasarYm\\_ve\\_Ynovasyon\\_Atolyesi\\_ProgramY.pdf](https://orgm.meb.gov.tr/meb_iys_dosyalar/2022_07/07180753_BYLSEM_Yaz_Okulu_TasarYm_ve_Ynovasyon_Atolyesi_ProgramY.pdf)
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Sage.

- Okulu, H. Z., Arabacioglu, S., & Unver, A. O. (2022). Enrichment of nature and science camps: Investigation of a STEM camp designed for gifted students. *Milli Eğitim Dergisi*, 51(235), 1983-2008. <https://doi.org/10.37669/milliegitim.910038>
- Roberts, A., & Cantu, D. (2012). Applying STEM instructional strategies to design and technology curriculum. In *PATT 26 Conference; Technology Education in the 21st Century*; (Vol. 73, pp. 111-118).
- Robinson, A., Dailey, D., Hughes, G., & Cotabish, A. (2014). The effects of a science-focused STEM intervention on gifted elementary students' science knowledge and skills. *Journal of Advanced Academics*, 25(3), 189-213. <https://doi.org/10.1177/1932202X14533799>
- Rotgans, J. I., & Schmidt, H. G. (2014). Situational interest and learning: Thirst for knowledge. *Learning and Instruction*, 32, 37-50. <https://doi.org/10.1016/j.learninstruc.2014.01.002>
- Seidel, T., Rimmel, R., & Prenzel, M. (2005). Clarity and coherence of lesson goals as a scaffold for student learning. *Learning and Instruction*, 15(6), 539-556. <https://doi.org/10.1016/j.learninstruc.2005.08.004>
- Shahali, E. H., Halim, L., Rasul, M. S., Osman, K., & Zulkifeli, M. A. (2016). STEM learning through engineering design: Impact on middle secondary students' interest towards STEM. *Eurasia Journal of Mathematics, Science And Technology Education*, 13, 1189-1211. <https://doi.org/10.12973/eurasia.2017.00667a>
- Sudrajat, U., Ardianto, D., & Permana, I. (2022). Engineering design process: A review and bibliometric analysis. *International Journal of STEM Education for Sustainability*, 2(2), 180-192. <https://doi.org/10.53889/ijses.v2i2.55>
- Ting, Y. L. (2016). STEM from the perspectives of engineering design and suggested tools and learning design. *Journal of Research in STEM Education*, 2(1), 59-71. <https://doi.org/10.51355/jstem.2016.22>
- Tröbst, S., Kleickmann, T., Lange-Schubert, K., Rothkopf, A., & Möller, K. (2016). Instruction and students' declining interest in science. *American Educational Research Journal*, 53(1), 162-193. <https://doi.org/10.3102/0002831215618662>
- Uzun, S. & Şen, N. (2023). The effects of a STEM-based intervention on middle school students science achievement and learning motivation. *Journal of Pedagogical Research*. 7(1), 228-242. <https://doi.org/10.33902/jpr.202319315>
- Wahono, B., Lin, P. L., & Chang, C. Y. (2020). Evidence of STEM enactment effectiveness in Asian student learning outcomes. *International Journal of STEM Education*, 7(1), 1-18. <https://doi.org/10.1186/s40594-020-00236-1>
- Winarno, N. (2020). The steps of the engineering design process (EDP) in science education: A systematic literature review. *Journal for the Education of Gifted Young Scientists*, 8(4), 1345-1360. <https://doi.org/10.17478/jegys.766201>
- Wu, Y. T., & Anderson, O. R. (2015). Technology-enhanced STEM (science, technology, engineering, and mathematics) education. *Journal of Computers in Education*, 2, 245-249. <https://doi.org/10.1007/s40692-015-0041-2>
- Yıldırım, A. & Şimşek, H. (2013). *Sosyal bilimlerde nitel araştırma yöntemleri (Qualitative research methods in social sciences)*. Seçkin.
- Zhan, Z., Shen, W., Xu, Z., Niu, S., & You, G. (2022). A bibliometric analysis of the global landscape on stem education (2004-2021): towards global distribution, subject integration, and research trends. *Asia Pacific Journal of Innovation and Entrepreneurship*, 16(2), 171-203. <https://doi.org/10.1108/apjie-08-2022-0090>

---

#### Author(s) Information

---

**Ibrahim Benek**

Van Edremit Science and Art Center  
Van, Türkiye  
Contact e-mail: [ibrahimbenek11@gmail.com](mailto:ibrahimbenek11@gmail.com)  
ORCID iD: <https://orcid.org/0000-0002-7124-4905>

**Ismail Donmez**

Muş Alparslan University  
Muş, Türkiye  
ORCID iD: <https://orcid.org/0000-0002-7792-0169>

**Sahin Idil**

The Scientific and Technological Research Council  
of Türkiye  
Ankara, Türkiye  
ORCID iD: <https://orcid.org/0000-0003-2366-913X>

---

## Use of Multi-Tier Concept Diagnostic Tests in Biology Education: A Systematic Review of the Literature

Tugce Duran, Musa Dikmenli

### Article Info

#### Article History

Published:  
01 October 2024

Received:  
13 April 2024

Accepted:  
24 August 2024

#### Keywords

Biology education,  
Multi-tier concept  
diagnostic test,  
Systematic review

### Abstract

This study aimed to comprehensively examine the articles in which multi-tier concept diagnostic tests, which are among the alternative assessment methods frequently used in recent years to identify misconceptions, were used in biology education between 2000 and 2022. For this purpose, systematic review steps were followed and summarized in the PRISMA diagram. In this process, four databases (Web of Science, SCOPUS, ERIC, and TR Index) were searched, articles were downloaded, elimination/inclusion criteria were applied, and the full texts of the remaining 71 articles were examined in detail. These articles were analyzed in terms of test type, year of publication, country of application, study group and number of participants, validity and reliability analysis, number of questions in the test and question types according to the test type, and it was revealed which subjects of biology they focused on. As a result of the analysis, it was seen that two-tier tests are mostly used in biology education (62%), followed by three-tier (30%) and four-tier tests (8%). It was seen that the reviewed articles were administered in a total of 21 countries. The subjects of the articles were grouped under 11 themes and 53 codes. Although they are new to the literature and have high validity and reliability, it was seen that especially four-tier concept diagnostic tests were rarely used in biology education. In this context, it was recommended to expand the use of four- and five-tier concept diagnostic tests in biology education.

## Introduction

Abstract concepts that students cannot directly experience through their sense organs are difficult to teach and learn. In this respect, measurement and assessment are important for science courses where abstract concepts are common. Alternative assessment methods are needed to reliably determine whether the science course has achieved its purpose, whether scientific concepts are understood by students, and whether there are missing information and misconceptions. According to the constructivist theory, students learn new information by comparing and associating it with existing knowledge. If there are deficiencies or misconceptions in prior knowledge, new information cannot be built correctly on prior knowledge. Therefore, students' prior knowledge should be tested before introducing a new subject, and if there are any errors or misconceptions, they should be intervened and corrected.

Students' prior knowledge is one of the most researched subjects in science education. The idea that students bring to science classes misconceptions that they find meaningful but contradict scientific facts is accepted by many researchers (Driver, 1988, pp. 33-39; Driver & Easley, 1978). Many studies in this field have shown that the misconceptions that students bring with them to the classroom affect their new learning (Duit & Treagust, 2003; Dykstra, Boyle & Monarch, 1992). Some of the information targeted to be acquired by students may contradict their prior knowledge, which may prevent students from learning the accurate information. Therefore, misconceptions in students' prior knowledge should be revealed and eliminated, and then teaching activities should be planned by taking this into consideration.

Concept diagnostic tests, one of the alternative assessment methods, are frequently used in the detection of misconceptions due to their easy administration and evaluation phases. Multiple-choice concept diagnostic tests were first developed by Tamir (1971). According to Tamir, in a single-tier multiple-choice concept diagnostic test, there is one correct answer in the options of the question and at least one of the remaining distractor options should contain a misconception. The biggest advantage of single-tier multiple-choice concept diagnostic tests is that they save time and can be applied at the same time to many people in a short time. In the preparation process of single-tier multiple-choice concept diagnostic tests, students' existing misconceptions should first be identified. In this context, in order to create distractors for multiple-choice concept diagnostic tests, it was

considered necessary for students to give the reasons for the statements they chose in the questions, which brought along the development process of two-tier concept diagnostic tests. Studies have shown that single-tier tests used to identify misconceptions may produce incorrect results due to contextual errors (Palmer, 1998). Due to the limitations of single-tier multiple-choice tests, multi-tier multiple-choice tests have been developed over time. Today, there are two-, three-, four- and five-tier multiple-choice tests. These tests are referred to as misconception diagnostic tests or concept diagnostic tests/concept tests in the literature.

The positive results obtained from the studies based on the justification of questions in multiple-choice concept diagnostic tests led to the development of two-tier multiple-choice concept diagnostic tests (Treagust, 2006). In the first tier of the two-tier tests, there is a multiple-choice question, and in the second tier, there is a question asking the reason for the answer given to the question in the first tier. In tests where the second tier is prepared as multiple-choice, an additional open-ended option can be given so that the student can express himself/herself more easily and not be limited only to the options given in the question (Mann & Treagust, 1998; Voska & Heikkinen, 2000). Although studies using two-tier tests initially seemed to yield positive results, over time it became clear that these tests had some disadvantages. For example, two-tier tests may cause students' misconception scores to be calculated higher than they actually are (Griffard & Wandersee, 2001). One of the most important disadvantages is that two-tier concept diagnostic tests are inadequate in distinguishing between students' misconceptions and lack of knowledge (Caleon & Subramaniam, 2010; Eryılmaz & Sürmeli, 2002).

In order to overcome the shortcomings of two-tier tests, three-tier concept diagnostic tests have been developed. Some studies have shown that three-tier tests are more reliable, valid and discriminative than two-tier tests (Lemma, 2012). In three-tier concept diagnostic tests, the first tier constitutes the content step. The main question to be asked is included in the content step. The second tier is called the reason step. In the reason step, the reason for the answer given to the question in the first step is asked to be explained. The third tier is the confidence step, in which students are questioned about their degree of certainty about their answers. The presence of a confidence step in the three-tier concept diagnostic tests eliminates the confusion about whether the student's error in the two-tier tests stems from misconceptions or lack of knowledge. If the student is sure of his/her answer, this indicates that he/she actually has a misconception that he/she is firmly committed to. Another advantage of three-tier tests is that they enable the analysis of false negatives and false positives (Hestenes & Halloun, 1995). Although three-tier concept diagnostic tests have great advantages in overcoming the shortcomings of two-tier concept diagnostic tests they also have some disadvantages. The most important of these disadvantages is that it is not known whether the answer given in the confidence step of three-tier concept diagnostic tests belongs to the content step or the reason step. This has been effective in the development and widespread use of four-tier concept diagnostic tests. The first step of four-tier concept diagnostic tests consists of the content step, similar to three-tier concept diagnostic tests. The second step is the confidence step. In this step, the student is asked to express whether he/she is sure of his/her answer to the question in the content step. The third step is the reason step, in which the reasons for the answers given to the question in the first step are questioned. Finally, the fourth step is the confidence step, in which the student is asked to express whether he/she is sure of his/her answer to the question in the reason step.

The four-tier concept diagnostic tests filled the gap of the three-tier concept diagnostic tests because they included a confidence step after both the content and the reason steps. The presence of a confidence step after both steps led to a clear distinction between lack of knowledge and misconception. In a recent study, five-tier concept diagnostic tests were created by adding a final microscopic drawing step to the four-tier concept diagnostic tests to increase their validity and reliability (Anam et al., 2019).

Biology, which is a sub-branch of science and examines all living things, is rich in both concrete and abstract concepts. Therefore, many difficulties are encountered in its learning process. When the studies conducted in biology education are examined, it is seen that students have difficulties in understanding and have misconceptions in genetics (Bahar, Johnstone & Hansell, 1999; Dikmenli, Cardak, & Kiray, 2011; Lewis & Wood-Robinson, 2000; Pearson & Hughes, 1988), cell division (Bahar, 2002; Cavallo & Schafer, 1994; Dikmenli, 2010), respiration (Haslam & Treagust, 1987; Songer & Mintzes, 1994), photosynthesis (Haslam & Treagust, 1987; Stavy, 1988; Waheed & Lucas, 1992), classification of living things (Ceylan & Umdu Topsakal, 2022) and ecology (Boyes, Stanisstreet & Papantoniou, 1999; Munson, 1994; Ola Adeniyi, 1985). In this regard, it is important to detect and eliminate misconceptions. Many methods are used today to detect misconceptions. Multi-tier concept diagnostic tests, which are among these methods and explained in detail above, are used in many fields such as physics (Chu, Treagust & Chandrasegaran, 2009; Hill et al., 2014; Irmak et al., 2023), chemistry (Al-Balushi, 2023; Rahayu, Treagust & Chandrasegaran, 2022). Ambusaidi, Al-Shuaili & Taylor, 2012; Artdej et al., 2010; Habiddin & Page, 2019), biology (Haslam & Treagust, 1987; Odom & Barrow, 1995), science (Kiray et al., 2015; Kiray & Simsek, 2021; Taban & Kiray, 2022), medicine (Badenhorst et al., 2016;

Versteeg, Wijnen-Meijer & Steendijk, 2019) and mathematics (Açıkgül, 2021; Kucam & Demir, 2020; Tabak, 2019; Yang & Sianturi, 2021). Although multi-tier misconception diagnostic tests first started with Tamir's (1971) sample application with 10th grade biology students in Israel, in which he showed that the options of multiple-choice questions could be written from student answers, today they have become more widespread in fields such as physics, chemistry and mathematics. The science of biology, by its nature, is prone to misconceptions as it contains many abstract concepts. Therefore, determination of students' misconceptions is very important for biology science. However, it is noteworthy that multi-tier misconception diagnostic tests, which have been among the important alternative assessment methods frequently used in detecting misconceptions in biology education literature in recent years, are used less than other disciplines.

The literature on the methods used to detect misconceptions was reviewed. Resbiantoro, Setiani, and Dwikoranto (2022) conducted an in-depth review of 72 articles on the methods used to detect misconceptions in physics, the causes of misconceptions and the ways to eliminate them. In their study, which was limited to articles published in ScienceDirect, SpringerLink, Taylor & Francis Online and Wiley Online Library databases between 2005 and 2020, they revealed that interviews, open-ended tests, multiple-choice tests, and multi-stage tests were generally used to detect misconceptions in physics. Rosida, Widarti, and Yahmin (2022) conducted a systematic analysis of studies that used multistage concept diagnostic tests to detect misconceptions about chemical equilibrium. In their study using the PRISMA diagram, they analyzed 30 scientific articles published between October 2010 and 2020, which they accessed using the Google Scholar search engine. They used articles, papers and theses published only in Indonesian journals indexed by SINTA. Önder Çelikkanlı and Kızılcık (2022) analyzed the studies using four-tier misconception diagnostic tests in physics education between 2010 and 2022. They included 69 studies accessed from Web of Science, ERIC, Scopus, EBSCOhost, Google Scholar and Turkey National Thesis Center databases. Soeharto et al. (2019) conducted a systematic analysis of studies on students' misconceptions in science. In their study using the PRISMA diagram, they analyzed a total of 111 scientific articles published between 2015 and 2019, which they accessed from some databases such as ERIC, EBSCO, SAGE, DOAJ, WILEY, JSTOR, ELSEVIER, SCOPUS and WOS. Kaltakci Gurel, Eryılmaz, and McDermott (2015) conducted a systematic analysis to determine the diagnostic tools used to detect misconceptions in science. They analyzed a total of 273 articles published between 1980 and 2014, which they accessed from a total of 9 databases, including ERIC, SSCI, SCI, JSTOR. When all these studies were evaluated, it was seen that there were quite a large number of literature reviews on the methods used in the detection of misconceptions in fields such as physics, chemistry and science. Even the studies of Rosida, Widarti, and Yahmin (2022) and Önder Çelikkanlı and Kızılcık (2022) were directly related to the use of multi-tier misconception diagnostic tests. However, as a result of the examinations, it was understood that there was no direct literature review specific to biology yet. Wulandari, Ramli, and Muzzazinah (2020) conducted a systematic analysis to determine the types of assessments for understanding the concept of biological materials and the properties and effects of materials for understanding biological concepts. In their study using the PRISMA diagram, they analyzed 20 articles published between 2004 and 2019, which they accessed from Google Scholar, ScienceDirect, CBE-life sciences education, and Taylor & Francis databases.

In a metasynthesis study, Kumandaş, Ateskan, and Lane (2019) compared and examined in depth the studies published in Türkiye on misconceptions in biology. In their study, they analyzed 67 articles obtained from the Institute for Scientific Information (ISI), Web of Science, Scopus, EBSCOhost, ULAKBIM and Academia Social Science (ASOS) Index databases. Most of these studies address various aspects of misconceptions in biology, but do not directly address the use of misconception diagnostic tests in biology. Therefore, it was seen that there was a need for an in-depth examination of the use of multi-tier misconception diagnostic tests in biology education in order to fill the gap in the literature. In this context, the aim of this study was to comprehensively examine the studies in which multi-tier concept diagnostic tests, were used in biology education between 2000 and 2022. In this respect, it was thought that the present study would make a significant contribution to the literature by providing an overview of the use of multi-tier concept diagnostic tests in biology education studies. On the other hand, deepening concept diagnostic tests will expand the use of these tests and contribute to revealing misconceptions in different subject areas.

## Research Problems

- 1) What is the distribution of the studies in which multi-tier concept diagnostic tests are used in biology education according to test types (two-tier, three-tier, etc.)?
- 2) What is the distribution of the studies in which multi-tier concept diagnostic tests are used in biology education according to years?

- 3) What is the distribution of test type in the studies in which multi-tier concept diagnostic tests are used in biology education according to years?
- 4) What is the distribution of the studies in which multi-tier concept diagnostic tests are used in biology education according to the countries in which they are administered?
- 5) What is the distribution of the studies in which multi-tier concept diagnostic tests are used in biology education according to the study group and number of participants?
- 6) Which subject areas do the studies using multi-tier concept diagnostic tests in biology education focus on?
- 7) What is the distribution of studies in which multi-tier concept diagnostic tests are used in biology education according to their validity and reliability analysis?
- 8) What is the distribution of studies in which multi-tier concept diagnostic tests are used in biology education according to the number of questions in the test?

## Method

The critical evaluation of the application of a method, technique or model in education is possible through systematic analysis. Systematic analysis, also known as systematic review, can be used for many purposes such as determining whether the current practice is based on evidence, determining the quality of evidence, and addressing any uncertainties that may occur during application (Munn et al., 2018). That is, if there are peer-reviewed empirical studies analyzing the effectiveness of a model, the field of study and effectiveness of this model can be detailed by a systematic review of the literature. In this context, in the present study, a systematic review was conducted to investigate the use and effectiveness of multi-tier concept diagnostic tests, which are among the important methods in determining misconceptions, in biology education. Systematic review is one of the important literature review methodologies that has existed since the late 1990s. Systematic review, which is widely used especially in health education fields such as nursing and medicine, provides access to all studies without bias and impartiality. As a result of the systematic review, not only the well-known studies in that field but also the entire literature is accessed, providing the reader with quality information. It is also very important to systematically detail the path followed in this process for the reliability and reproducibility of the study. Systematic reviews can provide guidance for future studies as they enable the identification of strengths and weaknesses in the literature. It is believed that the use of systematic review, which has been used in education as well as in the field of health in recent years, will add potential value to higher education (Bearman et al., 2012; Gough, 2007). In general, a five-tier systematic review process was adopted to answer the initial research questions.

### Phase 1: Article Collection, Review, and Initial Selection

#### *Databases and Search Terms, Article Collection*

To review the existing literature, four databases (Web of Science, SCOPUS, ERIC, and TR Index) were searched for articles. For each database, the title, subject, index, abstract and/or keywords and the entire text were searched in Turkish and English for the terms “aşamalı test/tier test”, “iki aşamalı/two tier”, “üç aşamalı/three tier”, “dört aşamalı/four tier”, “beş aşamalı/five tier” and two additional terms were searched. The first additional term used in the search was “Biology” and the second was “Diagnostic”. Therefore, each search contained “search term” + Biology + Diagnostic (e.g. “two tier” + biology + diagnostic). However, when an “advanced search” was performed in the TR Index database, it was seen that the results were limited incorrectly and the advanced search was insufficient. For this reason, since it was aimed to reach all studies, the search terms were directly searched in the TR Index database and no additional terms were added. The process of reviewing the databases and downloading the relevant studies was completed between January and March 2023.

#### *Article Review and Initial Selection*

The articles to be analyzed were uploaded to Zotero. In Zotero, the sources, which were foldered separately for each database in the form of collections and sub-collections, were put in a single folder called “Merged folder” when the review was completed, and duplicate articles were excluded. The remaining articles were assessed for eligibility according to the inclusion criteria. First, the titles, then the abstracts and finally the full texts of the studies were independently reviewed by two authors. Any conflicts between the two authors were resolved through discussion.

*Initial Inclusion Criteria*

For this review, six criteria were considered for the inclusion of the selected studies: (1) The selected studies were empirical, that is, the studies in which quantitative or qualitative data were collected, analyzed and reported were included, (2) They were not conference papers, books, book chapters or theses, (3) They were appropriate for the content of biology education, (4) Each selected study was written in English or Turkish, (5) They were studies in which multi-tier concept diagnostic tests were developed, (6) They were conducted between 2000 and 2022. In this context, studies that were not written in English or Turkish, were not conducted between 2000-2022 were not empirical, were not on subjects specific to biology education, and studies in which previously developed multi-tier concept diagnostic tests were re-applied were not included in the study. In addition, conference papers, books, book chapters and theses were also excluded. In determining empirical studies, finding the research question, specifying the study group, explaining the methodology and reporting the study results were accepted as criteria.

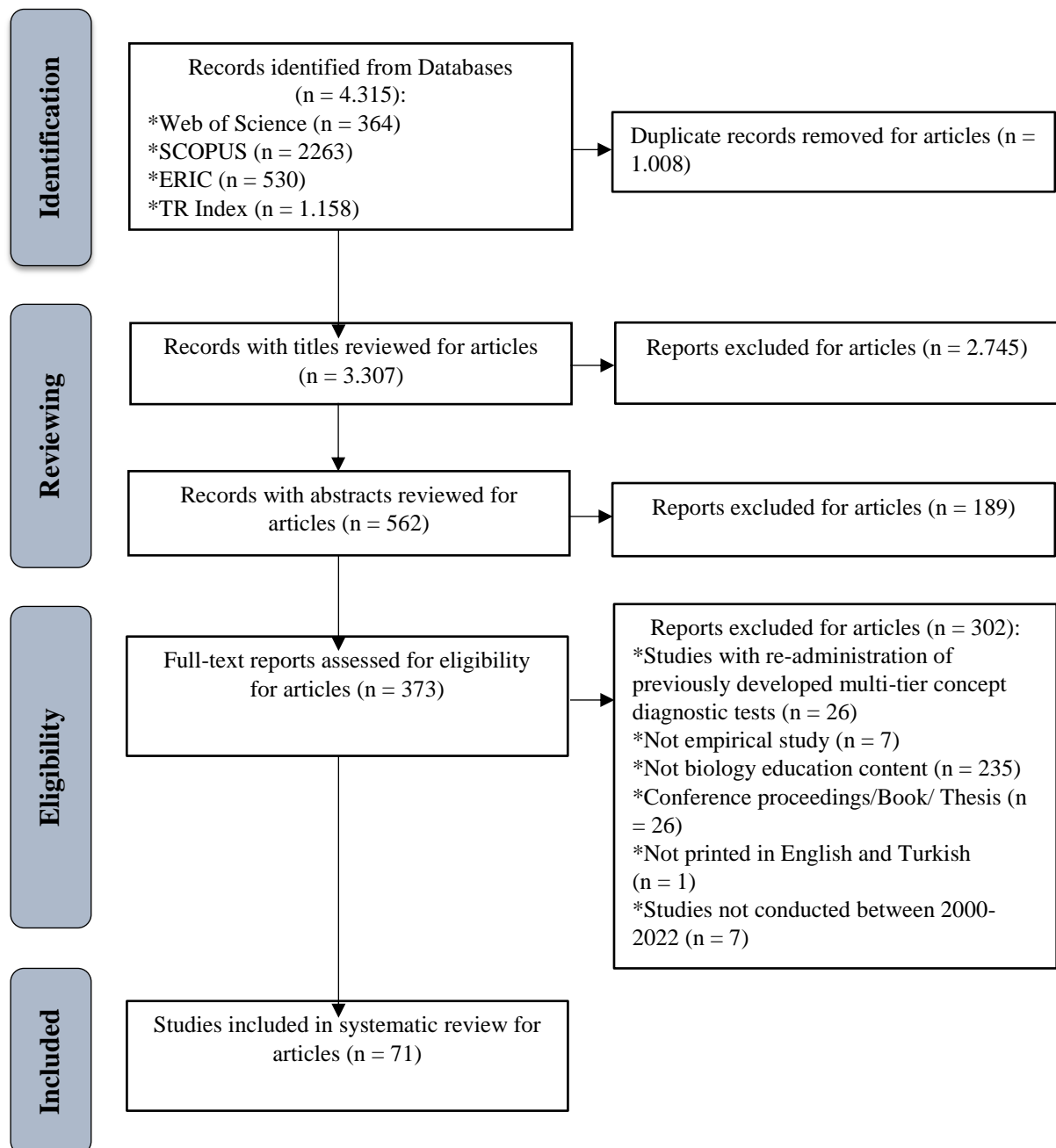


Figure 1. PRISMA diagram

This review adopted the inclusion and exclusion methods outlined by The PRISMA Group (Moher et al., 2009). The articles retrieved from Web of Science (n=364), SCOPUS (n=2,263), ERIC (n=530) and TR Index (n=1,158) databases were combined. Duplicate articles in different databases were excluded, leaving 3,307 articles. The titles of the remaining articles were analyzed and the articles determined to be outside the scope of the study (n=2,745) were excluded. The abstracts of the remaining articles were then examined in detail. At this tier, 189 articles determined to be outside the scope of the study were excluded.

Finally, the full texts of the remaining 373 articles were examined in detail and evaluated according to the eligibility criteria. As a result of this review, 302 articles that were determined to be outside the scope of the study were excluded. (1) 7 articles that were determined to be non-empirical and written for purposes such as review, introduction and meta-analysis, (2) 26 articles that were determined to be conference papers, books and book chapters, (3) 235 articles that were determined to be related to fields such as physics, chemistry, science, medicine and mathematics, (4) 1 article that was determined to be written in Thai, (5) 26 articles in which multi-tier tests previously developed in the field of biology education were re-applied to different study groups, (6) 7 articles that were determined to have been conducted in 1987, 1989, 1992, 1995, 1996 and 2023 were eliminated. This entire process is summarized in Figure 1.

As can be seen in the PRISMA diagram in Figure 1, initially 4,315 articles were found after reviewing the databases. Then, after eliminating duplicate articles in different databases, the titles of the remaining 3,307 articles were examined by both authors and 96% consensus was reached between the authors on the articles that should be included in the scope of the study. The abstracts of the remaining 562 articles were reviewed by both authors and 98% consensus was reached that 373 articles should be included in the study for full text review. The full-text articles were evaluated according to the eligibility criteria and as a result, 99% agreement was reached between the authors that 71 articles met the eligibility criteria. All these rates were calculated using Miles and Huberman's (2016) formula (i.e., Reliability = Agreement / Agreement + Disagreement) to ensure the reliability of the study.

## Phase 2: Final Article Selection

At this phase, 71 articles (Annex) that were decided to be included in the study after full text review were examined in detail and categorized. The studies in which previously developed multi-tier concept diagnostic tests were adapted from English to Turkish and from Turkish to English were included in the scope of the study because they were re-analyzed and their validity and reliability were calculated. It was determined that in some articles, a method called CRI (Certainty of Response Index) was used to determine the degree of confidence at one tier of multi-tier concept diagnostic tests. With this method, the degree of confidence in the answer given to the previous question is questioned by asking for a score from 0 to 5. Therefore, since it was equivalent to the confidence level, it was considered as a tier in multi-tier concept diagnostic tests.

Finally, all articles were categorized according to the number of tiers in the multi-tier concept diagnostic test. Articles using CRI and adaptation studies were also included in the classification according to the number of tiers. Although there were studies in different fields in the literature, it was seen that there was no five-tier concept diagnostic test in biology education. As a result, 71 articles included in the study were categorized as articles using CRI, adaptation studies into Turkish or English, two-tier tests, three-tier tests, four-tier tests (Figure 2), and separate collections were created in Zotero.

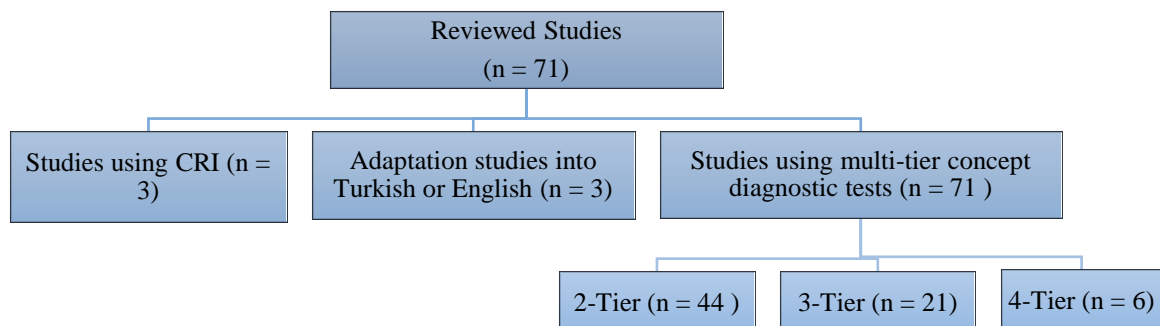


Figure 2. Study characteristics flowchart

*Phase 3: Data Determination*

The following characteristics were determined for the 71 articles included in the study: (1) year of publication, (2) study group and number of participants, (3) country of application, (4) name of the multi-tier concept diagnostic test, (5) subject on which multi-tier concept diagnostic test is administered, (6) number of questions in the multi-tier concept diagnostic test, (7) number of tiers in the multi-tier concept diagnostic test, (8) reliability analysis, (9) validity analysis. When determining the 9 features to be considered in the analysis of the studies, the authors independently listed the features. Then, the lists were compared and the discrepancies were reviewed and reconciled. Thus, the data for the final study were determined.

*Phase 4: Data Extraction and Audit*

After the selection of articles to be included in the study was completed (Phase 2) and the data were determined (Phase 3), all eliminated articles were removed from Zotero. At this tier, the two authors collaborated. In addition, a professor specialized in biology education was asked to check the data for accuracy. As a result, all of the data were effectively checked and the final data set was prepared.

*Phase 5: Analysis*

In this study, answers to 8 research problems were sought. Descriptive analysis was used to analyze the distribution of the studies in which multi-tier concept diagnostic tests were used in biology education according to test type, years, test type according to years, countries where they were administered, study groups and number of participants, number of questions in the test, validity and reliability analysis. In descriptive analysis, data are categorized and interpreted according to predetermined themes. Finally, content analysis was used to determine the subject areas of the studies in which multi-tier concept diagnostic tests were used in biology education. Content analysis aims to reach the concepts and relationships that explain the data. The data were coded in accordance with Strauss and Corbin's (1990) "coding made according to the concepts extracted from the data" type. Therefore, codes were generated directly from the data with inductive analysis. IBM-SPSS 24, Excell and MAXQDA 2018 programs were used to analyze and present the data.

## Findings

Descriptive statistics were used to determine the distribution of the analyzed studies according to test type. As a result of the analysis, frequencies and percentages were determined for test types and presented in Table 1.

Table 1. Distribution of the studies according to test type

Diagnosis Method	<i>f</i>	%
2-Tier	44	62
3-Tier	21	30
4-Tier	6	8
5-Tier	0	0
Total	71	100

According to these analyzed data, it was seen that two-tier tests (62%) were used the most in the studies in which multi-tier tests were used in biology education. They were followed by three-tier tests (30%). However, no five-tier tests were found. Descriptive statistics were used to determine the distribution of the analyzed studies by years and test type by years. As a result of the analysis, a column chart was created showing the distribution of test types by years and presented in Figure 3.

When Figure 3 was analyzed, it was seen that the number of three-tier tests gradually increased in the studies using multi-tier tests in biology education, and the number of four-tier tests increased in recent years due to the introduction of four-tier tests. It was determined that two-tier tests were found more in the early 2000s, and in the following years, with the increase in three-tier and four-tier tests, they fell behind three-tier and four-tier tests in 2019 and 2022. However, in 2020 and 2021, two-tier tests became more dominant again and the number of three-tier and four-tier tests decreased. When the data were analyzed in general, it was understood that the number of studies using multi-tier tests in biology education gradually increased over the years. The highest rate was in 2022 ( $f=11$ ; 15%), the last year included in the study. The years with the least number of studies were

2005, 2006 and 2010. In 2012, there were 5 studies, and despite a sudden decrease in 2014, this rate increased again in 2015 and reached 4 studies.

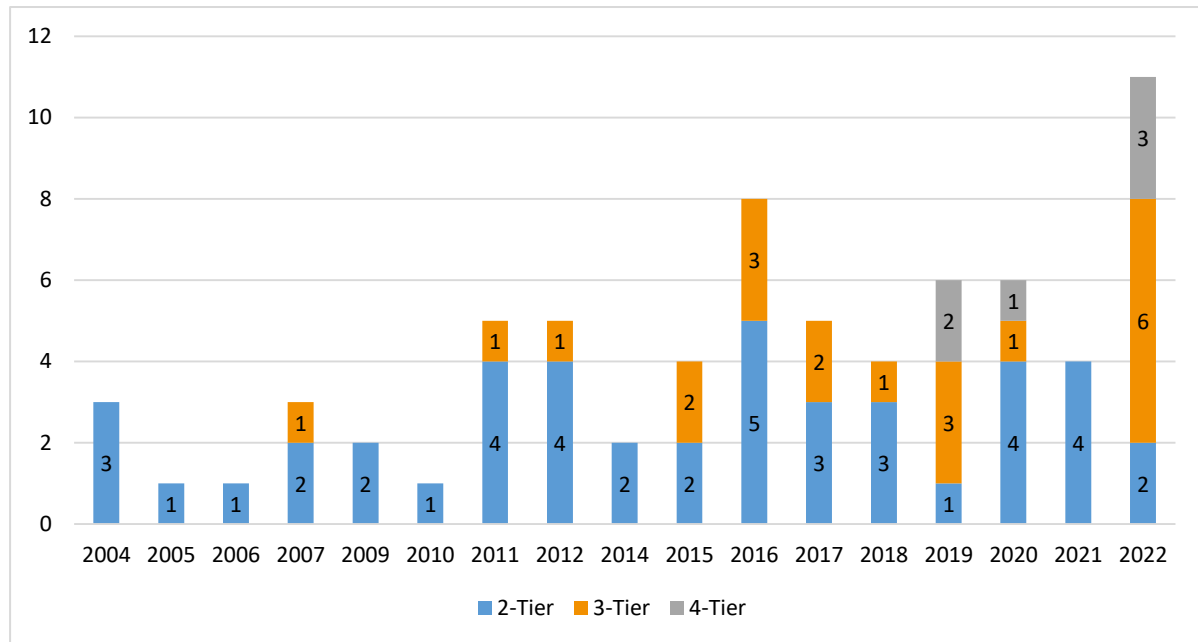


Figure 3. Distribution of test types according to years



Figure 4. Distribution of the studies according to the countries where they were administered

Descriptive statistics were used to determine the distribution of the analyzed studies according to the countries where they were administered. As a result of the analysis, a map was created on the world map showing the countries where these studies were administered and the frequencies in these countries (Figure 4). The multi-tier concept diagnostic test developed by Kılıç, Taber, and Winterbottom (2016) was administered separately in England and Türkiye. Similarly, Šorgo, Ambrožič-Dolinšek, Uşak, and Özel (2011) administered the multi-tier concept diagnostic test they developed in their study in Türkiye and Slovenia. In these studies conducted in two countries, frequencies for both countries were included in the analysis results. Therefore, the total frequency did not give the number of articles examined, but the total number of countries where the administrations were made ( $f = 73$ ).

According to Figure 4, the studies using multi-tier concept diagnostic tests in biology education were administered in 21 countries (the USA, Australia, China, Czech Republic, England, Germany, Greece, Indonesia, Malaysia, Mauritius, Morocco, Nigeria, Saudi Arabia, Serbia, Singapore, Slovenia, South Africa, Sri Lanka, Switzerland, Taiwan, Türkiye). Türkiye (37%) was the country with the highest number of studies. This was followed by Indonesia (12%), the USA (11%) and Taiwan (10%).

Similarly, descriptive statistics were used to determine the distribution of the studies according to the study group. As a result of the analysis, frequencies and percentages were determined for the study groups and presented in Table 2. It was observed that some studies were conducted in more than one different study group (Apaydın & Sürmeli, 2006; Çiğdemoglu & Arslan, 2017; Gürsel & Akçay, 2022; Liampa et al., 2019; Soeharto & Csapó, 2021; Wang, 2004; Yen, Yao & Chiu, 2004; Yen, Yao & Mintzes, 2007). Therefore, each study group was evaluated separately. The total number reached as a result of the analysis reflected the number of study groups, not the number of articles.

Table 2. Distribution of the studies according to study groups

Study Group	<i>f</i>	%
Primary school students	3	4
Secondary school students	18	21
High school students	23	27
University students	8	10
Students studying in the biology department of the university	4	5
Students studying in the pharmacy department of the university	1	1
Students studying in the preparatory class of the university's medicine department	1	1
Biology teacher candidates	8	10
Science teacher candidates	8	10
Primary school teacher candidates	4	5
Physics teacher candidates	1	1
Chemistry teacher candidates	1	1
Pre-school teacher candidates	1	1
Life and earth sciences teacher candidates	1	1
Teacher candidates studying in different branches at the university	2	2
Total	84	100%

It was observed that the study group of studies in which multi-tier concept diagnostic tests were used in biology education generally consisted of high school students (27%). This was followed by secondary school students (21%). When the study groups were evaluated by grouping them according to their education levels, it was observed that 4% of the study groups consisted of primary school students ( $f = 3$ ), 21% consisted of secondary school students ( $f = 18$ ), 27% consisted of high school students ( $f = 23$ ) and 48% consisted of university students ( $f = 40$ ). When the articles whose study group consisted of teacher candidates were evaluated within themselves, it was seen that the majority of studies were conducted with biology teacher candidates (31%) and science teacher candidates (31%). Descriptive statistics were used through SPSS to determine the distribution of articles according to the number of participants. The results of the analysis are given in Table 3.

Table 3. Data on the number of participants in the reviewed studies

	N	Minimum	Maximum	Mean	Std. Deviation
Number of Participants	71	17	1962	313.90	342.59

It was seen that the study group of the studies in which multi-tier concept diagnostic tests were used in biology education consisted of at least 17 participants and at most 1962 participants. When the number of participants in all reviewed articles was evaluated, it was determined that there were approximately 314 participants on average. Content analysis method was used to determine the subject areas of the studies in which multi-tier concept diagnostic tests were used in biology education. First, the subjects of the examined articles were listed. Some studies were found to be conducted in a general subject area such as science concepts or science (Soeharto, 2021; Soeharto & Csapó, 2021). These studies included physics, chemistry and biology subjects together. Due to the scope of the current study, only biology-related subjects in these studies were included in the analysis. The list of subjects to which the multi-tier concept diagnostic test was administered in the articles was coded separately by both researchers. A total of 53 codes were generated. At this tier, 99% agreement was seen among researchers (Miles & Huberman, 2016). The codes were grouped according to their similarities. These groups created the themes. More general subjects in biology education were used to create the themes, and it was ensured that the themes consisted of concepts that best represented the codes. The themes were (1)

physiology, (2) biochemistry, (3) genetics, (4) cell biology, (5) environmental education, (6) health, (7) evolution, (8) biodiversity and classification of living things, (9) socioscientific subjects, (10) general biology and (11) ecology. For the reliability of the study, two different experts were consulted to determine whether the 11 themes represented the codes. These experts were working as professors in the department of biology education at the faculty of education of a university. Two different lists were given to each faculty member. One of these lists consisted of codes arranged alphabetically ( $n=53$ ) and the other list consisted of 11 themes arranged alphabetically (supported by short explanations). Experts were asked to match all codes with these themes. Then, Miles and Huberman's (2016) formula ( $\text{Reliability} = \text{Agreement} / (\text{Agreement} + \text{Disagreement})$ ) was applied to determine reliability. The agreement between experts and researchers was 96% and 94% for two experts, respectively. The codes with disagreements were discussed between the experts and the researchers, and an agreement was reached.

The determined codes and themes were transferred to the MAXODA 2018 program. The codes were linked to the themes. A map was created using the "Code-Subcode-Sections model" through MAXMaps. In this model, previously determined codes represented sub-codes and themes represented codes. Since all the analyzed studies were related to biology education, all codes and sub-codes were combined in the "Biology Education" section. In the map created, the codes were shown in colored rectangles and each sub-code of that code was symbolized in the same color and connected to the codes. The codes were symbolically indicated by a triangle. Subcodes were symbolized with a label sign. The thickness of all connecting lines on the map (between subcodes and codes, between codes and section) was determined by frequency (Figure 5).

When Figure 5 was examined, it was seen that the subjects of the studies in which multi-tier concept diagnostic tests are used in biology education were grouped under 11 themes and that these themes consisted of 53 codes in total. The total frequency of these codes, 71, represented the number of articles examined. It was understood that the theme with the most codes was "physiology" ( $n = 16$ ). It was followed by "environmental education" ( $n=6$ ) and "biodiversity and classification of living things" ( $n=6$ ). The theme with the fewest codes was "socioscientific issues" ( $n=1$ ). Within the "Physiology" theme, the code with the highest frequency was diffusion and osmosis. Within the theme of "environmental education", climate change was the code with the highest frequency. Within the theme of "Biological diversity and classification of living things", each code had equal frequency. Similarly, the themes of "evolution" and "biochemistry" also consisted of codes with equal frequency. The codes with the highest frequency were genetics in the "genetics" theme, cell biology in the "cell biology" theme, ecological footprint in the "ecology" theme, microbes and diseases in the "health" theme, and general biology in the "general biology" theme. It was an expected situation that the codes with the highest frequency in the themes of genetics, cell biology and general biology were the same as the name of the theme. These were the main subjects in biology. Therefore, when the codes in the analyzed articles were grouped according to their similarities and themes were created, the themes were named after these subjects since they constituted the most general concept.

Descriptive statistics were used to determine the distribution of studies in which multi-tier concept diagnostic tests were used in biology education, according to their validity and reliability analysis. As a result of the analysis, frequencies and percentages were determined and presented in Table 4. Accordingly, it was seen that most of the articles (79%) were analyzed for both validity and reliability. It was determined that 7% of the articles did not provide any information about validity and reliability analysis. There were also articles in which only validity analysis (9%) or only reliability analysis (4%) was conducted.

Table 4. Distribution of the studies according to their validity and reliability analysis

Status of validity and reliability analysis	<i>f</i>	%
Articles with validity analysis only	6	9
Articles with reliability analysis only	3	4
Articles with both validity and reliability analyses	53	79
Articles with no information on validity and reliability analyses	5	7
Total	67	100

Descriptive statistics were used through SPSS to determine the distribution of studies using multi-tier concept diagnostic tests in biology education according to the number of questions in the test. In Wang's (2004) study, it was seen that he prepared three different tests separately for primary school, secondary school and high school. For this reason, the number of questions in each test in this study was treated as a separate test and included in the analysis. Therefore, the analysis values were determined based on a total of 73 tests in 71 articles. The results of the analysis are given in Table 5.

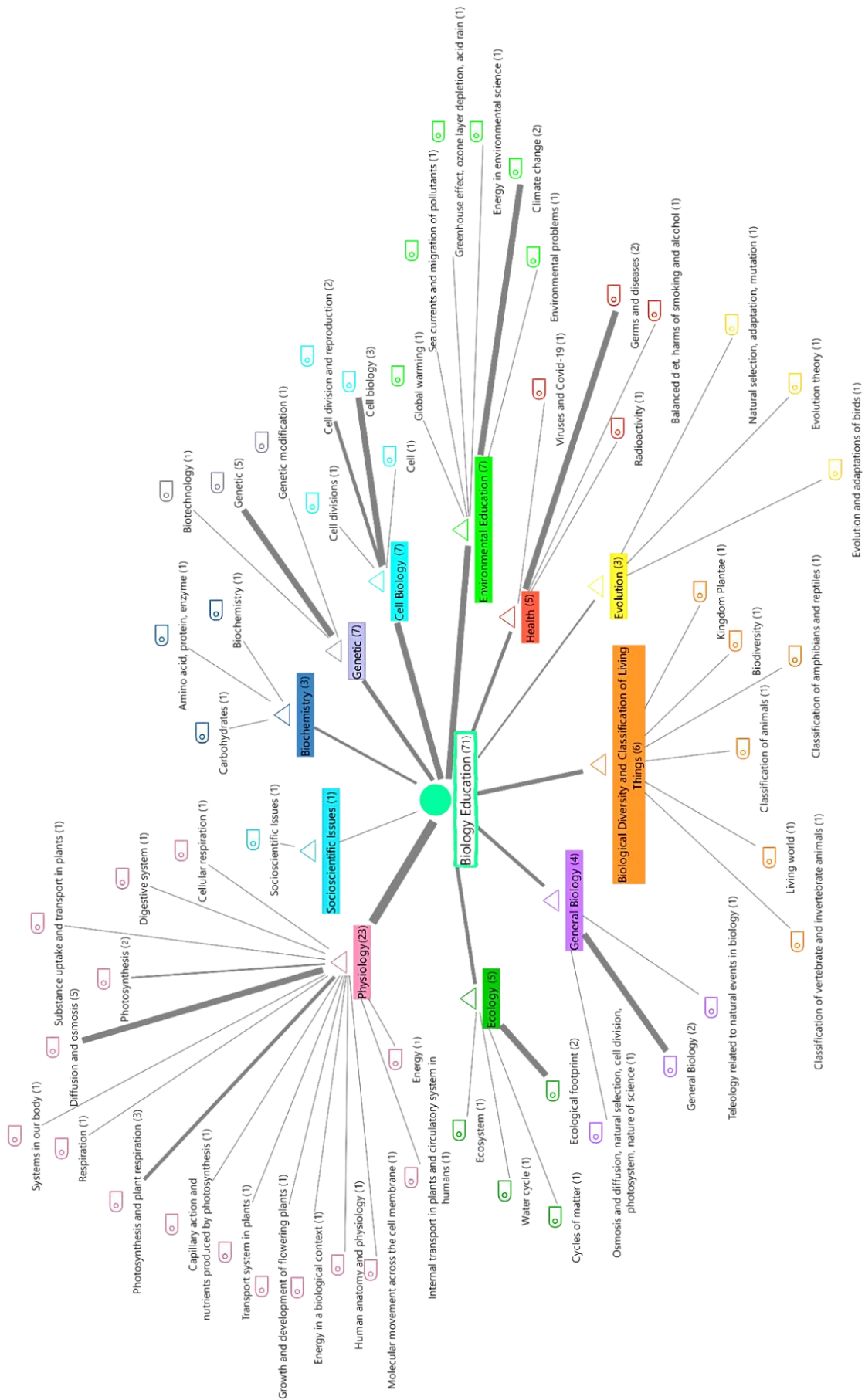


Figure 5. Distribution of the studies according to subjects

Table 5. Data of the studies regarding the number of questions in the test

	N	Minimum	Maximum	Mean	Std. Deviation
Number of Questions	73	2	49	15.78	8.29

In the studies in which multi-tier concept diagnostic tests were used in biology education, it was seen that the minimum number of questions in the test was 2 and the maximum number was 49. When the number of questions of multi-tier concept diagnostic tests in all the articles was evaluated, it was determined that there were 16 questions on average for each article.

## Conclusions and Recommendations

The aim of this study was to comprehensively examine the studies in which multi-tier concept diagnostic tests, one of the alternative assessment methods frequently used in recent years to identify misconceptions, were used in biology education between 2000 and 2022. This review was carried out by following the steps of a systematic review as summarized in the PRISMA diagram. In this process, four databases (Web of Science, SCOPUS, ERIC, and TR Index) were examined and articles were downloaded and eliminated by applying elimination/inclusion criteria at various tiers. The full text of the remaining 71 articles was subjected to detailed analysis. The 71 articles in which concept diagnostic tests were used were analyzed in terms of test type, year of publication, country of administration, study group and number of participants, validity and reliability analysis, number of questions in the test and it was tried to reveal which subjects of biology these studies focused on.

As a result of the analysis, it was seen that two-tier tests were mostly used in biology education (62%), followed by three-tier (30%) and four-tier (8%) tests. Five-tier tests were not used at all. This is probably due to the fact that five-tier concept diagnostic tests are new in the literature. Similarly, in a study examining the diagnostic assessment tools used to identify misconceptions in science between 2015 and 2019, it was determined that four-tier multiple-choice tests were used only in studies in the field of physics, and in addition, two-tier and three-tier diagnostic tests were used less in biology than in chemistry and physics (Soeharto et al., 2019). In this context, Kaltakci Gurel, Eryilmaz, and McDermott (2015) stated that the number of three- and four-tier multiple-choice tests in biology education is still low and should be increased. In a similar study in physics education, Resbiantoro and Setiani (2022) examined the studies on misconception identification between 2005 and 2020. According to their results, they determined that 10% of the concept diagnostic tests used in the studies were two-tier, 6% were three-tier and only 4% were four-tier. The findings of this study also support the opinion that four-tier concept diagnostic tests are used less than other multi-tier concept diagnostic tests.

It was understood that the number of three-tier tests gradually increased in studies in which multi-tier tests were used in biology education, and in recent years, the number of four-tier tests has increased due to the introduction of four-tier tests. It was found that two-tier tests were more common in the early 2000s, and with the increase in three-tier and four-tier tests in the following years, they fell behind three-tier and four-tier tests in 2019 and 2022. However, in 2020 and 2021, two-tier tests became more dominant again and the number of three- and four-tier tests decreased. While there were 5 studies in 2012, there was a sudden decrease in 2014, but this rate increased again in 2015 and reached 4 studies. Supporting this finding, a study reported that multi-tier concept diagnostic tests increased after 2015 (Soeharto et al., 2019).

It was observed that studies using multi-tier concept diagnostic tests in biology education were conducted in a total of 21 countries. Türkiye (37%) was the country with the highest number of studies. This was followed by Indonesia (12%), America (11%) and Taiwan (10%). Önder Çelikkanlı and Kızılcık (2022) stated that the most studies using four-tier concept diagnostic tests in physics education were conducted in Indonesia (76%), followed by Türkiye (21%) and Singapore (3%). It is thought that the aforementioned study differs from the findings of the current study because it was conducted in physics education.

As a result of the examinations, it was seen that the study group of the studies in which multi-tier concept diagnostic tests were used in biology education generally consisted of high school students. This was followed by secondary school students. When the study groups were grouped according to educational levels, it was understood that most of the studies were conducted with university students. When the articles in which the study group consisted of teacher candidates were evaluated within themselves, it was seen that most of the studies were conducted with biology teacher candidates and science teacher candidates. Önder Çelikkanlı and Kızılcık (2022) examined the articles in which four-tier concept diagnostic tests were used in physics education between 2010 and 2022 and determined that the study group of the articles they analyzed consisted mostly of high school students (41%). This finding is similar to the data of the present study. In a study conducted by

Resbiantoro and Setiani (2022), they analyzed the studies on misconception detection in physics education conducted between 2005 and 2020. They found that the study group was mostly undergraduate students (36%), followed by high school students (26%) and teacher candidates (20%). This difference in findings may be normal since this study was conducted in physics education and all studies on misconception detection, including multi-tier concept diagnostic tests, were examined.

In the reviewed articles, it was seen that the study group consisted of at least 17 participants and at most 1962 participants. When the number of participants in all the articles was evaluated, it was determined that there were approximately 314 participants on average. Factor analysis is always performed in studies where concept diagnostic tests are developed. For exploratory factor analysis, the sample size should be at least 300 or more (Seçer, 2013, pp. 119). It was determined that the sample size was below 300 in 67.6% of the studies. It is also known that in order to calculate Cronbach's alpha, if the eigenvalue is between 3.00 and 6.00, the sample should consist of at least 100 participants (Yurdugül, 2008). It was determined that the sample size was below 100 in 21.12% of the studies, which is a remarkable rate.

It was understood that the subjects of the studies in which multi-tier concept diagnostic tests were used in biology education were grouped under 11 themes. These themes were Physiology, Socioscientific issues, Biochemistry, Genetics, Cell biology, Environmental education, Health, Evolution, Biodiversity and classification of living things, General biology and Ecology, respectively. In their study, Wulandari, Ramli & Muzzazinah (2020) examined 20 articles published in international journals between 2004 and 2019 to determine the types of evaluations regarding the understanding of the concept of biological materials, the properties and effects of materials related to the understanding of biological concepts. According to their findings, they determined that studies to evaluate the understanding of a concept were generally carried out on Plantae, genetics, molecular biology, evolution, microbiology and photosynthesis. All of the subjects identified here are among the subjects identified in the current study. Therefore, the findings of the above-mentioned study support this study. The subjects that generally cause misconceptions in students in biology education are (1) Adaptations, habitat, biosphere, ecosystem, food chain and food web, functions of the ecosystem, biomass and biodiversity, (2) Osmosis and diffusion, (3) Plant transport, (4) Antibiotic resistance, (5) Acid rain, global warming, greenhouse effect and ozone layer depletion, (6) Water cycle, (7) Photosynthesis, (8) Nature of science, (9) Digestive system, (10) Energy and climate change, (11) Evolution of biology, (12) Human reproduction, (13) Human and plant transportation systems, (14) Global warming, (15) Ecological concepts (Soeharto, Csapó, Sarimanah, Dewi, & Sabri, 2019). The themes and codes we identified in our study overlap with the majority of the above subjects that were found to contain misconceptions. There are many studies in which multi-tier concept diagnostic tests have been developed, especially on the subject of osmosis and diffusion. When the basic subjects in biology education are examined in "Campbell Biology", which is widely used and accepted in biology education worldwide, it is seen that the basic subjects in biology education include living chemistry, cell, genetics, evolutionary history of biological diversity, plant structure and function, animal structure and function, evolutionary mechanism and ecology (Urry et al., 2022). In another similar study, "Life Science Biology", the main subjects include life science and the chemical basis of living things, cells, cells and energy, genes and heredity, models and processes of evolution, evolution of diversity, flowering plants: structure and function, animals: structure and function, genomes and ecology (Sadava et al., 2014). When the articles we examined in our study were evaluated according to their subjects, it was seen that there were not enough studies on the chemical basis of living things and genomes, which are among the basic biology subjects included in the aforementioned studies.

In test development studies, the measurement tool must be valid and reliable. A reliable measurement may not be valid. However, a valid measurement must be reliable (Özdamar, 2017, pp. 69). It was observed that both validity and reliability analyses were mostly performed in studies where multi-tier concept diagnostic tests were used in biology education. It was determined that 7% of the articles did not provide any information regarding validity and reliability analyses. There were also articles in which only validity analysis (9%) or only reliability analysis (4%) was conducted. In their study, Önder Çelikkanlı and Kızılcık (2022) examined the articles in which four-tier concept diagnostic tests were used in physics education between 2010 and 2022, and found that 29% of the articles did not include any explanation of how validity and 33% of the articles did not include any explanation of how reliability was determined. In the above-mentioned study, the rate of studies in which only validity and only reliability were determined is quite low compared to the findings of this study.

In the studies in which multi-tier concept diagnostic tests were used in biology education, it was seen that the minimum number of questions in the test was 2 and the maximum number was 49. The average number of questions in the test was 16. Similarly, in studies using four-stage concept diagnostic tests in physics education,

it was found that the number of questions in the test was at least 4 and at most 40, and the average number of questions was 18.73 (Önder Çelikkanlı & Kızılcık, 2022).

When evaluated in general, it can be said that the number of studies using multi-tier concept diagnostic tests is insufficient in biology education. Although they are new to the literature and have high validity and reliability, especially four-tier concept diagnostic tests are rarely used in biology education. Five-tier diagnostic tests were not encountered at all. In this context, it was recommended to expand the use of four and five-tier concept diagnostic tests in biology education. It is recommended that researchers who plan to develop multi-tier concept diagnostic tests in biology education should conduct studies on the subjects of "Chemical basis of living things" and "Genomes", which were found to be missing in the literature in the current study.

## Limitations

Studies that our reviewed was limited to the use of as Turkish and English the terms “aşamalı test/tier test”, “iki aşamalı/two tier”, “üç aşamalı/three tier”, “dört aşamalı/four tier”, “beş aşamalı/five tier” and two additional terms (Biology+Diagnostic) in article title, subject, index, abstract and/or keywords and the entire text title in four databases (Web of Science, SCOPUS, ERIC, and TR Index). Therefore, there is still potential that the review missed relevant articles which in other databases or expressed in different terms. Also according to the criteria we identified in the study studies that were not written in English or Turkish, were not conducted between 2000-2022, were not empirical, were not on subjects specific to biology education, and studies in which previously developed multi-tier concept diagnostic tests were re-applied were not included in the study. In addition, conference papers, books, book chapters and theses were also excluded. This is a limitation of the study.

## Scientific Ethics Declaration

The authors declare that the scientific ethical and legal responsibility of this article published in JESEH journal belongs to the authors.

## Acknowledgements or Notes

We would like to for contributing to the reliability of the study thank the professors specialized in biology education whose valuable opinions we received at "Phase 4: Data Extraction and Audit" and in the phase to determine whether the 11 themes represented the codes of our study.

## References

- Açıkgül, K. (2021). Developing a two-tier proportional reasoning skill test: Validity and reliability studies. *International Journal of Assessment Tools in Education*, 8(2), 357-375.
- Al-Balushi, S. M., Ambusaidi, A. K., Al-Shuaili, A. H., & Taylor, N. (2012). Omani twelfth grade students' most common misconceptions in chemistry. *Science Education International*, 23(3), 221-240.
- Anam, R. S., Widodo, A., Sopandi, W., & Wu, H. K. (2019). Developing a five-tier diagnostic test to identify students' misconceptions in science: an example of the heat transfer concepts. *Elementary Education Online*, 18(3), 1014-1029.
- Apaydin, Z., & Sürmeli, H. (2006). Üniversite öğrencilerinin doğal seçim, adaptasyon ve mutasyon ile ilgili görüşleri. *Ondokuz Mayıs Üniversitesi Eğitim Fakültesi Dergisi*, 22.
- Artdej, R., Ratanaroutai, T., Coll, R. K., & Thongpanchang, T. (2010). Thai Grade 11 students' alternative conceptions for acid-base chemistry. *Research in Science & Technological Education*, 28(2), 167-183.
- Badenhorst, E., Mamede, S., Abrahams, A., Bugarith, K., Friedling, J., Gunston, G., ... & Schmidt, H. G. (2016). First-year medical students' naïve beliefs about respiratory physiology. *Advances in Physiology Education*, 40(3), 342-348.
- Bahar, M. (2002). Students' learning difficulties in biology: Reasons and solutions. *Kastamonu Eğitim Dergisi*, 10(1), 73-82.
- Bahar, M., Johnstone, A. H., & Hansell, M. H. (1999). Revisiting learning difficulties in biology. *Journal of Biological Education*, 33(2), 84-86.

- Bearman, M., Smith, C. D., Carbone, A., Slade, S., Baik, C., Hughes-Warrington, M., & Neumann, D. L. (2012). Systematic review methodology in higher education. *Higher Education Research & Development*, 31(5), 625-640. <https://doi.org/10.1080/07294360.2012.702735>
- Boyes, E., Stanisstreet, M., & Papantoniou, V. S. (1999). The ideas of Greek high school students about the "ozone layer". *Science Education*, 83(6), 724-737.
- Caleon, I. S., & Subramaniam, R. (2010). Development and application of a three-tier diagnostic test to assess secondary students' understanding of waves. *International Journal of Science Education*, 32(7), 939-961.
- Cavallo, A. M. L., & Schafer, L. E. (1994). Relationships between students' meaningful learning orientation and their understanding of genetics topics. *Journal of Research in Science Teaching*, 31(4), 393-418.
- Ceylan, Ö., & Umdü Töpsakal, Ü. (2022). Canlı. Canan Laçın Şimşek (Ed.), *Fen öğretiminde kavram yanlışları, tespiti ve giderilmesi* 363-387. Pegem Akademi.
- Chu, H. E., Treagust, D. F., & Chandrasegaran, A. L. (2009). A stratified study of students' understanding of basic optics concepts in different contexts using two-tier multiple-choice items. *Research in Science & Technological Education*, 27(3), 253-265.
- Dikmenli, M. (2010). Misconceptions of cell division held by student teachers in biology: A drawing analysis. *Scientific Research and Essay*, 5(2), 235-247.
- Dikmenli, M., Cardak, O., & Kiray, S.A. (2011). Science student teachers' ideas about the 'gene' concept. *Procedia Social and Behavioral Sciences*, 15, 2609-2613.
- Driver, R. (1988). *The pupils as scientist?*, Open University Press, Milton Keynes, England & Philadelphia, USA.
- Driver, R., & Easley, J. (1978). Pupils and paradigms: A review of literature related to concept development in adolescent science students. *Studies in Science Education*, 5, 61-84.
- Duit, R., & Treagust, D. F. (2003). Conceptual change: A powerful framework for improving science teaching and learning. *International Journal of Science Education*, 25(6), 671-688.
- Dykstra, D. I., Boyle, C. F., & Monarch, I. A. (1992). Studying conceptual change in learning physics. *Science Education*, 76(6), 15-652.
- Eryılmaz, A., & Sürmeli E., (2002). Üç-aşamalı sorularla öğrencilerin ısı ve sıcaklık konularındaki kavram yanlışlarının ölçülmesi. <https://users.metu.edu.tr/eryilmaz/TamUcBaglant.pdf> adresinden erişilmiştir.
- Gough, D. (2007). Giving voice: Evidence-informed policy and practice as a democratizing process. In M. Reiss, R. DePalma, & E. Atkinson (Eds.), *Marginality and difference in education and beyond* (pp. 31-45). London: Trentham Books.
- Griffard, P. B., & Wandersee, J. H. (2001). The two-tier instrument on photosynthesis: What does it diagnose?. *International Journal of Science Education*, 23(10), 1039-1052.
- Gürsel, F. G., & Akçay, B. (2022). Developing two-tier diagnostic instrument to determine misconceptions on socioscientific issues. *Cumhuriyet International Journal of Education*, 11(1). <https://doi.org/10.30703/cije.1016641>
- Habiddin, H., & Page, E. M. (2019). Development and validation of a four-tier diagnostic instrument for chemical kinetics (FTDICK). *Indonesian Journal of Chemistry*, 19(3), 720-736.
- Haslam, F., & Treagust, D. F. (1987). Diagnosing secondary students' misconceptions of photosynthesis and respiration in plants using a two-tier multiple choice instrument. *Journal of Biological Education*, 21(3), 203-211.
- Hestenes, D., & Halloun, I. (1995). Interpreting the force concept inventory. A response to Huffman and Heller. *The Physics Teacher*, 33, 502-506.
- Hill, M., Sharma, M. D., O'Byrne, J., & Airey, J. (2014). Developing and evaluating a survey for representational fluency in science. *International Journal of Innovation in Science and Mathematics Education*, 22(6).
- Irmak, M., İnaltun, H., Ercan-Dursun, J., Yanış-Kelleci, H., & Yürük, N. (2023). Development and application of a three-tier diagnostic test to assess pre-service science teachers' understanding on work-power and energy concepts. *International Journal of Science and Mathematics Education*, 21(1), 159-185.
- Kaltakci Gurel, D., Eryılmaz, A., & McDermott, L. C. (2015). A review and comparison of diagnostic instruments to identify students' misconceptions in science. *Eurasia Journal of Mathematics, Science & Technology Education*. 11(5), 989-1008.
- Kılıç, D., Taber, K. S., & Winterbottom, M. (2016). A cross-national study of students' understanding of genetics concepts: Implications from similarities and differences in England and Turkey. *Education Research International*. Article ID 6539626.. <https://doi.org/10.1155/2016/6539626>
- Kiray, S. A., & Simsek, S. (2021). Determination and evaluation of the science teacher candidates' misconceptions about density by using four-tier diagnostic test. *International Journal of Science and Mathematics Education*, 19, 935-955. <https://doi.org/10.1007/s10763-020-10087-5>

- Kiray, S. A., Aktan, F., Kaynar, H., Kilinc, S. & Gorkemli, T. (2015). A descriptive study of pre-service science teachers' misconceptions about sinking–floating. *Asia-Pacific Forum on Science Learning and Teaching*, 16(2).
- Kucam, E. & Demir, E. (2020). Kavram yanlışları ile yanıtlayıcı davranışları arasındaki ilişkilerin matematik okuryazarlığı örneği üzerinde incelenmesi. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 53(1), 157-183.
- Kumandaş, B., Ateskan, A., & Lane, J. (2019). Misconceptions in biology: A meta-synthesis study of research, 2000-2014. *Journal of Biological Education*, 53(4), 350-364.
- Lemma, A. (2012). Diagnosing the diagnostics: Misconceptions of twelfth grade students on selected chemistry concepts in two preparatory Schools in eastern Ethiopia. *African Journal of Chemical Education*, 2(2), 16-31.
- Lewis, J., & Wood-Robinson, C. (2000). Genes, chromosomes, cell division and inheritance-do students see any relationship?. *International Journal of Science Education*, 22(2), 177-195.
- Liampa, V., Malandrakis, G. N., Papadopoulou, P., & Pnevmatikos, D. (2019). Development and evaluation of a three-tier diagnostic test to assess undergraduate primary teachers' understanding of ecological footprint. *Research in Science Education*, 49(3), 711-736. <https://doi.org/10.1007/s11165-017-9643-1>
- Mann, M., & Treagust, D. F. (1998). A pencil and paper instrument to diagnose students' conceptions of breathing, gas exchange and respiration. *Australian Science Teachers Journal*, 44(2), 55-59.
- Miles, M. B., & Huberman, A. M. (2016). *An extended sourcebook: Qualitative data analysis*. Pegem Akademi.
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *International Journal of Surgery*, 8(5), 336-341. <https://doi.org/10.1016/j.ijsu.2010.02.007>
- Munn, Z., Peters, M. D. J., Stern, C., Tufanaru, C., McArthur, A., & Aromataris, E. (2018). Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology*, 18, 1-7.
- Munson, B. H. (1994). Ecological misconceptions. *The Journal of Environmental Education*, 25(4), 30-34.
- Odom, A. L., & Barrow, L. H. (1995). Development and application of a two-tier diagnostic test measuring college biology students' understanding of diffusion and osmosis after a course of instruction. *Journal of research in Science Teaching*, 32(1), 45-61.
- Ola Adeniyi, E. (1985). Misconceptions of selected ecological concepts held by some Nigerian students. *Journal of Biological Education*, 19(4), 311-316.
- Önder Çelikkıran, N., & Kızılcık, H. Ş. (2022). A review of studies about four-tier diagnostic tests in physics education. *Journal of Turkish Science Education*, 19(4), 1291-1311.
- Özdamar, K. (2017). *Eğitim, sağlık ve davranış bilimlerinde ölçek ve test geliştirme yapısal eşitlik modellemesi-IBM SPSS, IBM SPSS AMOS ve MINITAB uygulamalı*. Nisan.
- Palmer, D. H. (1998). Measuring contextual error in the diagnosis of alternative conceptions in science, *Issues in Educational Research*, 8(1), 65-76.
- Pearson, J. T., & Hughes, W. J. (1988). Problems with the use of terminology in genetics education: 1, A literature review and classification scheme. *Journal of Biological Education*, 22(3), 178-182.
- Rahayu, S., Treagust, D. F. & Chandrasegaran, A.L. (2022). High school and preservice chemistry teacher education students' understanding of voltaic and electrolytic cell concepts: Evidence of consistent learning difficulties across years. *International Journal of Science and Mathematics Education*, 20, 1859-1882. <https://doi.org/10.1007/s10763-021-10226-6>
- Resbiantoro, G., Setiani, R. & Dwikoranto (2022). A review of misconception in physics: The diagnosis, causes, and remediation. *Journal of Turkish Science Education*, 19(2), 403-427.
- Rosida, N., Widarti, H. R., & Yahmin, Y. (2022). Diagnostic test on chemical equilibrium misconception: A literature review. *Jurnal Pembelajaran Kimia*, 7(2), 49-57.
- Sadava, D., Hillis, D.M., Heller, H.C., & Berenbaum, M.R. (2014). *Yaşam bilimi biyoloji*, Çeviri Editörleri: Ertunç Gündüz, İsmail Türkan, Palme Yayıncılık.
- Seçer, İ. (2013). *SPSS ve LISREL ile pratik veri analizi-Analiz ve raporlaştırma*. Ankara: Anı.
- Soeharto, S. (2021). Development of a diagnostic assessment test to evaluate science misconceptions in terms of school grades: A rasch measurement approach. *Journal of Turkish Science Education*, 18(3), 351-370.
- Soeharto, S., & Csapó, B. (2021). Evaluating item difficulty patterns for assessing student misconceptions in science across physics, chemistry, and biology concepts. *Heliyon*, 7(11). <https://doi.org/10.1016/j.heliyon.2021.e08352>
- Soeharto, S., Csapó, B., Sarimanah, E., Dewi, F. I., & Sabri, T. (2019). A review of students' common misconceptions in science and their diagnostic assessment tools. *Jurnal Pendidikan IPA Indonesia*, 8(2), 247-266.
- Songer, C. J., & Mintzes, J. J. (1994). Understanding cellular respiration: An analysis of conceptual change in college biology. *Journal of Research in Science Teaching*, 31(6), 621-637.

- Šorgo, A., Ambrožič-Dolinšek, J., Uşak, M., & Özel, M. (2011). Knowledge about and acceptance of genetically modified organisms among pre-service teachers: A comparative study of Turkey and Slovenia. *Electronic Journal of Biotechnology*, 14(4), 1-16. <https://doi.org/10.2225/vol14-issue4-fulltext-5>
- Stavy, R. (1988). Students' understanding of photosynthesis. *The American Biology Teacher*, 50(4), 208-212.
- Strauss, A. L. & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park, CA: Sage.
- Tabak, S. (2019). 6th, 7th and 8th grade students' misconceptions about the order of operations. *International Journal of Educational Methodology*, 5(3), 363-373.
- Taban, T., & Kiray, S. A. (2022). Determination of science teacher candidates' misconceptions on liquid pressure with four-tier diagnostic test. *International Journal of Science and Mathematics Education*, 20, 1791-1811. <https://doi.org/10.1007/s10763-021-10224-8>
- Tamir, P. (1971). An alternative approach to the construction of multiple choice test items. *Journal of Biological Education*, 5(6), 305-307.
- Treagust, D. F. (2006). Diagnostic assessment in science as a means to improving teaching, learning and retention. *UniServe Science-Symposium Proceedings: Assessment in science teaching and learning*, Uniserve Science, Sydney, Australia.
- Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. V., & Orr, R. B. (2022). *Campbell biyoloji*, Çeviri Editörleri: Ertunç Gündüz, İsmail Türkan, Palme Yayıncılık.
- Versteeg, M., Wijnen-Meijer, M., & Steendijk, P. (2019). Informing the uninformed: A multitier approach to uncover students' misconceptions on cardiovascular physiology. *Advances in Physiology Education*, 43(1), 7-14.
- Voska, K. W., & Heikkinen, H. W. (2000). Identification and analysis of student conception used to solve chemical equilibrium problems. *Journal of Research in Science Teaching*, 37(2), 160-176.
- Waheed, T., & Lucas, A. M. (1992). Understanding interrelated topics: photosynthesis at age photosynthesis at age14+. *Journal of Biological Education*, 26(3), 193-199.
- Wang, J. R. (2004). Development and validation of a two-tier instrument to examine understanding of internal transport in plants and the human circulatory system. *International Journal of Science and Mathematics Education*, 2, 131-157. <https://doi.org/10.1007/s10763-004-9323-2>
- Wulandari, T., Ramli, M., & Muzzazinah. (2020, June). The assessment of understanding biological concepts: a systematic review. In *Journal of Physics: Conference Series* (Vol. 1563, No. 1, p. 012030). IOP Publishing.
- Yang, D. C., & Sianturi, I. A. J. (2021). Sixth grade students' performance, misconception, and confidence on a three-tier number sense test. *International Journal of Science and Mathematics Education*, 19, 355-375. <https://doi.org/10.1007/s10763-020-10051-3>
- Yen, C. F., Yao, T. W., & Chiu, Y. C. (2004). Alternative conceptions in animal classification focusing on amphibians and reptiles: A cross-age study. *International Journal of Science and Mathematics Education*, 2, 159-174. <https://doi.org/10.1007/s10763-004-1951-z>
- Yen, C., Yao, T., & Mintzes, J. J. (2007). Taiwanese students' alternative conceptions of animal biodiversity. *International Journal of Science Education*, 29(4), 535-553. <https://doi.org/10.1080/09500690601073418>
- Yurdugül, H. (2008). Minimum sample size for Cronbach's coefficient alpha: A Monte-Carlo study. *Hacettepe University Journal of Education*, 35(1), 397-405.

---

### Author(s) Information

---

**Tuğçe Duran**

Department of Biology Education, Ahmet Keleşoğlu  
Faculty of Education, Necmettin Erbakan University,  
Konya, Türkiye  
Contact e-mail: [tugce.gulesir@erbakan.edu.tr](mailto:tugce.gulesir@erbakan.edu.tr)  
ORCID iD: <https://orcid.org/0000-0003-2428-7635>

**Musa Dikmenli**

Department of Biology Education, Ahmet Keleşoğlu  
Faculty of Education, Necmettin Erbakan University,  
Konya, Türkiye  
ORCID iD: <https://orcid.org/0000-0001-6501-9034>

---

## APPENDIX: List of Articles Reviewed

- 1) Akyurt, C., & Akaydin, G. (2009). Biyoloji öğretmen adaylarında bitkilerde madde taşınması konusundaki kavram yanlışları. *Kastamonu Üniversitesi Kastamonu Eğitim Dergisi*, 17(1).
- 2) AlHarbi, N. N. S., Treagust, D. F., Won, M., & Chandrasegaran, A. L. (2015). Influence of particle theory conceptions on pre-service science teachers' understanding of osmosis and diffusion. *Journal of Biological Education*, 49(3), 232-245. <https://doi.org/10.1080/00219266.2014.923488>
- 3) Andariana, A., Zubaidah, S., Mahanal, S., & Suarsini, E. (2020). Identification of biology students' misconceptions in human anatomy and physiology course through three-tier diagnostic test. *Journal for the Education of Gifted Young Scientists*, 8(3), 1071-1085. <https://doi.org/10.17478/jegys.752438>
- 4) Apaydin, Z., & Sürmeli, H. (2006). Üniversite öğrencilerinin doğal seçim, adaptasyon ve mutasyon ile ilgili görüşleri. *Ondokuz Mayıs Üniversitesi Eğitim Fakültesi Dergisi*, 22.
- 5) Arslan, H. O., Cigdemoglu, C., & Moseley, C. (2012). A three-tier diagnostic test to assess pre-service teachers' misconceptions about global warming, greenhouse effect, ozone layer depletion, and acid rain. *International Journal of Science Education*, 34(11), 1667-1686. <https://doi.org/10.1080/09500693.2012.680618>
- 6) Arslan, H. Ö., Geban, Ö., & Sağlam, N. (2015). Learning cycle model to foster conceptual understanding in cell division and reproduction concepts. *Journal of Baltic Science Education*, 14(5), 670-684. <https://doi.org/10.33225/jbse/15.14.670>
- 7) Artun, H., & Coştu, B. (2011). Unveiling primary student-teachers' misconceptions about diffusion and osmosis. *Journal of Turkish Science Education*, 8(4).
- 8) Assimi, E., Idrissi, R. J., Zerhane, R., & Boubih, S. (2022). The use of a three-tier diagnostic test to investigate conceptions related to cell biology concepts among pre-service teachers of life and earth sciences. *Journal of Biological Education*. <https://doi.org/10.1080/00219266.2022.2134175>
- 9) Atchia, S. M. C., Chummun, D., & Luckho, S. (2022). Use of design thinking as a strategy to identify and clear students' misconceptions in photosynthesis: A case study. *Journal of Biological Education*, 1-18. <https://doi.org/10.1080/00219266.2022.2100452>
- 10) Atmaca Aksoy, A. C., & Erten, S. (2022). A four-tier diagnostic test to determine pre-service science teachers' misconception about global warming. *Journal of Baltic Science Education*, 21(5), 747-761.
- 11) Aymen Peker, E., & Taş, E. (2019). Investigation of in-school factors affecting distributed leadership practices. *Journal of Computer and Education Research*, 7(14), <https://doi.org/10.18009/jcer.602223>
- 12) Bozdağ, H. C. (2017). Üç aşamalı kavramsal ölçme aracı ile öğrencilerin sindirim sistemi konusundaki kavram yanlışlarının tespiti. *Bartın Üniversitesi Eğitim Fakültesi Dergisi*, 6(3).
- 13) Bozdağ, H. C., & Ok, G. (2018). Dört aşamalı kavramsal ölçme aracı ile sekizinci sınıf öğrencilerinin hücre bölünmeleri konusundaki bilgi farkındalıkları ile kavram yanlışlarının belirlenmesi. *Sakarya University Journal of Education*, 8(2), <https://doi.org/10.19126/suje.389668>
- 14) Bozdağ, H. C., & Ok, G. (2019). Determination of the knowledge awareness and misconceptions of sixth grade students about the cell with four tier test. *Adıyaman University Journal of Educational Sciences*, 9(1).
- 15) Chabalengula, V. M., Sanders, M., & Mumba, F. (2012). Diagnosing students' understanding of energy and its related concepts in biological context. *International Journal of Science and Mathematics Education*, 10, 241-266. <https://doi.org/10.1007/s10763-011-9291-2>
- 16) Chang, C.-H., Pascua, L., & Ess, F. (2018). Closing the "Hole in the Sky": The use of refutation-oriented instruction to correct students' climate change misconceptions. *Journal of Geography*, 117(1), 3-16. <https://doi.org/10.1080/00221341.2017.1287768>
- 17) Choirina, Y., Ramli, M., & Rinanto, Y. (2019). Conceptual understanding of high school students on plantae. *Journal of Education and Learning (EduLearn)*, 13(1), 48-56.
- 18) Çakır Yıldırım, B., & Karaarslan Semiz, G. (2022). Üç aşamalı ekolojik ayak izi tanı testinin Türkçe'ye uyarlanması. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi*, 23(2), <https://doi.org/10.29299/kefad.979056>
- 19) Çakır, M., & Aldemir, B. (2011). İki aşamalı genetik kavramlar tanı testi geliştirme ve geçerlik çalışması. *Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 8(16).
- 20) Çiğdemoglu, C., & Arslan, H. Ö. (2017). Atmosfer ile ilgili çevre problemleri konularında kavram yanlışlarını tespit eden üç aşamalı tanı testinin Türkçe'ye uyarlanması. *Van Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 14(1).
- 21) da Rosa, J. A. (2022). Comparing climate science misconceptions with worldview and cognitive reflection suggests poor understanding and motivated reasoning among undergraduates. *Journal of Geoscience Education*, 70(4), 501-516. <https://doi.org/10.1080/10899995.2021.2006549>
- 22) Dikmenli, M., Türkmen, L., Çardak, O., & Kurt, H. (2005). Biyoloji öğretmen adaylarının genel biyoloji konularındaki kavram yanlışlarının iki aşamalı çoktan seçmeli bir araç ile belirlenmesi. *Dokuz Eylül Üniversitesi Buca Eğitim Fakültesi Dergisi*, 17.

- 23) Duda, H. J., & Adprijadi. (2020). Students' misconception in concept of biology cell. *Anatolian Journal of Education*, 5(1), 47-52.
- 24) Fikr, R. A., Suwono, H., & Susilo, H. (2022). Online three-tier diagnostic test to identify misconception about virus and COVID-19. *Jurnal Pendidikan Biologi Indonesia (JPBI)*, 8(2), 129-141.
- 25) Fisher, K. M., Williams, K. S., & Lineback, J. E. (2011). Osmosis and diffusion conceptual assessment. *CBE-Life Sciences Education*, 10(4), 418-429. <https://doi.org/10.1187/cbe.11-04-0038>
- 26) Gürsel, F. G., & Akçay, B. (2022). Developing two-tier diagnostic instrument to determine misconceptions on socioscientific issues. *Cumhuriyet International Journal of Education*, 11(1). <https://doi.org/10.30703/cije.1016641>
- 27) Kao, H. (2007). A study of aboriginal and urban junior high school students' alternative conceptions on the definition of respiration. *International Journal of Science Education*, 29(4), 517-533. <https://doi.org/10.1080/09500690601073376>
- 28) Karşlı Baydere, F., & Kurtoğlu, S. (2020). 5. Sınıf öğrencilerinin biyolojik çeşitlilik konusundaki kavramsal anlamalarına REACT stratejisinin etkisi. *Van Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 17(1), <https://doi.org/10.33711/yyuefd.800921>
- 29) Karşlı, F., & Saka, Ü. (2017). 5. Sınıf öğrencilerinin 'Besinleri Tanıyalım' konusundaki kavramsal anlamalarına bağlam temelli yaklaşımın etkisi. *İlköğretim Online (elektronik)*, 16(3). <https://doi.org/10.17051/ilkonline.2017.330230>
- 30) Keskin, B., & Özay Köse, E. (2017). Misconceptions of prospective biology teachers about theory of evolution. *Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education*, 11(2).
- 31) Kılıç, D., & Sağlam, N. (2009). Development of a two-tier diagnostic test to determine students' understanding of concepts in genetics. *Eurasian Journal of Educational Research*, 36.
- 32) Kılıç, D., Taber, K. S., & Winterbottom, M. (2016). A cross-national study of students' understanding of genetics concepts: Implications from similarities and differences in England and Turkey. *Education Research International*. <https://doi.org/10.1155/2016/6539626>
- 33) Li, M. (2022). An exploration on concept teaching knowledge of biology teachers in lower-secondary schools. *Journal of Baltic Science Education*, 21(1), 121-139. <https://doi.org/10.33225/jbse/22.21.121>
- 34) Liampa, V., Malandrakis, G. N., Papadopoulou, P., & Pnevmatikos, D. (2019). Development and evaluation of a three-tier diagnostic test to assess undergraduate primary teachers' understanding of ecological footprint. *Research in Science Education*, 49(3), 711-736. <https://doi.org/10.1007/s11165-017-9643-1>
- 35) Lim, H. L., & Poo, Y. P. (2021). A diagnosis of students' misconceptions of photosynthesis and plant respiration. *Asia Pacific Journal of Educators and Education*, 36(1), 155-176. <https://doi.org/10.21315/apjee2021.36.1.9>
- 36) Lin, S. W. (2004). Development and application of a two-tier diagnostic test for high school students' understanding of flowering plant growth and development. *International Journal of Science and Mathematics Education*, 2, 175-199. <https://doi.org/10.1007/s10763-004-6484-y>
- 37) Maier, U., Wolf, N., & Randler, C. (2016). Effects of a computer-assisted formative assessment intervention based on multiple-tier diagnostic items and different feedback types. *Computers & Education*, 95, 85-98. <https://doi.org/10.1016/j.compedu.2015.12.002>
- 38) Milenković, D. D., Hrin, T. N., Segedinac, M. D., & Saša, H. (2016). Development of a three-tier test as a valid diagnostic tool for identification of misconceptions related to carbohydrates. *Journal of Chemical Education*, 93(9), 1514-1520. <https://doi.org/10.1021/acs.jchemed.6b00261>
- 39) Oladipo, A. J., Osokoya, M., & Udeani, U. (2019). Conceptual understanding and application of diffusion and osmosis: An assessment of pre-degree students in a Nigerian University. *Journal of the International Society for Teacher Education*, 23(2), 82-92.
- 40) Ormancı, Ü., & Özcan, S. (2012). Fen ve teknoloji dersi vücudumuzda sistemler ünitesinde drama yönteminin etkililiği: İki aşamalı teşhis testi kullanımı. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi*, 6(2).
- 41) Özdemir, G., & Çalışkan, İ. (2018). Ortaokul 5. ve 6. Sınıf öğrencilerinin "Omurgalı ve Omurgasız Hayvanların Sınıflandırılması" konusuna ilişkin kavram yanılgıları. *İlköğretim Online (elektronik)*, 17(2).
- 42) Park, M., & Johnson, J. A. (2016). Evaluation of students' energy conception in environmental science. *International Journal of Environmental and Science Education*, 11(12), 5572-5590.
- 43) Park, M., & Liu, X. (2016). Assessing understanding of the energy concept in different science disciplines. *Science Education*, 100(3), 483-516. <https://doi.org/10.1002/sce.21211>
- 44) Ping, I. L. L., Halim, L., Osman, K., & Department of Teaching and Learning Innovation, Faculty of Education, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia. (2020). Rural students' conceptual understanding of diffusion and osmosis. *Asia Pacific Journal of Educators and Education*, 35(1), 93-109. <https://doi.org/10.21315/apjee2020.35.1.6>

- 45) Putica, K. B. (2022). Development and validation of a four-tier test for the assessment of secondary school students' conceptual understanding of amino acids, proteins, and enzymes. *Research in Science Education*, 53(3), 651-668. <https://doi.org/10.1007/s11165-022-10075-5>
- 46) Romine, W. L., Schaffer, D. L., & Barrow, L. (2015). Development and application of a novel raschbased methodology for evaluating multi-tiered assessment instruments: Validation and utilization of an undergraduate diagnostic test of the water cycle. *International Journal of Science Education*, 37(16), 2740-2768. <https://doi.org/10.1080/09500693.2015.1105398>
- 47) Saat, R. M., Fadzil, H. M., Aziz, N. A. A., Haron, K., Rashid, K. A., & Shamsuar, N. R. (2016). Development of an online three-tier diagnostic test to assess pre-university students' understanding of cellular respiration. *Journal of Baltic Science Education*, 15(4), 532-546.
- 48) Sesli, E., & Kara, Y. (2012). Development and application of a two-tier multiple-choice diagnostic test for high school students' understanding of cell division and reproduction. *Journal of Biological Education*, 46(4), 214-225. <https://doi.org/10.1080/00219266.2012.688849>
- 49) Shi, J., Knight, J. K., Chun, H., Guild, N. A., & Martin, J. M. (2017). Using pre-assessment and in-class questions to change student understanding of molecular movements. *Journal of Microbiology & Biology Education*, 18(1). <https://doi.org/10.1128/jmbe.v18i1.1195>
- 50) Soeharto, S. (2021). Development of a diagnostic assessment test to evaluate science misconceptions in terms of school grades: A rasch measurement approach. *Journal of Turkish Science Education*, 18(3), 351-370.
- 51) Soeharto, S., & Csapó, B. (2021). Evaluating item difficulty patterns for assessing student misconceptions in science across physics, chemistry, and biology concepts. *Heliyon*, 7(11). <https://doi.org/10.1016/j.heliyon.2021.e08352>
- 52) Sopandi, W., & Sukardi, R. R. (2020). Using four-tier diagnostic tests to understand the conceptions held by pre-service primary school teachers about sea pollutant migration. *Review of International Geographical Education Online*, 10(2), 13-29. <https://doi.org/10.33403/rigeo.629388>
- 53) Šorgo, A., Ambrožič-Dolinšek, J., Ušak, M., & Özel, M. (2011). Knowledge about and acceptance of genetically modified organisms among pre-service teachers: A comparative study of Turkey and Slovenia. *Electronic Journal of Biotechnology*, 14(4), 1-16. <https://doi.org/10.2225/vol14-issue4-fulltext-5>
- 54) Stern, F., Kampourakis, K., Huneault, C., Silveira, P., & Müller, A. (2018). Undergraduate biology students' teleological and essentialist misconceptions. *Education Sciences*, 8(3), 135. <https://doi.org/10.3390/educsci8030135>
- 55) Suwono, H., Prasetyo, T. I., Lestari, U., Lukiat, B., Fachrunnisa, R., Kusairi, S., Saefi, M., Fauzi, A., & Atho'illah, M. F. (2019). Cell Biology Diagnostic Test (CBD-Test) portrays pre-service teacher misconceptions about biology cell. *Journal of Biological Education*, 55(1), 82-105. <https://doi.org/10.1080/00219266.2019.1643765>
- 56) Svandova, K. (2014). Secondary school students' misconceptions about photosynthesis and plant respiration: Preliminary results. *Eurasia Journal of Mathematics, Science & Technology Education*, 10(1), 59-67. <https://doi.org/10.12973/eurasia.2014.1018a>
- 57) Tomažič, I., & Vidic, T. (2012). Future science teachers' understandings of diffusion and osmosis concepts. *Journal of Biological Education*, 46(2), 66-71. <https://doi.org/10.1080/00219266.2011.617765>
- 58) Tsui, C. Y., & Treagust, D. F. (2010). Evaluating secondary students' scientific reasoning in genetics using a two-tier diagnostic instrument. *International Journal of Science Education*, 32(8), 1073-1098. <https://doi.org/10.1080/09500690902951429>
- 59) Tunç, T., Akçam, H. K., & Dökme, İ. (2011). Üç aşamalı sorularla sınıf öğretmen adaylarının bazı temel fen kavramları hakkında sahip oldukları kavram yanılgıları. *Gazi Eğitim Fakültesi Dergisi*, 31(3).
- 60) Urey, M. (2018). Defining the relationship between the perceptions and the misconceptions about photosynthesis topic of the preservice science teachers. *European Journal of Educational Research*, 7(4), 813-826. <https://doi.org/10.12973/eu-jer.7.4.813>
- 61) Uzunhasanoğlu, Ö., Çakır, M., & Avcı, S. (2020). Biyoloji öğretmen adaylarının genel biyoloji kavram anlayışlarını ölçmek için iki aşamalı tanı testi geliştirilmesi ve uygulanması. *Turkish Studies - Educational Sciences*, 15(4). <https://doi.org/10.47423/TurkishStudies.44131>
- 62) Vitharana, P. R. K. A. (2015). Student misconceptions about plant transport-A Sri Lankan example. *European Journal of Science and Mathematics Education*, 3(3), 275-288.
- 63) Vlckova, J., Kubiato, M., & Usak, M. (2016). Czech high school students' misconceptions about basic genetic concepts: Preliminary results. *Journal of Baltic Science Education*, 15(6), 738-745.
- 64) Wahyono, P., & Susetyarini, E. (2021). Misconceptions of biology education students in biochemistry course during the COVID-19 pandemic. *Jurnal Pendidikan Biologi Indonesia (JPBI)*, 7(2), 104-110.
- 65) Wang, J. R. (2004). Development and validation of a two-tier instrument to examine understanding of internal transport in plants and the human circulatory system. *International Journal of Science and Mathematics Education*, 2, 131-157. <https://doi.org/10.1007/s10763-004-9323-2>

- 66) Wang, J. R. (2007). Students' thinking and alternative conceptions of transport systems in plants: A follow-up study. *International Journal of Science and Mathematics Education*, 5(2), 307-328. <https://doi.org/10.1007/s10763-006-9038-7>
- 67) Witzig, S. B., Rebello, C. M., Siegel, M. A., Freyermuth, S. K., Izci, K., & McClure, B. (2014). Building the BIKE: Development and testing of the biotechnology instrument for knowledge elicitation (BIKE). *Research in Science Education*, 44(5), 675-698. <https://doi.org/10.1007/s11165-014-9398-x>
- 68) Yen, C., Yao, T., & Mintzes, J. J. (2007). Taiwanese students' alternative conceptions of animal biodiversity. *International Journal of Science Education*, 29(4), 535-553. <https://doi.org/10.1080/09500690601073418>
- 69) Yen, C. F., Yao, T. W., & Chiu, Y. C. (2004). Alternative conceptions in animal classification focusing on amphibians and reptiles: A cross-age study. *International Journal of Science and Mathematics Education*, 2, 159-174. <https://doi.org/10.1007/s10763-004-1951-z>
- 70) Yeo, J. H., Yang, H. H., & Cho, I. H. (2022). Using a three-tier multiple-choice diagnostic instrument toward alternative conceptions among lower-secondary school students in Taiwan: Taking Ecosystems unit as an example. *Journal of Baltic Science Education*, 21(1), 69-83.
- 71) Yumuşak, A., Maraş, İ., & Şahin, M. (2016). Radyoaktivite konusunda kavram yanlışlarını belirlemeye yönelik iki aşamalı bir teşhis testinin geliştirilmesi. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 12(3).

# Pharmacy Undergraduates' Epistemic Cognition and Medication Beliefs as Predictors of Conceptual Learning and Academic Progress

Ilona Sodervik, Leena Hanski, Eero Laakkonen, Nina Katajavuori

## Article Info

### Article History

Published:  
01 October 2024

Received:  
04 June 2024

Accepted:  
03 September 2024

### Keywords

Epistemic belief  
Conceptual knowledge  
Pharmacy education

## Abstract

The importance of studies showing the impact of students' epistemic beliefs on their conceptual learning and academic progress is increasing. This study investigated pharmacy students' domain-specific epistemic cognition and topic-specific beliefs related to their level of conceptual prior knowledge, learning and study progress during the first academic year. Data were collected from 125 students using a pre-test/post-test design with a multiple-choice questionnaire, an open-ended case task and a measure of domain-specific epistemic cognition and topic-specific beliefs concerning medication. The results showed that students' prior knowledge predicted their academic performance and more sophisticated epistemic cognition was related to better conceptual understanding, faster study progress, and fewer anti-medication beliefs. Anti-medication beliefs hindered participants' success in the case task and were related to weaker study progress. Our study shows that epistemic cognition and topic-specific beliefs play a role in students' performance, learning, and study progress.

## Introduction

A challenge in European universities is that although a large proportion of the students thrive in their studies, too many struggle to reach a high-level understanding of their discipline's substance and methodologies, achieve appropriate skill levels, and proceed through the programme at the suggested pace (Jia & Maloney, 2015; Korhonen & Rautopuro, 2019). Previous research highlights the particular importance of the first year in higher education to students' future academic performance and achievement (Araujo et al., 2019; Constantini & Vitale, 2011; European Education and Culture Executive Agency, 2017; Jenert et al., 2015; Mennen & van der Klink, 2017; Trautwein & Bosse, 2017). Consequently, there is growing interest in knowing the factors relevant to predicting and influencing academic performance during students' first study year in order to advance learning and teaching in higher education (see Tinto, 1993).

According to current understanding, investigating the factors that influence high-level learning requires a domain-specific approach (Boshuizen & Schmidt, 2018), but such research is still relatively scarce. Therefore, the aim of this study is to add to the understanding of cognitive cornerstones of students' study success and professional development by investigating undergraduates' conceptual and epistemic understanding together with their topic-specific beliefs from domain-specific perspective in the pharmacy context and to compare them with respect to study progress during the first academic study year.

## Development of Conceptual Understanding

The single most significant aspect in students' knowledge building process and academic success, particularly in the early phases of study, is the level and quality of the students' prior knowledge (Binder et al., 2019; Bransford et al., 2000; Kuncel et al., 2001). The quality of prior knowledge has been shown to greatly influence how students interpret and learn from learning materials (e.g., Rittle-Johnson et al., 2009), how they use problem-solving strategies effectively (e.g. Koedinger & Roll, 2012), how they are able to use prior knowledge in novel contexts (e.g. Carpenter, et al., 1998), and how well they are prepared for future learning tasks (e.g., Bransford & Schwartz, 1999). Nevertheless, if students' prior knowledge contains information that is in conflict with current scientific understanding, learning requires profound changes in one's existing cognitive structures, i.e. conceptual change and becomes more laborious (Chi, 2013; Vosniadou, 2013). Furthermore, scientific concepts are often counterintuitive by nature and hence, learning of them requires intentional, high-level cognitive processing and sophisticated study strategies, particularly if misconceptions exist and conceptual change is needed. Based on extensive research, we now know that university students often have pre-existing,

alternative conceptions and robust misinterpretations related to the central course content and concepts to be studied (e.g. Flaig et al., 2018; Södervik et al., 2015, 2017, 2020).

However, a tricky question has bothered researchers for a long time: why are some students more likely to revise their prior conceptions and undergo conceptual change as a result of instructional intervention or regular teaching, whereas others with similar naïve conceptions are more resistant to changing their conceptions? An important reason for this is assumed to lie in the interplay between students' ideas concerning the nature of knowledge and knowing, their topic-specific beliefs and conceptual understanding (Abendroth & Richter, 2021; Amin & Levrini, 2018; Murphy & Alexander, 2013). Studies have shown that individuals' ideas about knowledge and knowing, i.e. epistemic cognition have an impact on conceptual understanding, conceptual change, and their acceptance of controversial theories (Greene et al., 2016; Hofer, 2018; Mason et al., 2008). Furthermore, belief-consistent information is more easily and better comprehended compared to belief-inconsistent information (Abendroth & Richter, 2021). However, few studies investigate these aspects in relation to academic achievement from discipline- and topic-specific perspectives, partly due to the lack of appropriate methodological tools (see e.g. Bråten et al., 2009; Hofer, 2000; Muis et al., 2006; Merk et al., 2018).

### **Domain-Specific Epistemic Cognition**

Researchers have argued that epistemic cognition is a key predictor of 21st-century learning outcomes (Cartiff et al., 2020; Greene et al., 2018, 1084). This is because in our society, scientific knowledge is growing faster than ever, and both students and professionals need to constantly evaluate the veracity and relevance of knowledge of their field. Individuals' domain-specific epistemic cognition, i.e. a set of mental processes that involve the development of one's conceptions of knowledge, plays a role in this, because it changes the premises on which students' knowledge and knowing are based (Hofer, 2016). Based on extensive earlier research, it influences reasoning, problem-solving, conceptual understanding, and argumentation and hence general academic performance (Greene et al., 2018; Hofer, 2016; Kuhn, 2005; Schommer-Aikins & Easter, 2006).

Whilst research related to epistemic cognition was previously conducted with an implicit assumption that these beliefs can be generalised across disciplines, it is now generally accepted that individuals hold both general, but also discipline- and even topic-specific beliefs about knowledge and knowing (see e.g. Bråten et al., 2008; Bråten et al., 2009; Hofer, 2000; Muis et al., 2006; Merk et al. 2018). This means that students may justify and construct knowledge in different ways depending on how they perceive the specific practices of the particular disciplines rather than general beliefs about knowledge, leading to a situation in which one may have beliefs about scientific knowledge in general that differ from one's beliefs on pharmaceutical knowledge and from their beliefs related to particular phenomenon. This may be true particularly in weighing information related to beliefs about complex and controversial issues, such as those related to drugs or vaccines in pharmaceuticals that are developed in the academic context, but also influenced by larger social-cultural contexts (Kolstø, 2001; Merk et al., 2018; Sinatra et al., 2014). All in all, a growing body of empiric evidence supports the view that domain-specific epistemic beliefs deserve further exploration, particularly in relation to academic performance and conceptual understanding.

### **Discipline of Pharmacy as a Study Context**

Pharmacy is a multidisciplinary field involving the natural sciences, health sciences and social sciences – disciplines that differ greatly in terms of research methodology. Furthermore, underlying theories related to health concepts are typically multifaceted, constantly evolving, and often disputed. Topics concerning unsettled or controversial issues related to people's health or safety may be well suited to eliciting epistemic cognition and topic-specific beliefs, because they often include disagreements regarding the validity or trustworthiness of the claims involved (Kolstø 2001; Sinatra et al., 2014).

Of all the scientific disciplines, knowledge related to health sciences has perhaps the greatest influence on individuals' daily lives and the choices that we make regarding nutrition, immunisation and medical treatment. Thus, pharmacists involved in patient-centred duties in pharmacies typically act as critical mediators and interpreters of scientific knowledge between the pharmaceutical industry, as drug developers, and customers, who may have adopted false or oversimplified information regarding drugs and their appropriate use (Kaufman et al., 2013). This is of particular concern when it comes to common medications such as vaccines and antibiotics, because the pharmacists need to be able to explain complex things plausibly to the customers. However, recent studies have shown severe problems in pharmacy students' understanding of elementary-level

knowledge about antibiotics and antibiotic resistance mechanisms, and beliefs related to the factors promoting the emergence and spread of antibiotic resistance (Inacio et al., 2017; Södervik et al., 2020). Therefore, better discipline- and topic-specific understanding of these aspects and their interrelations are needed.

Students, also in pharmacy, typically struggle to apply basic knowledge in practical problem-solving situations (e.g. Boshuizen et al., 2020; Persky & Murphy, 2019; Södervik et al., 2023). According to our current understanding, knowledge and the associated skills to use the knowledge develop simultaneously and interdependently (Boshuizen & Schmidt, 2018), and hence learning activities that require mental activities and processes like those necessary in real work life would be beneficial (Brown et al., 1989). The use of authentic, discipline-specific case tasks can have the potential to support learning effectively, particularly in the early stages of education, when a great part of learners' cognitive capacity is laden with the processing of less integrated conceptual knowledge, and authentic problems can be too demanding (see e.g. Boshuizen et al., 2020).

## The Study

Fostering the development of expertise in life sciences is a goal of the utmost importance in universities, since the rapidly changing society and the current global catastrophes and crises (such as pandemics and antibiotic resistance) challenge future experts to be able to evaluate and update their knowledge and skills in constantly evolving circumstances. Furthermore, the initial phase of study at university deserves more attention in terms of students' domain-specific conceptual and epistemic understanding. Therefore, in the present study we examined:

(1) How does first-year pharmacy students' conceptual understanding related to biosciences change during their first study year?

There is abundant evidence demonstrating the important role played by personal epistemology in students' academic performance, but rather limited understanding of this from domain- and topic-specific perspectives. Therefore, we also ask:

(2) How do students' domain-specific epistemic cognition and topic-specific beliefs predict their academic performance during the first study year?

## Methods

### Participants

Participating in this study was voluntary and informed consent was obtained from the participants in both data collection phases. The participants in the pre-test were 125 first-year pharmacy students (women  $n = 97$ ; men  $n = 26$ ; other  $n = 2$ ) from the University of Helsinki, and 81 of those (women  $n = 64$ ; men  $n = 16$ ; other  $n = 1$ ) took part in both data collection phases of the study.

### Measures and Procedure

#### *Design*

A pre-test/post-test design was utilised. The pre-test measuring students' level and quality of prior knowledge was conducted at the beginning of the academic studies (September 2018) and the post-test six months later (February 2019). Students answered individually in a Moodle electronic exam environment in a regular lecture hall, and they had 40 minutes to complete the tasks.

#### *Measures – Background Measures and Measure of Conceptual Understanding*

Background questions, such as the students' gender and commencement year, were asked during the pre-test. Students' entrance examination scores and accumulated number of study credits (information obtained from the student register) at the end of the first study year were used as background variables. The measure of conceptual

understanding in pre- and posttests consisted of 13 MCQs with the one-right/three-wrong answer model, and a case task with two related open-ended questions. The MCQs tested the students' basic knowledge of biochemistry, cell biology and molecular biology and referred to items in the phenomena to be handled in the case task. The MCQs were designed at a level that should have been mastered by the time students had completed Finnish high school. Students were permitted to leave the questions unanswered.

The case task was designed to require the application of conceptual knowledge, measuring conceptual change. It was the following 63-word description of a patient encounter with drug handling and antibiotic resistance (original task in Finnish): "A patient enters a pharmacy to pick up their prescribed broad-spectrum antibiotic. The patient tells you that this is their second prescribed antibiotic because the inflammation-relieving action of the first antibiotic was not effective, and in fact the symptoms got worse.

The patient seems dubious and tells you that he has heard several stories of the injurious side-effects of antibiotics, especially for intestinal bacteria. You suggest that the patient could take registered probiotics along with the antibiotics. Probiotics are used, for example, to prevent and alleviate antibiotic diarrhoea". After reading the case description, the following questions were asked: A) Name and define the essential science concepts which are related to this case (2–4 concepts). B) Describe the mechanisms related to this scientific phenomenon – both from the side of the human and the microbe. The case was designed to measure (A) the student's ability to identify and explain the central concepts regarding the phenomenon of drug-microbe interactions, covering antibiotic resistance and the effects of antibiotics on gut microbiota.

Table 1. Scoring of the case task

Case Task	Concepts and Contents	Grading Max: 16 points
<i>Conceptual level</i>	<u><i>Naming of the concept</i></u>	1 point for each correct concept and definition
Question A: Name and define the essential science concepts which are related to this case (up to four concepts)	<ul style="list-style-type: none"> <li>-Antibiotics (course),</li> <li>-Resistance/antibiotic resistance</li> <li>-Probiotics</li> <li>-Normal bacterial microbiota / intestinal microbiota</li> <li>-Microbe/bacterium/bacterial infection</li> <li>-Recombination/ genetic variation</li> <li>-Loss of microbiota</li> <li>-Superbug/ nosocomial infection;</li> <li>-Broad-spectrum / narrow spectrum antibiotics</li> <li>-Immunity</li> <li>-Antibiotic diarrhoea</li> </ul> <u><i>Defining of the concept</i></u> Defining the concept	
<i>Phenomenal level</i>	<u><i>Microbe's point of view:</i></u>	1 point for each correct mechanism
Question B: Describe the mechanisms underlying this scientific phenomenon – both from the side of human and the microbe	<ul style="list-style-type: none"> <li>-Resistance is a consequence of changes in genetic information</li> <li>-Resistance may occur as a structural change in the antibiotic target molecule</li> <li>-An ability of the microbe to cleave the antibiotic or to prevent it reaching its target</li> <li>-Under antibiotic pressure, mutations confer an advantage to the bacteria carrying them, leading to the flourishing of the mutated bacteria in the population</li> </ul> <u><i>Human body perspective:</i></u> <ul style="list-style-type: none"> <li>-Broad-spectrum antibiotics weaken the intestinal microbiota</li> <li>-Populations of <i>Clostridium difficile</i> and other non-susceptible microbes rise, causing antibiotic diarrhoea</li> <li>-Probiotics help to recover the normal microbiota</li> <li>-Probiotics mediate their action via their interactions with other microbes and the gut epithelium</li> </ul>	

Task (B) measured the student's understanding of the underlying scientific principles regarding the phenomenon, and the ability to apply and explain the causalities involved. Thus, success in the case task required a clear and profound understanding of the phenomenon on several conceptual levels. Misconceptions of these participants related to the case task have been published in a previous article (Södervik et al., 2020). This particular topic was chosen because antibiotics are commonly used to treat a range of diseases and resistance to antibiotics is a globally growing problem that presents real threats to public health and costs due to failure to treat and prevent infectious diseases (WHO, 2015). Thus, a comprehensive understanding of the safe use of antibiotics and their biomedical mechanisms is necessary for patient counselling. It was also anticipated that the topic would already be familiar to the students at the start of university. However, recent studies have shown severe problems and misconceptions in pharmacy students' understanding of elementary-level knowledge regarding antibiotics and antibiotic resistance mechanisms, and the factors that promote the emergence and spread of antibiotic resistance (Inacio et al., 2017; Södervik et al., 2020).

### *Measures of Epistemic Cognition and Beliefs regarding Medication*

The participants' epistemic cognition (EC) was studied through a 5-point Likert scale measure. We adapted the epistemic cognition items from a topic-specific measure of climate change from Bråten and his colleagues (2009). However, to capture the original idea of items, certain modifications were made to tap on the phenomenon (see Hofer, 2000; Merk et al., 2018). Additionally, items regarding students' beliefs of medication were included in the questionnaire. The participants were given the following written instruction before answering the questionnaire: 'Please answer the following questions according to your current ideas. There is no right or wrong answers. All items are presented as an Appendix.

## **Data Analysis**

### *Analysis of Conceptual Understanding*

For the MCQs, the frequencies of correct answers were calculated and summated. The answers to the case task were analysed and scored by three researchers, one of whom is a biologist and two of whom are experts in pharmacy (see Table 1). Task A was scored by identifying and defining relevant concepts for the case (one score for each concept/definition, max. eight points) and question B based on the level and quality of the explanation of the underlying life science-related phenomena (max. eight points). An inter-rater analysis was performed with 15% of the data, and correlation was found to be  $r = .82$  for the scoring of question A and  $r = .91$  for the scoring of question B.

### *Analysis of Epistemic Cognition and Beliefs regarding Medication*

Reliability analysis and principal components analysis were conducted using software IBM SPSS Statistics 27. First, the internal consistency of all 33 Likert scale items was evaluated with Cronbach alpha, and the item-total correlations were computed. The Cronbach alpha was .57 and six of the items had negative or low (less than .10) item-total correlations. These six items were left out of further analysis. With the remaining 27 items, the reliability analysis produced an alpha-coefficient of .70, indicating the inter-related dimensions of the epistemic beliefs. Next, principal component analyses were performed on the 27 EC items. The oblique rotation method was used (Promax with Kaiser Normalization), which allows calculation of the correlations between the components. The KMO measure showed that the sampling adequacy was at an acceptable level for PCA (KMO = .63). The number of components needed was first examined using the eigenvalues of the principal components. Nine of the components had an eigenvalue greater than 1.00 (Kaiser's criteria), with values ranging 4.11 to 1.09. Visual inspections of the Screen plot graph for the eigenvalues showed that the decrease of the eigenvalues flattened out slightly after three components and then after the fifth component (Cattell's criterion). Next, the five-component solution was extracted, which also corresponds to the number of the originally formulated dimensions.

However, the estimated solution with four main epistemic cognition components turned out to be unclear: many of the items loaded into the several components, with uniform and low loadings. In particular, the items of the Simplicity and Source dimensions did not perform consistently enough for the dimensions to be distinguishable in the solution. It is notable that, even in certain previous studies, it was somewhat problematic to distinguish the original factor structure and changes to the structure have been necessary (Bråten et al., 2009; Hofer, 2000;

Vančugovienė et al., 2024). Based on Hofer and Pintrich's original theory (1997), the measure consists of two larger entities, knowledge and knowing, and therefore the principal component analysis was next carried out with the number of components forced accordingly (two discipline-specific EC components plus one medication beliefs component). In this solution, there were five items in which the loadings were low (less than .30) or which loaded equally onto more than one component. These items were eliminated, and the solution was extracted again. With the remaining 22 items, the solution seemed clear and interpretable: all items loaded with a higher loading (greater than .39) into only one of the components, and the grouping of the items into the components was such that the components could be labelled as Justification, Certainty and Anti-medication beliefs. This solution explained 36.44 % of the total variance of the sample and the eigenvalues of the components were 3.66 for Justification, for 1.79 for Certainty and 2.57 for Anti-medication beliefs.

In sum, in the procedure described above, eleven of the 33 items were left out. The final dimensions and loadings, together with the items, are presented as an Appendix. Internal consistency measured with the Cronbach alpha for the three dimensions Justification, Certainty and Anti-medication beliefs were .73, .63 and .64 respectively. Based on this structure, the sum score variables were computed, and they were used in the subsequent analyses.

### Statistical Analyses

More statistical analyses were performed using the structural equation modelling framework with Mplus 8.1 software. Epistemic cognition items and topic-specific beliefs together with learning and academic performance were analysed using path-modelling. Univariate distributions of the variables used and missing data were first analysed, and afterwards the hypothesised path-model was fitted to the data. In the model estimations, full information maximum likelihood estimation with robust standard errors were used, and the goodness of fit of the model was evaluated using  $\chi^2$  test, comparative fit index (CFI), Tucker-Lewis index, RMSEA and SRMR. Model-fit evaluation was based on the recommendations by Hu & Bentler (1999).

## Results

### Students' Study Progress, Conceptual Understanding, Beliefs regarding Medication and Epistemic Cognition

During their first study year, the participants had completed 44.06 ECTS on average (Max: 67.00; Min: 18.00) by the end of their first academic study year (Table 2). In the pre-test, students received scores of approximately 6.67 out of 13.00 (Max: 12.00; Min: 1.00) from the MCQs measuring basic knowledge of biochemistry, cell biology and molecular biology, within a level that should have been mastered by the end of Finnish high school.

Table 2. Univariate descriptive statistics and correlations of the variables used in the path analysis model

	1.	2.	3.	4.	5.	6.	7.	8.	9.
Correlations									
1. EC: Justification	1.00								
2. EC: Certainty	.22	1.00							
3. Anti-medication	-.26	.06	1.00						
4. Entrance Exam	.23	.01	-.17	1.00					
5. Case task Pre	.30	.04	-.01	.17	1.00				
6. MCQ Pre	.27	.21	-.06	.23	.48	1.00			
7. Case task Post	.18	.07	-.37	.11	.45	.28	1.00		
8. MCQ Post	.13	.10	-.10	.23	.25	.54	.43	1.00	
9. Academic achievement	.15	.16	-.24	.19	.23	.29	.31	.41	1.00
Descriptive Statistics									
<i>M</i>	4.24	3.67	1.96	16.28	3.08	6.67	5.95	9.96	44.06
<i>SD</i>	.43	.55	.49	4.31	2.14	2.55	2.37	2.03	8.85
Skewness	-.35	-.09	.51	.42	-.01	.57	.14	-1.50	-.60
Kurtosis	-.60	-.44	-.31	-.43	.01	-.79	-.70	3.80	.25

Note. EC: Epistemic Cognition, sumscore variables based on the principal components in PCA. Case task: Open ended case task related to antibiotic resistance. MCQ: Test score based on multiple choice questions related to biochemistry, cell biology, molecular biology and immunology. Correlations greater or equal than |.18|, |.23| and |.30| are significant at levels  $p < .05$ ,  $p < .01$  and  $p < .001$ , respectively.

In the post-test, students received an average score of 9.96 out of 13.00 (Max: 13.00; Min: 1.00); hence, students' scores improved significantly from the pre-test to post-test [ $t(80) = -12.085, p < .001, d = 1.34$ ]. Scores on the case task in the pre-test were low, on average 3.79 out of 16.00 (Max: 10.00; Min: .00; SD: 2.15) and in the post-test on average 5.95 out of 16.00 (Max: 11.00; Min: 2.00). Thus, scores related to the case task improved significantly [ $t(63) = -10.00, p < .001, d = 1.25$ ]. The scores from the multiple-choice questions correlated with the case task scores in the pre-test ( $r = .48, p < .001$ ) and in the post-test ( $r = .34, p = .003$ ). The pre-test MCQ scores / case task scores correlated with the post-test MCQ scores / case task scores ( $r = .51, p < .001$  and  $r = .41, p = .001$  respectively).

Students' scores related to sumscales of Justification of knowledge were quite high, averaging 4.24 (Max: 5.00, Min: 3.13), scores related to Certainty averaged 3.67 (Max: 5.00; Min: 2.43), whereas scores related to Anti-medication beliefs were relatively low, on average 1.96 (Max: 3.43; Min: 1.14). Sumscales of Justification and Certainty correlated strongly with each other, and both correlated with MCQ pretest scores. Justification scores also correlated positively with variables measuring general academic success (Table 2). On the other hand, Anti-medication beliefs correlated negatively with several variables: Justification, post-test case task scores and academic study achievement measured via accumulated number of study credits.

### Epistemic Cognition and Beliefs regarding Medication in relation to Academic Performance

The interrelations of epistemic cognition, beliefs regarding medication and students' academic performance were studied using a longitudinal path-analysis model (Figure 1). At first, the distributions of the variables used were examined. According to visual evaluation (histograms and box-plots) and descriptive statistics (skewness and kurtosis), the distributional properties of the variables were within reasonable limits for the structural equation modelling (Curran, West & Finch, 1996). The missing data (assuming missing at random, MAR) was handled using the full information maximum likelihood method (MLR in Mplus) in the model estimations. The goodness-of-fit of the final path-model was good:  $\chi^2(6) = 7.68; p = .263$ , CFI = .98, TLI = .91, RMSEA = .05 (90% CI: .00 - .13), SRMR = .03. In the model, Justification was significantly related to the pre-test of the case task and the MSQs. Other EC variables were not linked to the pre-test variables. EC variables were not significantly connected to the entrance exam scores either. Anti-medication beliefs significantly predicted success in the case task in the post-test. Effects from the pre-test to post-test were also significant in both the case test and the MCQs, with estimates ( $\beta = .43, p = .001$  and  $\beta = .52, p < .001$ ) indicating moderate stability in the test scores. Academic performance was associated significantly with the MCQs. Indirect effects from EC variables to the post-tests of the case task and MSQ and academic performance were significant from Justification to MCQ post-test (total indirect effect .15;  $p = .025$ ), from Justification to the case task post-test (total indirect effect .14;  $p = .039$ ) and from Justification to Academic performance via pre- and post-test variables (total indirect effect .06;  $p = .043$ ).

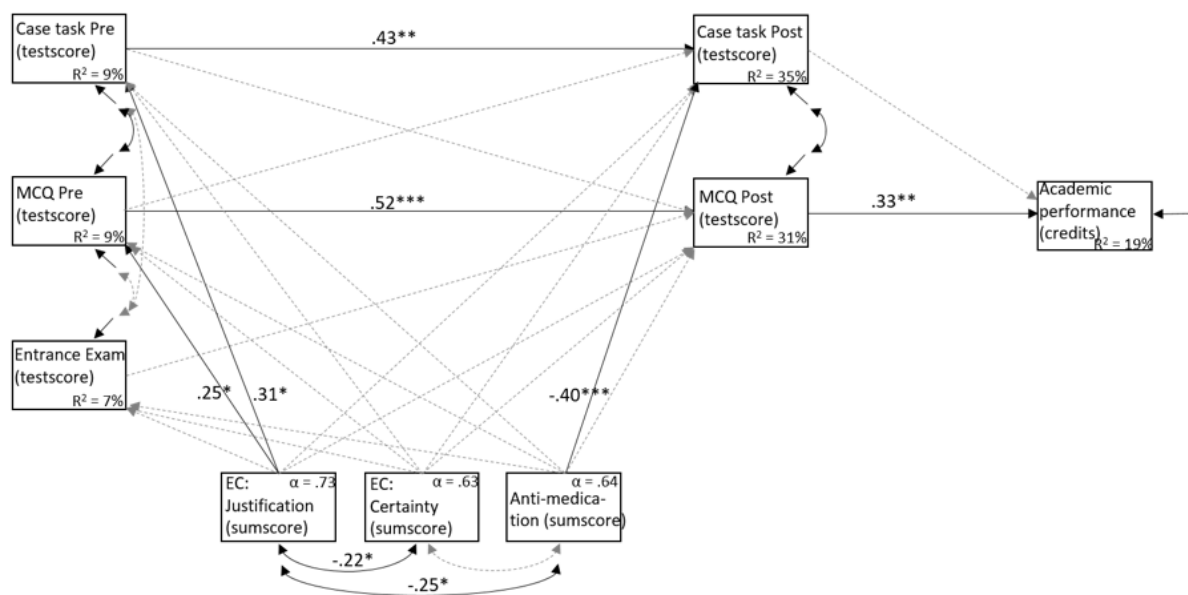


Figure 1. Path analysis model of epistemic cognition and medication beliefs predicting academic performance, with mediating effects of domain-specific conceptual understanding and entrance exam

## Discussion

The first academic study year is largely recognised as laying the foundation for students' academic performance, and understanding the factors that promote or hinder students' academic success is of wide interest. Our study focused on conceptual and epistemic understanding related to pharmaceutical knowledge, to be precise, on antibiotic resistance and medication together with underlying biological knowledge. Therefore, the aim of this study was to add understanding of the development of university students' conceptual understanding during their first study year in relation to their beliefs, epistemic cognition and study progress from a domain-specific perspective. The results of this study revealed relatively serious challenges in students' conceptual understanding concerning elementary level knowledge. This is a worrying result considering that the multiple-choice questions were designed at a level that should have been mastered by the time of completing Finnish high school and challenges in basic conceptual understanding may have serious consequences during subsequent study (Södervik et al., 2020). However, the results of our study state that early identification of students with a higher risk of experiencing study delays or even dropping out is possible by using a relatively light prior knowledge test. These students would benefit from early support such as educational remediation. All in all, adequate first-year support should be provided, when necessary, because in the worst cases, delays may lead to dropping out from undergraduate education (Araujo et al., 2019).

We were also interested in students' underlying epistemic cognition of pharmaceutical knowledge and topic-specific beliefs on medication as such, and in relation to their conceptual understanding. The role of epistemic cognition in learning has been investigated extensively and it is suggested that it could be related to how learners integrate new knowledge with prior knowledge (Jacobson & Spiro, 1995). Therefore, a domain-specific epistemic cognition and beliefs measure was designed and tested. A new questionnaire was designed based on the assumption that pharmaceutical knowledge of medication, including antibiotics and vaccines, involves controversial and even disputed issues (Sinatra et al., 2014) and hence, measures that are more general might have failed to capture the phenomenon (Muis et al., 2006).

Our results showed that more sophisticated beliefs, particularly those related to justification of pharmaceutical knowledge, were related to general academic success, which was seen as a better score in the multiple choice questions and both in the baseline and in the follow-up measurement points in the case task and as a higher accumulated number of study credits after the first study year. This result is in line with previous findings highlighting the importance of justification in relation to academic achievement (for a review, see Cartiff et al., 2020). Conversely, a higher anti-medication beliefs score correlated with a lower accumulated number of study credits, less sophisticated beliefs related to justification of pharmaceutical knowledge, and weaker success in the problem-solving case task in the follow-up. Unscientific beliefs, such as anti-medication, presumably result in higher resistance towards conceptual change and reaching an adequate conceptual understanding among students, and further studies in higher education are needed. The results indicate that understanding how domain-specific knowledge claims can be warranted, how individuals evaluate the use of evidence, and how they assess the authority and expertise of science (Sinatra et al., 2014, 129) is an important factor influencing academic performance during the first study year (see also Bråten & Strømsø, 2009).

## Limitations of the Study

This study is constrained by certain factors that need to be considered when the results outlined are interpreted and generalised. First, our study was conducted at one university, among one cohort of students, resulting in a quite small sample size. Additionally, the sample size was regrettably small in the post-test, which is typical problem in the studies with a follow-up design. Thus, it would be important to repeat the study with a larger sample in other contexts, as well as adapting the measures to other central topics. Additionally, the loadings of epistemic cognition items did not fully follow the theoretical assumption and we were able to distinguish only two dimensions instead of four. Several other previous studies have also found inconsistent results regarding earlier versions of epistemic cognition measures and have not been able to identify the four dimensions in Hofer and Pintrich's (1997) theoretical framework (see e.g. Bråten et al., 2009; Vančugovienė et al., 2024).

## Conclusion

Our study showed that it may be possible to identify students at higher risk of delayed progress through their study with a relatively light prior knowledge test. Furthermore, our study adds to the empirical body of research, which states that domain-specific epistemic cognition and unscientific beliefs play a role in students' learning

and study progress. Although this study was undertaken in pharmacy, the findings contribute to the larger-scale discussion about discipline-specific conceptual and epistemic understanding in higher education. The rapidly expanding amount of scientific knowledge and the continuous changes in work-related matters and society mean that we do not know the specific set of skills and knowledge today that will be necessary for future experts to succeed in the decades to come. However, we do know that students need support in learning to seek, evaluate and use scientific knowledge effectively and critically, and update their personal professional knowledge accordingly, to become adaptive experts with adaptive epistemic cognition.

## Recommendations

To conclude, on one hand, students should understand that scientific explanations — such as theories and hypotheses — are subject to critical evaluation by the scientific community, in which explanations can be revised. At the same time, they also need to be able to discriminate, to know which findings have been well substantiated (Hofer, 2018). This requires students to understand the nature, foundations and principles underlying scientific research and the empirical grounds for acceptance. Thus, to improve the understanding of science, the students need to reconstruct both their conceptual understanding and their epistemic cognitive processes (Sinatra & Chinn, 2011). Hence, students' conceptual understanding requires not only the restructuring of students' alternative conceptions but also the restructuring of their modes of learning and reasoning (Vosniadou, 2013). For that, support from instruction is needed: university teachers should be aware that their students might hold unscientific beliefs, such as anti-medication beliefs among pharmacy students, that can have harmful consequences for students' learning and general study success — an aspect that may surprise many instructors, who may expect that beliefs of this type do not exist in their study programme. According to previous studies, even short interventions may be beneficial for supporting epistemic change (Cartiff et al., 2020). To be precise, instruction at universities should work towards building not only students' conceptual understanding, but also their epistemic competence, by supporting students to evaluate the source and worthiness of evidence (Hofer, 2016; Hofer, 2018; Kienhues, Jucks & Bromme, 2020; Murphy & Alexander, 2016).

## Scientific Ethics Declaration

The authors declare that the scientific ethical and legal responsibility of this article published in JESEH journal belongs to the authors.

## Acknowledgements

This work was supported by the University of Helsinki via the project of Cultivating Expertise in Learning of Life Sciences, CELLS (Research Funds of the University of Helsinki, HY/716/05.01.07/2018).

## References

- Abendroth, J., & Richter, T. (2021). How to understand what you don't believe: Metacognitive training prevents belief-biases in multiple text comprehension. *Learning and Instruction*, 71, 101394. <https://doi.org/10.1016/j.learninstruc.2020.101394>
- Amin, T.G., & Levrini, O. (2018). *Converging perspectives on conceptual change*. Routledge.
- Araujo, A. M., Leite, C., Costa, P. & Costa, M. J. (2019). Early identification of first year students at risk of dropping out of high school entry medical school: the usefulness of teachers' ratings of class participation. *Advances in Health Sciences Education*, 24 (2), 251–268. <https://doi.org/10.1007/s10459-018-9863-9>
- Binder, T., Sandmann, A., Sures, B., Friege, G., Theyssen, H. & Schmiemann, P. (2019) Assessing prior knowledge types as predictors of academic achievement in the introductory phase of biology and physics study programmes using logistic regression. *International Journal of STEM Education*, 6(1), 1–14. <https://doi.org/10.1007/s10459-018-9863-9>
- Boshuizen, H. P. A., Gruber, H. & Strasser, J. (2020). Knowledge restructuring through case processing: The key to generalise expertise development theory across domains? *Educational Research Review*, 29, Article 100310. <https://doi.org/10.1186/s40594-019-0189-9>

- Boshuizen, H. P. A., & Schmidt, H. G. (2018). The development of clinical reasoning expertise. In J. Higgs, G. Jensen, S. Loftus, & N. Christensen (Eds.), *Clinical reasoning in the health professions*, 4th ed. (pp. 57–65). Elsevier.
- Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education*, 24, 61–100. <https://doi.org/10.2307/1167267>
- Bransford, J. D., Brown, A. L. & Cocking, R. R. (2000). *How people learn: Brain, mind, experience and school*: Expanded Edition of Sciences. National Academy Press.
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32–41. <https://doi.org/10.4324/9780203990247-34>
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). Guilford Press.
- Bråten, I. & Strømsø, H.I. (2009). Effects of task instruction and personal epistemology on the understanding of multiple texts about climate change. *Discourse Processes*, 47(1), 1–31. <https://doi.org/10.1080/01638530902959646>
- Bråten, I., Strømsø, H. I., & Samuelstuen, M. S. (2008). Are sophisticated students always better? The role of topic-specific personal epistemology in the understanding of multiple expository texts. *Contemporary Educational Psychology*, 33, 814–840. <https://doi.org/10.1016/j.cedpsych.2008.02.001>
- Bråten, I., Gill, L., Strømsø, H. I. & Vidal-Abarca, E. (2009). Personal epistemology across cultures: exploring Norwegian and Spanish university students' epistemic beliefs about climate change. *Social Psychology Education*, (12), 529–560. <https://doi.org/10.1007/s11218-009-9097-z>
- Carpenter, T., Franke, M., Jacobs, V., Fennema, E., & Empson, S. (1998). A longitudinal study of invention and understanding in children's multidigit addition and subtraction. *Journal for Research in Mathematics Education*, 29(1), 3–20. <https://doi.org/10.2307/749715>
- Cartiff, B. M., Duke, R. F. & Greene, J.A. (2020). The effect of epistemic cognition interventions on academic achievement: A meta-analysis. *Journal of Educational Psychology*. <http://dx.doi.org/10.1037/edu0000490>
- Chi, M. T. H. (2013). Two kinds and four sub-types of misconceived knowledge, ways to change it, and the learning outcomes. In S. Vosniadou (Ed.), *International handbook of research on conceptual change*, (2nd ed., pp. 49–70). Routledge.
- Constantini, P., & M. Vitale. 2011. “Analyzing undergraduate student graduation delay: A longitudinal perspective.” In: M. Attanasio and C. Vincenza (Eds.), *Statistical methods for the evaluation of university systems* (pp. 145–159). Springer.
- Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, 1(1), 16–29. <https://doi.org/10.1037/1082-989X.1.1.16>
- European Education and Culture Executive Agency: Eurydice and Crosier, D., (2017). *Modernisation of higher education in Europe – Academic staff Publications Office of the European Union*, <https://data.europa.eu/doi/10.2797/806308>
- Flaig, M., Simonsmeier, B. A., Mayer, A.-K., Rosman, T., Gorges, J., & Schneider, M. (2018). Conceptual change and knowledge integration as learning processes in higher education: A latent transition analysis. *Learning and Individual Differences*, 62, 49–61. <https://doi.org/10.1016/j.lindif.2018.07.001>
- Greene, J. A., Sandoval, W.A. & Bråten, I. (2016). An introduction to epistemic cognition. In J.A. Greene, W. A. Sandoval & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 1–15). Routledge.
- Greene, J. A., Cartiff, B. M., & Duke, R. F. (2018). A meta-analytic review of the relationship between epistemic cognition and academic achievement. *Journal of Educational Psychology*, 110(8), 1084–1111. <https://doi.org/10.1037/edu0000263>
- Hofer, B. K. (2000). Dimensionality and disciplinary differences in personal epistemology. *Contemporary Educational Psychology*, 25(4), 378–405. <https://doi.org/10.1006/ceps.1999.1026>
- Hofer, B. K. (2006). Domain specificity of personal epistemology: Resolved questions, persistent issues, new models. *International Journal of Educational Research*, 45, (1), 85–95. <https://doi.org/10.1016/j.ijer.2006.08.006>
- Hofer, B. K. (2016). Epistemic cognition as a psychological construct: Advancements and challenges. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 19–38). Routledge.
- Hofer, B. K. (2018). Identifying the role of epistemic cognition and metacognition in conceptual change. In T. G. Amin & O. Levrini (Eds.): *Converging perspectives on conceptual change – Mapping an emerging paradigm in the learning sciences* (pp. 229–236). Routledge.
- Hofer, B. K. & Pintrich, P. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67, 88–140. <https://doi.org/10.3102/00346543067001088>

- Hu, L.-T. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Inácio, J., Barnes, L.M., Jeffs, S., Castanheira, P., Wiseman, M., Inacio S., Bowler, L., Lansley, A. (2017). Master of pharmacy students' knowledge and awareness of antibiotic use, resistance and stewardship. *Currents in Pharmacy Teaching and Learning*, 9(4), 551–559. <https://doi.org/10.1016/j.cptl.2017.03.021>
- Jacobson, M. J., & Spiro, R. J. (1995). Hypertext learning environments, cognitive flexibility, and the transfer of complex knowledge: An empirical investigation. *Journal of Educational Computing Research*, 12, 301–333.
- Jenert, T., Postareff, L., Brahm, T., & Lindblom-Ylänne, S. (2015). Editorial: enculturation and development of beginning students. *Zeitschrift Für Hochschulentwicklung*, 10(4), 9–21. <https://doi.org/10.3217/zfhe-10-04/01>
- Jia, P. & Maloney, T. (2015). Using predictive modelling to identify students at risk of poor university outcomes. *Higher Education*, 70, 127–149. <https://doi.org/10.1007/s10734-014-9829-7>
- Kaufman, D. R., Keselman, A., & Patel, V. (2013). Conceptual understanding in the domain of health. In: S. Vosniadou (Ed.), *International handbook on conceptual change research: second edition* (pp. 240–252). Routledge.
- Kienhues, D. Jucks, R. & Bromme, R. (2020). Sealing the gateways for post-truthism: Re-establishing the epistemic authority of science, *Educational Psychologist*, 55(3), 144–154. <https://doi.org/10.1080/00461520.2020.1784012>
- Koedinger, K. R., & Roll, I. (2012). Learning to think: Cognitive mechanisms of knowledge transfer. In: K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford handbook of thinking and reasoning* (pp. 789 – 806). Oxford University Press.
- Kolstø, S. D. (2001). Scientific literacy for citizenship: Tools for dealing with the science dimension of controversial socioscientific issues. *Science Education*, 85(3), 291–310. <https://doi.org/10.1002/sce.1011>
- Korhonen, V. & Rautopuro, J. (2019): Identifying problematic study progression and “at-risk” students in higher education in Finland. *Scandinavian Journal of Educational Research*, 63(7), 1056–1069. <https://doi.org/10.1080/00313831.2018.1476407>
- Kuhn, D. (2005). *Education for thinking*. Harvard University Press.
- Kuncel, N. R., Hazlett, S. A., & Ones, D. R. (2001). A comprehensive meta-analysis of the predictive validity of the Graduate Record Examinations: Implications for graduate student selection and performance. *Psychological Bulletin*, 127, 162–181. <https://doi.org/10.1037/0033-2909.127.1.162>
- Mason, L., Gava, M., & Boldrin, A. (2008). On warm conceptual change: The interplay of text, epistemological beliefs, and topic interest. *Journal of Educational Psychology*, 100(2), 291–309. <https://doi.org/10.1037/0022-0663.100.2.291>
- Mennen, J. & van der Klink, M. (2017). Is the first-year predictive for study success in subsequent years? Findings from a academy of music. *Music Education Research* 19(3), 339–351. <https://doi.org/10.1080/14613808.2016.1189524>
- Merk, S., Rosman, T., Muis, K. R., Kelava, A. & Bohl, T. (2018). Topic specific epistemic beliefs: Extending the theory of integrated domain in personal epistemology. *Learning and Instruction*, 56, 84–97. <https://doi.org/10.1016/j.learninstruc.2018.04.008>
- Muis, K. R., Bendixen, L. D. & Haerle, F. C. (2006). Domain generality and domain specificity in personal epistemology research: Philosophical and empirical reflections in the development of a theoretical model. *Educational Psychology Review*, 18, 3–54. <https://doi.org/10.1007/s10648-006-9003-6>
- Murphy, P. K. & Alexander, P. A. (2013). Situating text, talk, and transfer in conceptual change – concluding thoughts. In: S. Vosniadou (Ed.), *International handbook on conceptual change research: 2nd edition* (pp. 603–621). Routledge.
- Murphy, P. K. & Alexander, P. A. (2016). Interrogating the relation between conceptual change and epistemic beliefs. In J. A. Greene, W. A. Sandoval & I. Bråten (Eds.) *Handbook of Epistemic Cognition* (pp. 439–459). Routledge.
- Persky, A. & Murphy, K (2019). Investigating whether transfer of learning in pharmacy students depends more on knowledge storage or accessibility. *American Journal of Pharmaceutical Education*, 83 (6) 6809. <https://doi.org/10.5688/ajpe6809>
- Rittle-Johnson, B., Star, J. R., & Durkin, K. (2009). The importance of prior knowledge when comparing examples: Influences on conceptual and procedural knowledge of equation solving. *Journal of Educational Psychology*, 101(4), 836–852. <https://doi.org/10.1037/a0016026>

- Schommer-Aikins, M., & Easter, M. (2006). Ways of knowing and epistemological beliefs: Combined effect on academic performance. *Educational Psychology*, 26, 411–423. <https://doi.org/10.1080/01443410500341304>
- Sinatra, G., Kienhues, D. & Hofer, B. K. (2014). Addressing challenges to public understanding of science: epistemic cognition, motivated reasoning, and conceptual change. *Educational Psychologist*, 49 (2), 123–138. <https://doi.org/10.1080/00461520.2014.916216>
- Sinatra, G. & Chinn, C. A. (2011). Thinking and reasoning in science: Promoting epistemic conceptual change. In: K. Harris, C. B. McCormick, G. M. Sinatra, & J. Sweller (Eds.), *Critical theories and models of learning and development relevant to learning and teaching*, (pp. 257–282). APA Publisher.
- Södervik, I., Hanski, L., Boshuizen, H. P. A., & Katajavuori, N. (2023). Clinical reasoning in pharmacy: what do eye movements and verbal protocols tell us about the processing of a case task? *Advances in Health Sciences Education*, 29, 45–65. <https://doi.org/10.1007/s10459-023-10242-x>
- Södervik, I., Hanski, L., & Katajavuori, N. (2020). First-year pharmacy students' prior knowledge correlates with study progress and reveals different dynamics of misconceptions. *Pharmacy Education*, 20(1), 94–102. <https://doi.org/10.46542/pe.2020.201.94102>
- Södervik, I., Vilppu, H., Mikkilä-Erdmann, M. & Österholm, E. (2017). Medical students' biomedical and clinical knowledge: combining longitudinal design, eye tracking and comparison with residents' performance. *Learning and Instruction*, 52, 139–147. <https://doi.org/10.1016/j.learninstruc.2017.06.002>
- Södervik, I., Virtanen, V., & Mikkilä-Erdmann, M. (2015). Challenges in understanding photosynthesis in a university introductory biosciences class. *International Journal of Science and Mathematics Education*, 13(4), 733–750. <https://doi.org/10.1007/s10763-014-9571-8>
- Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition*. University of Chicago Press.
- Trautwein, C. & Bosse, E. (2017). The first year in higher education—critical requirements from the student perspective. *Higher Education*, 73, 371–387. <https://doi.org/10.1007/s10734-016-0098-5>
- Vančugovienė, V., Södervik, I., Lehtinen, E. & McMullen, J. (2024). Individual differences in secondary school students' conceptual knowledge: Latent profile analysis of biology concepts. *Learning and Individual Differences*, 111, 102436. <https://doi.org/10.1016/j.lindif.2024.102436>
- Vosniadou, S. (2013). Conceptual change in learning and instruction – The Framework Theory Approach. In: S. Vosniadou (Ed.), *International Handbook on Conceptual Change Research*: 2nd edition (pp. 11–30). Routledge.
- WHO (2015). *Global action plan on antimicrobial resistance*. Accessed: 8.2.2024. [https://apps.who.int/iris/bitstream/handle/10665/193736/9789241509763\\_eng.pdf?sequence=1&angeli](https://apps.who.int/iris/bitstream/handle/10665/193736/9789241509763_eng.pdf?sequence=1&angeli)

---

### Authors Information

---

**Iiona Södervik**

Centre for University Teaching and Learning,  
Siltavuorenpenger 5B, 00014,  
University of Helsinki, Finland  
Contact e-mail: [iiona.sodervik@helsinki.fi](mailto:iiona.sodervik@helsinki.fi)  
ORCID iD: <https://orcid.org/0000-0001-8575-487X>

**Leena Hanski**

Faculty of Pharmacy  
Viikinkaari 5E, 00014,  
University of Helsinki, Finland  
ORCID iD: <https://orcid.org/0000-0002-3121-8542>

**Eero Laakkonen**

Department of Teacher Education,  
Assistentinkatu 5, 20014  
University of Turku, Finland  
ORCID iD: <https://orcid.org/0009-0000-6763-5877>

**Nina Katajavuori**

Centre for University Teaching and Learning,  
Viikinkaari 9, 00140,  
University of Helsinki, Finland  
ORCID iD: <https://orcid.org/0000-0001-5093-2402>

---